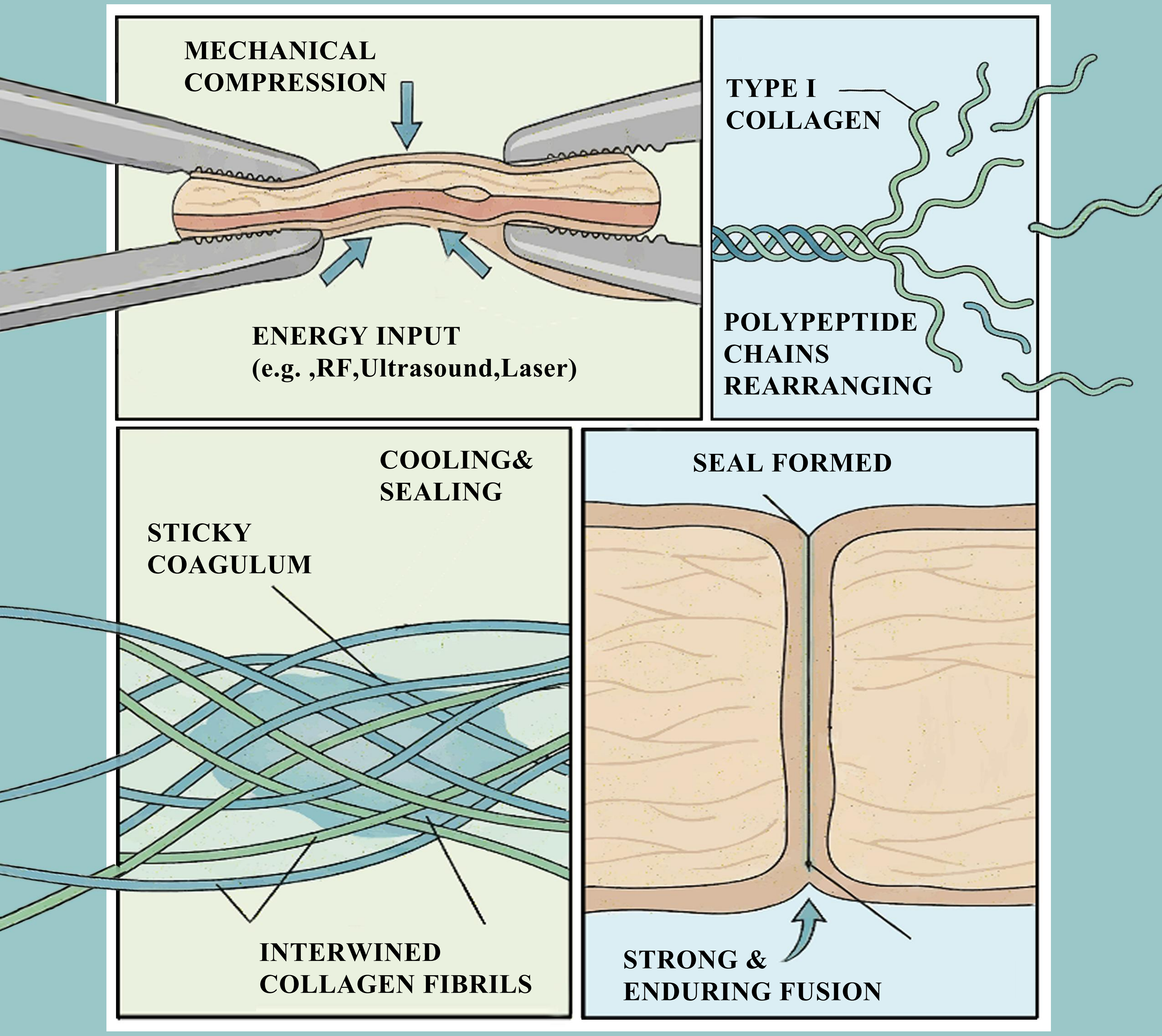
Zentime
SITIA
Open Access

PROGRESS IN MEDICAL DEVICES

ISSN: 2957-5478
Editor-in-Chief: Haipo Cui



Volume 4 , Issue 2 · 2026
Published by ZENTIME PUBLISHING CORPORATION

Progress in
Medical Devices

RESEARCH ARTICLE

Optimization of multilayer shell structures for wearable sensors based on polyvinylidene fluoride

Ke Wang, Rongguo Yan, Wenjing Du, Shoucheng Chen

School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Rongguo Yan.

Address correspondence to: Rongguo Yan, School of Health Science and Engineering, University of Shanghai for Science and Technology, No. 516, Jun Gong Road, Yangpu District, Shanghai 200093, China.
E-mail: yanrongguo@usst.edu.cn.

Received November 18, 2025; Accepted January 16, 2026; Published June 3, 2026

DOI: 10.61189/931077ergknd

Abstract

Objective: Accurately capturing signals from healthcare devices worn on the body is vital for health monitoring, a field of growing interest. A flexible polyvinylidene fluoride (PVDF)-based piezoelectric sensor has been chosen for wearable applications. This study aims to optimize PVDF sensor performance by focusing on structural design. **Methods:** According to piezoelectric theory, this paper presents a mathematical expression system centered on the first-order piezoelectric equation, which describes the relationship between the electric displacement, the quantity of charge, and the output voltage of the PVDF film under strain. The study focuses on the features of weak physiological signals in the human body to establish a multi-layer shell structure finite element model using COMSOL Multiphysics. The analysis covers the electric displacement field mode, output charge, electric potential, and uniformity of stress distribution. We systematically investigated the impact of PVDF piezoelectric layer area, shape, thickness, and encapsulation materials on sensor performance using the controlled variable method. **Results:** As the area increases, the sensitivity and stress uniformity improve. Rectangular shapes exhibit more uniform stress distribution than circular shapes, and raising the aspect ratio within a certain range further enhances stress uniformity while maintaining comparable sensitivity. Increasing the thickness of the piezoelectric layer raises the electrical potential. Flexible encapsulation materials can provide higher sensitivity than rigid materials but are more susceptible to stress concentration. **Conclusion:** This paper finds the best structural form and parameter adjustment range of the PVDF sensor, which can provide a theoretical basis and numerical references for designing high-performance wearable sensors.

Keywords: Polyvinylidene fluoride flexible sensor, Wearable sensing technology, COMSOL simulation, Piezoelectric effect

Highlights

- Established potential displacement-charge-voltage relation for polyvinylidene fluoride by means of the first-order piezoelectric equation.
- Applied COMSOL multilayer shell model and multiphysics coupling for calculating the interlayer stress and electric displacement field.
- Structurally explored how changes to structural parameters impacted sensor performance through the control variable method.

1 INTRODUCTION

Due to the increasing social demand for daily health monitoring, wearable sensing technology is widely recognized and is gradually becoming a major direction of development in medical diagnosis. Among the different types of wearable sensors, piezoelectric, piezoresistive, capacitive, and triboelectric sensors have the best application prospects [1, 2]. Piezoelectric materials have obvious self-powered characteristics and low energy consumption, which can meet the needs of long-term health monitoring, exercise physiology parameter monitoring, and non-invasive monitoring for chronic disease patients [3-5]. Sensors made with piezoelectric material, as shown in **Figure 1A**, can convert mechanical energy into electrical energy via the piezoelectric effect when subjected to mechanical force [6, 7]. Traditional piezoelectric materials mostly use rigid ceramics such as lead zirconate titanate and lead titanate. Though highly sensitive, these devices are fragile, and there is a biocompatibility problem that affects comfort; results can also be inaccurate if they do not adhere properly [8, 9]. To tackle these shortcomings, flexible sensors have become an important direction for development. In 1969, Kawai discovered the piezoelectricity of polyvinylidene fluoride (PVDF), whose polarized film has a piezoelectric coefficient as high as 6-7 pC/N—ten times that of any other polymer [10]. As can be seen from **Figure 1B**, PVDF is becoming a popular material for flexible wearable sensors because of its good piezoelectric property, flexibility, biocompatibility, and cost-effectiveness [11-13].

In recent years, a lot of research has proved that PVDF can be used for wearable applications. Ghosh et al. developed a Pt-PVDF piezoelectric nanogenerator that served as a static strain-sensitive wearable tactile sensor [14]. Yang et al. created a self-powered, breathable, high-performance flexible wearable sensor using an electrospun polarized PVDF-barium titanate membrane and a Ni fabric electrode [15]. The physiological signals of the human body are mostly weak and easily interfered with, so high sensor sensitivity is required. To improve the piezoelectric properties of PVDF, most studies aim at increasing the β-phase content [16]. Electrospinning is a commonly used method. He et al. found that by controlling process, solution, and environmental parameters, PVDF molecular dipoles can be oriented to significantly increase the β-phase fraction and crystallinity [17]. Adding nanofillers can also significantly increase the proportion of β-phase. Parangusan et al. noticed that increasing the amount of Co-ZnO nanofillers in the PVDF-hexafluoropropylene matrix raised the quantity of crystalline β-phase, leading to a dramatic increase in the output voltage of the flexible nanogenerator [18]. Islam et al. used 3D printing to fabricate PVDF-MoS_2 composites and achieved in-situ dipole alignment through shear stress, which increased the piezoelectric coefficient eightfold relative to pure PVDF, correlating with an increased β-phase volume fraction [19]. Structural optimization has also been studied. Liu et al. built a multiscale porous structure, as shown in **Figure 1C**, and combined it with corona polarization to produce an electret effect and improve sensor performance [20]. Xue et al. designed a cylindrical arch structure to optimize PVDF film deformation and enhance the sensor's signal acquisition capacity [21]. Kim et al. developed an air-expansion approach to create dome-shaped PVDF films (**Figure 1D**) and fabricated a dome-shaped 4×4 array tactile sensor, finding that dome-shaped sensors are more sensitive than planar ones [22]. Yan et al. built a basic PVDF simulation model and changed the length, width, and height of the PVDF structure to study the effect of structural parameters on sensor sensitivity, as shown in the stress distribution results in **Figure 1E** [23]. Wang et al. proposed external strategies like electrode design and stress modification, offering new ways to improve PVDF sensor performance [24]. These studies provide practical technical pathways for developing the next generation of comfortable and efficient wearable health monitoring devices.

However, despite some achievements in material composites and structural design in current research, systematic studies on structural optimization bottleneck issues in the wearable field remain insufficient. Therefore, based on the demands of wearable health monitoring, and in accordance with the structural characteristics of PVDF piezoelectric properties, this paper proposes an innovative method: by establishing a multilayer shell finite element model in COMSOL Multiphysics, the impact of sensor area, shape, thickness, and encapsulation material on stress distribution and electric displacement output is investigated. This method explores an optimized sensor structure suitable for wearable applications. Compared with traditional trial-and-error experimentation, it greatly reduces research and development costs and cycle time, and more efficiently identifies the optimal structural configuration. Moreover, it reveals the coupling relationship between piezoelectric effects and structure parameters, providing a theoretical reference for designing high-performance wearable sensors.

2 MATERIALS AND METHODS

2.1 Piezoelectric effect of PVDF film sensors

The finding of the piezoelectric effect is traced back to when the Curie brothers first found an unusual electromechanical phenomenon while working with a quartz crystal. When an electric field is applied to the crystal, it deforms; conversely, when a mechanical load is applied, surface charges are generated. As for the features of PVDF thin-film sensors, their general structure is a layered thin film. The thickness direction (usually the 3-axis direction) is the polarization direction, with electrodes deposited on the top and bottom surfaces. When subjected to a load, charge separation along the polarization direction generates equal but opposite charges on the upper and lower surfaces. This is known as the direct piezoelectric effect. As shown in **Figure 2**, polarization occurs along the thickness

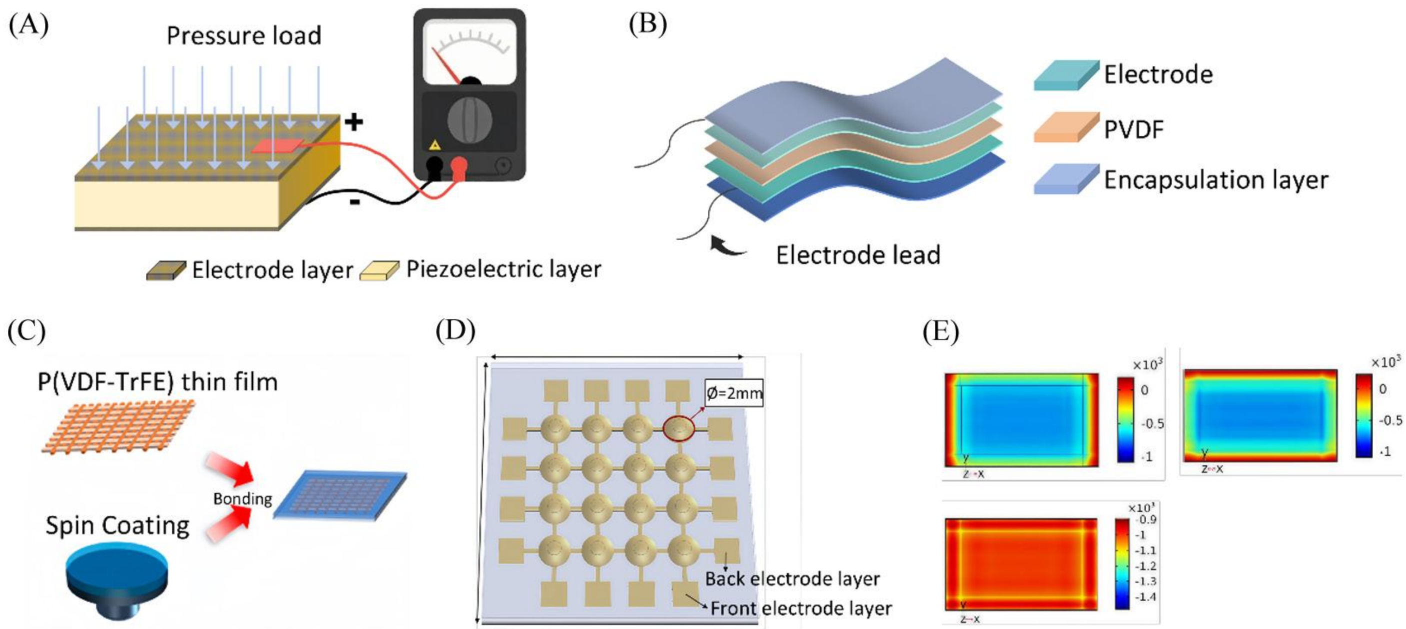


Figure 1. Schematic diagram related to the development of PVDF sensors. (A) Schematic of the piezoelectric sensor principle; (B) Flexural characteristics of PVDF piezoelectric film; (C) Preparation of multi-scale porous P(VDF-TrFE) film; (D) Structure of a 4×4 array tactile sensor; (E) Stress distribution along the length, width, and height of the PVDF film. PVDF, polyvinylidene fluoride; P(VDF-TrFE), poly(vinylidene fluoride-trifluoroethylene).

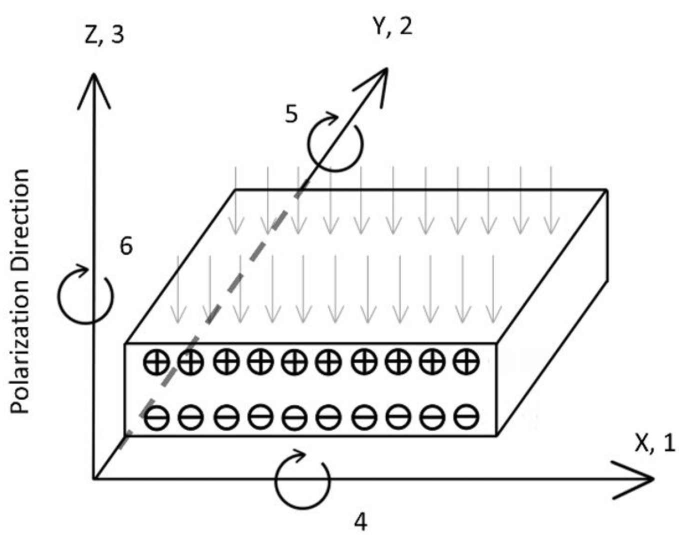


Figure 2. Schematic of the positive piezoelectric effect in PVDF film. PVDF, polyvinylidene fluoride.

direction; applying a load results in positive charge accumulation on the upper surface and negative charge on the lower surface.

To quantitatively characterize this electromechanical conversion mechanism, a comprehensive mathematical framework must be established. Under the assumption of linear material response, this can be expressed via matrix equations, typically employing the first-order piezoelectric equation [25-27]:

$$\begin{bmatrix} D \\ x \end{bmatrix} = \begin{bmatrix} \varepsilon^X & d \\ d^T & s^E \end{bmatrix} \begin{bmatrix} E \\ X \end{bmatrix} \tag{1}$$

Where D (C/m^2) is the electric displacement vector, x is the strain tensor; E (V/m) is the electric field intensity vector, X (Pa) is the stress tensor; ε^X represents the dielectric constant matrix under constant stress, s^E represents the elastic modulus matrix under constant electric field, d denotes the piezoelectric strain coefficient matrix, and d^T is the transpose of d. This equation describes the bidirectional energy conversion mechanism of piezoelectric materials.

When PVDF is applied to thin-film sensors, no external electric field is present. Under such conditions, the general equation (1) reduces to the following form [26]:

$$D = d{:}X \tag{2}$$

Expanding equation (2) using Voigt notation yields [27]:

$$\begin{bmatrix} D_1 \\ D_2 \\ D_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & d_{15} & 0 \\ 0 & 0 & 0 & d_{24} & 0 & 0 \\ d_{31} & d_{32} & d_{33} & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \\ X_3 \\ X_4 \\ X_5 \\ X_6 \end{bmatrix} \tag{3}$$

In the equation, D_1, D_2, D_3 denote the components of the electric displacement vector, describing the charge density distribution along the corresponding axes; X_1, X_2, X_3 denote the normal stress components, characterizing tensile or compressive stress along each principal direction; X_4, X_5, X_6 denote the shear stress components, describing the in-plane shear deformation forces acting on the material.

The thickness of PVDF film sensors is far smaller than their in-plane dimensions, making direct measurement of in-plane displacement components D_1 and D_2 challenging. Conversely, D_3 can be directly acquired via the upper and lower electrodes, representing an effective physical quantity directly correlated with the sensor's output signal. Considering measurement feasibility and practical application requirements, we focus on D_3 and simplify it as [23]:

$$D_3 = d_{31}X_1 + d_{32}X_2 + d_{33}X_3 \tag{4}$$

The surface charge density $\sigma = D_3$ on the electrode plane, and the total accumulated surface charge Q can be expressed as:

$$Q = \int \sigma dA = D_3 A \tag{5}$$

From a circuit standpoint, the PVDF film structure can be seen as a parallel-plate capacitor, whose capacitance C is a function of the dielectric properties and dimensions of the material:

$$C = \frac{\varepsilon_0 \varepsilon_r A}{h} \tag{6}$$

Where ε_0 is the vacuum permittivity, ε_r is the relative permittivity of PVDF, A is the film area, and h is the film thickness. This yields an expression for the output voltage U across the electrodes:

$$U = \frac{Q}{C} = \frac{D_3 h}{\varepsilon_0 \varepsilon_r} \tag{7}$$

According to the piezoelectric mechanism, structural configuration, and mathematical modeling of the PVDF film sensor, a clear relationship among the electrical displacement, accumulated charge, and output voltage has been established.

2.2 COMSOL simulation preparation

2.2.1 Constitutive equation simulation implementation

To convert the above theoretical model into a practical simulation method in COMSOL Multiphysics, we need to specify the form of the governing equations to be implemented. For applications dominated by the direct piezoelectric effect, like PVDF thin-film sensor, the "strain-charge" formulation should be chosen. The fundamental governing system consists of three physical equations [28, 29].

Dynamic equilibrium equation:

$$\rho\left(\frac{\partial^2 u}{\partial t^2}\right) = \nabla \cdot S + F_v \tag{8}$$

Where ρ is the density of PVDF, $\frac{\partial^2 u}{\partial t^2}$ is the acceleration vector, $\nabla \cdot S$ is the divergence of the stress tensor S, and F_v is the volume force vector. This equation describes the mechanical vibration response of the film and serves as the foundation for analyzing deformation due to load.

Electrostatic Gauss's law:

$$\nabla \cdot D = \rho_v \tag{9}$$

Where $\nabla \cdot D$ is the divergence of the electric displacement vector D and ρ_v is the free charge density. This equation governs how the electric displacement field distributes in space with respect to free charges.

Electromechanical coupling constitutive relations:

$$\begin{cases} S = c:\varepsilon - e^T:E \\ D = e:\varepsilon + \varepsilon_0 \varepsilon_r E \end{cases} \tag{10}$$

Where c is the elastic stiffness tensor, e is the piezoelectric stress constant tensor, and ε is the total strain tensor. This is the basic system of equations for electromagnetic coupling. The former refers to the mechanical response "strain → stress", whereas the latter one is the piezoelectric effect "strain → electric displacement". This creates a direct correlation between PVDF film deformation and electrical signal output, thereby transforming mechanical load into an electrical quantity.

According to the derivation and theory of the above piezoelectric equation, the material parameters of PVDF must be defined to establish a successful simulation model in COMSOL. These parameters include density, elastic modulus matrix, piezoelectric strain constant matrix, and free relative permittivity. They govern PVDF's electromechanical response characteristics, and must satisfy the complete tensor description for an orthotropic anisotropic material [28]. The selected parameter values are shown in **Table 1**.

2.2.2 Construction of the multilayer shell model

PVDF sensors typically feature a multilayer planar structure. As shown in **Figure 3**, the configuration comprises five layers: upper and lower encapsulation layers, a PVDF piezoelectric layer, and two intermediate electrode layers. The electrode layers serve to conduct charge [11, 30]. Given that this study focuses on the influence of piezoelectric-layer structural parameters and encapsulation-material selection on sensor performance, and considering that the electrode-layer thickness is extremely

Table 1. Key PVDF material parameters used in the simulation

Density ρ (kg/m^3)	Piezoelectric constant d_{31} (C/N)	Piezoelectric constant d_{32} (C/N)	Piezoelectric constant d_{33} (C/N)	Relative permittivity ε_r
1,780	1.358×10^{-11}	1.476×10^{-12}	-3.380×10^{-11}	7.74

Note: PVDF, polyvinylidene fluoride.

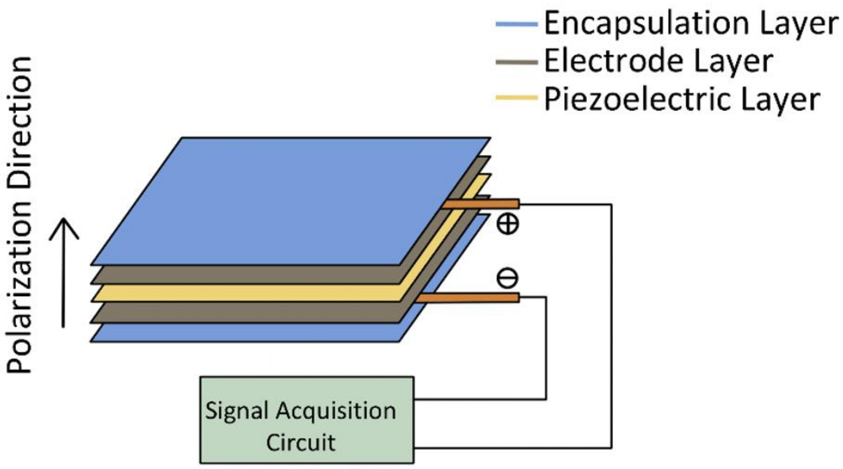


Figure 3. Multilayer structure of the PVDF sensor. PVDF, polyvinylidene fluoride.

Table 2. Fundamental parameters of different encapsulation materials

Material	Density (kg/m^3)	Young's modulus (MPa)	Poisson's ratio
PET	1,400	3,100	0.34
PDMS	970	7.50	0.49
Nylon	1,150	2,000	0.40
PI	1,400	3,100	0.34
PVA	1,300	3,000	0.35

Note: PET, polyethylene terephthalate; PDMS, polydimethylsiloxane; PI, polyimide; PVA, polyvinyl alcohol.

thin (typically 6 μm), its effect on the overall stress distribution can be neglected. The simulation model can thus be simplified to a three-layer structure consisting of the upper and lower encapsulation layers with the PVDF piezoelectric layer in between.

To construct this configuration in the COMSOL environment, a multilayer shell model is employed. The theoretical basis for such modeling primarily includes layered theory and equivalent single-layer theory. Since this study requires accurate calculation of stress and electric displacement in each layer, layered theory is adopted as basic modeling way. Layered theory solves equations in the direction of thickness, making it suitable for the analysis of composite shell structures of medium thickness and above. It uses a solid-like formulation that distributes the degrees of freedom over three displacement components along the thickness direction [31]. This feature allows it to fully assess stress and strain distributions within each layer, correctly forecasting interlaminar stresses and delamination phenomena. For the present study, this approach enables a detailed examination of stress transfer and interactions between the PVDF piezoelectric layer and the encapsulation layers, as well as within the PVDF layer itself. It also provides insight into how each material deforms under loading. Such comprehensive data support the analysis of how the electric displacement field correlates with interlayer stress.

2.2.3 Sensor design parameters and boundary condition settings

In this study, it is found that area, shape, thickness and encapsulation material of the sensor's piezoelectric layer are the most important parameters. Area is chosen between 1 and 25 cm^2. An excessively small area hampers the capture of broad-range signals like breathing and body motion, while an overly large area could cause non-uniform strain because of insufficient bonding space and increase wearing discomfort. Common sensor shapes include circular and rectangular geometries. Several standard PVDF thicknesses are available: 28 μm, 52 μm, and 110 μm, among others. Thinner films (e.g., 28 μm) offer greater comfort, whereas thicker ones (e.g., 110 μm) tend to provide higher sensitivity. Common encapsulation materials are polyethylene terephthalate (PET), polydimethylsiloxane (PDMS), nylon, polyimide (PI), and polyvinyl alcohol (PVA) [1]. Different encapsulation materials can significantly affect sensor performance. The material parameters used for the five encapsulation materials in this study are listed in **Table 2**.

After determining the sensors' parameters and selecting the encapsulation materials, the simulation boundary conditions are established to simulate real-world conditions. A fixed constraint is applied to the bottom surface to simulate skin contact, restricting both translational and rotational movement. A constant pressure of 1,000 Pa is applied on the top surface, corresponding approximately to the minimum pressure detectable by human fingers [32].

2.3 COMSOL simulation workflow

After the above preparations are completed, the model can be built. First, select "Geometry" and draw rectangles or circles to make up a composite structure. Under "Materials", add multilayer material. Assign PVDF to the middle layer and the

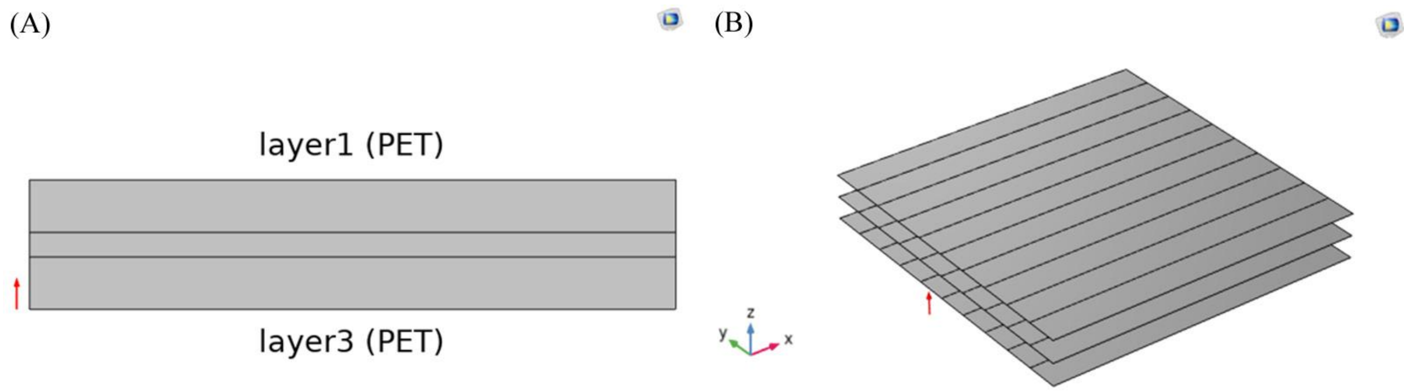


Figure 4. Multilayer shell structure preview. (A) Layer cross-section; (B) Layer stacking. PET, polyethylene terephthalate.

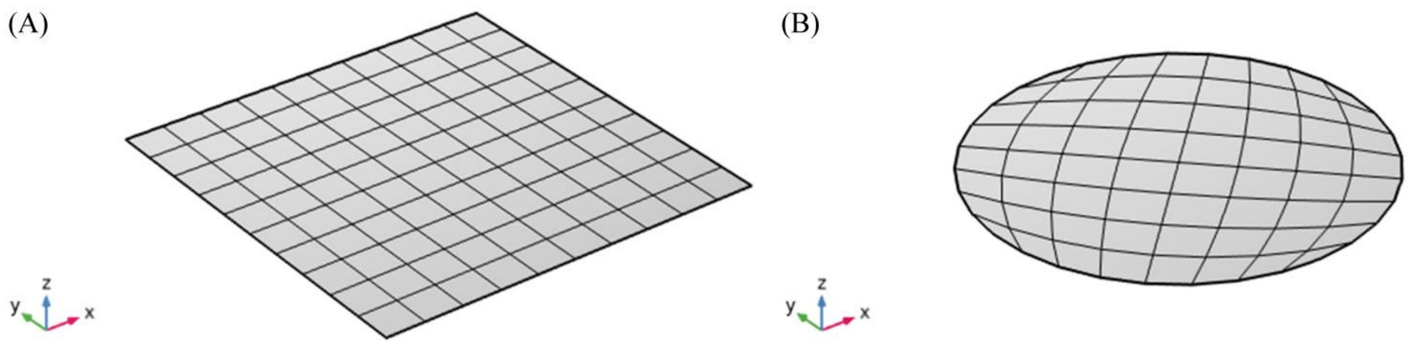


Figure 5. Mesh partitioning of multilayer shell model. (A) Rectangular; (B) Circular.

chosen encapsulation material to the top and bottom layers. Parameters for each layer like its thickness, elastic modulus, Poisson's ratio, and piezoelectric constant are then set. Finally, the "Layer Cross-Section Preview" and "Layer Stacking Preview" functions can be used to visualize the multilayer shell structure, as shown in **Figure 4**.

The model setup is completed and boundary conditions are applied. First, the "Multilayer Shell" and "Current in Multilayer Shell" physics interfaces are added to form the basis for multiphysics coupling. Under the "Multilayer Shell" node, the following features are added: "Piezoelectric Material", "Fixed Constraint", and "Surface Load". For "Piezoelectric Material", "Layer 2" is selected with the constitutive relationship set to "Strain-Charge". The option "Maintain out-of-plane material orientation unchanged" is enabled to keep polarization aligned along the Z-axis. A "Fixed Constraint" is applied to "Layer 3" (upward direction). For "Surface Load", "Downward on Layer 1" is chosen and a value of 1,000 Pa is entered. Under the "Current in Multilayer Shell" node, "Insulating Layer" and "Piezoelectric Layer" are added. "Layer 1" and "Layer 3" are assigned as "Insulating Layer", while "Layer 2" is defined as the "Piezoelectric Layer". If required, a "Ground" condition can be added under the "Piezoelectric Layer" node. When calculating surface potential and stress, "Layer 2-Layer 3" is set as grounded. The surface displacement field modulus and output charge are then computed. For charge-related outputs, both "Layer 1-Layer 2" and "Layer 2-Layer 3" interfaces are grounded.

After applying the boundary conditions, the mesh is generated. Mesh quality is directly related to the accuracy of the simulation. In this study, a mapped mesh is used, as shown in **Figure 5**. Mapped meshing allows precise control of element size and distribution, resulting in high-quality, uniformly distributed elements [33].

After completing the mesh generation, a steady-state study is selected to perform the simulation and obtain the required performance metrics. Under the "Dataset" node, "Multilayer Material" is added, and the calculation location is set to "Interface". After completing the interface selection, the "Multilayer Material" dataset is chosen under the "Derived Values" node. By inputting the calculation expressions shown in **Table 3**, the average, maximum, and minimum values of the electric displacement field modulus, electric potential, and stress for each layer can be precisely computed, along with the output charge—i.e., the integral value of the electric displacement field modulus.

Table 3. Computational expressions for performance metrics

Performance metric	Expression	Unit
Electric displacement field modulus	ecis.normD	C/m^2
Electric potential	V2	mV
Normal stress X_1	lshell.sp1	N/m^2
Normal stress X_2	lshell.sp2	N/m^2
Normal stress X_3	lshell.sp3	N/m^2

3 RESULTS

3.1 Simulation results of the basic model

To prove the rationality of the simulation model and methodology, a basic sensor model is set up using 9 cm^2 area, a 28 μm thickness and 60 μm thick PET encapsulating material. According to the "COMSOL simulation workflow" described above, the electric displacement field distribution (**Figure 6A**), electric potential distribution (**Figure 6B**), and stress distribution on the upper surface of the multilayer shell piezoelectric layer (**Figure 6C**) under static condition are obtained through COMSOL simulation. As shown in **Figure 6A**, under stress the piezoelectric layer has identical electric displacement field modulus on both the upper and lower surfaces. The arrow direction indicates the direction of the electric displacement field on the surface of the piezoelectric layer, which is generally downward. Near the edges, however, affected by the boundary of the multilayer shell and the material interface, the direction of the electric displacement field is not parallel to the electric field direction but has a certain deflection angle. As can be seen in **Figure 6B**, the potential of the top and bottom insulating layers remains constant. The lower surface of the piezoelectric layer is grounded, so its potential is almost 0 mV. On the upper side, polarization charges accumulate, creating a potential difference relative to the lower side, reaching up to 9.54 millivolts. **Figure 6C** shows a non-uniform stress distribution in the piezoelectric layer. Stress components in different directions exhibit distinct color gradients in the plane, with a pronounced difference between the edge and center regions. The average values of the surface electric displacement field modulus, output charge, electric potential, and stress components X_1, X_2, and X_3 are summarized in **Table 4**.

Substituting the simulated stress values X_1, X_2, X_3 and piezoelectric constants into equation (4), the electric displacement field modulus is calculated as:

$$D_3 = d_{31}X_1 + d_{32}X_2 + d_{33}X_3 = 2.283\times10^{-8}\ \mathrm{C/m^2} \tag{11}$$

Compared with the simulated electric displacement field modulus of 2.253×10^{-8} C/m^2, the calculation *error rate* is:

$$error\ rate = \left|\frac{2.283\times10^{-8} - 2.253\times10^{-8}}{2.283\times10^{-8}}\right| \times 100\% = 1.31\% \tag{12}$$

Using equations (5), (6), and (7), the calculated charge output is 20.547 pC and the potential is 9.335 mV, with relative errors of 1.29% and 0.45%, respectively—both below 5%. These results indicate that the simulation model accurately captures the coupling between the piezoelectric effect and the stress field, validating the model and analysis methodology.

To analyze the linearity of the sensor under varying loads, the load was successively set to 0, 100, 200, 300, 400, 500, 600, 700, 800, 900, and 1,000 Pa. The surface potential of the piezoelectric layer was measured at each load step, yielding the results shown in **Figure 7**. A linear fit of upper surface potential versus load gives the relationship y=0.0095x, with R^2=0.99. Indicating a very high degree of linear correlation between the sensor output and the applied external load.

3.2 Parameter optimization of multilayer shell structures

Based on the simulation results of the baseline model, sensitivity, stress distribution uniformity, and linearity are key metrics for evaluating sensor performance. Sensitivity is characterized by three indicators: the electric displacement field magnitude on the upper surface of the piezoelectric layer, the output charge quantity, and the electric potential. The electric displacement field pattern represents the charge generation capacity per unit stressed area, the output charge quantity represents the total charge output, and the electric potential represents the voltage signal strength formed by the accumulated charges [34, 35]. Stress distribution uniformity refers to the uniformity of the stress distribution on the upper surface of the piezoelectric layer under external load. Uniform stress distribution can ensure consistent piezoelectric conversion efficiency and help avoid the problems such as signal distortion, sensitivity change, and structural fatigue failure caused by concentrated stress. Maintaining uniform stress is therefore critical for sensor accuracy, stability and durability [36, 37]. Since the piezoelectric effect of this sensor is primarily determined by X_1, X_2, and X_3, the uniformity of stress distribution can be quantified by calculating the standard deviation of these three components. A smaller standard deviation indicates more uniform distribution of stress in that direction and better overall sensor performance. Linearity shows the proportional connection between the output signal and the input load, which is important for ensuring measurement accuracy [34].

Based on the parameters outlined in Section 2.2.3 "Sensor design parameters and boundary condition settings"—namely piezoelectric-layer area, shape, thickness, and encapsulation material—an optimization strategy can be formulated. To evaluate the individual effect of each structural parameter on sensor performance, a controlled-variable approach is adopted. This involves keeping all variables constant except the one under investigation. Simulations are conducted sequentially by varying the area, shape, thickness, and encapsulation material,

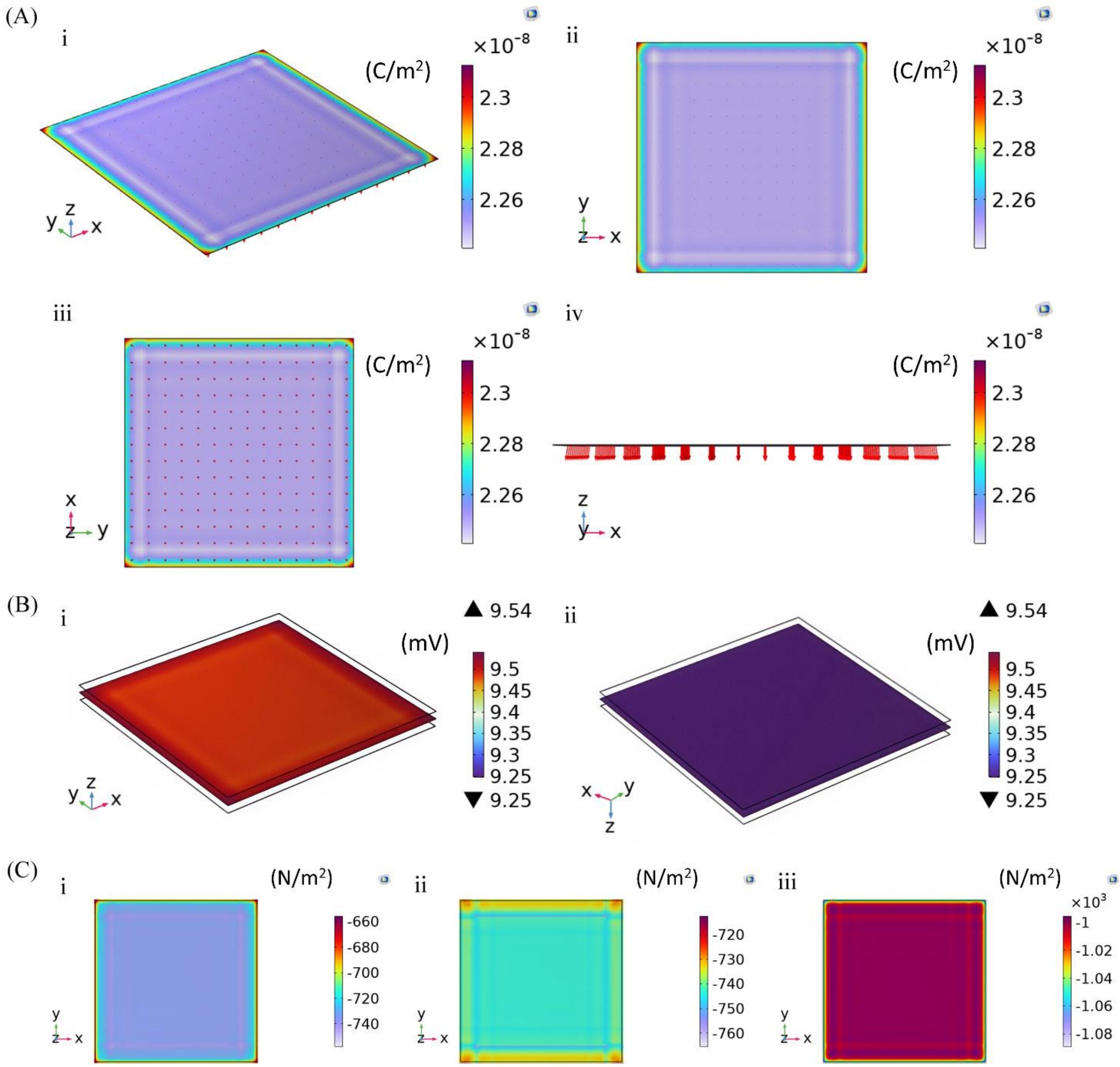


Figure 6. Simulation results for the base model. (A) Electric displacement field distribution of the multilayer shell piezoelectric layer: (i) Front view; (ii) Upper surface; (iii) Lower surface; (iv) Side view. (B) Electric potential distribution of the multilayer shell piezoelectric layer: (i) Upper surface; (ii) Lower surface. (C) Stress distribution on the upper surface of the multilayer shell piezoelectric layer: (i) Stress X_1; (ii) Stress X_2; (iii) Stress X_3.

while recording the corresponding changes in performance metrics. A comparative analysis of these results identifies the optimal range for each parameter, providing a quantitative basis for structural optimization of the sensor. Furthermore, subsequent simulations reveal that adjusting these structural parameters does not significantly alter the sensor's linear characteristic, which remains consistently high. This confirms the stability of linearity within the current structural framework and establishes a reliable benchmark for subsequently optimizing sensitivity and evenness of the stress distribution.

3.2.1 Analysis of piezoelectric layer structure impact on sensor performance

To investigate the influence of area on sensor performance, the thin film area is changed according to the model in the Section

Table 4. Surface-related data of the piezoelectric layer

Electric displacement field modulus (C/m²)	Output charge (pC)	Potential (mV)	Stress X_1 (N/m²)	Stress X_2 (N/m²)	Stress X_3 (N/m²)
2.253×10^{-8}	20.281	9.293	732.120	741.430	1001.900

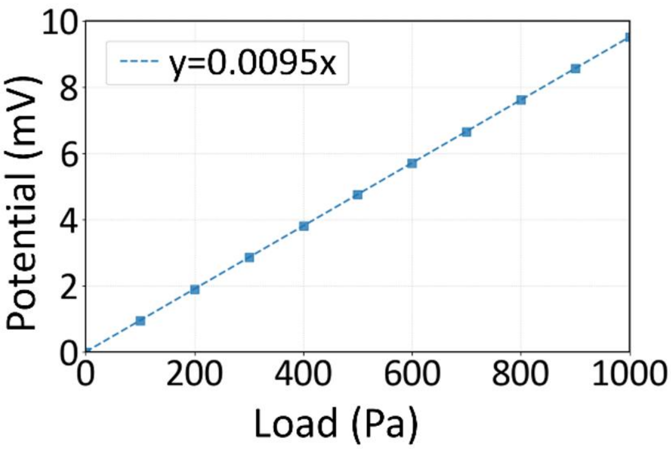


Figure 7. Sensor linear characteristic curve.

3.1 "Simulation results of basic model", while keeping all other simulation conditions unchanged. Effects of different areas on sensor sensitivity and stress distribution are summarized in **Table 5**. As for the sensitivity, the electric displacement field modulus and the electric potential decrease only slightly as the area increased, while the output charge increase linearly with the area, as seen in the theoretical derivation in equation (5). For the stress distribution uniformity aspect, increasing the area greatly improves uniformity: the standard deviations of stress in the three directions decrease significantly with increasing area. This shows that a larger area distributes the load more broadly and reduces stress concentration. Overall, increasing the area benefits both sensitivity and stress distribution uniformity.

To explore the impact of sensor shape, a circular model and an equilateral rectangular model—each with an area of approximately 9 cm^2 (radius =1.693 cm)—are simulated under identical conditions. The stress distribution of the circular piezoelectric layer is depicted in **Figure 8**, and the sensitivity and stress distribution results for both shapes are displayed in **Table 6**. A comparison of the stress distribution in **Figure 8** with that of the rectangular layer in **Figure 6C** reveals stress concentration at the edges of both shapes. According to the color legend data, it can be known that the stress values in all three directions are higher for the circular layer and are more dispersed, indicating poorer stress uniformity in the circular geometry. Regarding sensitivity, **Table 6** shows that the electric displacement field modulus and electric potential are nearly identical for both shapes, while the rectangular sensor has a slightly higher output charge. The overall differences are negligible, suggesting that shape change does not have a significant effect on the sensitivity of the piezoelectric conversion. In terms of stress distribution uniformity, the standard deviations of stress in all three directions are smaller for the rectangular sensor than for the circular sensor, by factors ranging from 1 to 8. This shows that under external loading, the circular sensor exhibits much poorer directional consistency in stress distribution and is more prone to issues such as local stress concentration.

Building upon the rectangular shape, the aspect ratio is varied to investigate its influence on sensitivity and stress distribution; the results are presented in **Table 7**. Concerning sensitivity, the modulus of the electric displacement field is constant across different aspect ratios. Both the output charge and the electric potential have a slight decreasing trend as the aspect ratio increases, although the numerical differences are minor. This means that changes in aspect ratio do not significantly affect the piezoelectric conversion sensitivity of the sensor. For the stress distribution uniformity, the standard deviations of stresses in three directions reduce continuously with increasing aspect ratios, with the most pronounced decrease found in X_1. These results suggest that properly enhancing the aspect ratio improves the evenness of the stress distribution and helps mitigate local stress concentration.

To investigate the effect of thickness on sensor performance, simulations are carried out under fixed conditions while varying the thickness of piezoelectric layer. **Table 8** shows the data result on the effect of thickness on sensitivity and stress distribution. In terms of sensitivity, the electric displacement field modulus remains consistent across different thicknesses. The output charge quantity increases slightly but not significantly with thickness. Electric potential shows an obvious increase as the thickness increases, which is consistent with the theoretical derivation in equation (7). As for the stress distribution uniformity, it is found that the standard deviations of stress in all three directions have an obvious increasing trend with greater thickness. This shows that the thickening of piezoelectric layer brings a less uniform distribution of stress, which enlarges the possibility of local stress concentration and adversely affects the stability of the output signal.

3.2.2 Analysis of encapsulation material effects on sensor performance

To examine the influence of Encapsulation materials on sensor performance, multilayer shells using different materials are simulated under the same conditions. The results in **Table 9** show the effect of encapsulation material on sensitivity and stress distribution.

Table 5. Effects of different areas on sensor sensitivity and stress distribution

Area (cm^2)	Electric displacement field modulus (C/m^2)	Output charge (pC)	Potential (mV)	Stress X_1 standard deviation (N/m^2)	Stress X_2 standard deviation (N/m^2)	Stress X_3 standard deviation (N/m^2)
1	2.265×10^{-8}	2.265	9.341	55.526	21.004	42.964
4	2.256×10^{-8}	9.024	9.303	29.211	10.178	24.755
9	2.253×10^{-8}	20.281	9.293	16.561	5.772	14.874
16	2.252×10^{-8}	36.039	9.288	10.133	4.134	9.620
25	2.252×10^{-8}	56.298	9.286	6.605	3.304	6.636

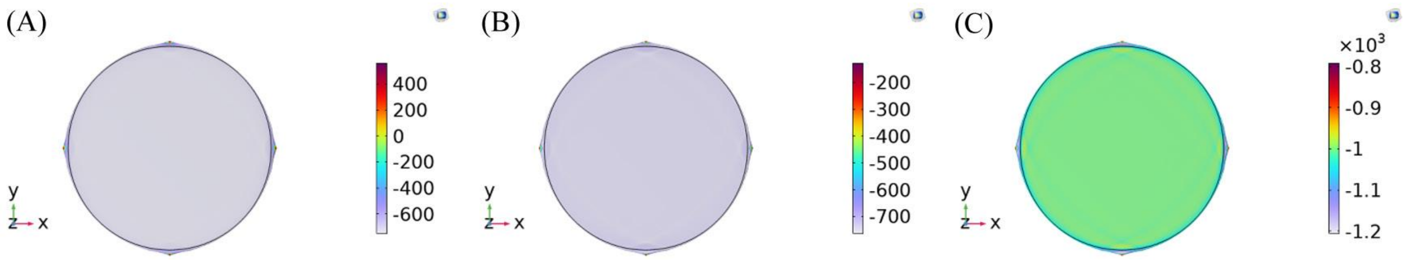


Figure 8. Stress distribution in circular piezoelectric layer. (A) Stress X_1; (B) Stress X_2; (C) Stress X_3.

In terms of sensitivity, sensors encapsulated with PET, PI, PVA, and nylon exhibit the identical values for the electric displacement field modulus, output charge, and electric potential. In contrast, PDMS encapsulation yields significantly higher values for these metrics. This difference is probably due to different Young's moduli of the encapsulation materials. The data suggest that a lower Young's modulus corresponds to higher sensitivity. Rigid materials such as PET, PI, PVA, and nylon impose more constraints on the mechanical deformation of the piezoelectric layer, restricting its strain response to external excitation and decreasing piezoelectric conversion efficiency. On the contrary, flexible PDMS exerts less constraint, maximizing the mechanical deformation of the piezoelectric layer. Under the same external excitation, a piezoelectric layer embedded in PDMS therefore produces a larger effective strain output.

As to the degree of stress distribution uniformity, the standard deviations of stresses in the three directions are in creasing successively for PET, PI, PVA and nylon encapsulation. This shows that among these materials, PET has the best stress distribution uniformity, and nylon gives the worst. However, for the PDMS encapsulation case, the standard deviations of stress in the X_1 and X_2 directions increase sharply, while only the X_3 direction is comparable to other materials. This shows that the PDMS seriously degrades the uniformity of the stress distribution and is prone to the formation of local stress concentration. Rigid encapsulation materials undergo only limited deformation and can transmit external stress more uniformly to the piezoelectric layer. Their stiffness helps prevent large localized deformation in the piezoelectric layer, thereby lowering the risk of stress concentration. In contrast, flexible PDMS deforms non-uniformly. Under external loading, it not only transmits stress unevenly, but its own deformation can also induce uneven stress distribution among various parts of the piezoelectric layer, leading to pronounced localized stress concentration and an increase in stress standard deviation.

4 DISCUSSION

4.1 Reliability and rationality of the simulation model

This study uses a multi-layer shell structure finite element model built by COMSOL, and the experiment shows good validation accuracy. Theoretical calculations on the electric displacement field, output charge quantity, and electric potential all show less than 5% error during simulation, showing that the model is correct in simulating the piezoelectric effect and charge generation of PVDF film under stress. A strict controlled variable approach is used to build the simulation framework. All the simulations use the same boundary conditions, such as the size of the external load and the method of restricting sensors. This minimizes interference from confounding factors, allowing clear attribution of performance variations to the specific parameter being studied. Consequently, the approach provides a reliable basis for subsequent parameter optimization. Trends observed in the simulation data closely align with key aspects of mechanical transmission and piezoelectric theory. For instance, the variation in sensitivity with piezoelectric-layer thickness follows the theoretical piezoelectric equations. Concerning the encapsulation material, flexible PDMS has far more permissive deformations of the piezoelectric layer, leading to a larger effective strain release and significantly higher sensitivity metrics compared to rigid PET materials. But because it is prone to non-uniform deformation, the stress standard deviation in the X_1 and X_2 directions increases, which is in line with the stress-transmission characteristics of flexible materials [38]. The data show no logical contradictions

Table 6. Effects of different shapes on sensor sensitivity and stress distribution

Shape	Electric displacement field modulus (C/m²)	Output charge (pC)	Potential (mV)	Stress X_1 standard deviation (N/m²)	Stress X_2 standard deviation (N/m²)	Stress X_3 standard deviation (N/m²)
Rectangular	2.253×10^{-8}	20.281	9.293	16.561	5.772	14.874
Circular	2.254×10^{-8}	20.256	9.294	94.868	42.031	25.088

Table 7. Effects of different aspect ratios on sensor sensitivity and stress distribution

Aspect ratio	Electric displacement field modulus (C/m²)	Output charge (pC)	Potential (mV)	Stress X_1 standard deviation (N/m²)	Stress X_2 standard deviation (N/m²)	Stress X_3 standard deviation (N/m²)
30:30	2.253×10^{-8}	20.281	9.293	16.561	5.772	14.874
36:25	2.253×10^{-8}	20.277	9.291	11.995	5.125	11.357
45:20	2.253×10^{-8}	20.274	9.290	9.354	5.025	9.681
50:18	2.253×10^{-8}	20.273	9.289	7.642	4.656	8.430
60:15	2.253×10^{-8}	20.272	9.289	5.595	4.360	6.942

Table 8. Effects of different thicknesses on sensor sensitivity and stress distribution

Thickness (μm)	Electric displacement field modulus (C/m²)	Output charge (pC)	Potential (mV)	Stress X_1 standard deviation (N/m²)	Stress X_2 standard deviation (N/m²)	Stress X_3 standard deviation (N/m²)
28	2.253×10^{-8}	20.281	9.293	16.561	5.772	14.874
52	2.255×10^{-8}	20.294	17.268	24.937	8.739	22.339
110	2.259×10^{-8}	20.333	36.598	45.999	17.199	41.982

Table 9. Effects of different encapsulation materials on sensor sensitivity and stress distribution

Encapsulation material	Electric displacement dield modulus (C/m²)	Output charge (pC)	Potential (mV)	Stress X_1 standard deviation (N/m²)	Stress X_2 standard deviation (N/m²)	Stress X_3 standard deviation (N/m²)
PET	2.253×10^{-8}	20.281	9.293	16.561	5.772	14.874
PI	2.256×10^{-8}	20.305	9.303	27.638	11.719	17.879
PVA	2.256×10^{-8}	20.307	9.305	29.323	12.406	18.690
Nylon	2.260×10^{-8}	20.340	9.319	44.322	20.577	22.636
PDMS	2.988×10^{-8}	26.896	12.303	1235.419	628.479	15.409

Note: PET, polyethylene terephthalate; PI, polyimide; PVA, polyvinyl alcohol; PDMS, polydimethylsiloxane.

or anomalous fluctuations, further verifying the reliability of the simulation results.

4.2 Influence of structural parameters on sensor performance

The influence of structural parameters on the performance of wearable sensors involves a compromise between "optimal performance" and "diversity of wearable applications", achieved by adjusting stress transfer efficiency, material usage adaptability, and piezoelectric conversion. Wearable applications include motion posture monitoring, physiological signal collection, and human-computer interaction, which all have different requirements for sensor sensitivity, portability, and fit. Therefore, structural parameters must be tuned to balance general applicability with scenario-specific demands.

Increasing the sensor area distributes external loads more widely and reduces stress concentration, significantly lowering the standard deviations of stresses in all three directions. The output charge is proportional to the area, allowing simultaneous optimization of output uniformity and total charge output. This feature might have unique benefits in motion monitoring situations, like joint activity watching, where it allows for more complete capture of big-range deformation signals. However, in applications close to the body surface—such as wrist pulse or chest auscultation monitoring—or in miniaturized wearable devices such as smart bracelet, an excessively large area creates non-functional regions outside the effective signal range. These regions cannot bear sufficient load to generate useful charge, thereby reducing the charge output density per unit area and impairing the response to weak physiological signals. Moreover, a larger area consumes more piezoelectric and encapsulation material, which increases production costs. And it also increases the sensor's overall size, reducing portability and compromising conformity to curved body surfaces. For example, on small areas like fingertips, an oversized sensor cannot conform to the surface curvature, leading to displacement and preventing stable long-term monitoring.

Likewise, the shape and aspect ratio of sensors must align with the effective signal range and the form factor of the wearable carrier. Circular sensors not only exhibit poorer stress uniformity but also offer limited spatial adaptability for elongated wearables such as smart belts. Therefore, excluding circular shapes, exploring different rectangular aspect ratios is more significant. The increasing of the aspect ratio can progressively reduce stress standard deviation by optimizing stress transferring path, making it suitable for observing linear deformation. However, an excessively high aspect ratio causes the sensor edges to extend beyond the target signal range, creating inactive stress zones that degrade performance. Considering the general requirements across wearable applications, size and aspect ratio should be adjusted flexibly according to specific applications. In scenarios like motion monitoring that require large-sacle signals, moderately increasing the area and aspect ratio helps achieve uniform stress distribution and full signal capture. On the contrary, for surface physiological signal acquisition or miniaturized devices, controlling the area and aspect ratio ensures that inactive zones do not compromise sensitivity and that the sensor conforms to the surface geometry and device carrier size.

Although increasing the piezoelectric layer thickness raises the output potential, the electric displacement field modulus and output charge show only minor fluctuations. This shows that thickness does not substantially change the core efficiency of piezoelectric conversion. In engineering applications for wearable sensors, output charge is typically converted to voltage using charge amplifiers; thus, a thicker piezoelectric layer is not necessary to achieve higher potential. More importantly, wearable devices are in direct, prolonged contact with the skin. A thicker piezoelectric layer can lead to signal instability and introduce additional inertial interference during movement. It also requires more piezoelectric material like PVDF, raising cost, while reducing adhesion to the skin surface. Excessively thick piezoelectric layers cannot conform well to body curvatures, causing wrinkles, creating a pronounced foreign-body sensation, and reducing comfort during extended wear. At actively moving joints such as the neck and elbow, excessive thickness would severely compromise wearability. Therefore, a 28 μm thin piezoelectric layer can maintain low stress variance, meet the signal-acquisition requirements of wearable applications, control cost, and improve wearing comfort simultaneously, representing the optimal thickness choice.

The effect of the encapsulation material on the performance of the wearable sensor mainly depends on Young's modulus, which must also be flexible, biocompatible, and environmentally tolerant (e.g., to sweat and temperature variation). Flexible PDMS imposes less constraint on the deformation of the piezoelectric layer, which allows for a greater effective strain to be released. This increases the modulus of the electric displacement field, the charge quantity, and the electric potential by about 32% with respect to a rigid material such as PET, which is particularly advantageous for high-sensitivity situations like micro-motion detection. However, PDMS is susceptible to non-uniform deformation, leading to an extremely large standard deviation of X_1, X_2 stress, severe stress concentration, and significant signal fluctuation. Therefore, it is only appropriate for extremely sensitive scenarios where stability can be compromised. Rigid materials such as PET, PI, PVA, and nylon do not increase sensitivity but provide uniform stress distribution through strong mechanical constraint, making them suitable for daily wearables that prioritize stability. Among these, PET has the smallest stress standard deviation, along with favorable flexibility, chemical resistance, and biocompatibility, meeting the general requirements of most wearable sensors [39]. Consequently, PET should be the choice of encapsulation material for everyday wearables like fitness trackers. For applications with extreme sensitivity requirements, PDMS may be considered, but material or structure optimization is required to address the problem of stress concentration, while maintaining stability and wearability.

5 CONCLUSION

This study is to optimize PVDF sensors for wearable health monitoring by researching the effect of multilayer shell structural parameters on sensor performance. The following conclusions can be drawn. Increasing the sensor area improves sensitivity and stress distribution uniformity. Rectangular structures have higher stress uniformity than circular structures; a moderate increase in aspect ratio maximizes uniformity while maintaining sensitivity. Thinner piezoelectric layers are better, as they maintain lower stress variance and better wearing comfort without significantly compromising output potential. Flexible encapsulation materials such as PDMS yield higher sensitivity, but they introduce a risk of stress concentration and signal instability. This study identifies an optimal structural configuration and practical parameter ranges for PVDF-based wearable sensors, giving a theoretically grounded approach for designing high-performing wearable devices and supporting the development of effective health-monitoring technologies.

DECLARATIONS

Author contributions

Ke Wang was responsible for simulation experiments, data analysis, and drafting the initial manuscript. Rongguo Yan provided guidance, reviewed and revised the manuscript, and engaged in constructive discussions regarding analytical work. Wenjing Du and Shoucheng Chen contributed data and drafted the initial manuscript.

Funding

This research received no external funding.

Data availability

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.

REFERENCES

[1] Liu T, Mao Y, Dou H, Zhang W, Yang J, Wu P, et al. Emerging wearable acoustic sensing technologies. Adv Sci (Weinh). 2025 Feb;12(6):2408653. https://doi.org/10.1002/advs.202408653

[2] Ochonska D, Brzychczy-Wloch M, Talaga-Cwiertnia K. Different faces of the stethoscope: history, usefulness, evolution, contamination, and disinfection practices. Med Stud. 2025; 41(1):53-63. https://doi.org/10.5114/ms.2024.144160

[3] Jin H, Gao X, Ren K, Liu J, Qiao L, Liu M, et al. Review on piezoelectric actuators based on high-performance piezoelectric materials. IEEE Trans Ultrason Ferroelect Freq Contr. 2022 Nov;69(11):3057-3069. https://doi.org/10.1109/tuffc.2022.3175853

[4] Habib M, Lantgios I, Hornbostel K. A review of ceramic, polymer and composite piezoelectric materials. J Phys D Appl Phys. 2022 Oct 20;55(42):423002. https://doi.org/10.1088/1361-6463/ac8687

[5] Wang Z, Wang Y, Yang C, Zheng T, Luo R, Wang Y. Applications of piezoelectric materials in biomedical engineering. Macromol Biosci. 2025 Aug;25(8):2500033. https://doi.org/10.1002/mabi.202500033

[6] Wu Y, Ma Y, Zheng H, Ramakrishna S. Piezoelectric materials for flexible and wearable electronics: A review. Mater Des. 2021 Dec 1;211:110164. https://doi.org/10.1016/j.matdes.2021.110164

[7] Chen Y, Zhang X, Lu C. Flexible piezoelectric materials and strain sensors for wearable electronics and artificial intelligence applications. Chem Sci. 2024 Oct;15(40):16436-16466. https://doi.org/10.1039/d4sc05166a

[8] Zhang J, Wang J, Zhong C, Zhang Y, Qiu Y, Qin L. Flexible electronics: Advancements and applications of flexible piezoelectric composites in modern sensing technologies. Micromachines. 2024 Aug;15(8):982. https://doi.org/10.3390/mi15080982

[9] Li M, Aoyama J, Inayoshi K, Zhang H. Wearable PZT piezoelectric sensor device for accurate arterial pressure pulse waveform measurement. Adv Electron Mater. 2025 June;11(9):2400852. https://doi.org/10.1002/aelm.202400852

[10] Kawai H. The piezoelectricity of poly (vinylidene Fluoride). Jpn J Appl Phys. 1969;8(7):975. https://doi.org/10.1143/JJAP.8.975

[11] Qi F, Xu L, He Y, Yan H, Liu H. PVDF-based flexible piezoelectric tactile sensors: Review. Cryst Res Technol. 2023 Oct;58(10):2300119. https://doi.org/10.1002/crat.202300119

[12] Duan S, Wu J, Xia J, Lei W. Innovation strategy selection facilitates high-performance flexible piezoelectric sensors. Sensors. 2020 May;20(10):2820. https://doi.org/10.3390/s20102820

[13] Khan B, Khalid RT, Masrur MH, Awais M, Imdad N, Hilal ME, et al. Advances in polyvinylidene fluoride (PVDF) for self-powered wearable physiological monitoring and energy harvesting applications. Nano Energy. 2025 Oct;143:111296. https://doi.org/10.1016/j.nanoen.2025.111296

[14] Ghosh SK, Mandal D. Synergistically enhanced piezoelectric output in highly aligned 1D polymer nanofibers integrated all-fiber nanogenerator for wearable nano-tactile sensor. Nano Energy. 2018 Nov;53:245-257. https://doi.org/10.1016/j.nanoen.2018.08.036

[15] Yang J, Wang M, Meng Y, Niu Z, Hao Y, Liu H, et al. High-performance flexible wearable triboelectric nanogenerator sensor by β-phase polyvinylidene fluoride polarization. ACS Appl Electron Mater. 2024 Feb 13;6(2):1385-1395. https://doi.org/10.1021/acsaelm.3c01678

[16] Kalimuldina G, Turdakyn N, Abay I, Medeubayev A, Nurpeissova A, Adair D, et al. A review of piezoelectric PVDF film by electrospinning and its applications. Sensors. 2020 Sep;20(18):5214. https://doi.org/10.3390/s20185214

[17] He Z, Rault F, Lewandowski M, Mohsenzadeh E, Salaun F. Electrospun PVDF nanofibers for piezoelectric applications: A review of the influence of electrospinning parameters on the β phase and crystallinity enhancement. Polymers. 2021 Jan 6;13(2):174. https://doi.org/10.3390/polym13020174

[18] Parangusan H, Ponnamma D, Al-Maadeed MA. Stretchable electrospun PVDF-HFP/Co-ZnO nanofibers as piezoelectric nanogenerators. Sci Rep. 2018 Jan 15;8(1):754. https://doi.org/10.1038/s41598-017-19082-3

[19] Islam MN, Rupom RH, Adhikari PR, Demchuk Z, Popov I, Sokolov AP, et al. Boosting piezoelectricity by 3D printing PVDF-MoS_2 composite as a conformal and high-sensitivity piezoelectric sensor. Adva Funct Mater. 2023 Oct;33(42): 2302946. https://doi.org/10.1002/adfm.202302946

[20] Liu B, Han L, Pan L, Li H, Zhao J, Dong Y, et al. Flexible multiscale pore hybrid self-powered sensor for heart sound detection. Sensors. 2021 Jul;21(13):4508. https://doi.org/10.3390/s21134508

[21] Xue J, Xin Y. Development and property of flexible respiratory sensor based on PVDF. New Chem Mater. 2019;47(12):104-108.

[22] Kim MS, Ahn HR, Lee S, Kim C, Kim YJ. A dome-shaped piezoelectric tactile sensor arrays fabricated by an air inflation technique. Sensor Actuat A-Phys. 2014 Jun;212:151-158. https://doi.org/10.1016/j.sna.2014.02.023

[23] Yan L, Yan R, Guo X. Piezoelectric simulation and structural optimization of PVDF films. Prog Biomed Eng. 2023;44(02):119-127.

[24] Wang F, Song Z, Cai X, Guo K, Pan X, Ren C, et al. External strategies for enhanced sensing performance of self-powered

polyvinylidene fluoride-based sensors. Nanoscale. 2025 Mar;17(12):6981-6992. https://doi.org/10.1039/d4nr05200e

[25] Chen C, Zhao S, Pan C, Zi Y, Wang F, Yang C, et al. A method for quantitatively separating the piezoelectric component from the as-received “Piezoelectric” signal. Nat Commun. 2022 Mar 16;13(1):1391. https://doi.org/10.1038/s41467-022-29087-w

[26] Gao X, Yang J, Wu J, Xin X, Li Z, Yuan X, et al. Piezoelectric actuators and motors: Materials, designs, and applications. Adv Mater Technol. 2020 Jan;5(1):1900716. https://doi.org/10.1002/admt.201900716

[27] Covaci C, Gontean A. Piezoelectric energy harvesting solutions: A review. Sensors. 2020 Jun;20(12):3512. https://doi.org/10.3390/s20123512

[28] Thompson ML. On the material properties and constitutive equations of piezoelectric poly vinylidene fluoride (PVDF) [dissertation]. Philadelphia (PA): Drexel University; 2002. 184 p.

[29] COMSOL. Structural mechanics module user’s guide (version 5.4) [Internet]. 2018 [updated 2018; cited 2026 Mar 6]. Available from: https://doc.comsol.com/5.4/doc/com.comsol.help.sme/StructuralMechanicsModuleUsersGuide.pdf

[30] Zhang W, Wu G, Zeng H, Li Z, Wu W, Jiang H, et al. The preparation, structural design, and application of electroactive poly(vinylidene fluoride)-based materials for wearable sensors and human energy harvesters. Polymers. 2023 Jul;15(13):2766. https://doi.org/10.3390/polym15132766

[31] COMSOL. Piezoelectricity in a layered shell [Internet]. 2022 [updated 2022; cited 2026 Mar 6]. Available from: https://cn.comsol.com/model/piezoelectricity-in-a-layered-shell-78301

[32] Zhou Z, Du X, Luo J, Yao L, Zhang Z, Yang H, et al. Coupling of interface effects and porous microstructures in translucent piezoelectric composites for enhanced energy harvesting and sensing. Nano Energy. 2021 Jun;84:11. https://doi.org/10.1016/j.nanoen.2021.105895

[33] Krusell E. Best practices for meshing domains with different size settings [Internet]. COMSOL; 2019 Jul 9 [updated 2019 Jul 9; cited 2026 Mar 6]. Available from: https://www.comsol.com/blogs/best-practices-for-meshing-domains-with-different-size-settings

[34] Xu Z, Wu D, Chen Z, Wang Z, Cao C, Shao X, et al. A flexible pressure sensor with highly customizable sensitivity and linearity via positive design of microhierarchical structures with a hyperelastic model. Microsyst Nanoeng. 2023 Jan 4;9:5. https://doi.org/10.1038/s41378-022-00477-w

[35] Li G, Chen D, Li C, Liu W, Liu H. Engineered microstructure derived hierarchical deformation of flexible pressure sensor induces a supersensitive piezoresistive property in broad pressure range. Adv Sci. 2020 Sep;7(18):2000154. https://doi.org/10.1002/advs.202000154

[36] Xia Q, Shi T, Liu S, Wang M. Optimization of stresses in a local region for the maximization of sensitivity and minimization of cross-sensitivity of piezoresistive sensors. Struct Multidiscip Optim. 2013 Nov;48(5):927-938. https://doi.org/10.1007/s00158-013-0997-3

[37] Xia Q, Shi T, Liu S, Wang M. Shape and topology optimization for tailoring stress in a local region to enhance performance of piezoresistive sensors. Comput Struct. 2013 Jan;114:98-105. https://doi.org/10.1016/j.compstruc.2012.10.020

[38] Jian A, Tang X, Feng Q, Duan Q, Ji J, Ma Z, et al. A PDMS surface stress biosensor with optimized micro-membrane: Fabrication and application. Sens Actuators B Chem. 2017 Apr;242:969-976. https://doi.org/10.1016/j.snb.2016.09.157

[39] Gramada AM, Stoica AE, Niculescu AG, Bîrca AC, Vasile BS, Holban AM, et al. Zinc oxide-loaded recycled PET nanofibers for applications in healthcare and biomedical devices. Polymers. 2025 Jan;17(1):45. https://doi.org/10.3390/polym17010045

Progress in
Medical Devices

Zentime

RESEARCH ARTICLE

A correlation study of paraspinal muscle functions in adolescent idiopathic scoliosis

Rong Pang[1], Chen He[1], Huidong Wu[2]

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.
[2]Department of Prosthetic and Orthotic Engineering, School of Rehabilitation, Kunming Medical University, Kunming 650032, Yunnan, China.

Corresponding author: Chen He.

Address correspondence to: Chen He, School of Health Science and Engineering, University of Shanghai for Science and Technology, No. 516 Jungong Road, Yangpu District, Shanghai 200093, China.
E-mail: hechen@usst.edu.cn.

Received November 21, 2025; Accepted April 16, 2026; Published June 18, 2026

DOI: 10.61189/126256lnkxbu

Abstract

Objective: To explore the correlation among paraspinal muscle functions electromyography (EMG), muscle stiffness, and pain threshold in patients with adolescent idiopathic scoliosis (AIS). **Methods:** Eighteen patients with AIS were recruited. A Noraxon system equipped with four wireless EMG sensors was used to collect EMG data on the paraspinal muscles in relaxed standing and weight-bearing standing states. Muscle stiffness and pain threshold were measured using a muscle tonometer. The differences in mean EMG amplitude, muscle stiffness, and pain threshold between the concave and convex sides of the scoliosis were analyzed. **Results:** Among patients with different scoliosis locations, Cobb angles, ages, and brace treatment durations, the mean EMG amplitude of the paraspinal muscles on the convex side of scoliosis was significantly higher than that on the concave side ($P<0.05$). The muscle stiffness and pain threshold of the paraspinal muscles on the convex side were also significantly higher than those on the concave side (both $P<0.05$). There was a low correlation between the mean EMG amplitude of the paraspinal muscles, muscle stiffness, and pain threshold ($R<0.5$, $P>0.05$). **Conclusion:** In AIS patients, the electromyographic activity, muscle stiffness, and pain threshold of the paraspinal muscles on the convex side of scoliosis were all higher than those on the concave side, and the correlation among the three indicators was low.

Keywords: Scoliosis, Paraspinal muscles, Electromyographic activity, Muscle stiffness, Pain threshold

Highlights

- Electromyographic activity, muscle stiffness, and pain threshold on the convex side of the scoliotic curve exhibited significantly higher than those on the concave side.
- In adolescent idiopathic scoliosis patients, there was a weak correlation between electromyographic activity, muscle stiffness, and pain threshold of the paraspinal muscles.

1 INTRODUCTION

Scoliosis is a three-dimensional spinal deformity with an incidence rate of 0.23%–2.40% [1]. The vast majority of scoliosis patients are diagnosed with adolescent idiopathic scoliosis (AIS). The etiology of AIS remains unclear, but it may be related to factors such as genetics, growth and development, abnormal neurological function, and melatonin secretion [2]. Among the many influencing factors, paraspinal muscle dysfunction is considered to be the main cause of the occurrence and progression of scoliosis [3, 4]. The paraspinal muscles are mainly composed of the superficial erector spinae muscles and the deep multifidus muscles on both sides of the spine [5]. The two work together to maintain the upright posture of the human body and support trunk movement [6]. Patients with scoliosis have paraspinal muscle imbalance, which is manifested by differences in muscle volume and different degrees of muscle activation during movement [7]. This muscle imbalance is both a cause and a result of scoliosis. Symmetrical paraspinal muscles have regulatory and balancing mechanisms that can correct mild scoliosis and further control the progression of the disease [8].

Surface electromyography (EMG), muscle stiffness, and muscle pain threshold are important indicators for assessing the paraspinal muscle status in individuals with AIS. They can reflect muscle imbalance, adaptive changes in muscle activity, muscle activation patterns, and functional limitations in scoliosis patients. These three parameters were selected because they characterize paraspinal muscle function from complementary dimensions, covering neuromuscular activation status, biomechanical properties, and sensory pain characteristics, respectively. Their combined application allows for a multidimensional assessment of paraspinal muscle dysfunction in AIS. Surface EMG can reveal muscle activation patterns, detect muscle activity imbalances, and assess the functional performance of paraspinal muscles under different activity states. In the field of scoliosis research, surface EMG is particularly suitable for identifying asymmetric muscle activation patterns and compensatory muscle recruitment patterns caused by changes in spinal alignment. Previous studies have found that surface EMG signal levels in scoliosis patients are higher than in healthy individuals, and the signal level on the convex side of the scoliosis is higher than on the concave side [9-11].

Muscle stiffness refers to the resistance generated when the paraspinal muscles are passively stretched by external force [12]. Changes in muscle stiffness can reflect changes in muscle tissue characteristics, including muscle tension, stiffness, and other biomechanical and structural features. It can also reflect potential changes in muscle tissue composition caused by long-term asymmetrical loading. Liu et al. found that the stiffness of the paraspinal muscles on the concave side was higher than that on the convex side in AIS patients [5]. This difference can affect the alignment of the spinal force line and thus aggravate spinal deformity.

Paraspinal muscle pain threshold assessment is the measurement of the pressure threshold that the human body can withstand when it begins to feel pain [13]. Monitoring the muscle pain threshold can help understand the pain sensitivity and pain perception of the paraspinal muscles in scoliosis patients. According to statistics, 23% of scoliosis patients have symptoms of lower back pain [14]. A comprehensive exploration of the physiological characteristics of the paraspinal muscles in scoliosis patients, including EMG signals, muscle stiffness, and pain threshold, can provide an important reference for elucidating the pathogenesis of scoliosis. At the same time, the above indicators can be applied in clinical practice to provide a basis for individualized diagnosis, treatment plan formulation, and clinical efficacy evaluation of scoliosis.

Therefore, this study aims to explore the symmetry and correlation among surface EMG of the paraspinal muscles, muscle stiffness, and muscle pain threshold in individuals with AIS. The findings are expected to deepen our understanding of musculoskeletal dysfunction in scoliosis at the research context and provide a reference for developing individualized intervention programs for specific muscle problems in scoliosis patients at the clinical level.

2 MATERIALS AND METHODS

2.1 Study design

All participants were recruited from the Department of Rehabilitation, First Affiliated Hospital of Kunming Medical University. This study was approved by the Medical Ethics Committee of Kunming Medical University (Approval No: KMMU2023MEC088), and the recruitment period was from May 2023 to February 2024.

Inclusion criteria: (1) AIS patients with a Cobb angle $<60°$; (2) Age 10-18 years old; (3) Risser sign ≤ 4.

Exclusion criteria: (1) Non-AIS; (2) History of spinal surgery; (3) Comorbid neuromuscular diseases and spinal lesions affecting muscle function or pain sensitivity.

Surface EMG signals of the paraspinal muscles were collected using a surface EMG system (TeleMyo 2400 G2, Noraxon, USA). Participants were placed in a prone position with their arms relaxed. After cleaning the skin of the paraspinal muscles with a damp cotton ball, the electrode pads were placed on the erector spinae muscles at the apical level, 2 cm from the midline of the spine. For patients with unilateral curvature, two pairs of electrodes were placed at the apical level; for patients with bilateral curvature, four pairs of electrodes were placed at two apical levels (**Figure 1**). The sensor was connected to the electrodes and fixed 1 cm from them. Surface EMG signals were collected for 10 seconds each in a naturally standing, resting state and under a 2 kg load, with a 30-second rest interval between collections. All measurements were repeated three times. The ratio of the average EMG amplitude on the convex side to the concave side was calculated as the paraspinal muscle symmetry index; the closer the value was to 1, the higher the symmetry of the muscles on both sides; the farther the value deviated from 1, the lower the symmetry.

The paraspinal muscle stiffness and pain threshold were measured using a muscle state detection system (Itochu, Japan, model OE-220). Subjects were placed in a prone position with arms relaxed. The erector spinae muscles at the apical level

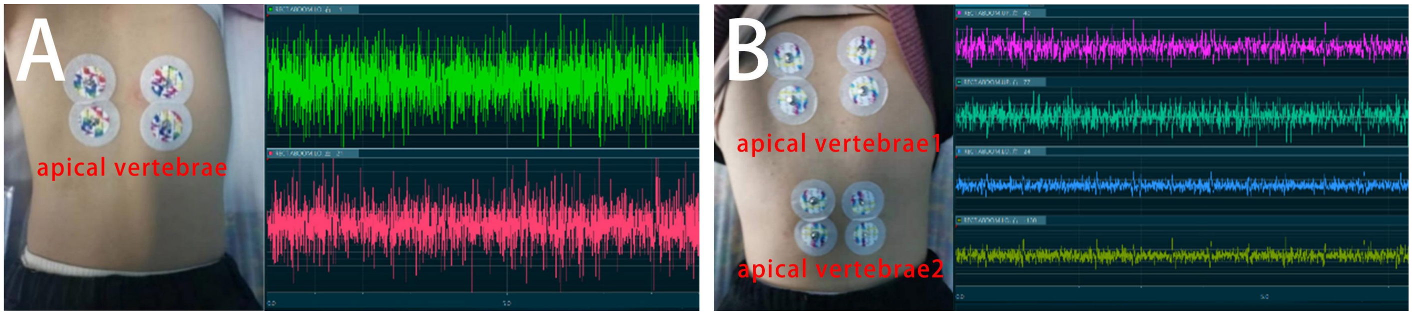


Figure 1. Electrodes placement and real-time sEMG. (A) Single curve; (B) Double curve. EMG, electromyography.

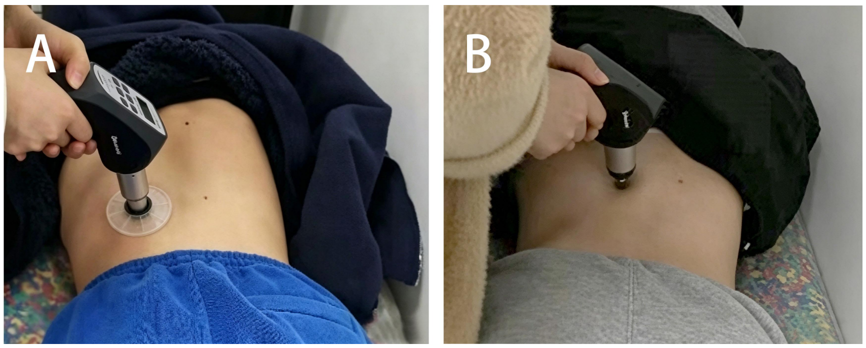


Figure 2. Assessment of paraspinal muscle. (A) Measurement of muscle hardness; (B) Measurement of pain threshold.

Table 1. Demographic data

Number	Age	BMI	Cobb angle	Risser sign	Apex
n=18	13.56±3.00	17.26±2.46	32.54±10.83	0~4	T8-L2

Note: BMI, body mass index; Apex, apical vertebra.

were marked as detection sites; two detection sites were marked for patients with unilateral curvature, and four detection sites were marked for patients with bilateral curvature.

For paraspinal muscle stiffness assessment, the rapid muscle condition detection device was aligned to the detection site, and pressure was slowly applied vertically (**Figure 2**). Once the device screen displayed the test results, the single test ended and the data was saved. Each subject underwent the test three times. The ratio of muscle stiffness on the convex side to the concave side was defined as the paraspinal muscle symmetry index; the closer the value was to 1, the higher the symmetry between the two sides; the further the value deviated from 1, the lower the symmetry.

For paraspinal muscle pain threshold test, the muscle condition detection device was applied vertically to the designated detection site, and pressure was gradually increased through the detection probe. The subject held the braking device throughout the test. As the pressure slowly increases, the subject actively pressed the remote control to terminate the test when the pain threshold was reached, and the data displayed on the device was recorded. Each test was repeated three times, with a 30-second rest interval between tests. The ratio of the pain threshold on the convex side to the concave side was used as the paraspinal muscle symmetry index; the closer the value was to 1, the higher the symmetry between the two sides; the further the value deviated from 1, the lower the symmetry.

2.2 Statistical analysis

Data analysis was performed using IBM SPSS Statistics 26.0 statistical software. Paired t-tests were used to compare the differences in mean surface EMG amplitude, muscle stiffness, and muscle pain threshold of the paraspinal muscles on the concave and convex sides of scoliosis. Analysis of variance with Bonferroni post-hoc tests was used to analyze the effects of different scoliosis types, Cobb angles, age, and treatment duration of scoliosis braces on mean surface EMG amplitude. Pearson correlation coefficients were used to analyze the correlation among mean surface EMG amplitude, muscle stiffness, and muscle pain threshold of paraspinal muscles in scoliosis patients.

3 RESULTS

3.1 Demographic data

This study included 18 participants, of whom 13 had received or were undergoing scoliosis exercise therapy, and 9 were undergoing orthotic therapy. Demographic data of the participants are shown in **Table 1**.

3.2 Surface EMG signals

The mean EMG amplitudes of the paraspinal muscles on the convex and concave sides of the spine are shown in **Table 2**. The results showed that the mean EMG amplitude on the convex side of the paraspinal muscles was significantly higher than

Table 2. Mean surface electromyographic amplitude of paraspinal muscles in patients with scoliosis

Position	Mean EMG convex (mV)	Mean EMG concave (mV)	Concave-convex ratio	P value
Relax standing (n=18)	16.57±7.52	10.83±3.99	1.56±0.49	P<0.001
Weight-bearing standing (n=18)	83.64±55.41	52.84±46.99	1.94±0.96	P<0.001
P value	P<0.001	P<0.001	P<0.001	

Note: EMG, electromyography.

Table 3. Symmetry of mean surface electromyographic amplitude of paraspinal muscles in different subgroups

	Relax standing	P value	Weight-bearing standing	P value
Double curves (n=10)	1.40±0.41	0.128	1.97±1.04	0.586
Single curve (n=8)	1.73±0.16		2.22±1.06	
Cobb angle <30° (n=7)	1.64±0.60	0.467	2.02±1.01	0.725
Cobb angle ≥30° (n=11)	1.50±0.40		1.88±0.94	
Age <15 years (n=11)	1.48±0.40	0.807	1.80±0.83	0.267
Age ≥15 years (n=7)	1.54±0.53		2.27±1.22	
Orthosis <6 m (n=9)	1.31±0.30	0.260	2.31±1.19	0.706
Orthosis ≥6 m (n=9)	1.73±0.65		2.64±1.12	

Table 4. Symmetry of paraspinal muscle stiffness and pain threshold

Measurement index	Convex side	Concave side	Ratio	P value
Muscle hardness m/N (n=18)	47.63±12.72	44.41±13.17	1.10±0.20	0.047
Muscle pain threshold N (n=18)	32.00±12.45	29.28±11.50	1.12±0.24	0.030

Table 5. Correlation analysis of paraspinal muscle surface electromyographic signal, muscle stiffness, and pain threshold symmetry

Measurement index	Pearson correlation coefficient	P value
Surface EMG & muscle hardness	0.418	0.734
Surface EMG & pain threshold	0.356	0.390
Muscle hardness & pain threshold	0.394	0.618

that on the concave side (P<0.05). The mean EMG amplitude under weight-bearing standing conditions was higher than that under relaxed standing conditions (P<0.05). Furthermore, the dispersion of the mean EMG amplitude data under weight-bearing standing conditions was significantly greater.

Patients were grouped according to the location of the scoliosis (single curve, double curve), Cobb angle (<30°, ≥30°), age (<15 years, ≥15 years), and duration of orthotic treatment (<6 months, ≥6 months). The influence of each variable on the symmetry of the paraspinal muscles was analyzed (**Table 3**). Pairwise comparisons between groups showed no statistically significant difference in the symmetry of the mean surface EMG amplitude of the paraspinal muscles between any two groups, regardless of whether the patient was standing relaxed or with loading (P>0.05).

3.3 Muscle stiffness and pain threshold

The results of paraspinal muscle stiffness and pain threshold tests are shown in **Table 4**. Muscle stiffness results showed that the stiffness of the paraspinal muscles on the convex side was significantly greater than that on the concave side in AIS patients (P=0.047). Pain threshold results revealed showed that the pain threshold on the convex side of the paraspinal muscles was significantly higher than that on the concave side (P=0.030).

3.4 Correlation analysis

Pearson correlation analysis was used to explore the correlation among the symmetry of mean surface EMG amplitude, muscle stiffness, and pain threshold symmetry of the paraspinal muscles on the convex and concave sides in AIS patients (**Table 5**). The results indicated symmetry of mean surface EMG amplitude was weakly correlated with muscle stiffness symmetry (R=0.418, P=0.734); symmetry of mean surface EMG amplitude was weakly correlated with pain threshold symmetry (R=0.356, P=0.390); and symmetry of muscle stiffness was also weakly correlated with pain threshold symmetry (R=0.394, P=0.618).

4 DISCUSSION

Paraspinal muscles play a crucial role in the pathogenesis and progression of scoliosis. This study confirmed that, regardless of curve location, Cobb angle, patient age, or orthotic treatment duration, the surface EMG signal on the convex side was significantly higher than that on the concave side, consistent with previous findings [15, 16]. This asymmetry may stem from atrophy of the concave paraspinal muscle fibers accompanied by excessive stretching of the convex fibers [17, 18]. Mechanistically, metabolic processes in the paraspinal muscles may disrupt the balance between melatonin and calmodulin. Increased calmodulin levels on the convex side can enhance contractile force, thereby amplifying EMG activity [19].

Significant differences in mean EMG signals were observed under weight-bearing standing conditions, likely due to increased contraction amplitude during weight-bearing and individual variability in neuromuscular control strategies. Weight-bearing conditions require greater postural control, leading patients to develop individualized compensatory activation patterns to maintain trunk stability. These differentiated force exertion patterns, prevalent in scoliosis populations, increase the dispersion of surface EMG results.

The study also found that different curve locations, Cobb angles, ages, and orthotic treatment durations did not significantly affect the symmetry of paraspinal surface EMG signals.

This study found that muscle stiffness and pain threshold on the convex side were higher than those on the concave side, a finding that differs from previous studies reporting higher stiffness on the concave side. This discrepancy may stem from differences in detection methods, study population characteristics, and clinical conditions. Severity of scoliosis, age, and treatment history may also affect long-term adaptive changes.

Patients who rely predominantly on the convex side for postural compensation may develop increased muscle tension and stiffness on that side due to chronic overactivation. Clinical observation indicates that most scoliosis patients habitually use the convex side as the primary force-generating muscle group [20]. Long-term continuous contraction of convex-side muscles leads to changes in muscle stiffness. Increased muscle stiffness is also associated with chronic musculoskeletal dysfunction, suggesting possible pathological changes such as tissue edema, hypoxia, and acidosis [21].

Asymmetry in pain threshold is primarily related to individual differences in pain sensitivity and subjective perception. Scoliosis orthoses apply corrective force to the apical vertebra on the convex side, placing muscles under prolonged stress, which may increase their tolerance to external forces and raise the pain threshold. Studies have confirmed that scoliosis rehabilitation exercises can effectively retrain dominant muscle group force exertion habits through diverse specific movements, thereby relieving pain and improving paraspinal muscle asymmetry [22].

Assessing muscle stiffness and pain threshold helps comprehensively evaluate neuromuscular function and provides a basis for developing personalized rehabilitation programs and pain management strategies. Rehabilitation training should focus on improving bilateral muscle symmetry and correcting compensatory force exertion habits on the convex side. For patients with high muscle stiffness, targeted local relaxation and stretching exercises are recommended; for those with low pain threshold, gentler intervention methods should be considered.

Changes in muscle stiffness and pain threshold before and after treatment can serve as important indicators for evaluating efficacy and adjusting treatment plans. Future research should compare these indicators before and after rehabilitation exercise interventions to achieve more accurate and objective evaluation of treatment efficacy.

This study found a low correlation between the surface EMG, muscle stiffness, and pain threshold in scoliosis patients. Muscle stiffness is defined as the degree of deformation of muscle morphology under specific pressure, while mean EMG signal reflects changes in EMG amplitude during paraspinal muscle contraction [23].

Leonard et al. showed that muscle stiffness and surface EMG were positively correlated in non-scoliosis patients, whereas the present study found a low correlation in scoliosis patients [24]. This difference may be explained by pathological abnormalities in scoliosis disrupting the original correlation between these indicators. This disruption is more pronounced in patients wearing orthotics, whose muscle activity is restricted and who are subjected to long-term pressure.

Only a few subjects reported significant back muscle pain symptoms in daily life. Such symptoms are more common in adult scoliosis patients but have a relatively low incidence in those with AIS [25]. Other studies have shown that rehabilitation exercises have no significant effect on muscle stiffness and pain sensitivity in healthy individuals [26]. Further research is needed to determine whether specific rehabilitation exercises can improve muscle stiffness and reduce pain sensitivity in scoliosis patients.

This study has several limitations. First, the sample size was relatively small (n=18), primarily focusing on adolescents with mild to moderate scoliosis, limiting generalizability to patients with severe scoliosis or other age groups. Second, as a cross-sectional study, causality cannot be inferred; it remains unclear whether abnormal muscle characteristics induce scoliosis progression or whether spinal deformity leads to secondary adaptive muscle changes. Third, previous rehabilitation exercises and orthotic treatments may have confounded the results, although subgroup analysis based on treatment duration did not reveal significant differences.

Future studies should conduct large-scale, stratified longitudinal follow-up studies based on treatment history to clarify the causal relationship and dynamic changes in paraspinal muscle function in AIS.

5 CONCLUSION

In AIS patients, EMG activity, muscle stiffness, and pain threshold are all higher on the convex side than on the concave side, and the correlation among these three indicators is low. In-depth research on paraspinal muscle function can provide important references for developing individualized diagnosis and treatment plans for AIS.

DECLARATIONS

Author contributions

Rong Pang participated in data curation, software development, investigation, writing the original draft, resource provision and visualization; Chen He was responsible for conceptualization, methodology, software development, investigation, formal analysis and writing the original draft; Huidong Wu contributed to data curation, resource provision and investigation.

Funding

This research received no external funding.

Data availability

Not applicable.

Ethics approval and consent to participate

This study was approved by the Institutional Review Board (IRB) at the Kunming Medical University (Approval No: KMMU2023MEC088). Written informed consent was obtained from all participants for publication of relevant data.

Consent for publication

All authors have given their consent for publication of this manuscript.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

We thank all the participants included in this study.

REFERENCES

[1] Zheng Y, Dang Y, Wu X, Yang Y, Reinhardt J, He C, et al. Epidemiological study of adolescent idiopathic scoliosis in Eastern China. J Rehabil Med. 2017 Jun 28;49(6):512-519. https://doi.org/10.2340/16501977-2240

[2] Burwell RG. Aetiology of idiopathic scoliosis: Current concepts. Pediatr Rehabil. 2003 Jul-Dec;6(3-4):137-170. https://doi.org/10.1080/13638490310001642757

[3] Tang Y, Yang S, Chen C, Luo K, Chen Y, Wang D, et al. Assessment of the association between paraspinal muscle degeneration and quality of life in patients with degenerative lumbar scoliosis. Exp Ther Med. 2020 Jul;20(1):505-511. https://doi.org/10.3892/etm.2020.8682

[4] Modi H, Suh S, Yang JH, Hong JY, Venkatesh K, Muzaffar N. Spontaneous regression of curve in immature idiopathic scoliosis - does spinal column play a role to balance? An observation with literature review. J Orthop Surg Res. 2010 Nov 4;5:80. https://doi.org/10.1186/1749-799x-5-80

[5] Liu Y, Pan A, Hai Y, Li W, Yin L, Guo R. Asymmetric biomechanical characteristics of the paravertebral muscle in adolescent idiopathic scoliosis. Clin Biomech (Bristol). 2019 May;65:81-86. https://doi.org/10.1016/j.clinbiomech.2019.03.013

[6] Banno T, Arima H, Hasegawa T, Yamato Y, Togawa D, Yoshida G, et al. The effect of paravertebral muscle on the maintenance of upright posture in patients with adult spinal deformity. Spine Deform. 2019 Jan;7(1):125-131. https://doi.org/10.1016/j.jspd.2018.06.008

[7] Fidler M, Jowett R. Muscle imbalance in the aetiology of scoliosis. J Bone Joint Surg Br. 1976 May;58-B(2):200-201. https://doi.org/10.1302/0301-620x.58b2.932082

[8] Wong C, Gosvig K, Sonne-Holm S. The role of the paravertebral muscles in adolescent idiopathic scoliosis evaluated by temporary paralysis. Scoliosis Spinal Disord. 2017 Oct 10;12:33. https://doi.org/10.1186/s13013-017-0138-7

[9] Feipel V, Aubin C, Ciolofan O, Beauséjour M, Labelle H, Mathieu PA. Electromyogram and kinematic analysis of lateral bending in idiopathic scoliosis patients. Med Biol Eng Comput. 2002 Sep;40(5):497-505. https://doi.org/10.1007/BF02345446

[10] Farahpour N, Younesian H, Bahrpeyma F. Electromyographic activity of erector spinae and external oblique muscles during trunk lateral bending and axial rotation in patients with adolescent idiopathic scoliosis and healthy subjects. Clin Biomech (Bristol). 2015 Jun;30(5):411-417. https://doi.org/10.1016/j.clinbiomech.2015.03.018

[11] Kwok G, Yip J, Cheung MC, Yick KL. Evaluation of myoelectric activity of paraspinal muscles in adolescents with idiopathic scoliosis during habitual standing and sitting. Biomed Res Int. 2015 Jan;2015:958450. https://doi.org/10.1155/2015/958450

[12] Murayama M, Watanabe K, Kato R, Uchiyama T, Yoneda T. Association of muscle hardness with muscle tension dynamics: A physiological property. Eur J Appl Physiol. 2011 Apr 17; 112(1):105-112. https://doi.org/10.1007/s00421-011-1959-3

[13] Bisset LM, Evans K, Tuttle N. Reliability of 2 protocols for assessing pressure pain threshold in healthy young adults. J Manipulative Physiol Ther. 2015 May;38(4):282-287. https://doi.org/10.1016/j.jmpt.2015.03.001

[14] Calvo-Muñoz I, Gómez-Conesa A, Sánchez-Meca J. Prevalence of low back pain in children and adolescents: A meta-analysis. BMC Pediatr. 2013 Jan 26;13(1):14. https://doi.org/10.1186/1471-2431-13-14

[15] Liang Q, Xu Y, Wang C, Yang L, Lin K. The patterns of time domain indexes of surface electromyographic signals of adolescent idiopathic scoliosis patients' paraspinal muscles. Chin J Rehabil Med. 2009;24(11):989-991.

[16] Yuan W, Chen L, Shen J, Wang H, Yu K, Liu Y, et al. Correlation of surface electromyogram signals of the apex vertebral paraspinal muscle with cobb angle and axial trunk rotation angle in adolescent idiopathic scoliosis patients. Chin J Tissue Eng Res. 2019;23(24):3824-3828. https://doi.org/10.3969/j.issn.2095-4344.1223

[17] Butterworth TR Jr, James C. Electromyographic studies in idiopathic scoliosis. South Med J. 1969 Aug;62(8):1008-1010. https://doi.org/10.1097/00007611-196908000-00026

[18] Shafaq N, Suzuki A, Matsumura A, Terai H, Toyoda H, Yasuda H, et al. Asymmetric degeneration of paravertebral muscles in patients with degenerative lumbar scoliosis. Spine.

2012 Jul 15;37(16):1398-1406. https://doi.org/10.1097/BRS.0b013e31824c767e

[19] Acaroglu E, Bobe R, Enouf J, Marcucio R, Moldovan F, Moreau A. The metabolic basis of adolescent idiopathic scoliosis: 2011 report of the "metabolic" workgroup of the fondation yves cotrel. Eur Spine J. 2012 Jun;21(6):1033-1042. https://doi.org/10.1007/s00586-012-2245-8

[20] Chwała W, Koziana A, Kasperczyk T, Walaszek R, Płaszewski M. Electromyographic assessment of functional symmetry of paraspinal muscles during static exercises in adolescents with idiopathic scoliosis. Biomed Res Int. 2014 Sep 2;2014:573276. https://doi.org/10.1155/2014/573276

[21] Sakai F, Ebihara S, Akiyama M, Horikawa M. Pericranial muscle hardness in tension-type headache: A non-invasive measurement method and its clinical application. Brain. 1995 Apr; 118(2):523-531. https://doi.org/10.1093/brain/118.2.523

[22] Fan Y, Ren Q, To M, Cheung J. Effectiveness of scoliosis-specific exercises for alleviating adolescent idiopathic scoliosis: A systematic review. BMC Musculoskelet Disord. 2020 Jul 27; 21(1):495. https://doi.org/10.1186/s12891-020-03517-6

[23] Ashina M, Bendtsen L, Jensen R, Sakai F, Olesen J. Muscle hardness in patients with chronic tension-type headache: Relation to actual headache state. Pain. 1999 Feb 1;79(2):201-205. https://doi.org/10.1016/S0304-3959(98)00167-5

[24] Leonard C, Brown J, Price T, Queen S, Mikhailenok E. Comparison of surface electromyography and myotonometric measurements during voluntary isometric contractions. J Electromyogr Kinesiol. 2004 Dec;14(6):709-714. https://doi.org/10.1016/j.jelekin.2004.06.001

[25] Winter R, Lonstein J, Denis F. Pain patterns in adult scoliosis. Orthop Clin North Am. 1988 Apr;19(2):339-345. https://doi.org/10.1016/S0030-5898(20)30313-8

[26] Andersen H, Ge H, Arendt-Nielsen L, Danneskiold-Samsøe B, Graven-Nielsen T. Increased trapezius pain sensitivity is not associated with increased tissue hardness. J Pain. 2010 May; 11(5):491-499. https://doi.org/10.1016/j.jpain.2009.09.017

Progress in
Medical Devices

Zentime

REVIEW ARTICLE

Advances in endoscopic closure devices for postoperative gastrointestinal defects

Yuxiao Li, Junjie Shen, Yuxuan Hou, Shilong Li, Chengli Song, Lin Mao

Shanghai Institute for Minimally Invasive Therapy, School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Lin Mao.

Address correspondence to: Lin Mao, Shanghai Institute for Minimally Invasive Therapy, School of Health Science and Engineering, University of Shanghai for Science and Technology, No. 516 Jungong Road, Yangpu District, Shanghai 200093, China. Tel: +86-21-55572159.
E-mail: linmao@usst.edu.cn.

Received November 28, 2025; Accepted February 7, 2026; Published June 17, 2026

DOI: 10.61189/030332bzapdu

Abstract

Advanced endoscopic therapeutic techniques, such as endoscopic mucosal resection, endoscopic submucosal dissection, and third-space endoscopy, have transformed digestive endoscopy into a minimally invasive surgical platform. This transformation, nonetheless, poses a major challenge, namely the safe and efficient closure of the postoperative gastrointestinal (GI) defects that arise, as inadequately closed defects may lead to serious postoperative complications. This review provides a comprehensive overview of the historical background, the basic principles of closure, and the technical changes of the mainstream devices used for closure of GI defects, viewed mostly from the perspective of medical devices and engineering design. We have divided the most significant closure devices currently in use into endoscopic clipping devices and endoscopic suturing devices, and we have outlined their respective historical development and working principles. In addition, this review has identified and explained advanced clip-based closure methods that employ multiple through-the-scope clips (TTSCs) or are combined with auxiliary devices to overcome the inherent limitations of single-TTSC use. This review aims to provide endoscopists with a deeper understanding of the existing closure devices so as to facilitate their optimal clinical use, and to offer insights that may help in the development of next-generation closure technologies.

Keywords: Endoscopy, Postoperative defects, Through-the-scope clip, Over-the-scope clip, Endoscopic suturing device

Highlights

- Current innovation in gastrointestinal defect management falls into two main categories: clip-based mechanical compression devices and advanced high-precision endoscopic suturing systems.
- Through-the-scope clips combined with auxiliary devices offer a strategic solution to overcome the size limitations inherent to single-device closure strategies for larger or complex defects.
- Future technological development should focus on enhancing reliability, operability, cost-effectiveness, and overall user-friendliness of advanced endoscopic closure devices to broaden their clinical applicability.

1 INTRODUCTION

Digestive endoscopy has evolved beyond tissue sampling and excision to become a surgical platform for delivering therapeutic devices to treat a variety of gastrointestinal (GI) disorders [1]. The emergence of techniques like endoscopic mucosal resection (EMR) and endoscopic submucosal dissection (ESD) has made the resection of early GI tumors possible [2, 3]. The

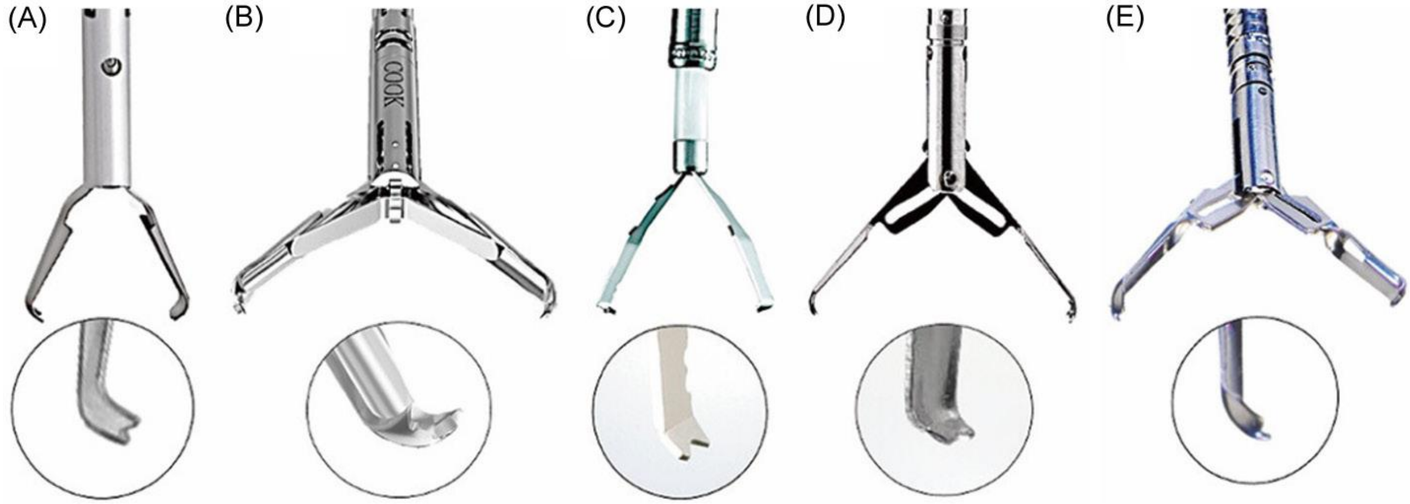


Figure 1. Different types of through-the-scope endoscopic clips. (A) Resolution 360 (Boston Scientific); (B) Instinct Plus (Cook Medical); (C) EZ Clip (Olympus); (D) SureClip (Micro-Tech); (E) DuraClip (ConMed). This figure is cited from [136].

popularization of endoscopic full-thickness resection (EFTR) and of third-space endoscopy techniques such as peroral endoscopic myotomy, have established digestive endoscopy as a natural orifice transluminal endoscopic surgery platform [4-6].

While these advanced techniques offer immense clinical benefits, a significant challenge arises: the management of complicated postprocedural defects [7-9]. Endoscopic resection procedures inevitably create large, deep, or even full-thickness defects [10, 11]. Postoperative defects, especially iatrogenic endoscopic perforations, can lead to extremely severe complications [11, 12]. Therefore, the timely closure of defects can effectively reduce postoperative adverse events [13, 14]. Traditional closure devices are often insufficient for complex defects. To meet these challenges, endoscopic closure devices and related technologies have evolved in parallel with advancements in endoscopic surgery [15]. The endoscopist's armamentarium has expanded from the early through-the-scope clips (TTSCs) for closing small defects, to over-the-scope clips (OTSCs) with greater closure strength, and subsequently to endoscopic suturing devices capable of achieving "surgical-style" closure.

To date, many reviews have introduced commercially available closure devices from a clinical perspective, providing strategies for device selection based on different types of GI defects [16-20]. This review, however, focuses on the development process and working principles of closure devices. It reviews the closure principles and development history of endoscopic clipping devices and endoscopic suturing devices, and also introduces advanced clip-based closure methods that combine TTSCs with auxiliary devices. This review aims not only to guide endoscopists in the rational selection and utilization of their armamentarium but also to provide insights for both clinicians and engineers to develop next-generation closure devices and methods based on current instrumentation.

2 ENDOSCOPIC CLIPPING DEVICES

2.1 TTSCs

Since their first report by Hayashi et al. in 1975, TTSCs have evolved for approximately 50 years [21]. As indications have continuously expanded, TTSCs are now the most common and widely used closure devices in diagnostic and therapeutic digestive endoscopy (**Figure 1**) [22]. A regular TTSC comprises three key components: a metallic clip for tissue approximation, a delivery catheter based on tendon-sheath transmission, and an operating handle to control clip deployment [23, 24]. All TTSCs are passed through the instrument channel and exit at the distal end of the endoscope to reach the target lesion.

The basic working principle of TTSCs is to compress tissue by using the mechanical force generated upon closure, thereby achieving hemostasis or the approximation of defect edges. Different brands and models of TTSCs possess specific geometrical parameters, physical properties, or additional features, such as rotatability, compression strength, and compatibility with magnetic resonance imaging (**Table 1**) [25, 26].

Early applications of TTSCs demonstrated favorable therapeutic results, but a series of limitations hindered their clinical adoption, such as cumbersome manipulation, low clip-retention rate, and the inability to rotate the clip to the designated axial direction [27-29]. In the 1990s, Hachisu et al., in collaboration with Olympus (Tokyo, Japan), developed a prototype and a clinical model of the rotatable TTSC, which improved the device's maneuverability [30]. In the 21st century, TTSC development has concentrated on integrating

Table 1. Specifications and characteristics of currently available TTSCs

Product name	Manufacturer	Minimum compatible instrument channel diameter (mm)	Maximum opening width (mm)	Working length (mm)	MR status	Delivery system type	Repeated opening	Rotational capability
QuickClip Pro	Olympus (Tokyo, Japan)	2.8	11	1,650 (Upper GI), 2,300 (Lower GI)	MR conditional	Disposable	Yes	Yes
Instinct Plus	Cook Medical (NC, USA)	2.8	16	2,300	MR conditional	Disposable	Yes	Yes
Resolution 360	Boston Scientific (Massachusetts, USA)	2.8	11	1,550 2,350	MR conditional	Disposable	Yes	Yes
SureClip	Micro-Tech (Nanjing, China)	2.8	8, 11, 16, 17	2,350	MR conditional	Disposable	Yes	Yes
Lockado	Micro-Tech (Nanjing, China)	2.8	11, 16, 22	2,350	MR conditional	Disposable	Yes	Yes
DuraClip	ConMed (NY, USA)	2.8	11, 16	2,350	MR conditional	Disposable	Yes	Yes
EZ Clip	Olympus (Tokyo, Japan)	2.8	7, 9.5, 11	1,650 2,300	MR conditional	Reusable	No	Yes
MANTIS	Boston Scientific (Massachusetts, USA)	2.8	11	2,350	MR conditional	Disposable	Yes	Yes
Dual action tissue	Micro-Tech (Nanjing, China)	3.2	15	1,650 1,950 2,350	MR conditional	Disposable	Yes	Yes

Note: TTSC, through-the-scope clip; MR, magnetic resonance; GI, gastrointestinal.

enhanced functional characteristics, including pre-loading, rotatability, re-opening capability, and reliable deployment [31]. Representative examples include the QuickClip series (Olympus, Tokyo, Japan), the Resolution Clip (Boston Scientific, Massachusetts, USA), the Instinct Clip (Cook Medical, NC, USA), and the SureClip (Micro-Tech, Nanjing, China) [25, 32, 33].

Most TTSCs consist of two prongs, whereas in 2003 Cook Medical (Winston-Salem, NC, USA) introduced the three-pronged TriClip, designed to more easily grasp protruding lesions and potentially reduce the number of clips used [34]. However, due to its higher cost and reportedly inferior hemostatic effect compared to regular TTSCs, two-pronged TTSCs remain the preferred choice [35].

Although the clip components themselves are typically disposable, TTSC systems can be categorized based on the reusability of the delivery system into single-use disposable systems and reusable systems with a reloadable applicator. Common reusable TTSCs include the EZ Clip (Olympus Medical Systems Corp., Tokyo, Japan) and the ZEOCLIP (Zeon Medical, Tokyo, Japan). The EZ Clip was one of the earliest reloadable, rotatable TTSCs, while the ZEOCLIP reportedly achieves a larger opening width faster and easier than the EZ Clip [36]. Furthermore, the ZEOCLIP can be modified into a fluorescent marking clip for tumor localization [37]. The recently reported re-openable SB clip (Sumitomo Bakelite Co., Ltd., Tokyo, Japan) also belongs to the category of reusable TTSCs [38]. Its deployment, rotation, and price are comparable to the EZ Clip, and it can close large mucosal defects via hold-and-drag closure [39].

The above-mentioned TTSC systems, whether disposable or reloadable, typically deploy only one clip before the catheter must be reloaded or replaced. In contrast, multi-firing TTSCs are less common, such as the InScope Multi-Clip Applier (Ethicon Endo-Surgery, Cincinnati, Ohio, USA) and the Clipmaster3 (Medwork, GmbH, Aisch, Germany) [40, 41]. These devices make it possible to insert multiple clips during one single operation, and thus a withdrawal and reload are not necessary for every clip [40, 41]. However, a related article reported that multi-firing TTSCs might not actually shorten the time of the procedure, probably due to the limited experience, and their effectiveness still needs to be confirmed by more clinical trials [41].

Regular TTSCs are known for their cost-efficiency, ease of operation, and minimal limitations on the maneuverability of the endoscope [42]. Their opening width and closure strength, however, are limited, so they are mainly used for closing small luminal defects [43, 44]. In general, it is advisable to use advanced clip-based methods combined with other devices for defects larger than 10 mm [15, 45].

Currently, TTSCs are designed with several sophisticated mechanical and functional features. For example, the through-the-scope twin clip, made by Micro-Tech (Nanjing, China), comprises two pairs of prongs which can be independently opened and closed (**Figure 2**) [46]. Each prong can individually

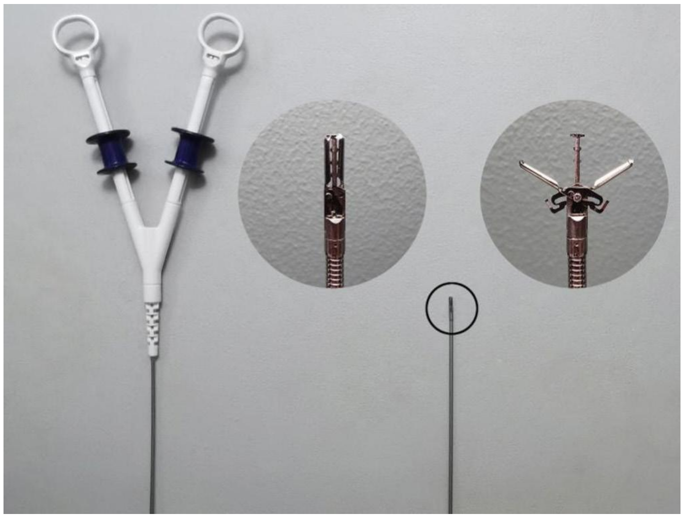


Figure 2. Photograph of the through-the-scope twin clip with an enlarged view of its distal end. This figure is cited from [47].

hold one side of the defect, and thus the area of large defects is effectively reduced and final closure becomes easier [47]. Another notable innovation is the MANTIS Clip (Boston Scientific, Marlborough, MA, USA), specifically designed for the hold-and-drag closure technique [48]. The clip features a mantis-like claw design to enable not only secure tissue engagement but also retraction [48]. The MANTIS Clip, as opposed to regular TTSCs, facilitates full-thickness closure, shortens closure time, and increases post-closure strength, while minimizing the created submucosal dead space [49]. Relevant studies even suggest that its closure strength is better than that of OTSCs in the management of 30 mm perforations [42].

2.2 OTSCs

Due to the size restrictions imposed by the endoscopic instrument channel diameter and the mechanical structure of TTSCs, as well as their inability to achieve full thickness closure, the limited closure strength of TTSCs remains a challenge. Theoretically, closure effectiveness can be improved by increasing the number of deployed TTSCs, but this method will correspondingly increase the cost of defect closure, especially since TTSCs may fail during anchoring or misfire [50]. In an effort to eliminate the limit of TTSCs, OTSCs were subsequently developed, and the first clinical application was reported in 2007 [51].

As an over-the-scope technique, OTSC systems do not pass through the instrument channel but are mounted on the outer side of the endoscopic distal end before insertion. Because of the shape memory properties of the material, the OTSC, which is usually made of nitinol, is held in an open configuration before release; after deployment, it returns to its original closed shape, thus providing continuous mechanical closure force, comparable to that of manual surgical suturing [52-54].

Figure 3. Ovesco OTSC (left) and the Padlock clip (right) in their deployed configurations. This figure is cited from [137]. OTSC, over-the-scope clip.

The principle for defect closure across the most common commercially available OTSC systems, such as the OTSC System (Ovesco Endoscopy AG, Tuebingen, Germany) and the Padlock Clip System (Aponos Medical Co., Kingston, NH, USA), remains consistent: the edges of the defect are drawn into an application cap, where the tissue is compressed, achieving complete defect closure with the OTSC (**Figure 3**) [51, 55]. In terms of structural design, the Ovesco OTSC includes teeth similar to a bear claw, thus achieving strong tissue approximation [56]. In addition, the Ovesco OTSC has three different tooth geometries (type a, t, gc) with different uses: type a with blunt teeth for compression, type t with sharp teeth for compression and anchoring, and type gc for gastric wall closure (**Figure 4**) [57, 58]. In contrast, the Padlock Clip has a different structure that, upon release, creates a hexagonal ring, thereby providing 360-degree radial puncture force to the tissue through six inner needles pointing toward the center [59]. Furthermore, the tips are equipped with tissue controllers so that the depth of the puncture is not excessive, and the space between the needles allows blood to flow into the tissue, thus accelerating the healing process [59].

As for the mounting procedure, the preparation of OTSC systems is more complicated than that of TTSCs, as it requires assembling the components outside the body before insertion. For both OTSC systems, the components are fitted onto the endoscopic distal end by means of an application cap [55, 60]. However, there is a difference in the release mechanism between these two systems: the Ovesco OTSC's trigger wire is located inside the instrument channel and is activated by a hand wheel installed on the endoscopic grip section (**Figure 5**), whereas the Padlock Clip's trigger wire is external to the endoscope and is released by the push-button Lock-It delivery system (Aponos Medical, Kingston, NH, United States) (**Figure 6**) [55, 59, 60].

After the installation of the OTSC system on the endoscope, the device is moved to the site of the defect. At the target defect, the edges of the lesion are drawn into the distal application cap either by suction or with the help of suitable auxiliary devices

Figure 4. OTSC clip types. (A) Type a (atraumatic); (B) Type t (traumatic); (C) Type gc (gastric closure). This figure is cited from [58].

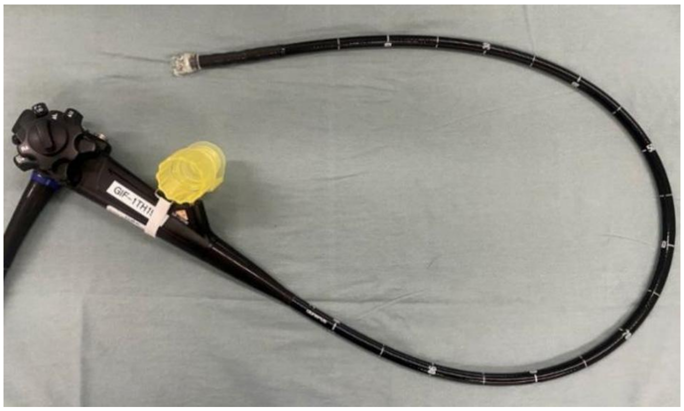


Figure 5. Ovesco OTSC loaded on a therapeutic upper endoscope. The distal end of the endoscope is attached to an application cap loaded with the OTSC, and the grip section is fitted with the hand wheel. This figure is cited from [138].

[57]. Subsequently, the endoscopist deploys the OTSC from the application cap by activating the specific trigger mechanism, rotating the hand wheel for the Ovesco OTSC or pressing the button on the Lock-It delivery system for the Padlock Clip, thereby compressing the tissue and achieving complete defect closure [57, 61].

Successful OTSC closure is predicated on the effective and secure capture of the target lesion into the application cap. Therefore, choosing an optimal tissue approximation strategy is crucial and requires consideration of the defect size, the duration since onset, and the specific clinical indication [62]. Common approximation methods include simple suction, twin-grasper assistance, and anchor assistance [62-64]. Simple suction offers clear advantages including convenience, rapid application, and low cost. However, its maximum size limit for complete closure is typically no more than 10 mm, and it may sometimes result in failure to achieve full-thickness closure [62]. Twin-grasper assistance and anchor assistance require specific auxiliary instruments and are typically used for large or full-thickness defects [63, 65]. While twin-grasper assistance exhibits limitations when managing large, fibrotic defects,

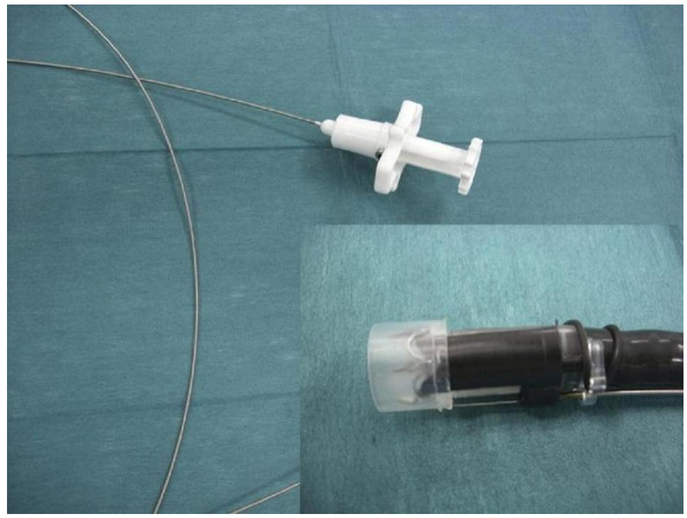


Figure 6. The Lock-It delivery system. The system consists of an application cap at the distal end of the endoscope and a trigger wire (parallel to the scope) that connects to a handle. This figure is cited from [139].

anchor assistance proves highly effective for capturing fibrotic or indurated tissue [64, 66].

Compared to TTSCs, OTSCs provide a more durable and forceful full-thickness closure. OTSCs are a safe and effective treatment method for refractory GI conditions such as perforations, leaks, severe acute GI bleeding, and recurrent ulcer bleeding [53, 55, 66-68]. However, OTSCs are more expensive than regular TTSCs, have a lower margin for error (allowing only one attempt), are difficult and costly to remove, and are less adaptable for very large defects, despite their efficacy for small- to medium-sized perforations [8, 42].

3 ENDOSCOPIC SUTURING DEVICES

3.1 Through-the-scope suturing devices

Due to the diameter constraints of the endoscopic instrument channel, the size of through-the-scope suturing devices is lim-

Figure 7. The SutuArt through-the-scope endoscopic needle holder. This figure is cited from [140].

ited, making it difficult to achieve continuous suturing with complex mechanical structures. Thus, early devices primarily utilized a hollow needle and T-tag to achieve tissue approximation rather than double full-thickness surgical closure [69-71]. When utilizing a single-channel endoscope, the suturing system typically comprised a semiflexible hollow needle catheter, two metal T-tags attached to a polypropylene thread placed inside the hollow needle, and a thread-locking device to eliminate the need for traditional knot tying, enabling the successful closure of full-thickness defects after EFTR [69]. Alternatively, when a double-channel endoscope is employed, the suturing system could consist of two flexible hollow needle catheters and three T-tags interconnected by a Y-shaped thread [70]. Both hollow needles were used to traverse the tissue on the two opposite sides of the defect and to deploy the corresponding T-tags [72]. Subsequently, the proximal movable T-tag is pushed along the Y-shaped thread by the outer sheath that is advanced to close the defect margins for final closure [72]. Although the first device models had relatively simple mechanical structures, they were conceptually limited to performing interrupted rather than continuous suturing. Furthermore, they lacked robust clinical data that could confirm their stability and safety.

In an effort to achieve continuous suturing endoscopically, endoscopists tried to adapt the needle-and-thread suturing method of traditional surgery for use in flexible endoscopy. In 2014, Goto et al. were the first to report the utilization of the SutuArt (Olympus Medical Systems Corp, Tokyo, Japan) prototype and a semicircular needle attached to an absorbable barbed suture for endoscopic hand suturing (EHS), verifying the method's feasibility and effectiveness in ex vivo experiments [73]. SutuArt is a through-the-scope, flexible, and rotatable needle driver. The handle of the SutuArt controls the opening and closing of the distal jaws, which are used for a stable needle grasp, and also has a locking mechanism (**Figure 7**) [73].

EHS is capable of securing defect closure with high safety and strength. Therefore, it is a suitable candidate for large defects (measuring 20 mm to 30 mm or more) at various locations and depths. Consequently, it shows enormous potential in the field of advanced gastrointestinal endoscopic surgery [15, 74, 75]. However, the dissemination of this technique is limited by several practical problems, most prominently, the prolonged time required for the procedure and the technical complexity of EHS, which together necessitate a high degree of operator skill and a long period of training [74, 76, 77].

Recent advancements in endoscopic technology have given rise to new through-the-scope suturing devices. One of these innovations is the X-Tack (Apollo Endosurgery, Inc, Austin, Tex, USA), a through-the-scope continuous suturing device designed for efficient and secure closure of large or irregularly shaped GI defects (**Figure 8**) [78, 79]. The main parts of the X-Tack system comprise an applicator that converts the linear motion of the handle slider into rotational motion, four 5-mm-long surgical steel laser-cut helix tacks connected to the suture, and a cinch device used to tighten and secure the suture [80].

During the procedure, the endoscopist uses the preloaded applicator to rotate and anchor the first helix tack into the tissue at the defect margin, thereby securing the initial end of the suture. The applicator catheter is then withdrawn and reloaded externally with the subsequent helical tacks, and the anchoring process is repeated at the defect edges. The subsequent helix tacks are preloaded on the suture and can slide along it. Therefore, a final cinch device is used to tighten the suture, thereby approximating the anchored helix tacks and achieving defect closure. After securing the other end of the suture with a T-tag deployed from the cinch device, the entire closure construct is stabilized, thus preventing re-opening and ensuring efficient and reliable closure [79].

Similar to EHS, X-Tack has been reported to have a high rate of clinical success [81]. However, as reported in a related study, several technical limitations may restrict its overall effectiveness [42]. For instance, difficulty in achieving full-thickness suturing results in a closure strength that is lower than that of TTSCs.

Figure 8. X-Tack suturing device. (A) X-Tack handle; (B) Pre-loaded helix tack. This figure is cited from [141].

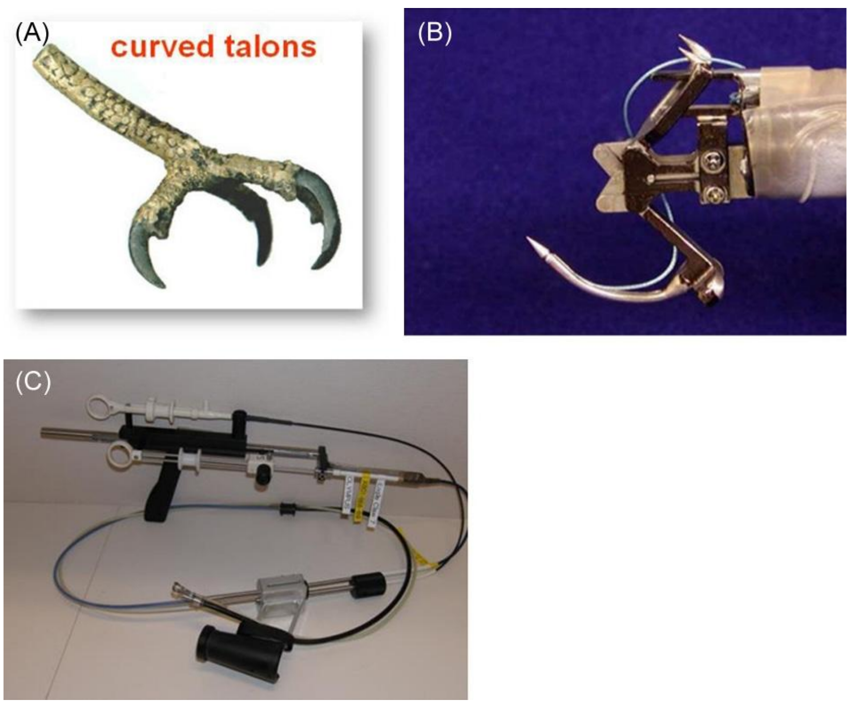


Figure 9. The eagle claw endoscopic suturing device. (A) Biological inspiration: curved talons; (B) Distal working end of the eagle claw device; (C) The complete system. This figure is cited from [141].

3.2 Over-the-scope suturing devices

As early as 1986, Swain et al. had already reported a chain-stitch suturing machine that was primarily made of rod-shaped transparent acrylic glass [82]. This suturing device is quite cumbersome and is fixed at the distal end of a double-channel endoscope. The tissue was drawn into the device's chamber by the vacuum. Then a suture-loaded needle and a flexible control cable for the catch mechanism were carefully guided through the two instrument channels to perform the tissue approximation [82, 83].

In 1994, Swain et al. reported a tag-based suturing device used for fixing biosensors to the GI tract wall [84]. The two most significant features of this device were a specially designed tag and a hollow needle. In 1996, Kadirkamanathan et al. further developed a new endoscopic suturing device based on this principle for anti-reflux gastroplasty procedures [85]. This device connected the T-tag to a nylon thread which, after deployment from the hollow needle, was received by a dedicated reception chamber, thereby facilitating the continuation of the suturing process following external reloading [85]. Although the above devices were not developed with the direct purpose of suturing GI tract defects, the mechanical structures, suturing methods, and knot-tying techniques proposed during their development provided a reference for the research and development of future over-the-scope suturing devices.

Thereafter, the Apollo Group and Olympus Corporation designed prototype suturing devices that could be visualized endoscopically, the Eagle Claw series (**Figure 9**) [86]. Comprising a curved needle and opposable jaws, the devices were mounted on the exterior of the endoscopic distal end. The needle-tip of Eagle Claw II was non-detachable and had a hole for the thread to pass through. When the jaws closed, the tissue was punctured, and the thread passed through the tissue with the needle. A retrieval hook then retrieved one side of the suture through an extra channel provided by the device. After extracorporeal knotting, the knot was advanced back into the body using a second endoscope fitted with a knot-pushing cap [87]. The Eagle Claw V improved upon the previous generation by adding a detachable needle-tip with a thread and a catching cartridge attached to the other end of the suture, thereby avoiding extracorporeal knots [86]. When the detachable needle-tip, carrying the suture thread, punctured the tissue, it was captured by a catching cartridge [88]. Subsequent iterations of the device retained the basic mecha-

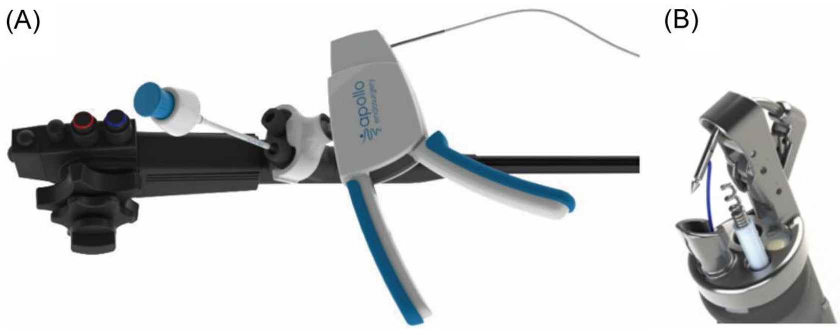


Figure 10. Double-channel OverStitch suturing device with a tissue-capturing helix. (A) The needle driver handle mounted on the endoscope; (B) The Endcap mounted on the endoscopic distal end. This figure is cited from [141].

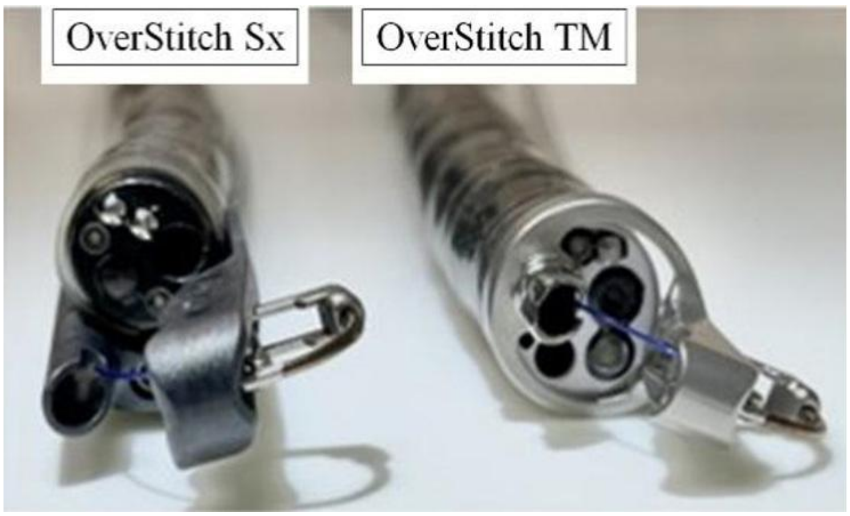


Figure 11. Comparison of the OverStitch Sx and OverStitch endoscopic suturing systems. This figure is cited from [142].

nism and principle of the Eagle Claw V, featuring no significant alterations, and its feasibility and efficacy for hemostasis and gastrotomy closure were validated in a porcine gastric model [89, 90].

Building upon the Eagle Claw, the OverStitch (Apollo Endosurgery, Inc, Austin, Tex, USA) received Food and Drug Administration approval in 2008 after years of development [91]. The OverStitch is capable of safely and reliably achieving continuous suturing of the GI tract, enabling its use for endoscopic sleeve gastroplasty, full-thickness GI wall suturing, closure of post-ESD/EMR defects, management of fistulas or leaks, stent fixation, and various other indications [92-95]. It has become one of the most widely used over-the-scope suturing devices [96]. In contrast to its predecessors, the Eagle Claw series, OverStitch is a cap-based device, and is therefore more compact and maneuverable (**Figure 10**) [97]. This design eliminates the need for the large external catheter assembly attached to the outside of the endoscope body, thereby maximizing the inherent endoscopic flexibility [97]. However, its first and second iterations were only compatible with specific double-channel Olympus endoscopes, which consequently impeded its widespread clinical adoption [98].

The continuous suturing mechanism of the OverStitch is engineered to emulate the operational mechanism of conventional manual suturing techniques for endoscopic application [91, 97]. The process begins as the unloaded needle driver closes and securely docks with the anchor exchange catheter within the instrument channel. The needle driver then reopens to retrieve the suture-attached anchor, thereby forming the functional operating unit. Following anchor retrieval, a helix device is deployed via the secondary instrument channel to capture and retract full-thickness tissue into the device's jaw. Subsequently, closure of the needle driver advances the anchor to puncture and traverse the captured tissue. Upon full traversal, the deployed anchor is captured by the anchor exchange catheter, which simultaneously facilitates its separation from the needle driver, concluding the successful double full-thickness passage of the suture thread. Thereafter, the needle driver is reopened, and the suture thread, having passed through the tissue, is returned to the instrument channel, completing a single stitch. This entire sequence may then be repeated in vivo to establish a continuous running suture line. To secure the continuous suture line, the suture-attached anchor is released while the needle driver remains open, serving as the initial T-tag to fasten the distal end. In contrast, a dedicated cinch device is required proximally to tighten the suture thread and deploy a second T-tag for definitive fixation.

Compared to the OverStitch system, the OverStitch Sx (Apollo Endosurgery, Inc, Austin, Tex, USA) features two independent, external working channels that effectively substitute the two channels required by a double-channel endoscope (**Figure 11**) [99]. This innovative design significantly enhances device versatility by enabling compatibility with a wide range of standard single-channel flexible endoscopes, thereby promoting broader accessibility and application [99]. However, the addition of this external device may slightly restrict the overall flexibility and maneuverability of the endoscope [100].

4 ADVANCED CLIP-BASED CLOSURE METHODS

The major drawback of regular TTSCs is their maximum opening width which is limited by the endoscopic working channel diameter. Whereas the clip span can be slightly increased by elongating the clip prongs, this physical limitation still consid-

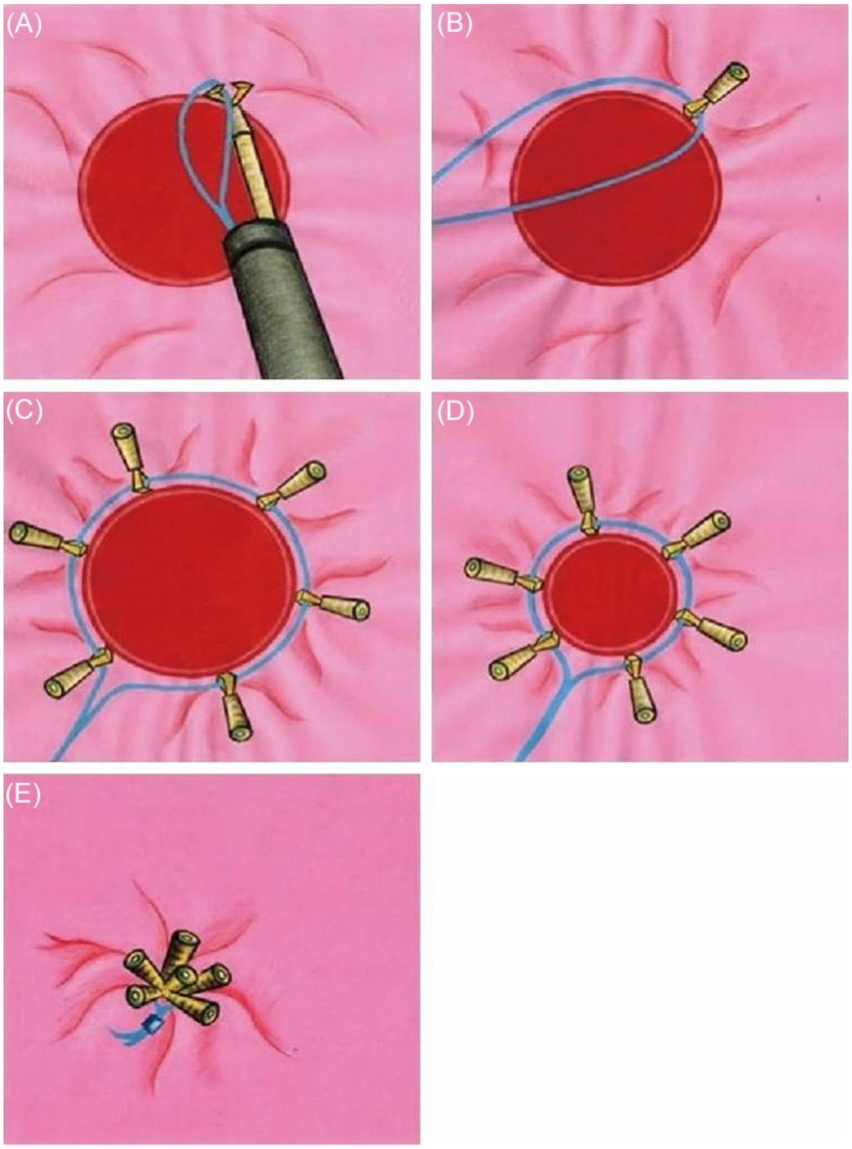


Figure 12. Schematic illustration of the EPSS procedure for closure of a gastrointestinal defect. (A, B) An initial clip is applied and an endoloop is positioned at the proximal margin of the defect; (C) Additional clips are deployed to secure the endoloop along the proximal edge; (D) The defect margins are approximated by tightening the endoloop; (E) Complete closure of the defect is achieved. This figure is cited from [102]. EPSS, endoscopic purse-string suture.

erably hampers their ability to close large mucosal or full-thickness defects. Compared with regular TTSCs, OTSCs provide a higher mechanical closure force, but they are not suitable for very large defects. Although endoscopic suturing devices can be used for larger defects, their higher economic cost and steep learning curve limit their widespread adoption in regular clinical practice.

As a result, over years of continuous clinical practice, endoscopists have come up with a multitude of advanced techniques to specifically surpass this limitation. These advanced clip-based closure techniques are systematically reviewed and presented in this paper. Depending on the main auxiliary instruments used for defect approximation, this review distinguishes clip-and-snare closure methods, loop-clip closure methods, clip-only closure methods, and clip-and-line closure methods.

4.1 Clip-and-snare closure methods

Clip-and-snare methods make use of regular clips in combination with detachable snares or similar auxiliary devices to achieve defect closure. In this approach, the method exerts continuous radial tension on the defect edges, which leads to the effective approximation of large defects.

The first use of a clip and snare for defect closure, in combination with a double-channel colonoscope, was reported by Matsuda et al. in 2004 [101]. An endoloop (Olympus, Tokyo, Japan) was placed around the defect, and regular clips were used to fix the two ends of the snare separately to the margins of the defect. Subsequent tightening of the endoloop provided the required radial tension for preliminary defect closure.

By inventing the endoscopic purse-string suture (EPSS) method in 2014, Zhang et al. significantly advanced the clip-and-snare closure method (**Figure 12**) [102]. The method involved placing multiple clips around the edges of the defect to fix the endoloop. The edges of the defect were drawn together by tightening the snare, and the method was confirmed effective for closing gastric wall defects less than 40 mm in diameter [102].

The initial dependence on a double-channel endoscope to allow simultaneous handling of both instruments limited the widespread adoption of the clip-and-snare method. However, Nomura et al. described an innovative single-channel endoscope method in 2018 [103]. Firstly, the operation begins with fastening the separated endoloop to one side of the clip using a surgical thread. The clip-connected endoloop is then withdrawn into the clip's outer sheath, passed through the endoscope channel to the defect site, deployed from the sheath, and fixed to one side of the defect. The closure is completed using the standard purse-string steps.

After the EPSS had been established, endoscopists made various alterations to the technique. For example, placing the endoloop and the clips as a preventive measure before perform-

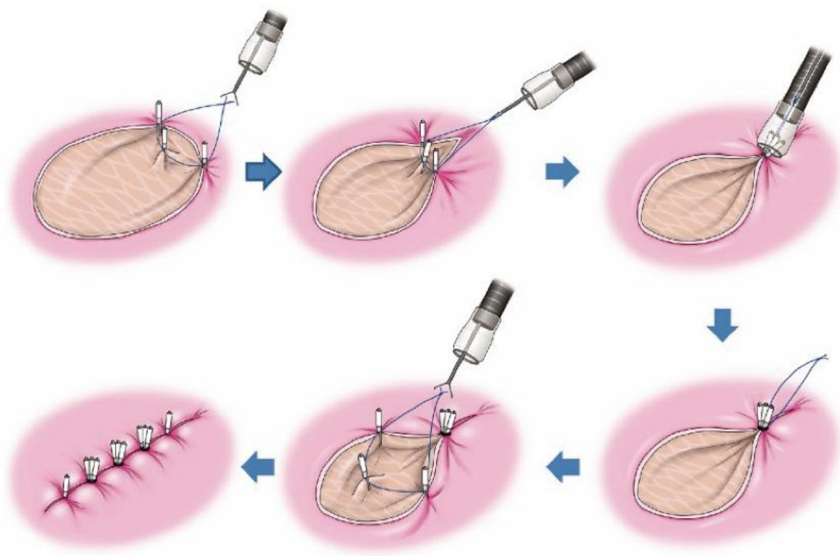


Figure 13. Schematic illustration of the endoscopic ligation with O-ring closure procedure. This figure is cited from [109].

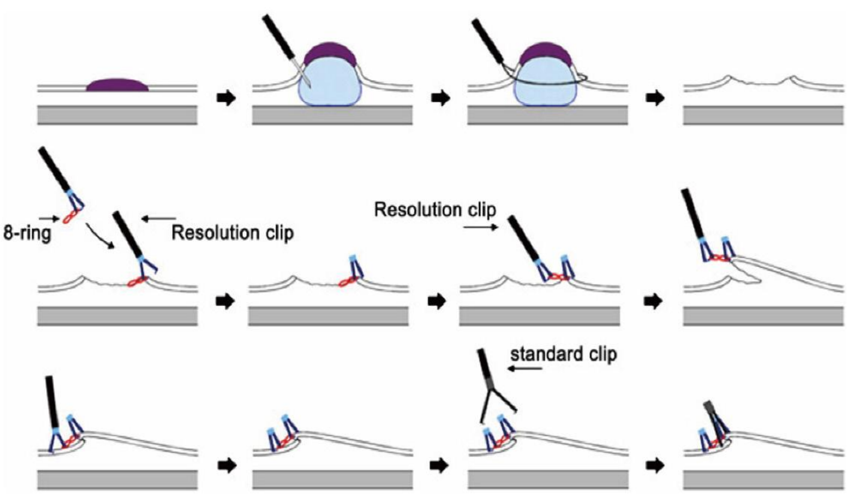


Figure 14. EMR and subsequent closure of the resulting mucosal defect using a figure-of-8-shaped metallic loop in combination with two resolution clips. This figure is cited from [111]. EMR, endoscopic mucosal resection.

ing EFTR helps the subsequent closure of the created defect, thereby shortening the procedure time [104, 105]. Additionally, dental floss can be used for clip traction to prevent the clip from inverting during endoloop tightening. Alternatively, an additional endoloop can be placed over the EPSS for circumferential bundling and securing of all clips, thus effectively closing defects that EPSS cannot adequately address [106-108].

For more secure closure, Nishiyama et al. developed a novel technique named endoscopic ligation with O-ring closure (**Figure 13**) [109, 110]. In this method, the surgical nylon loop used for clip tightening replaces the detachable snare, and the endoscopic variceal ligation device is employed to gather the clips [109]. This process is then repeated to achieve total closure of large defects.

4.2 Loop-clip closure methods

Loop-clip closure methods use clips that are either custom-made or commercially available and feature an integrated loop structure to achieve preliminary closure of large defects. Although terminology varies across the literature, this review refers to these devices collectively as "loop-clips".

By way of an example in 2007, Fujii et al. utilized a regular clip to grasp a specially designed figure-of-8-shaped metallic loop, advancing it through the instrument channel to close a post-EMR defect (**Figure 14**) [111]. This method laid the foundation for subsequent loop-clip development. In 2008, Sakamoto et al. first described the loop-clip as a new closure device, which involved both a metallic and a nylon loop [112]. The operation of loop-clip methods is inherently simpler and more convenient than clip-and-snare methods. The process involves deploying the loop-clip at the defect margin, subsequently inserting a regular clip through the instrument channel to grasp the loop. The loop is then pulled to the opposing normal mucosa and the second clip is deployed, thus achieving preliminary closure of the large defect [113].

The loop structure is highly versatile, with its material often consisting of metal, nylon, or rubber [111, 112, 114]. This variability accommodates diverse clinical scenarios, such as the closure method with an elastic ring reported by Kowazaki et al., which utilizes a clip with an attached elastic loop and also belongs to this category [114]. Regardless of the specific designs, the loop-clip essentially helps to turn a large defect into several smaller ones, which can then be closed individually with regular clips. In recent years, the loop-clip has been utilized to raise the mucosal layer and provide traction to the lesion, thus visually exposing more of the tissue that the endoscopist can dissect during ESD operations [115-117].

The most notable difference between loop-clip and clip-and-snare closure methods lies in how they apply force. Loop-clip closure methods use the loop as a medium. A second clip grips the loop of the first-deployed loop-clip and pulls it toward the opposing mucosa, thereby creating the traction. On the other hand, clip-and-snare closure methods utilize the loop structure

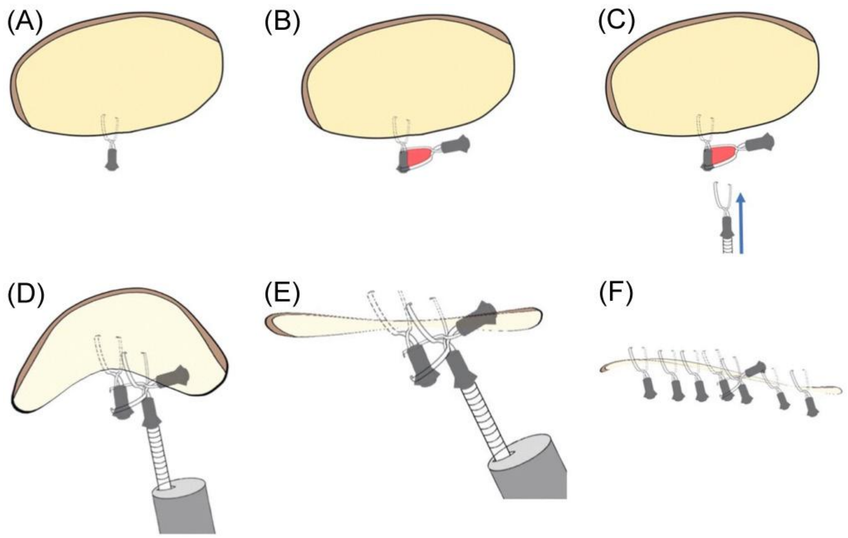


Figure 15. Schematic illustration of the CCCM. (A) A first clip is deployed on normal mucosa adjacent to the defect margin; (B) A second clip is attached to the stem of the first clip; (C, D) A third clip is advanced with one prong passing through the space between the jaws of the second clip to serve as a traction point; (E) The third clip is then secured to the contralateral normal mucosa to approximate the defect margins; (F) Complete closure is achieved with CCCM and additional clips. This figure is cited from [123]. CCCM, lip-on-clip closure method.

itself, which is fastened to the defect rims by several clips. The radial contraction force resulting from loop tightening thus brings the defect rims toward the center.

4.3 Clip-only closure methods

Clip-only closure methods refer to techniques for closing large defects using only multiple clips, without auxiliary devices or tools. The double-layered closure method for closing large mucosal defects after ESD and EMR was introduced by Tanaka et al. in 2012 [118]. They deploy a set of clips along the long axis of the defect in the submucosal layer, thus reducing the defect area, and then, another set of clips is interwoven in the mucosal layer to achieve complete defect closure [118]. In 2017, Banno introduced the mucosa–submucosal clip closure method [119]. The principle involves orienting a regular clip parallel to the defect's short axis, using its two prongs to hook the mucosa and submucosa separately, and then deploying it [119]. The process is carried out incrementally along the defect edge to bring the mucosal margins closer together, thus reducing the defect area that is finally closed with clips. Subsequent investigations, however, revealed that this method was limited to defects smaller than 50 mm and to those that were technically challenging [120, 121].

In 2018, the clip-on-clip closure method was proposed as an effective, straightforward, and safe technique (**Figure 15**) [122, 123]. First of all, a clip is deployed on the normal mucosa outside the defect circumference. Then, a second clip is placed on the first clip's stem. To finish, one prong of the third clip is inserted into the space between the two prongs of the second clip, pulled to the opposite mucosa, and released. After confirming stability, additional clips are deployed to fully close the mucosal defect. In principle, the clip-on-clip closure method is functionally analogous to loop-clip methods, because the second clip is used to establish the traction bridge instead of a loop structure. Notably, to date, there have been no direct comparative ex vivo or clinical studies among different types of clip-only closure methods.

4.4 Clip-and-line closure methods

Clip-and-line closure methods achieve defect closure through the synergistic action of regular clips and a line, typically a suture or dental floss. This technique relies on linear tension exerted on the line, causing the clips anchored at the defect margin to converge, thereby achieving tissue approximation.

In 2016, Yamasaki et al. reported the line-assisted complete closure (LACC) method [124]. This method is effective for closing mucosal defects and perforations, significantly reducing the incidence of post-endoscopic submucosal dissection coagulation syndrome [125]. LACC is performed with only one line and regular clips. After binding the line to one clip prong, the clip is deployed on the side of the defect. A second clip is subsequently used to grasp the line and deploy it on the opposite side. If more clips are needed, a zigzag pattern is used along the margin [126]. Continuous tension is then applied to the line, approximating the defect for final closure with additional regular clips. Further research has found that LACC can be a viable option for large perforations and defects [127-129].

In 2021, Nomura et al. introduced the reopenable-clip over the line method (ROLM), which employs a line and reopenable clips to close defects without creating submucosal dead space [130, 131]. Compared to LACC, in ROLM the suture for the second and subsequent clips must be threaded extracorporeally through the tooth hole of the reopenable clip before insertion through the instrument channel [132]. This design ensures the line and the clip remain continuously connected, which eliminates the need for challenging intraprocedural line grasping and provides a more robust and stable connection for subsequent tension and defect approximation. It is important to note that both LACC and ROLM require the use of additional clips for complete closure of the defect after manual line tensioning, as the defect will reopen if tension is released. Furthermore, completion of the entire procedure necessitates the endoscopic transection of the line, which is typically performed using specialized accessories such as loop cutters or scissor forceps.

To overcome this limitation, Nomura et al. successfully combined ROLM with a modified locking-clip technique in 2022 [133]. The locking-clip technique is a technique that uses special TTSCs with specific mechanical structures [134]. The clip's root is where the line is pre-threaded. When the clip is fully opened, the root of the clip and the line become firmly engaged. Therefore, endoscopists can first use ROLM with regular clips to gradually reduce defect area and approximate the tissue, and then they can apply modified locking-clip technique to tighten the line and fix it to the surrounding normal mucosa. After this important anchoring step, the line can be manually severed, thus allowing it to remain in vivo even under continuous high tension [133, 135].

5 CONCLUSION

The advancements in endoscopic closure devices, driven by clinician-engineer collaboration, have successfully transformed post-procedural GI defect management. Current innovation can be broadly categorized into two major directions: TTSCs and OTSCs and highly accurate instruments intended to imitate surgical suturing techniques. Within this landscape, the strategic use of TTSCs in combination with auxiliary devices effectively bypasses the mechanical constraints of single instruments, broadening their clinical applicability. Nonetheless, optimal clinical utilization demands a tailored approach, where endoscopists must integrate established procedural guidelines with defect complexity to judiciously select the most appropriate single device or combined closure strategy. Looking forward, although present closure devices, such as clips and emerging suturing systems, have shown functionality in various GI defects, the direction of future technological development should be that of enhanced reliability, operability, cost-effectiveness, and general ease of use of next-generation endoscopic closure devices.

DECLARATIONS

Author contributions

Yuxiao Li contributed to the conceptualization and design of the review, performed the comprehensive literature search and data collection, prepared all tables, and was responsible for writing and revising the entire manuscript as the first author; Junjie Shen and Yuxuan Hou contributed to the literature review, data organization, and revision of sections of the manuscript; Shilong Li assisted with the literature screening, data verification, and proofreading of the manuscript for clarity and accuracy; Chengli Song provided critical feedback on the structure and content, and critically revised the manuscript before submission; Lin Mao, as the corresponding author, provided the overall project concept and methodology guidance, validated the findings, and provided final critical review and editing.

Funding

This research received no external funding.

Data availability

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.

REFERENCES

[1] Ponsky JL, Strong AT. A history of flexible gastrointestinal endoscopy. Surg Clin North Am. 2020 Dec;100(6):971-992. https://doi.org/10.1016/j.suc.2020.08.013

[2] Rashid MU, Alomari M, Afraz S, Erim T. EMR and ESD: Indications, techniques and results. Surg Oncol. 2022 Aug; 43:101742. https://doi.org/10.1016/j.suronc.2022.101742

[3] Pimentel-Nunes P, Libânio D, Bastiaansen BAJ, Bhandari P, Bisschops R, Bourke MJ, et al. Endoscopic submucosal dissection for superficial gastrointestinal lesions: European Society of Gastrointestinal Endoscopy (ESGE) Guideline - Update 2022. Endoscopy. 2022 Jun;54(6):591-622. https://doi.org/10.1055/a-1811-7025

[4] Inoue H, Maydeo A. Peroral endoscopic myotomy (POEM) opens the door of third-space endoscopy. Endoscopy. 2019 Nov;51(11):1010-1012. https://doi.org/10.1055/a-1019-1865

[5] Wang K, Gao P, Cai M, Song B, Zhou P. Endoscopic full-thickness resection, indication, methods and perspectives. Dig Endosc. 2023 Jan;35(2):195-205. https://doi.org/10.1111/den.14474

[6] Cecinato P, Sinagra E, Laterza L, Pianigiani F, Grande G, Sassatelli R, et al. Endoscopic removal of gastrointestinal lesions by using third space endoscopy techniques. Best Pract Res Cl Ga. 2024 Aug;71:101931. https://doi.org/10.1016/j.bpg.2024.101931

[7] Oka S, Tanaka S, Kanao H, Ishikawa H, Watanabe T, Igarashi M, et al. Current status in the occurrence of postoperative bleeding, perforation and residual/local recurrence during colonoscopic treatment in Japan. Dig Endosc. 2010 Oct;22(4):376-380. https://doi.org/10.1111/j.1443-1661.2010.01016.x

[8] Lee JH, Kedia P, Stavropoulos SN, Carr-Locke D. AGA clinical practice update on endoscopic management of perforations in gastrointestinal tract: Expert review. Clin Gastroen-

terol Hepatol. 2021 Nov;19(11):2252-2261.e2252. https://doi.org/10.1016/j.cgh.2021.06.045
[9] Dalal I, Andalib I. Advances in endoscopic resection: A review of endoscopic submucosal dissection (ESD), endoscopic full thickness resection (EFTR) and submucosal tunneling endoscopic resection (STER). Transl Gastroenterol Hepatol. 2022 Apr 25;7:19. https://doi.org/10.21037/tgh-2020-10
[10] Yoshida N, Wakabayashi N, Kanemasa K, Sumida Y, Hasegawa D, Inoue K, et al. Endoscopic submucosal dissection for colorectal tumors: Technical difficulties and rate of perforation. Endoscopy. 2009 Sep 10;41(9):758-761. https://doi.org/10.1055/s-0029-1215028
[11] Fujihara S, Mori H, Kobara H, Nishiyama N, Matsunaga T, Ayaki M, et al. Management of a large mucosal defect after duodenal endoscopic resection. World J Gastroenterol. 2016 Aug 7;22(29):6595-6609. https://doi.org/10.3748/wjg.v22.i29.6595
[12] Sun H, Cao T, Zhang F, Tao K, Xu H. Gastric defect closure after endoscopic full-thickness resection: The closing while dissecting technique. Surg Endosc. 2023 Jan;37(1):234-240. https://doi.org/10.1007/s00464-022-09457-7
[13] Paspatis GA, Arvanitakis M, Dumonceau JM, Barthet M, Saunders B, Turino SY, et al. Diagnosis and management of iatrogenic endoscopic perforations: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement - Update 2020. Endoscopy. 2020 Sep;52(9):792-810. https://doi.org/10.1055/a-1222-3191
[14] Zhang Q, Han B, Xu J, Gao P, Shen Y. Clip closure of defect after endoscopic resection in patients with larger colorectal tumors decreased the adverse events. Gastrointest Endosc. 2015 Nov;82(5):904-909. https://doi.org/10.1016/j.gie.2015.04.005
[15] Spychalski M, Orzeszko Z, Kasprzyk P. Advancements in endoscopic closure: Embracing a new era of managing complications and postprocedural defects after endoscopic submucosal dissection. World J Gastrointest Endosc. 2025 Oct 16;17(10): 107840. https://doi.org/10.4253/wjge.v17.i10.107840
[16] Granata A, Martino A, Ligresti D, Zito FP, Amata M, Lombardi G, et al. Closure techniques in exposed endoscopic full-thickness resection: Overview and future perspectives in the endoscopic suturing era. World J Gastrointest Surg. 2021 Jul 27;13(7):645-654. https://doi.org/10.4240/wjgs.v13.i7.645
[17] Kobara H, Tada N, Fujihara S, Nishiyama N, Masaki T. Clinical and technical outcomes of endoscopic closure of postendoscopic submucosal dissection defects: Literature review over one decade. Dig Endosc. 2023 Jan;35(2):216-231. https://doi.org/10.1111/den.14397
[18] Staudenmann D, Choi KKH, Kaffes AJ, Saxena P. Current endoscopic closure techniques for the management of gastrointestinal perforations. Ther Adv Gastrointest Endosc. 2022 Feb 27;15:26317745221076705. https://doi.org/10.1177/26317745221076705
[19] Medas R, Rodrigues-Pinto E. Technical review on endoscopic treatment devices for management of upper gastrointestinal postsurgical leaks. Gastroenterol Res Pract. 2023 Jun 12; 2023(1):9712555. https://doi.org/10.1155/2023/9712555
[20] Papaefthymiou A, Norton B, Telese A, Murray C, Murino A, Johnson G, et al. Endoscopic suturing and clipping devices for defects in the GI tract. Best Pract Res Cl Ga. 2024 Jun;70:101915. https://doi.org/10.1016/j.bpg.2024.101915
[21] Hayashi T, Yonezawa M, Kuwabara T, kudoh I. The study on stanch clip for the treatment by endoscopy. Gastrointest Endosc. 1975 Feb 20;17(1):92-101. https://doi.org/10.11280/gee1973b.17.92
[22] Gong R, Wang S, Song J, He Z, Li P, Zhang S, et al. Closure methods for large defects after gastrointestinal endoscopic submucosal dissection. J Gastroenterol Hepatol. 2024 Dec; 39(12):2511-2521. https://doi.org/10.1111/jgh.16722
[23] Conway JD, Adler DG, Diehl DL, Farraye FA, Kantsevoy SV, Kaul V, et al. Endoscopic hemostatic devices. Gastrointest Endosc. 2009 May;69(6):987-996. https://doi.org/10.1016/j.gie.2008.12.251
[24] McCarty TR, Jirapinyo P. Endoscopic closure: Tools and techniques. Gastrointest Endosc Clin N Am. 2023 Jan;33(1):169-182. https://doi.org/10.1016/j.giec.2022.08.003
[25] Wang TJ, Aihara H, Thompson AC, Schulman AR, Thompson CC, Ryou M. Choosing the right through-the-scope clip: A rigorous comparison of rotatability, whip, open/close precision, and closure strength (with videos). Gastrointest Endosc. 2019 Jan;89(1):77-86.e71. https://doi.org/10.1016/j.gie.2018.07.025
[26] Shin DY, Park S, Kim A, Kim ES, Jeon HH. Compatibility of endoclips in the gastrointestinal tract with magnetic resonance imaging. Sci Rep. 2020 Oct 6;10(1):16537. https://doi.org/10.1038/s41598-020-73726-5
[27] Hajiro K, Matsui H, Tsujimura D. Endoscopic hemostasis with hemoclips, local injection and other new techniques: The Japanese experience. Endoscopy. 1986;18:65-67. https://doi.org/10.1055/s-2007-1018433
[28] Hachisu T. Evaluation of endoscopic hemostasis usingan improved clipping apparatus. Surg Endosc. 1988 Mar;2(1):13-17. https://doi.org/10.1007/BF00591392
[29] Hachisu T, Miyazaki S, Hamaguchi K. Endoscopic clip-marking of lesions using the newly developed HX-3L clip. Surg Endosc. 1989 Sep;3(3):142-147. https://doi.org/10.1007/BF00591360
[30] Hachisu T, Yamada H, Satoh SI, Kouzu T. Endoscopic clipping with a new rotatable clip-device and a long clip. Dig Endosc. 1996 Apr;8(2):127-133. https://doi.org/10.1111/j.1443-1661.1996.tb00428.x
[31] Sonthalia N, Puri R, Sethi S. Role of clips, devices and sprays in endoscopic hemostasis: An update (with videos). Indian J Gastroenterol. 2025 Oct;44(5):595-604. https://doi.org/10.1007/s12664-025-01815-2
[32] Technology Assessment Committee, Chuttani R, Barkun A, Carpenter S, Chotiprasidhi P, Ginsberg GG, Hussain N, et al. Endoscopic clip application devices. Gastrointest Endosc. 2006 May;63(6):746-750. https://doi.org/10.1016/j.gie.2006.02.042
[33] Daram SR, Tang S, Wu R, To SD. Benchtop testing and comparisons among three types of through-the-scope endoscopic clipping devices. Surg Endosc. 2013 May;27(5):1521-1529. https://doi.org/10.1007/s00464-012-2679-4

[34] Romagnuolo J. Endoscopic clips: Past, present and future. Can J Gastroenterol. 2009 Mar;23(3):158-160. https://doi.org/10.1155/2009/515937

[35] Lin H. Can the triclip be successfully applied in clinical practice? Endoscopy. 2007 Apr;39(4):371. https://doi.org/10.1055/s-2007-966201

[36] Shinozaki S, Sunada K, Otake T, Lefor AT, Yamamoto H. Utility of a new reusable clip device for endoscopic treatment. Dig Endosc. 2015 Jan;27(1):106-112. https://doi.org/10.1111/den.12329

[37] Narihiro S, Yoshida M, Ohdaira H, Sato T, Suto D, Hoshimoto S, et al. A novel fluorescent marking clip for laparoscopic surgery of colorectal cancer: A case report. Int J Surg Case Rep. 2019;64:170-173. https://doi.org/10.1016/j.ijscr.2019.10.024

[38] Kato H, Kobayashi M, Yano M. Temporary sutures using reopenable clips and barbed threads. Dig Endosc. 2023 Jan; 35(1):e18-e19. https://doi.org/10.1111/den.14459

[39] Takeuchi Y, Shigita K, Asayama N, Aoyama T, Fukumoto A, Mukai S, et al. Clinical usefulness of the hold-and-drag closure using the SB clip for large mucosal defects after colorectal endoscopic submucosal dissection. Surg Endosc. 2023 Jul;37(7): 5719-5725. https://doi.org/10.1007/s00464-023-10146-2

[40] Raju GS, Ahmed I, Xiao S, Brining D, Poussard A, Tarcin O, et al. Controlled trial of immediate endoluminal closure of colon perforations in a porcine model by use of a novel clip device (with videos). Gastrointest Endosc. 2006 Dec;64(6):989-997. https://doi.org/10.1016/j.gie.2006.06.058

[41] Verlaan T, Ponds FA, Bastiaansen BA, Bredenoord AJ, Fockens P. Single clips versus multi-firing clip device for closure of mucosal incisions after peroral endoscopic myotomy (POEM). Endosc Int Open. 2016 Oct;4(10):E1052-E1056. https://doi.org/10.1055/s-0042-113126

[42] Walradt T, Szvarca D, Zarrella S, Ryou M. New generation endoscopic closure devices for full-thickness defects: A comparative pre-clinical study (with videos). Surg Endosc. 2025 Jan; 39(1):649-656. https://doi.org/10.1007/s00464-024-11434-1

[43] Raju GS. Endoscopic clip closure of gastrointestinal perforations, fistulae, and leaks. Dig Endosc. 2014 Jan;26(S1):95-104. https://doi.org/10.1111/den.12191

[44] Paspatis GA, Dumonceau JM, Barthet M, Meisner S, Repici A, Saunders BP, et al. Diagnosis and management of iatrogenic endoscopic perforations: European Society of Gastrointestinal Endoscopy (ESGE) Position Statement. Endoscopy. 2014 Aug; 46(8):693-711. https://doi.org/10.1055/s-0034-1377531

[45] Rogalski P, Daniluk J, Baniukiewicz A, Wroblewski E, Dabrowski A. Endoscopic management of gastrointestinal perforations, leaks and fistulas. World J Gastroenterol. 2015 Oct 7;21(37):10542-10552. https://doi.org/10.3748/wjg.v21.i37.10542

[46] Zhang Q, Wang Z, Bai Y. A novel through-the-scope twin endoclip for a large mucosal closure in a live pig model. Endoscopy. 2019 Dec;51(12):E372-E373. https://doi.org/10.1055/a-0948-5252

[47] Zhang Q, Jin H, Shen Z, Ma X, Li C, Tang Z, et al. Novel through-the-scope twin clip for the closure of GI wounds: The first experimental survival study in pigs (with videos). Gastrointest Endosc. 2021 Oct;94(4):850-858.e2. https://doi.org/10.1016/j.gie.2021.04.027

[48] Shiomi D, Tanabe M, Uragami N, Toshimori A, Kawamata N, Muraoka M, et al. Clinical utility of a novel anchor pronged clip for mucosal defect closure after colorectal endoscopic submucosal dissection (with video). Endosc Int Open. 2024 Oct 10;12(10):E1127-E1133. https://doi.org/10.1055/a-2409-3622

[49] Inada T, Sumida Y, Matsumoto T, Fukuda SI, Maehara K, Akiho H, et al. Muscle inversion closure technique with novel clips minimizes submucosal dead space after colorectal endoscopic submucosal dissection. Endoscopy. 2025 Dec;57(S 01): E717-E718. https://doi.org/10.1055/a-2612-3542

[50] Baron TH, Song LM, Ross A, Tokar JL, Irani S, Kozarek RA. Use of an over-the-scope clipping device: Multicenter retrospective results of the first U.S. experience (with videos). Gastrointest Endosc. 2012 Jul;76(1):202-208. https://doi.org/10.1016/j.gie.2012.03.250

[51] Kirschniak A, Kratt T, Stüker D, Braun A, Schurr MO, Königsrainer A. A new endoscopic over-the-scope clip system for treatment of lesions and bleeding in the GI tract: First clinical experiences. Gastrointest Endosc. 2007 Jul;66(1):162-167. https://doi.org/10.1016/j.gie.2007.01.034

[52] Schurr MO, Arezzo A, Ho CN, Anhoeck G, Buess G, Di Lorenzo N. The OTSC clip for endoscopic organ closure in NOTES: Device and technique. Minim Invasiv Ther Allied Technol. 2008;17(4):262-266. https://doi.org/10.1080/13645700802275114

[53] Voermans RP, van Berge Henegouwen MI, Bemelman WA, Fockens P. Novel over-the-scope-clip system for gastrotomy closure in natural orifice transluminal endoscopic surgery (NOTES): An ex vivo comparison study. Endoscopy. 2009 Dec;41(12):1052-1055. https://doi.org/10.1055/s-0029-1215231

[54] Renteln DV, Schmidt A, Vassiliou MC, Rudolph HU, Gieselmann M, Caca K. Endoscopic closure of large colonic perforations using an over-the-scope clip: A randomized controlled porcine study. Endoscopy. 2009 Jun;41(6):481-486. https://doi.org/10.1055/s-0029-1214759

[55] Armellini E, Crinò SF, Orsello M, Ballarè M, Tari R, Saettone S, et al. Novel endoscopic over-the-scope clip system. World J Gastroenterol. 2015 Dec 28;21(48):13587-13592. https://doi.org/10.3748/wjg.v21.i48.13587

[56] Mönkemüller K, Toshniwal J, Zabielski M, Vormbrock K, Neumann H. Utility of the “bear claw”, or over-the-scope clip (OTSC) system, to provide endoscopic hemostasis for bleeding posterior duodenal ulcers. Endoscopy. 2012;44(S 02):E412-E413. https://doi.org/10.1055/s-0032-1325737

[57] Weiland T, Fehlker M, Gottwald T, Schurr MO. Performance of the OTSC System in the endoscopic closure of iatrogenic gastrointestinal perforations: A systematic review. Surg Endosc. 2013 Jul;27(7):2258-2274. https://doi.org/10.1007/s00464-012-2754-x

[58] Bartell N, Bittner K, Kaul V, Kothari TH, Kothari S. Clinical efficacy of the over-the-scope clip device: A systematic review.

World J Gastroenterol. 2020 Jun 28;26(24):3495-3516. https://doi.org/10.3748/wjg.v26.i24.3495
[59] Goenka MK, Rodge GA, Tiwary IK. Endoscopic management with a novel over-the-scope Padlock clip system. Clin Endosc. 2019 Nov;52(6):574-580. https://doi.org/10.5946/ce.2019.122
[60] Repici A, Arezzo A, De Caro G, Morino M, Pagano N, Rando G, et al. Clinical experience with a new endoscopic over-the-scope clip system for use in the GI tract. Dig Liver Dis. 2009 Jun;41(6):406-410. https://doi.org/10.1016/j.dld.2008.09.002
[61] Dinelli M, Omazzi B, Andreozzi P, Zucchini N, Redaelli A, Manes G. First clinical experiences with a novel endoscopic over-the-scope clip system. Endosc Int Open. 2017 Mar;5(3): E151-E156. https://doi.org/10.1055/s-0043-101692
[62] Kobara H, Mori H, Fujihara S, Nishiyama N, Chiyo T, Yamada T, et al. Outcomes of gastrointestinal defect closure with an over-the-scope clip system in a multicenter experience: An analysis of a successful suction method. World J Gastroenterol. 2017 Mar 7;23(9):1645-1656. https://doi.org/10.3748/wjg.v23.i9.1645
[63] Matthes K, Jung Y, Kato M, Gromski MA, Chuttani R. Efficacy of full-thickness GI perforation closure with a novel over-the-scope clip application device: An animal study. Gastrointest Endosc. 2011 Dec;74(6):1369-1375. https://doi.org/10.1016/j.gie.2011.07.057
[64] Singhal S, Changela K, Papafragkakis H, Anand S, Krishnaiah M, Duddempudi S. Over the scope clip: Technique and expanding clinical applications. J Clin Gastroenterol. 2013 Oct;47(9): 749-56. https://doi.org/10.1097/MCG.0b013e318296ecb9
[65] von Renteln D, Schmidt A, Vassiliou MC, Gieselmann M, Caca K. Natural orifice transluminal endoscopic surgery gastrotomy closure with an over-the-endoscope clip: A randomized, controlled porcine study (with videos). Gastrointest Endosc. 2009 Oct;70(4):732-739. https://doi.org/10.1016/j.gie.2009.03.010
[66] Kobara H, Mori H, Nishiyama N, Fujihara S, Okano K, Suzuki Y, et al. Over-the-scope clip system: A review of 1517 cases over 9 years. J Gastroenterol Hepatol. 2019 Jan;34(1):22-30. https://doi.org/10.1111/jgh.14402
[67] Haito-Chavez Y, Law JK, Kratt T, Arezzo A, Verra M, Morino M, et al. International multicenter experience with an over-the-scope clipping device for endoscopic management of GI defects (with video). Gastrointest Endosc. 2014 Oct;80(4):610-622. https://doi.org/10.1016/j.gie.2014.03.049
[68] Schmidt A, Gölder S, Goetz M, Meining A, Lau J, von Delius S, et al. Over-the-scope clips are more effective than standard endoscopic therapy for patients with recurrent bleeding of peptic ulcers. Gastroenterology. 2018 Step;155(3):674-686.e6. https://doi.org/10.1053/j.gastro.2018.05.037
[69] Ikeda K, Fritscher-Ravens A, Mosse CA, Mills T, Tajiri H, Swain CP. Endoscopic full-thickness resection with sutured closure in a porcine model. Gastrointest Endosc. 2005 Jul;62(1):122-129. https://doi.org/10.1016/S0016-5107(05)00517-1
[70] Sumiyama K, Gostout CJ, Rajan E, Bakken TA, Deters JL, Knipschield MA. Endoscopic full-thickness closure of large gastric perforations by use of tissue anchors. Gastrointest Endosc. 2007 Jan;65(1):134-139. https://doi.org/10.1016/j.gie.2006.01.050
[71] Sporn E, Miedema BW, Bachman SL, Astudillo JA, Loy TS, Calaluce R, et al. Endoscopic colotomy closure after full thickness excision: Comparison of T fastener with multiclip applier. Endoscopy. 2008 Jul;40(7):589-594. https://doi.org/10.1055/s-2008-1077377
[72] Guarner-Argente C, Córdova H, Martínez-Pallí G, Navarro-Ripoll R, Rodríguez-d'Jesús A, de Miguel CR, et al. Gastrotomy closure with a new tissue anchoring device: A porcine survival study. World J Gastroenterol. 2011 Apr 7;17(13):1732-1738. https://doi.org/10.3748/wjg.v17.i13.1732
[73] Goto O, Sasaki M, Ishii H, Horii J, Uraoka T, Takeuchi H, et al. A new endoscopic closure method for gastric mucosal defects: Feasibility of endoscopic hand suturing in an ex vivo porcine model (with video). Endosc Int Open. 2014 Jun 24;2(2): E111-E116. https://doi.org/10.1055/s-0034-1377180
[74] Niu C, Zhang J, Farooq U, Zahid S, Okolo PI. Endoscopic suturing for mucosal defect closure following endoscopic submucosal dissection: Systematic review and meta-analysis. Endosc Int Open. 2024 Oct 10;12(10):E1150-E1159. https://doi.org/10.1055/a-2411-8724
[75] Orzeszko Z, Kasprzyk P, Zawada U, Szura M, Spychalski M. Endoscopic hand suturing after advanced endoscopic procedures: Early outcomes of 31 cases in the upper and lower gastrointestinal tract (with video). Wideochir Inne Tech Maloinwazyjne. 2025 Feb 10;20(1):44-50. https://doi.org/10.20452/wiitm.2025.17928
[76] Scheppach MW, Nagl S, Muzalyova A, Classen J, Messmann H, Ebigbo A. Feasibility of a new endoscopic suturing device: A first Western experience (with video). Gastrointest Endosc. 2025 Jan;101(1):207-212. https://doi.org/10.1016/j.gie.2024.08.001
[77] Uozumi T, Abe S, Mizuguchi Y, Sekiguchi M, Toyoshima N, Takamaru H, et al. Endoscopic hand suturing using a modified through-the-scope needle holder for mucosal closure after colorectal endoscopic submucosal dissection: Prospective multicenter study (with video). Dig Endosc. 2024 Nov;36(11):1245-1252. https://doi.org/10.1111/den.14808
[78] Mahmoud T, Wong Kee Song LM, Stavropoulos SN, Alansari TH, Ramberan H, Fukami N, et al. Initial multicenter experience using a novel endoscopic tack and suture system for challenging GI defect closure and stent fixation (with video). Gastrointest Endosc. 2022 Feb;95(2):373-382. https://doi.org/10.1016/j.gie.2021.10.018
[79] Hernandez-Lara A, Garcia Garcia de Paredes A, Rajan E, Storm AC. Step-by-step instruction: Using an endoscopic tack and suture device for gastrointestinal defect closure. VideoGIE. 2021 Apr 5;6(6):243-245. https://doi.org/10.1016/j.vgie.2021.02.015
[80] Hernandez A, Marya NB, Sawas T, Rajan E, Gades NM, Wong Kee Song LM, et al. Gastrointestinal defect closure using a novel through-the-scope helix tack and suture device compared to endoscopic clips in a survival porcine model (with video). Endosc Int Open. 2021 Apr;9(4):E572-E577. https://doi.org/10.1055/a-1370-9256

[81] Suresh Kumar VC, Singh S, Loganathan P, Mohan BP, Aswath G, Akbar Khan HM, et al. Comparing the safety and efficacy of over-the-scope suturing, through-the-scope suturing, and endoscopic hand suturing for closure of GI defects after endoscopic resection: Systematic review and meta-analysis. Gastrointest Endosc. 2025 Sep;102(3):326-336.e325. https://doi.org/10.1016/j.gie.2025.03.002

[82] Swain CP, Mills TN. An endoscopic sewing machine. Gastrointest Endosc. 1986 Feb;32(1):36-38. https://doi.org/10.1016/s0016-5107(86)71727-6

[83] Swain CP. Endoscopic sewing and stapling machines. Endoscopy. 1997 Mar;29(3):205-210. https://doi.org/10.1055/s-2007-1004165

[84] Swain CP, Brown GJ, Gong F, Mills TN. An endoscopically deliverable tissue-transfixing device for securing biosensors in the gastrointestinal tract. Gastrointest Endosc. 1994 Nov-Dec;40(6):730-734. https://doi.org/10.1016/S0016-5107(94)70117-2

[85] Kadirkamanathan SS, Evans DF, Gong F, Yazaki E, Scott M, Swain CP. Antireflux operations at flexible endoscopy using endoluminal stitching techniques: An experimental study. Gastrointest Endosc. 1996 Aug;44(2):133-143. https://doi.org/10.1016/S0016-5107(96)70130-X

[86] Hu B, Chung SC, Sun LC, Lau JY, Kawashima K, Yamamoto T, et al. Developing an animal model of massive ulcer bleeding for assessing endoscopic hemostatic devices. Endoscopy. 2005 Sep;37(9):847-851. https://doi.org/10.1055/s-2005-870226

[87] Hu B, Chung SC, Sun LC, Kawashima K, Yamamoto T, Cotton PB, et al. Eagle Claw II: A novel endosuture device that uses a curved needle for major arterial bleeding: A bench study. Gastrointest Endosc. 2005 Aug;62(2):266-270. https://doi.org/10.1016/S0016-5107(05)00375-5

[88] Hu B, Chung SC, Sun LC, Lau JY, Kawashima K, Yamamoto T, et al. Endoscopic suturing without extracorporeal knots: A laboratory study. Gastrointest Endosc. 2005 Aug;62(2):230-233. https://doi.org/10.1016/S0016-5107(05)00331-7

[89] Chiu PW, Hu B, Lau JY, Sun LC, Sung JJ, Chung SS. Endoscopic plication of massively bleeding peptic ulcer by using the Eagle Claw VII device: A feasibility study in a porcine model. Gastrointest Endosc. 2006 Apr;63(4):681-685. https://doi.org/10.1016/j.gie.2005.10.030

[90] Chiu PW, Lau JY, Ng EK, Lam CC, Hui M, To KF, et al. Closure of a gastrotomy after transgastric tubal ligation by using the Eagle Claw VII: A survival experiment in a porcine model (with video). Gastrointest Endosc. 2008 Sep;68(3):554-559. https://doi.org/10.1016/j.gie.2008.03.1110

[91] ASGE Technology Committee, Banerjee S, Barth BA, Bhat YM, Desilets DJ, Gottlieb KT, et al. Endoscopic closure devices. Gastrointest Endosc. 2012 Aug;76(2):244-251. https://doi.org/10.1016/j.gie.2012.02.028

[92] Pauli EM, Delaney CP, Champagne B, Stein S, Marks JM. Safety and effectiveness of an endoscopic suturing device in a human colonic treat-and-resect model. Surg Innov. 2013 Dec;20(6):594-599. https://doi.org/10.1177/1553350613479204

[93] Stavropoulos S, Modayil R, Friedel D. Current applications of endoscopic suturing. World J Gastrointest Endosc. 2015 Jul 10;7(8):777-789. https://doi.org/10.4253/wjge.v7.i8.777

[94] Abu Dayyeh BK, Bazerbachi F, Vargas EJ, Sharaiha RZ, Thompson CC, Thaemert BC, et al. Endoscopic sleeve gastroplasty for treatment of class 1 and 2 obesity (MERIT): A prospective, multicentre, randomised trial. Lancet. 2022 Aug 6;400(10350):441-451. https://doi.org/10.1016/S0140-6736(22)01280-6

[95] Callahan ZM, Su B, Kuchta K, Conaty E, Novak S, Linn J, et al. Endoscopic suturing results in high technical and clinical success rates for a variety of gastrointestinal pathologies. J Gastrointest Surg. 2020 Feb;24(2):278-287. https://doi.org/10.1007/s11605-019-04485-6

[96] Maselli R, Palma R, Traina M, Granata A, Juzgado D, Bisello M, et al. Endoscopic suturing for GI applications: Initial results from a prospective multicenter European registry. Gastrointest Endosc. 2022 Nov;96(5):780-786. https://doi.org/10.1016/j.gie.2022.06.004

[97] Moran EA, Gostout CJ, Bingener J. Preliminary performance of a flexible cap and catheter-based endoscopic suturing system. Gastrointest Endosc. 2009 Jun;69(7):1375-1383. https://doi.org/10.1016/j.gie.2009.01.032

[98] Kantsevoy SV. The development of the overstitch system and its potentials. Gastrointest Endosc Clin N Am. 2020 Jan;30(1):107-114. https://doi.org/10.1016/j.giec.2019.08.004

[99] Keihanian T, Othman MO. OverStitch Sx Endoscopic suturing system in minimally invasive endoscopic procedures: Overview of its safety and efficacy and comparison to overstictch™. Expert Rev Med Devices. 2022 Jan;19(1):11-23. https://doi.org/10.1080/17434440.2022.2019579

[100] Chou CK, Chen SS, Lee CY, Toh DE, Chuah YY, Tsai KF, et al. Comprehensive defect closure with overStitch Sx after endoscopic resection of gastric and duodenal subepithelial lesions. J Formos Med Assoc. 2025 Oct 28;S0929-6646(25)00470-X. https://doi.org/10.1016/j.jfma.2025.09.005

[101] Matsuda T, Fujii T, Emura F, Kozu T, Saito Y, Ikematsu H, et al. Complete closure of a large defect after EMR of a lateral spreading colorectal tumor when using a two-channel colonoscope. Gastrointest Endosc. 2004 Nov;60(5):836-838. https://doi.org/10.1016/S0016-5107(04)02033-4

[102] Zhang Y, Wang X, Xiong G, Qian Y, Wang H, Liu L, et al. Complete defect closure of gastric submucosal tumors with purse-string sutures. Surg Endosc. 2014 Jun;28(6):1844-1851. https://doi.org/10.1007/s00464-013-3404-7

[103] Nomura T, Kobayashi M, Morikawa T, Horiki N. Clip-fixed endoloop: An efficacious new method for mucosal defect closure. Endoscopy. 2018 Jun;50(6):E126-E127. https://doi.org/10.1055/s-0044-101025

[104] Wu N, Liu S, Chen M, Zeng X, Wang F, Zhang J, et al. The prepurse-string suture technique for gastric defect after endoscopic full-thickness resection (with video). Medicine (Baltimore). 2018 Sep;97(36):e12118. https://doi.org/10.1097/MD.0000000000012118

[105] Li M, Zhang Y, Sun F, Huai M, Zhang F, Qu C, et al. Feasibility and efficacy of endoscopic purse-string suture-assisted closure for mucosal defects induced by endoscopic manipulations. World J Gastroenterol. 2023 Jan 28;29(4):731-743. https://doi.org/10.3748/wjg.v29.i4.731
[106] Li B, Huang S, Liao S, He M, Xiao Y, Zhong E, et al. Modified endoscopic purse-string suture with dental floss traction in the management of a duodenal defect. Endoscopy. 2025 Dec;57(S 01):E491-E492. https://doi.org/10.1055/a-2590-8450
[107] Wang S, Ji M, Huang X, Li Y, Yu S, Zhao Y, et al. Double-nylon purse-string suture in closing postoperative wounds following endoscopic resection of large (≥ 3 cm) gastric submucosal tumors. World J Gastroenterol. 2024 Mar 7;30(9):1143-1153. https://doi.org/10.3748/wjg.v30.i9.1143
[108] Walia A, Trasolini RP, Shahidi N. Double-nylon purse-string suture technique: Another addition to the endoscopist's toolbox for full-thickness defect closure. World J Gastroenterol. 2024 Jul 7;30(25):3152-3154. https://doi.org/10.3748/wjg.v30.i25.3152
[109] Nishiyama N, Kobara H, Kobayashi N, Chiyo T, Fujihara S, Yachida T, et al. Novel endoscopic ligation with O-ring closure involving muscle layer of a gastric artificial defect. Endoscopy. 2020 Nov;52(11):E413-E414. https://doi.org/10.1055/a-1149-8597
[110] Nishiyama N, Kobara H, Kobayashi N, Chiyo T, Tada N, Kozuka K, et al. Efficacy of endoscopic ligation with O-ring closure for prevention of bleeding after gastric endoscopic submucosal dissection under antithrombotic therapy: A prospective observational study. Endoscopy. 2022 Nov;54(11):1078-1084. https://doi.org/10.1055/a-1782-3448
[111] Fujii T, Ono A, Fu KI. A novel endoscopic suturing technique using a specially designed so-called "8-ring" in combination with resolution clips (with videos). Gastrointest Endosc. 2007 Dec;66(6):1215-1220. https://doi.org/10.1016/j.gie.2007.05.054
[112] Sakamoto N, Beppu K, Matsumoto K, Shibuya T, Osada T, Mori H, et al. "Loop Clip", a new closure device for large mucosal defects after EMR and ESD. Endoscopy. 2008 Sep;40:E97-E98. https://doi.org/10.1055/s-2007-995604
[113] Mori H, Sakamoto N, Osada T, Ritsuno H, Hidaka Y, Nakae K, et al. The 'Loop Clip' is useful for closing large mucosal defects after colorectal endoscopic submucosal dissection: A preliminary clinical study. Dig Endosc. 2011 Oct;23(4):330-331. https://doi.org/10.1111/j.1443-1661.2010.01099.x
[114] Kowazaki Y, Fukuda H, Saito I, Sawano T, Kurokawa T, Kanzaki N, et al. A novel method for endoscopic closure of endoscopic submucosal dissection-induced defects in the colorectum: The closure method with an elastic-rubber ring (CMER). Endoscopy. 2021 Aug;54(08):E407-E408. https://doi.org/10.1055/a-1550-2246
[115] Ge PS, Aihara H. A novel clip-band traction device to facilitate colorectal endoscopic submucosal dissection and defect closure. VideoGIE. 2020 May 31;5(5):180-186. https://doi.org/10.1016/j.vgie.2020.01.012
[116] Jinushi R, Tashima T, Terada R, Miyaguchi K, Katsuda H, Ogawa T, et al. Effectiveness of a multi-loop traction device for colorectal endoscopic submucosal dissection performed by trainees: A pilot study. Scientific Reports. 2022 Jun 17;12(1): 10197. https://doi.org/10.1038/s41598-022-14407-3
[117] Muramatsu T, Tashima T, Ishikawa T, Terada R, Kawasaki T, Itoi T, et al. A novel triangle traction method using a clip-band traction device for endoscopic submucosal dissection of early gastric cancer involving the pyloric ring. Endoscopy. 2022 Dec;55(S 01):E207-E208. https://doi.org/10.1055/a-1956-1349
[118] Tanaka S, Toyonaga T, Obata D, Ishida T, Morita Y, Azuma T. Endoscopic double-layered suturing: A novel technique for closure of large mucosal defects after endoscopic mucosal resection (EMR) or endoscopic submucosal dissection (ESD). Endoscopy. 2012;44(S 02):E153-E154. https://doi.org/10.1055/s-0031-1291673
[119] Banno S, Nishizawa T, Kinoshita S, Mori H, Uraoka T. Endoscopic mucosa–submucosal clip closure method. Endoscopy. 2017 Dec;49(12):E307-E308. https://doi.org/10.1055/s-0043-119349
[120] Kinoshita S, Nishizawa T, Fujimoto A, Mori H, Nakazato Y, Kikuchi M, et al. Efficacy of mucosa-submucosa clip closure method after gastric endoscopic submucosal dissection. World J Gastrointest Endosc. 2020 Jan 16;12(1):17-22. https://doi.org/10.4253/wjge.v12.i1.17
[121] Nishizawa T, Banno S, Kinoshita S, Mori H, Nakazato Y, Hirai Y, et al. Feasibility of endoscopic mucosa-submucosa clip closure method (with video). Endosc Int Open. 2018 Aug;6(8): E1070-E1074. https://doi.org/10.1055/a-0630-0566
[122] Nomura T, Kamei A, Sugimoto S, Oyamada J. New closure method for a mucosal defect after endoscopic submucosal dissection: The clip-on-clip closure method. Endoscopy. 2018 May;50(5):547-548. https://doi.org/10.1055/s-0044-100486
[123] Nomura T, Matsuzaki I, Sugimoto S, Oyamda J, Kamei A, Kobayashi M. Clip-on-clip closure method for a mucosal defect after colorectal endoscopic submucosal dissection: A prospective feasibility study. Surg Endosc. 2020 Mar;34(3):1412-1416. https://doi.org/10.1007/s00464-019-07195-x
[124] Yamasaki Y, Takeuchi Y, Kato M, Uedo N, Ishihara R. Line-assisted complete closure of large gastric mucosal defects by use of multiple clip-and-line technique. VideoGIE. 2016 Sep 3; 1(3):49-50. https://doi.org/10.1016/j.vgie.2016.08.008
[125] Yamasaki Y, Takeuchi Y, Iwatsubo T, Kato M, Hamada K, Tonai Y, et al. Line-assisted complete closure for a large mucosal defect after colorectal endoscopic submucosal dissection decreased post-electrocoagulation syndrome. Dig Endosc. 2018 Sep;30(5):633-641. https://doi.org/10.1111/den.13052
[126] Ominami M, Nagami Y, Tanaka C, Sakai T, Fukunaga S, Watanabe T, et al. Endoscopic technique for closure of a large gastric tube perforation by using endoclips with line-assisted complete closure. Endoscopy. 2019 Mar;51(3):E49-E50. https://doi.org/10.1055/a-0800-8502
[127] Nakahira H, Takeuchi Y, Garcia JS, Morita Y, Uedo N, Ishihara R, et al. Line-assisted endoscopic complete closure of a large perforation during colonic endoscopic submucosal dissection. Endoscopy. 2018 Jan;50(1):E32-E33. https://doi.org/10.1055/s-0043-121137

[128] Yamasaki Y, Takeuchi Y, Uedo N, Hanaoka N, Higashino K, Ishihara R, et al. Line-assisted complete closure of duodenal mucosal defects after underwater endoscopic mucosal resection. Endoscopy. 2017 Feb;49(S 01):E37-E38. https://doi.org/10.1055/s-0042-120707

[129] Nakahira H, Takeuchi Y, Kanesaka T, Uedo N, Ishihara R. Wide-field underwater EMR followed by line-assisted complete closure for a large duodenal adenoma. VideoGIE. 2019 Oct;4(10):469-471. https://doi.org/10.1016/j.vgie.2019.07.004

[130] Nomura T, Sugimoto S, Tsuda N, Matsushima R, Oyamada J, Kamei A. Mucosal defect closure after duodenal endoscopic submucosal dissection using the reopenable-clip over the line method. JGH Open. 2021 Jul;5(7):831-833. https://doi.org/10.1002/jgh3.12577

[131] Nomura T, Sugimoto S, Temma T, Oyamada J, Ito K, Kamei A. Reopenable clip-over-the-line method for closing large mucosal defects following gastric endoscopic submucosal dissection: Prospective feasibility study. Dig Endosc. 2023 May;35(4):505-511. https://doi.org/10.1111/den.14466

[132] Nomura T, Sugimoto S, Temma T, Oyamada J, Ito K, Kamei A, et al. Gastric antral defect closure after endoscopic submucosal dissection with the reopenable clip-over-the-line method using a multibending scope. Endoscopy. 2023 Dec;55(S 01):E866-E867. https://doi.org/10.1055/a-2107-2474

[133] Nomura T, Sugimoto S, Kawabata M, Oyamada J, Kamei A. Large colorectal mucosal defect closure post-endoscopic submucosal dissection using the reopenable clip over line method and modified locking-clip technique. Endoscopy. 2022 Feb;54(2):E63-E64. https://doi.org/10.1055/a-1381-6435

[134] Nomura T, Kamei A, Sugimoto S, Harada T, Oyamda J. Thread-assisted mucosal defect closure using the locking-clip technique after gastric endoscopic submucosal dissection. Endoscopy. 2018 Dec;50(12):E342-E343. https://doi.org/10.1055/a-0725-7754

[135] Nomura T, Kamei A, Sugimoto S. Clip line traction method using locking clip technique for colorectal endoscopic submucosal dissection. Dig Endosc. 2019 May;31(3):e72-e73. https://doi.org/10.1111/den.13364

[136] Nabi Z, Reddy DN. Endoscopic closure of gastrointestinal defects: Devices and techniques. Expert Rev Gastroenterol Hepatol. 2025 Sep;19(9):973-983. https://doi.org/10.1080/17474124.2025.2539793

[137] Al-Bawardy B, Rajan E, Wong Kee Song LM. Over-the-scope clip-assisted endoscopic full-thickness resection of epithelial and subepithelial GI lesions. Gastrointest Endosc. 2017 May;85(5):1087-1092. https://doi.org/10.1016/j.gie.2016.08.019

[138] Wei MT, Ahn JY, Friedland S. Over-the-scope clip in the treatment of gastrointestinal leaks and perforations. Clin Endosc. 2021 Nov;54(6):798-804. https://doi.org/10.5946/ce.2021.250

[139] Guarner-Argente C, Córdova H, Martínez-Pallí G, Navarro R, Cuatrecasas M, Rodríguez de Miguel C, et al. Yes, we can: Reliable colonic closure with the Padlock-G clip in a survival porcine study (with video). Gastrointest Endosc. 2010 Oct;72(4):841-844. https://doi.org/10.1016/j.gie.2010.06.054

[140] Troya J, Fuchs KH, Hann A, Meining A. Technical aspects of endoscopic internal drainage procedure, secured by endoscopic suture fixation: Experimental study. Endosc Int Open. 2026 Jan 14;14:a27779441. https://doi.org/10.1055/a-2777-9441

[141] McKhann C, Lee LS. Saving patients one stitch at a time? iGIE. 2023 Jun;2(2):202-208. https://doi.org/10.1016/j.igie.2023.04.002

[142] Keihanian T, Zabad N, Khalaf M, Abdel-Jaber W, Kim YI, Jawaid SA, et al. Safety and efficacy of a novel suturing device for closure of large defects after endoscopic submucosal dissection (with video). Gastrointest Endosc. 2023 Sep;98(3):381-391. https://doi.org/10.1016/j.gie.2023.04.006

Progress in
Medical Devices

REVIEW ARTICLE

Research progress on postoperative bleeding after endoscopic submucosal dissection and its treatment

Jin Xu, Shiju Yan

School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Shiju Yan.

Address correspondence to: Shiju Yan, School of Health Science and Engineering, University of Shanghai for Science and Technology, No. 334 Jungong Road, Yangpu District, Shanghai 200093, China. E-mail: yanshj99@aliyun.com.

Received December 3, 2025; Accepted April 30, 2026; Published June 18, 2026

DOI: 10.61189/828857qvjtwr

Abstract

Endoscopic submucosal dissection (ESD) is now considered the standard endoscopic resection technique for patients with early gastric cancer. ESD provides a higher rate of complete resection and a lower local recurrence rate. However, ESD results in larger and deeper ulcers, and post-ESD bleeding is a common complication. Bleeding after ESD cannot be completely avoided, especially in patients with large-sized gastric ulcers, those on anticoagulant therapy, and elderly patients. Most bleeding can be controlled by endoscopic hemostatic methods during the procedure, such as the application of hemostatic clips for ulcer closure and hemostatic powder for ulcer shielding. In addition, we also found the potential value of using new materials such as self-assembling peptides for hemostasis. This review first revisits the definition of endoscopic resection of the digestive tract. Then, we discuss post-ESD bleeding and the influence of risk factors such as the location, size, and depth of the surgical lesion, anticoagulant medication use, and the patient's age and lifestyle. Finally, we review the treatment methods for post-ESD bleeding, including intraoperative ulcer closure, ulcer shielding, and the application of thrombin and adrenaline injection.

Keywords: Digestive tract tumors, Digestive tract endoscopy, Bleeding risk factors, Hemostasis in gastric endoscopic submucosal dissection

1 INTRODUCTION

Esophageal cancer and gastric cancer are common malignant tumors of the digestive system. Patients who can be diagnosed and treated in the early stage generally have a relatively good prognosis. Atypical hyperplasia of the esophageal and gastric mucosa is a precancerous lesion of esophageal cancer and gastric cancer. Early diagnosis and treatment of atypical hyperplasia of the esophageal and gastric mucosa can block the development of precancerous lesions into esophageal cancer and gastric cancer.

Endoscopic resection is now widely accepted as a less invasive treatment for local resection of early gastric cancer, with a negligible risk of lymph node metastasis [1]. A suitable endoscopic resection technique should be safe, effective, and applicable to various clinical situations. However, endoscopic resection is associated with various complications, the most important of which are bleeding and perforation [2].

Endoscopic resection for digestive tract lesions has the advantages of minimal surgical trauma, fewer postoperative complications, and good postoperative recovery, making it the preferred minimally invasive surgical treatment for digestive tract mucosal and submucosal lesions [3].

Endoscopic mucosal resection (EMR) and endoscopic submucosal dissection (ESD) are two endoscopic resection procedures for gastrointestinal lesions. Compared with standard polypectomy techniques, EMR and ESD provide wider and deeper resection margins and allow for more detailed pathological evaluation of the overall resected lesion, which has advantag-

Table 1. Effect of lesion site on bleeding after endoscopic submucosal dissection

Author	Lower-third (%)	Middle-third (%)	Upper-third (%)	P-value
Nam et al. [18]	3.8	2.0	0.0	<0.05
Tsuji et al. [6]	9.6	3.2	4.6	<0.05
Furuhata et al. [14]	7.3	4.5	3.1	<0.01
Sato et al. [16]	5.3	5.4	4.1	>0.05
Libânio et al. [2]	5.5	5.6	5.2	>0.05
Matsumura et al. [13]	6.1	3.3	4.6	>0.05
Toyokawa et al. [15]	5.2	4.9	5.0	>0.05
Mukai et al. [12]	16.9	13.5	0.0	>0.05

Note: This table is cited from [17].

es in treating early neoplastic gastrointestinal lesions [3]. Although EMR is convenient, larger lesions cannot be completely resected in a single EMR procedure. Therefore, doctors cannot make an accurate diagnosis of the entire pathological specimen to make more appropriate treatment decisions for the patient's condition, which increases the risk of local tumor recurrence or inappropriate treatment. ESD can directly resect tumors in the submucosa. Due to the large wound size and the difficulty of implementing the technique, there is a higher incidence of bleeding and perforation after ESD [4].

This review focuses on post-ESD bleeding by summarizing relevant literature on ESD, including surgical procedures, risk factors for delayed bleeding, prevention and treatment strategies, and related medical devices.

2 DEFINITIONS, EPIDEMIOLOGY, AND RISK FACTORS

2.1 Definitions and pathophysiology

In this review, the literature search was carried out through the Web of Science database, and the search time range was set from the establishment of the database to September 2025. The core search terms included "Endoscopic Submucosal Dissection", "Gastrointestinal Hemorrhage", "Immediate Bleeding", "Delayed Bleeding", and "Risk Factors". According to the onset time, bleeding complications can be subdivided into immediate (intraoperative) bleeding that occurs during the procedure and delayed bleeding that occurs after the procedure. There are differences in risk factors and treatment logic between these two types of hemorrhage. Immediate hemorrhage is mainly affected by anatomical location and procedural factors, while delayed hemorrhage is more related to the patient's general condition and healing environment. Immediate bleeding is infrequent with EMR techniques, but is quite common with ESD. The management of immediate bleeding plays a key role in the successful completion of ESD. It is difficult to accurately measure the amount of bleeding during EMR and ESD. Therefore, in the definitions of upper gastrointestinal bleeding (GIB) in different countries, bleeding after ESD is usually defined as a condition with any clinical signs of bleeding. Clinical signs include hematemesis, melena, or a decrease in haemoglobin of >2 g/dL from the previous laboratory test, and the condition requires endoscopic hemostasis [5]. The definition of delayed bleeding can range from clinical symptoms of bleeding such as hematemesis and melena to bleeding that requires blood transfusion or even bleeding that requires endoscopic treatment. Delayed bleeding usually occurs within 24 hours after ESD and has been reported to be related to lesion location, size, patient age, and anticoagulant drug use [6-9]. In formulating treatment plans, immediate bleeding focuses on intraoperative visual field maintenance and precise electrocoagulation; for delayed bleeding, single prophylactic thermocoagulation cannot significantly reduce the rate of delayed bleeding and may even aggravate thermal damage to gastric tissue [10]. Therefore, the treatment strategy for patients should also consider more complex procedural approaches, such as ulcer shielding methods and ulcer closure methods, for prevention.

2.2 Gastric lesions and ESD

The size of the gastric lesion is a risk factor for bleeding after ESD. Theoretically, the larger the lesion, the greater the risk of bleeding after ESD [5]. In addition, there is controversy about whether tumor location affects bleeding after ESD. The ideal dissection plane is the avascular layer below the vascular network of the gastric wall and above the muscular layer [11]. Several studies reported that the bleeding rate was significantly higher for tumors located in the lower third and middle/lower third of the stomach, while other studies found that an upper location of the lesion in the stomach was a more influential risk factor [2, 6, 12-16]. Although the correlation between lesion location and bleeding after ESD is inconsistent in the literature, studies tend to support that ESD in the lower third of the stomach is associated with bleeding after ESD (**Table 1**) [17]. Therefore, endoscopists should more closely monitor patients with lesions located in the lower third of the stomach.

2.3 Patient-related risk factors

Studies have shown that the use of anticoagulants has a significant impact on post-ESD bleeding in the stomach. For patients receiving anticoagulant therapy, delayed bleeding after ESD remains an unpreventable adverse event [19].

The use of aspirin is also a risk factor for delayed bleeding, but there is controversy regarding the relationship between the risk of post-ESD bleeding and the aspirin treatment regimen during the procedure. A prospective study by Cho et al. on 328 patients showed that the postoperative bleeding rate in the group that continued to use aspirin was 15.6%, significantly higher than the rate of 4.2% in the group that stopped taking the drug (odds ratio [OR]=4.12, P=0.001) [20]. The authors confirmed through multivariate analysis that aspirin is an independent risk factor for post-ESD bleeding. A meta-analysis by Yang et al. covering 15 studies with 4,876 patients confirmed that the use of aspirin increases the risk of bleeding (OR=2.34, P<0.001), but subgroup analysis showed that the risk was lower with low-dose aspirin (≤100 mg/d) (OR=1.87) [21]. Similarly, a systematic review by Jaruvongvanich et al. that included a total of 2,143 patients showed through meta-analysis that the continued use of aspirin did not significantly increase the risk of bleeding (OR=1.25, P=0.20) [22]. These studies reveal the complexity of risk assessment: the absolute risk of bleeding increases by about 4–6% in patients at high risk of thrombosis who continue to take aspirin, while the risk of thrombotic events due to stopping the drug may be even higher. Therefore, when treating patients with different medication statuses, doctors need to comprehensively consider risk factors such as aspirin dosage, lesion size, and procedure time to develop an appropriate ESD plan.

Gastric ESD is increasingly being used in elderly patients, yet bleeding remains the main complication. Park et al. used the Korean National Health Insurance Service database, which included patients who underwent gastric ESD from November 2011 to December 2022 [8]. The elderly group (aged >75 years) accounted for 18.6% of all cases. The study used a mixed-effects logistic regression model to analyze and identify the risk factors for bleeding and analyzed the risk factors for bleeding in the elderly and non-elderly groups, respectively, to evaluate the differences in their patterns. In the multivariate logistic regression, significant risk factors included younger age, male sex, a higher comorbidity index, gastric cancer, multiple lesions, several comorbidities (such as renal failure), antithrombotic drug use, being underweight, smoking, and heavy drinking habits. The impact of some factors on the risk of bleeding differed significantly between age groups. In the elderly group, the risk was increased in association with a high comorbidity index (≥5). Although the incidence of bleeding was higher in the elderly group, this may be due to the higher prevalence of comorbidities and more frequent use of antithrombotic drugs in this population, rather than age itself [8]. The correlation between age and the risk of bleeding is consistent with previous studies, and doctors need to consider age when developing appropriate treatment plans for patients.

3 BLEEDING TREATMENT OPTIONS

3.1 Observation without immediate intervention

For small polyps with a wound diameter of less than 2 cm after electrocoagulation or electroresection, if there is no bleeding, perforation, or other complication, generally no treatment is required. For some patients with early gastric cancer or early esophageal cancer and larger wounds with a diameter greater than 2 cm, if the muscularis propria is intact after ESD, hemostasis is effective during the procedure, the risk of perforation is low, and further intervention is difficult after the procedure, wound observation without immediate intervention can also be considered [23]. After the procedure, patients should fast, stay in bed, and receive enhanced acid suppression treatment.

Epinephrine has the effect of constricting blood vessels and achieving local hemostasis and is often added to the submucosal injection solution [24]. The injection needle consists of an outer sheath (made of plastic, polytetrafluoroethylene, or stainless steel) and an inner hollow needle. Using the handle at the end of the needle sheath, the operator uses a syringe connected to the handle to inject liquid agents into the target tissue. The injection of various solutions achieves hemostasis through mechanical tamponade and chemical mechanisms. A study showed that preoperative submucosal injection of epinephrine during EMR for colonic polyps significantly reduced the risk of bleeding after colonic polyp resection compared with no intervention, and the cost was significantly lower than that of mechanical hemostasis methods such as hemostatic clips. Hemocoagulase for injection is a hemostatic agent extracted from snake venom that can promote the formation of fibrin polymers, thereby achieving hemostasis, and is commonly used in surgical procedures.

3.2 ESD ulcer closure

Endoscopic ulcer closure and hemostasis include the use of hemostatic clips, the combination of hemostatic clips and nylon ropes for wound closure, and the use of over-the-scope clips (OTSCs). Hemostatic clips are one of the most widely used endoscopic closure devices after treatment. In areas where gastrointestinal ligation devices are difficult to use for closure, hemostatic clips are generally selected. Clinically, for small wounds, a single hemostatic clip can completely close the wound; for larger wounds, multiple hemostatic clips are needed for continuous closure similar to sutures, or a combination of hemostatic clips and nylon ropes can be used for purse-string closure. The literature reports that endoscopic purse-string suturing using hemostatic clips and nylon ropes to treat large lesions in the gastric antrum (single lesion with a diameter greater than 3 cm) is safe and feasible [25]. The OTSC system with a rake-shaped metal clip can effectively close large-diam-

Table 2. Recently developed hemostatic powders and their mechanisms of action

Name	Composition	Mechanism of action
Hemospray	Mineral	Absorption of water; Concentration of platelets and clotting factors
EndoClot	Polysaccharide	Absorption of water; Concentration of platelets and clotting factors
Nexpowder	Biocompatible natural polymer	Modification of water absorption capacity using coating technology; Reversible cross-linking of amine and aldehyde groups
CGGEL	Bioadhesive gels containing recombinant human EGF	Absorption of water; Concentration of platelets and clotting factors

Note: EGF, epidermal growth factor. This table is cited from [27].

eter gastrointestinal perforations or refractory gastrointestinal fistulas. Prophylactic closure of artificial ulcer surfaces with the rake-shaped metal clip can effectively prevent postoperative perforation.

For postoperative bleeding after ESD, through-the-scope clipping (TTSC), OTSC, and suturing represent the main techniques for tissue apposition [25]. TTSC is the most commonly used hemostatic clip in clinical practice at present. It is released through the endoscopic working channel and is suitable for clamping small mucosal or submucosal bleeding points. It has the advantages of easy operation and low cost. However, its clamping force is limited, and it can usually only close superficial tissues. It has poor closure effects for deeper or larger wounds, perforations, or full-thickness defects. In contrast, OTSC is a cap-like device mounted at the front end of the endoscope. It achieves strong occlusion through a spring-loaded mechanism and can simultaneously clamp the mucosal layer, muscular layer, and even the serosal layer. It is suitable for closing full-thickness gastrointestinal wall defects (such as perforations and fistulas) or treating massive bleeding. OTSC has a larger clamping range, deeper grasping ability, and stronger closing force, but the operation is relatively complex, the cost is high, and it usually cannot be released repeatedly. TTSC is suitable for superficial hemostasis during routine ESD, while OTSC is more commonly used for the emergency treatment of complex complications, such as perforations, delayed massive bleeding, or fistulas. The two are complementary in clinical application.

3.3 ESD ulcer shielding

3.3.1 Hemostatic powders

Polyglycolic acid (PGA) is a high-molecular-weight polymer. It is a soft material with certain ductility and tensile strength and can be hydrolyzed into carbon dioxide and water. Therefore, it is safe for application in the human body. The earliest application of PGA in the treatment of wounds after digestive tract ESD was reported in 2012. Takimoto et al. used PGA combined with fibrin glue to prevent delayed perforation after ESD of duodenal tumors with a diameter of 20 mm [26]. The PGA sheet was placed over the ulcer with biopsy forceps and fixed with sprayed fibrin glue, and it was absorbed spontaneously within four to fifteen weeks. This shows that the endoscopic tissue shielding method using PGA sheets and fibrin glue has potential in preventing bleeding after ESD.

In recent years, several hemostatic powders have been developed and clinically used to treat bleeding after gastric ESD. These hemostatic powders can be used clinically (**Table 2**) [27]. In the study by Zhang et al., patients (n=173) who underwent ESD from January 2009 to June 2012 were included in a prospective randomized study [28]. The remaining 171 patients (excluding two patients whose procedures were aborted) were randomly divided into two groups: Group A (medical adhesive group, n=89) and Group B (control group, n=82). There were no significant differences in the average treatment time (59.4 min and 55.0 min, respectively) and the incidence of intraoperative perforation (10.1% in Group A and 9.8% in Group B) between Group A and Group B. There was a significant difference in the average length of hospital stay between Group A and Group B (8.89 days and 9.90 days), and the incidence of delayed bleeding after the procedure in Group A was significantly lower than that in Group B (0.00% and 4.88%).

One of the advantages of applying hemostatic powder is that it requires minimal technical expertise. Since the powder only needs to be sprayed onto the surface of the ulcer after ESD through a catheter, regardless of the anatomical location, hemostatic powder has certain natural advantages in difficult areas [29]. Therefore, beginners in ESD can easily apply hemostatic powder to prevent bleeding after ESD-induced iatrogenic ulcers [27]. However, there is a learning curve for those who use hemostatic powder for the first time. In addition, hemostatic powder can be quickly and easily applied to large lesions. Moreover, due to its non-contact application, hemostatic powder does not cause additional secondary tissue damage [30]. Therefore, applying hemostatic powder is an effective treatment option for protecting the ulcer surface after ESD and reducing postoperative scarring [31]. However, hemostatic powder may be affected by gravity. Most powders adhere well to the surface due to the moisture of the ulcer after ESD. After the endoscopist sprays the powder onto the target lesion, the remaining powder can accumulate in a gravity-dependent manner in the target area.

The hemostatic material TC-325 (Hemospray®, Cook Medical, Winston-Salem, NC, USA) has been reported as an endoscopic hemostatic option for gastrointestinal bleeding, including bleeding after endoscopic procedures [32, 33]. Aziz et al. studied and evaluated the efficacy of Hemospray in patients with non-variceal upper GIB [33]. The primary outcomes were clinical and technical success; the secondary outcomes were clustered rebleeding, early rebleeding, delayed rebleeding, refractory bleeding, mortality, and treatment failure. The final analysis included a total of 20 studies and 1,280 patients. Technical success of Hemospray was observed in 97% of the cases, and a significant trend of increasing technical success was observed during the publication period from 2011 to 2019. The clinical success rate of Hemospray was 91%, while the clinical success rate of other hemostatic measures was 87%. After using Hemospray, the secondary outcomes of clustered rebleeding, early rebleeding, delayed rebleeding, refractory rebleeding, mortality, and treatment failure occurred in 27%, 20%, 9%, 8%, 8%, and 31% of the cases, respectively. Hemospray is safe and effective in treating non-variceal upper GIB, is not inferior to traditional hemostatic measures, and can be used as an alternative option.

Polysaccharide hemostatic materials have been proven in multiple clinical trials to be applicable to the treatment of post-ESD bleeding. Yu et al. used a carbon dioxide insufflator (CO_2 Regulation Unit, Olympus) to insufflate the hemostatic material [34]. This device uses high airflow during the catheter insertion stage to disperse moisture. After aiming the catheter at the ulcer surface, the assistant changes the mode to low airflow to evaluate the efficacy of the newly designed polyethylene oxide (PEO) adhesive in preventing delayed bleeding. The rate of delayed bleeding in the PEO group was relatively lower than that in the control group, but the difference was not significant (P=0.094). No bleeding events occurred within 48 hours after ESD in the PEO group, while the bleeding rate after 48 hours in the control group was 60.9% (14/23). Cha et al. used the hemostatic powder UI-EWD (Nexpowder™, Nextbiomedical, Incheon, South Korea) to study the endoscopic hemostasis success rate, UI-EWD-related adverse events, and rebleeding rate [35]. The incidence of endoscopic hemostasis in the UI-EWD group for lesions located at the hepatic flexure (7.3% vs. 0.0%, p=0.011) and lesions larger than 4 cm (25.5% vs. 8.0%, p=0.002) was significantly higher than that in the conventional treatment group. The cumulative rebleeding rate within 28 days was 5.5% (3/55), significantly lower than that in the conventional treatment group (17.0%, 19/112; p=0.039). For lower GIB with anatomical or technical obstacles to endoscopic access, Cha et al. believe that UI-EWD should be considered first.

3.3.2 New treatment option: self-assembling peptide TDM

The synthetic self-assembling peptide TDM is derived from non-biological materials and consists of 16 amino acid residues. It is usually in the form of an aqueous solution and can easily pass through a catheter. The self-assembling peptide can form a gel under specific physiological conditions [36]. Therefore, the self-assembling peptide TDM can be used not only in open surgery but also in flexible endoscopic treatment.

The first generation of self-assembling peptides (TDM-621, 3-D Matrix, Tokyo, Japan) is currently used as a hemostatic agent in Europe. TDM-621 has some drawbacks. Its gel formation rate is slow, and its retention on the wound is poor. Therefore, the second-generation synthetic peptide TDM-623 was developed. TDM-623 has been developed as the second generation of self-assembling peptide materials, which can accelerate gel formation and improve tissue retention. Yoshida et al. tested the feasibility of the synthetic self-assembling peptide TDM-621 for hemostasis after endoscopic treatment of gastric tumors and evaluated its hemostatic effect in 12 patients [37]. No secondary bleeding was observed in any of the 12 patients, and no adverse reactions considered to be related to TDM-621 were observed. However, this study not only has a small sample size but also lacks a direct efficacy comparison with traditional mechanical clipping or prophylactic electrocoagulation, and further verification is needed. Kubo et al. conducted an animal experiment using the self-assembling peptide TDM-623 and performed endoscopic hemostatic treatment with the self-assembling peptide on 23 bleeding lesions [36]. Seventeen of the 23 cases (74%) achieved complete hemostasis on the anterior wall of the stomach. Kubo et al. believe that the new self-assembling peptide (TDM-623) shows a strong hemostatic effect, and TDM-623 has potential practical value for upper gastrointestinal endoscopic procedures [36]. The success rate of the new-generation material in animal models (74%) is lower than that of the first-generation material in human single-arm studies (100%). This illustrates the limitations of the current study. There are huge differences in hemodynamics and microenvironment between artificial wound hemorrhage in animal models and real human postoperative ESD hemorrhage; at the same time, the remaining 26% hemostatic failure rate also indicates that the procedural protocol relying solely on self-assembling peptides to deal with high-pressure arterial hemorrhage is unreliable. On the whole, the current literature on self-assembling peptides mainly remains at the stage of small-sample case series or animal experiments. Before self-assembling peptides can be routinely applied to upper gastrointestinal endoscopic treatment, large-sample, multicenter randomized controlled trials with positive controls (such as TC-325 hemostatic powder or traditional hemostatic clips) need to be carried out to clarify their true clinical benefits and applicable boundaries.

4 SUMMARY

With the development of endoscopic technology, endoscopic treatment techniques for gastric ESD provide a less invasive and effective treatment method than traditional surgery. Endoscopy is becoming the first-line method for treating precancerous lesions of gastric cancer. Although significant progress has been made in research on post-ESD bleeding in the stomach, there is still room for further research. With population aging and the increasing prevalence of chronic diseases, the use of aspirin and other anticoagulant drugs in clinical practice has increased significantly, and they have become important means for the primary and secondary prevention of cardiovascular diseases. In the face of an aging population and patients receiving anticoagulant therapy, in addition to traditional mechanical closure and thermocoagulation therapy, the development of tissue shielding technology and the emergence of new aqueous hemostatic materials provide a new solution to make up for the limitations of traditional devices in treating large-area mucosal defects.

A review of the treatment of bleeding after ESD in recent years shows that clinical practice urgently needs to shift from simple post-hemorrhage remediation to risk-based procedural management. Doctors not only need to accurately evaluate the use of antithrombotic drugs before the procedure but also need to carry out personalized prevention and treatment according to the characteristics of the lesions during the procedure. Future research should focus on developing dedicated hemostatic protocols for patients receiving anticoagulant therapy and verifying the cost-effectiveness of novel hemostatic materials through multicenter clinical trials to ultimately minimize ESD complications, improve the clinical success rate, and reduce procedural risk [38].

DECLARATIONS

Author contributions

Jin Xu was responsible for the literature search, literature screening, analysis and summary of relevant studies, and drafting of the manuscript. Shiju Yan provided academic guidance, supervised the overall design and structure of the review, and critically revised the manuscript for important intellectual content. All authors read and approved the final manuscript.

Funding

This research received no external funding.

Data availability

Not applicable.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.

REFERENCES

[1] Oda I, Suzuki H, Nonaka S, Yoshinaga S. Complications of gastric endoscopic submucosal dissection. Dig Endosc. 2013 Mar;25 Suppl 1:71-78. https://doi.org/10.1111/j.1443-1661.2012.01376.x

[2] Libânio D, Costa MN, Pimentel-Nunes P, Dinis-Ribeiro M. Risk factors for bleeding after gastric endoscopic submucosal dissection: A systematic review and meta-analysis. Gastrointest Endosc. 2016 Oct;84(4):572-586. https://doi.org/10.1016/j.gie.2016.06.033

[3] Jeon MY, Park JC, Hahn KY, Shin SK, Lee SK, Lee YC. Long-term outcomes after noncurative endoscopic resection of early gastric cancer: The optimal time for additional endoscopic treatment. Gastrointest Endosc. 2018 Apr;87(4):1003-1013.e2. https://doi.org/10.1016/j.gie.2017.10.004

[4] Zhao Y, Wang C. Long-term clinical efficacy and perioperative safety of endoscopic submucosal dissection versus endoscopic mucosal resection for early gastric cancer: An updated meta-analysis. BioMed Res Int. 2018 Jan 11;2018:3152346. https://doi.org/10.1155/2018/3152346

[5] Kataoka Y, Tsuji Y, Sakaguchi Y, Minatsuki C, Asada-Hirayama I, Niimi K, et al. Bleeding after endoscopic submucosal dissection: Risk factors and preventive methods. World J Gastroenterol. 2016 Jul 14;22(26):5927-5935. https://doi.org/10.3748/wjg.v22.i26.5927

[6] Tsuji Y, Ohata K, Ito T, Chiba H, Ohya T, Gunji T, et al. Risk factors for bleeding after endoscopic submucosal dissection for gastric lesions. World J Gastroenterol. 2010 Jun 21;16(23):2913-2917. https://doi.org/10.3748/wjg.v16.i23.2913

[7] Mujtaba S, Chawla S, Massaad JF. Diagnosis and management of non-variceal gastrointestinal hemorrhage: A review of current guidelines and future perspectives. J Clin Med. 2020 Feb 2;9(2):402. https://doi.org/10.3390/jcm9020402

[8] Park JY, Kim MS, Kim BJ, Kim JG. Analysis of risk factors for bleeding after gastric endoscopic submucosal dissection with an emphasis on age and sex differences: A nationwide population-based study. Surg Endosc. 2025 Dec;39:8418-8431. https://doi.org/10.1007/s00464-025-12201-6

[9] Koh R, Hirasawa K, Yahara S, Oka H, Sugimori K, Morimoto M, et al. Antithrombotic drugs are risk factors for delayed post-operative bleeding after endoscopic submucosal dissection for gastric neoplasms. Gastrointest Endosc. 2013 Sep;78(3):476-483. https://doi.org/10.1016/j.gie.2013.03.008

[10] Ono H, Yao K, Fujishiro M, Oda I, Uedo N, Nimura S, et al. Guidelines for endoscopic submucosal dissection and endoscopic mucosal resection for early gastric cancer (second edition). Dig Endosc. 2021 Jan;33(1):4-20. https://doi.org/10.1111/den.13883

[11] Toyonaga T, Nishino E, Man-I M, East JE, Azuma T. Principles of quality controlled endoscopic submucosal dissection with ap-

propriate dissection level and high quality resected specimen. Clin Endosc. 2012 Nov;45(4):362-374. https://doi.org/10.5946/ce.2012.45.4.362

[12] Mukai S, Cho S, Kotachi T, Shimizu A, Matuura G, Nonaka M, et al. Analysis of delayed bleeding after endoscopic submucosal dissection for gastric epithelial neoplasms. Gastroenterol Res Pract. 2012;2012:875323. https://doi.org/10.1155/2012/875323

[13] Matsumura T, Arai M, Maruoka D, Okimoto K, Minemura S, Ishigami H, et al. Risk factors for early and delayed post-operative bleeding after endoscopic submucosal dissection of gastric neoplasms, including patients with continued use of antithrombotic agents. BMC Gastroenterol. 2014 Oct 3;14:172. https://doi.org/10.1186/1471-230X-14-172

[14] Furuhata T, Kaise M, Hoteya S, Iizuka T, Yamada A, Nomura K, et al. Postoperative bleeding after gastric endoscopic submucosal dissection in patients receiving antithrombotic therapy. Gastric Cancer. 2017 Jan;20(1):207-214. https://doi.org/10.1007/s10120-015-0588-7

[15] Toyokawa T, Inaba T, Omote S, Okamoto A, Miyasaka R, Watanabe K, et al. Risk factors for perforation and delayed bleeding associated with endoscopic submucosal dissection for early gastric neoplasms: Analysis of 1123 lesions. J Gastroenterol Hepatol. 2012 May;27(5):907-912. https://doi.org/10.1111/j.1440-1746.2011.07039.x

[16] Sato C, Hirasawa K, Koh R, Ikeda R, Fukuchi T, Kobayashi R, et al. Postoperative bleeding in patients on antithrombotic therapy after gastric endoscopic submucosal dissection. World J Gastroenterol. 2017 Aug 14;23(30):5557-5566. https://doi.org/10.3748/wjg.v23.i30.5557

[17] Liu L, Liu H, Feng Z. A narrative review of postoperative bleeding in patients with gastric cancer treated with endoscopic submucosal dissection. J Gastrointest Oncol. 2022 Feb;13(1):413-425. https://doi.org/10.21037/jgo-21-466

[18] Nam HS, Choi CW, Kim SJ, Kim HW, Kang DH, Park SB, et al. Risk factors for delayed bleeding by onset time after endoscopic submucosal dissection for gastric neoplasm. Sci Rep. 2019 Feb 25;9(1):2674. https://doi.org/10.1038/s41598-019-39381-1

[19] Tomida H, Yoshio T, Igarashi K, Morita Y, Oda I, Inoue T, et al. Influence of anticoagulants on the risk of delayed bleeding after gastric endoscopic submucosal dissection: A multicenter retrospective study. Gastric Cancer. 2021 Jan;24(1):179-189. https://doi.org/10.1007/s10120-020-01105-0

[20] Cho SJ, Choi IJ, Kim CG, Lee JY, Nam BH, Kwak MH, et al. Aspirin use and bleeding risk after endoscopic submucosal dissection in patients with gastric neoplasms. Endoscopy. 2012 Feb;44(2):114-121. https://doi.org/10.1055/s-0031-1291459.

[21] Yang C, Qiu Y, Li X, Shi R. Bleeding after endoscopic submucosal dissection of gastric lesions. J Dig Dis. 2020 Mar;21(3):139-146. https://doi.org/10.1111/1751-2980.12850

[22] Jaruvongvanich V, Sempokuya T, Wijarnpreecha K, Ungprasert P. Continued versus interrupted aspirin use and bleeding risk after endoscopic submucosal dissection of gastric neoplasms: A meta-analysis. Ann Gastroenterol. 2018 May-Jun;31(3):344-349. https://doi.org/10.20524/aog.2018.0251

[23] Paspatis GA, Arvanitakis M, Dumonceau JM, Barthet M, Saunders B, Turino SY, et al. Diagnosis and management of iatrogenic endoscopic perforations: European society of gastrointestinal endoscopy (ESGE) position statement - update 2020. Endoscopy. 2020 Sep;52(9):792-810. https://doi.org/10.1055/a-1222-3191

[24] Conway JD, Adler DG, Diehl DL, Farraye FA, Kantsevoy SV, Kaul V, et al. Endoscopic hemostatic devices. Gastrointest Endosc. 2009 May;69(6):987-996. https://doi.org/10.1016/j.gie.2008.12.251

[25] Papaefthymiou A, Norton B, Telese A, Murray C, Murino A, Johnson G, et al. Endoscopic suturing and clipping devices for defects in the GI tract. Best Pract Res Clin Gastroenterol. 2024 Jun;70:101915. https://doi.org/10.1016/j.bpg.2024.101915

[26] Takimoto K, Toyonaga T, Matsuyama K. Endoscopic tissue shielding to prevent delayed perforation associated with endoscopic submucosal dissection for duodenal neoplasms. Endoscopy. 2012;44 Suppl 2 UCTN:E414-E415. https://doi.org/10.1055/s-0032-1325739

[27] Jung DH, Park JC. Strategies that reduce post-endoscopic submucosal dissection bleeding. Korean J Helicobacter Up Gastrointest Res. 2021;21(3):194-202. https://doi.org/10.7704/kjhugr.2021.0028

[28] Zhang Y, Chen Y, Qu C, Zhou M, Ni Q, Xu L. Effects of medical adhesives in prevention of complications after endoscopic submucosal dissection. World J Gastroenterol. 2013 May;19(17):2704-2708. https://doi.org/10.3748/wjg.v19.i17.2704

[29] Baracat FI, de Moura DTH, Brunaldi VO, Tranquillini CV, Baracat R, Sakai P, et al. Randomized controlled trial of hemostatic powder versus endoscopic clipping for non-variceal upper gastrointestinal bleeding. Surg Endosc. 2020 Jan;34(1):317-324. https://doi.org/10.1007/s00464-019-06769-z

[30] Karna R, Deliwala S, Ramgopal B, Mohan BP, Kassab L, Becq A, et al. Efficacy of topical hemostatic agents in malignancy-related GI bleeding: A systematic review and meta-analysis. Gastrointest Endosc. 2023 Feb;97(2):202-208.e8. https://doi.org/10.1016/j.gie.2022.07.033

[31] Beye B, Barret M, Alatawi A, Beuvon F, Nicco C, Pratico CA, et al. Topical hemostatic powder promotes reepithelialization and reduces scar formation after extensive esophageal mucosal resection: Prevention of esophageal stricture. Dis Esophagus. 2016 Aug;29(6):520-527. https://doi.org/10.1111/dote.12378

[32] Pittayanon R, Khongka W, Linlawan S, Thungsuk R, Aumkaew S, Teeratorn N, et al. Hemostatic powder vs standard endoscopic treatment for gastrointestinal tumor bleeding: A multicenter randomized trial. Gastroenterology. 2023 Sep;165(3):762-772.e2. https://doi.org/10.1053/j.gastro.2023.05.042

[33] Aziz M, Weissman S, Mehta TI, Hassan S, Khan Z, Fatima R, et al. Efficacy of Hemospray in non-variceal upper gastrointestinal bleeding: A systematic review with meta-analysis. Ann Gastroenterol. 2020 Mar-Apr;33(2):145-154. https://doi.org/10.20524/aog.2020.0448

[34] Yu Y, Hu T, Kuai X, Liu X, Li R, Zhou C. Propensity score matching analysis to evaluate efficacy of polyethylene oxide ad-

hesive on preventing delayed bleeding after gastric endoscopic submucosal dissection. Sci Rep. 2022 Mar 16;12(1):4538. https://doi.org/10.1038/s41598-022-08499-0

[35] Cha B, Lee D, Shin J, Park JS, Kwon GS, Kim H. Hemostatic efficacy and safety of the hemostatic powder UI-EWD in patients with lower gastrointestinal bleeding. BMC Gastroenterol. 2022 Apr 7;22(1):170. https://doi.org/10.1186/s12876-022-02247-4

[36] Kubo Y, Kobayashi S, Yamamoto K, Nakagawa Y, Yamashita K, Saito T, et al. Endoscopic application of novel, infection-free, advanced hemostatic material: Its usefulness to upper gastrointestinal oozing. DEN Open. 2021 Aug 22;2(1):e25. https://doi.org/10.1002/deo2.25

[37] Yoshida M, Goto N, Kawaguchi M, Koyama H, Kuroda J, Kitahora T, et al. Initial clinical trial of a novel hemostat, TDM-621, in the endoscopic treatments of the gastric tumors. J Gastroenterol Hepatol. 2014 Dec;29 Suppl 4:77-79. https://doi.org/10.1111/jgh.12798

[38] Cereatti F, Grassia R, Drago A, Conti CB, Donatelli G. Endoscopic management of gastrointestinal leaks and fistulae: What option do we have? World J Gastroenterol. 2020 Aug 7;26(29):4198-4217. https://doi.org/10.3748/wjg.v26.i29.4198

Progress in
Medical Devices

Zentime

RESEARCH ARTICLE

Cardiac function state recognition model based on bimodal time–frequency representation

Mingzhi Zhang, Piding Li

School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Piding Li.

Address correspondence to: Piding Li, School of Health Science and Engineering, University of Shanghai for Science and Technology, No. 516 Jungong Road, Yangpu District, Shanghai 200093, China. E-mail: lpdbyusst@163.com.

Received November 28, 2025; Accepted February 27, 2026; Published June 24, 2026

DOI: 10.61189/784716ypyhmm

Abstract

Objective: This study uses dual-modality signals, including phonocardiogram (PCG) and electrocardiogram (ECG), together with machine learning methods to distinguish cardiac function states in subjects. **Methods:** We developed a model based on time–frequency representations. The model includes data preprocessing, a time–frequency conversion module, a feature extraction module, and a feature-fusion classifier module. The system uses complete ensemble empirical mode decomposition with adaptive noise to remove noise from the PCG and applies filters to reduce noise in the ECG. The system extracts Mel-frequency cepstral coefficients from the PCG and uses Fourier synchrosqueezed transform for the ECG. This study also improves VGG16 and ResNet18 as feature extractors by inserting a variant attention mechanism into the feature extraction networks. Finally, the system feeds the feature vector into a support vector machine for classification. **Results:** The dual-modality time–frequency method achieves 95.4% accuracy and 97.4% sensitivity for positive cases on public datasets, demonstrating strong performance in cardiac function classification. **Conclusion:** This research shows that the approach improves both diagnostic accuracy and sensitivity. The system provides valuable support for the preliminary screening of cardiac dysfunction.

Keywords: Multi-modal, Phonocardiogram signal, Electrocardiogram signal, Feature encoding, Heart disease screening

Highlights

- We use two types of cardiac physiological signals together. They complement each other and help improve the final classification accuracy.
- This study converts phonocardiograms and electrocardiograms into time–frequency images, which helps increase the positive detection rate and enables automatic learning of modality-specific features through a neural network.
- This study modifies the baseline model to achieve a more streamlined neural network architecture and incorporates an attention mechanism to better focus on information correlations.

1 INTRODUCTION

Cardiovascular disease remains one of the leading causes of death and illness worldwide, currently accounting for 16% of all deaths [1]. The prevalence of cardiovascular disease in China continues to rise, with an estimated 330 million people now affected [2]. Phonocardiogram (PCG) and electrocardiogram (ECG) serve as critical physiological parameters for evaluating cardiac function. PCG mainly reflect the mechanical and acoustic features of valve movement and blood flow, while ECG records the heart's electrophysiological activity [3, 4].

Current research has widely examined single-modality classification of PCG and ECG [5, 6]. Advances in sensing and signal-

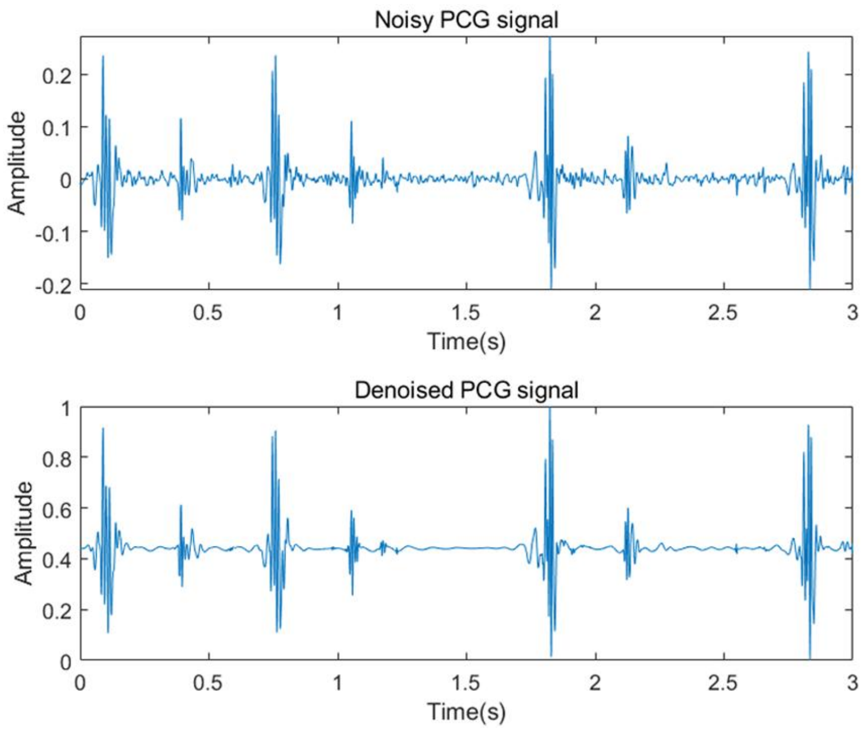


Figure 1. Comparison of PCG signals before and after preprocessing. PCG, phonocardiogram.

processing technologies have made multimodal technology an emerging direction, as the two modalities together provide complementary information [7].

Chakir et al. extracted ten time-domain statistical features from ECG and PCG [8]. Li et al. divided the PCG signal into four frequency bands and used convolutional neural networks and long short-term memory for feature extraction [9]. Sun et al. improved coronary artery disease detection by decoupling ECG and PCG to create coupled signals, and extracted entropy features and recursive graph depth features [10]. Li et al. enhanced the Dempster-Shafer evidence theory to fuse classification results, which improved the accuracy of the final decision [11]. Zhang et al. proposed an end-to-end Co-learning-assisted Progressive Dense Fusion Network with a three-branch interleaved architecture, which showed strong performance on both public and private datasets [12]. Zhu et al. proposed Dual-Scale Deep Residual Network, which uses a dual-scale feature aggregation module to merge low-level features from different scales [13]. Liu et al. used the Vision Transformer to detect coronary heart disease [14]. However, several challenges remain, including low positive detection rates for medical conditions, limited feature representations, and complex network architectures.

This study proposes a cardiac function state recognition method based on dual-modality time–frequency representations. The time–frequency maps provide a comprehensive view of the cardiac signals. A neural network module then automatically extracts features from these maps, and finally, a machine learning model performs classification.

2 MATERIALS AND METHODS

2.1 Overall architecture

To address the aforementioned issues, this study proposes a cardiac function state recognition model based on dual-modality PCG and ECG data.

The proposed model framework consists of four main modules: the signal preprocessing module, the time–frequency information conversion module, the feature vector extraction module, and the feature fusion and classification module.

2.2 Data preprocessing

Given the characteristics of the signals and the presence of noise, this study first performed preprocessing and denoising on the PCG and ECG signals separately.

Noise in PCG signals usually comes from friction sounds caused by a subject's breathing or movements [15]. To denoise the PCG signals, this study adopted the Complete Ensemble Empirical Mode Decomposition with Adaptive Noise proposed in the literature [16]. First, the method was used to extract multi-timescale intrinsic mode function (IMF) components. Next, detrend fluctuation analysis metrics were calculated for these IMF components. Finally, adaptive dual-threshold wavelet analysis was applied to the IMF components based on these metrics to perform denoising and signal reconstruction, producing the cleaned PCG signal. Following preprocessing, the PCG signal was normalized to the 0–1 range.

Figure 1 shows the PCG signals before and after noise reduction. The noise was effectively removed without altering the original waveform shape.

The main sources of noise in ECG signals are baseline drift caused by a subject's respiration and power-line interference from electrical equipment [17]. To remove baseline drift, this study applied a fourth-order Butterworth high-pass filter with a cutoff frequency of 0.5 Hz. To eliminate 60 Hz power-line interference and its 120 Hz harmonic, a second-order infinite impulse response comb filter was used on the ECG signal. **Figure 2** shows the ECG signals before and after noise reduction. The noise was also eliminated.

The total duration of normal sample recordings was 3,770 seconds, while abnormal samples totalled 8,897 seconds. To bal-

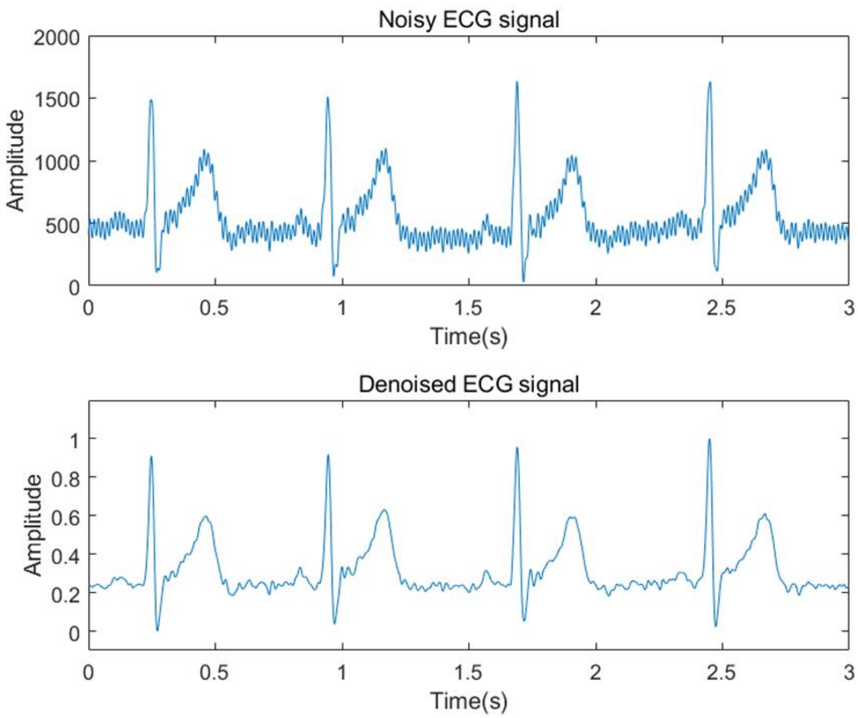


Figure 2. Comparison of ECG signals before and after preprocessing. ECG, electrocardiogram.

ance the number of positive and negative samples, this study used overlapping sliding windows to segment the raw signals. Each window spanned 3 seconds, with a stride of 1.25 seconds for normal samples. After segmentation, there were 2,809 normal segments and 2,786 abnormal segments.

2.3 Time-frequency information conversion module

A spectrogram is a time–frequency representation of a signal, converting a one-dimensional time-domain signal into a two-dimensional image that contains both time and frequency information. This allows for a more comprehensive analysis of non-stationary signals. Compared with one-dimensional time-series signals, time–frequency plots preserve both temporal and frequency details, helping models capture the dynamic evolution of signals more effectively. Unlike traditional spectral analysis, which provides only global frequency information, time–frequency plots show when specific frequencies occur or change, offering greater flexibility and precision.

The time–frequency plot shows time on the horizontal axis and frequency on the vertical axis. Color or brightness represents the signal's power or energy, with brighter colors indicating higher energy at a specific time and frequency. This representation provides neural networks with input that carries rich physical meaning.

2.3.1 Mel-frequency cepstral coefficients (MFCC) transformation of PCG signals

PCG signals are non-stationary, with frequencies that change over time. This study uses MFCC as the time–frequency representation for one-dimensional PCG signals. MFCC has the advantage of simulating the human ear's sensitivity to low frequencies using Mel filter banks, which map the linear spectrum onto the Mel scale.

The logarithmic Mel-spectrogram is created by first filtering a signal through a set of Mel-scale filters, followed by logarithmic compression applied to the resulting power spectrum.

First, the power spectral density of the signal $P(k)$ is calculated:

$$P(k) = \frac{1}{N}\left| X(k) \right|^2 \tag{1}$$

where $X(k)$ denotes the frequency component of the signal in the k band, and N represents the number of sample points.

Next, the Mel filter bank is calculated.

The Mel-scale frequency intervals for the Mel filter bank are:

$$\Delta\phi = (\phi_{\max} - \phi_{\min})/(M+1) \tag{2}$$

where M denotes the number of Mel filters. The Mel filter bank covers a minimum frequency of $\phi_{\min}$ and a maximum frequency of $\phi_{\max}$.

Therefore, the centre frequencies of the Mel scale for the Mel filters in Group M are:

$$\phi_c(m) = m\Delta\phi,\ \ 0 \le m \le M \tag{3}$$

The formula for converting frequency from hertz to the Mel scale is:

$$\phi = 2595\log_{10}(1 + \frac{f}{700}) \tag{4}$$

The formula for converting frequency from the Mel scale back to hertz is:

$$f = 700(10^{\frac{\phi}{2595}} - 1) \tag{5}$$

The frequency response formula for an equal-height triangular filter bank is:

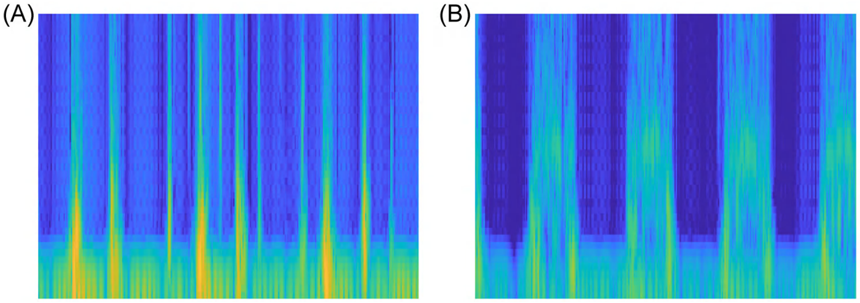


Figure 3. Logarithmic Mel-spectrograms of normal and abnormal PCG. (A) Normal PCG; (B) Abnormal PCG. ECG, electrocardiogram; PCG, phonocardiogram.

$$H_m(k) = \begin{cases} 0, & k < f(m-1) \\ \dfrac{k - f(m-1)}{f(m) - f(m-1)}, & f(m-1) \le k \le f(m) \\ \dfrac{f(m+1) - k}{f(m+1) - f(m)}, & f(m) < k \le f(m+1) \\ 0, & k > f(m+1) \end{cases} \tag{6}$$

where $f(m)$ denotes the centre frequency of the m-th filter bank in hertz.

Next, the logarithmic spectral energy distribution of the Mel frequencies for each frame is calculated:

$$S(m) = \log_e\left(\sum_{k=0}^{N-1} P(k) \cdot H_m(k)\right), \ 0 \le m \le M \tag{7}$$

In this study, the MFCC window length was set to 25 milliseconds to balance temporal and frequency resolution. The overlap length was set to 15 milliseconds to reduce information loss between frames and to smooth time–frequency variations. The number of Mel filters was set to 40, providing higher resolution in the low-frequency range while still covering the high-frequency range. The fast Fourier transform length was set to 512 to balance efficiency and accuracy. Since PCG signals have frequency components between 0 and 500 Hz, the frequency range was set to 0–500 Hz [18].

The MFCCs of normal and abnormal PCG signals are shown in **Figure 3A** and **3B**, respectively.

The spectrogram of normal PCG signals shows higher energy in the low-frequency range and displays more distinct frequency bands.

2.3.2 Fourier synchrosqueezing transform (FSST) transformation of ECG signals

ECG signals are also non-stationary. FSST allows visualization of their frequency information over time. When ECG signals have abrupt changes or transient frequency components, FSST can identify their instantaneous frequency characteristics [19]. By synchronously compressing the results of the short-time Fourier transform (STFT), FSST concentrates dispersed energy toward the instantaneous frequency curve, producing a clearer time–frequency representation.

First, the STFT of the input signal is computed:

$$V(t, f) = \int_{-\infty}^{+\infty} x(\tau) \cdot g(\tau - t) \cdot e^{-j2\pi f\tau} d\tau \tag{8}$$

where f represents frequency, τ represents the time parameter, and $g(\tau\text{-}t)$ represents the window function.

The STFT is limited by the principle of time–frequency uncertainty. Signal energy spreads in the time–frequency plane due to the window function, causing frequency blurring. To estimate instantaneous frequency, the partial derivative with respect to time is calculated:

$$\hat{\omega}(t, f) = \mathrm{Re}\left(\frac{1}{2\pi i} \times \frac{\partial_t(V(t, f))}{V(t, f)}\right) \tag{9}$$

The coefficients $V(t, f)$ obtained from the STFT are first mapped along the frequency axis to the estimated instantaneous frequency $\hat{\omega}(t, f)$. Synchronous compression is then applied to redistribute the energy, producing the amplitude of the FSST:

$$T(t, \omega) = \frac{1}{g(0)} \int_{-\infty}^{+\infty} V(t, f) \delta(\omega - \hat{\omega}(t, f)) dt \tag{10}$$

where ω represents the frequency after synchronous compression, $g(0)$ denotes the window function at time zero, and δ represents the Dirac delta function.

The FSST tracings for normal and abnormal ECG signals are shown in **Figure 4A** and **4B** respectively.

Normal ECG signals show cleaner and more concentrated energy in the FSST, whereas abnormal ECG signals display chaotic and dispersed energy across the spectrum.

In this study, the FSST used a Kaiser window function with a length of 64. Since ECG signals contain frequency components between 0 and 250 Hz, the frequency range was set to 0–250 Hz accordingly [14].

2.4 Feature vector extraction module

2.4.1 PCG feature vector extraction module

This study uses VGG16 as the baseline model for fine-tuning and modification. The network architecture and parameters are shown in **Table 1** and **Figure 5**.

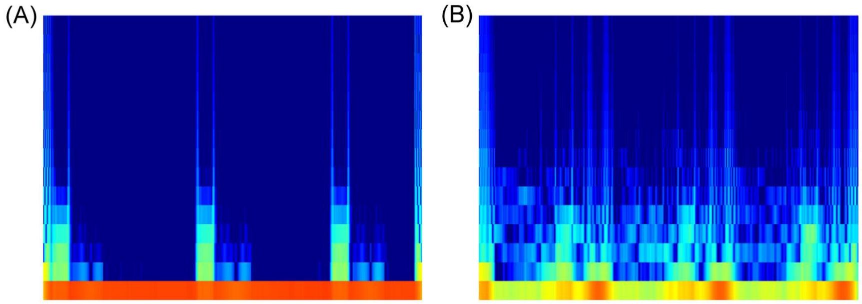


Figure 4. FSST plots of normal and abnormal ECG signals. (A) Normal ECG signal; (B) Abnormal ECG signal. ECG, electrocardiogram; PCG, phonocardiogram.

Table 1. Parameters of the PCG feature vector extraction module

Layer	Structure
Conv	3×3, Cin=Cout, S=1
Conv1	3×3, Cin=3, Cout=16, S=1
Conv2	3×3, Cin=16, Cout=32, S=1
Conv3	3×3, Cin=32, Cout=64, S=1
Conv4	3×3, Cin=64, Cout=256, S=1
Conv5	3×3, Cin=256, Cout=512, S=1
Dropout1	p=0.1
Dropout2	p=0.5
MaxPool	2×2, S=2
DA Module	3×3, d_model=512, H=8, W=8
Linear	In=512×4×4, Out=64
Fully connected	In=64, Out=2

Note: PCG, phonocardiogram; Cin, the number of input channels for the convolutional layer; 3×3, the size of the convolutional kernel; Cout, the number of output channels; p, the dropout probability for the Dropout layer; S, stride; d_model, the number of feature channels; H, the height of the feature map; W, width; In, the input dimension size for the linear layer; Out, the output dimension size; DA Module, Dual Relation-Aware Attention Network Module; MaxPool, max pooling layer; Conv, convolutional layer; Linear, linear layer; Fully Connected, fully connected layer; Dropout, dropout layer.

Additionally, this study incorporates the Dual Relation-Aware Attention Network Module (DA Module) to improve classification accuracy [20]. The architecture of the DA Module is shown in **Figure 6**.

The DA Module contains three components: the Position Attention Module, the Channel Attention Module, and the additive fusion layer. Both the input size and the output size are C×H×W.

The structures of the Position Attention Module and Channel Attention Module are shown in **Figure 7A** and **7B** respectively.

Within the Position Attention Module, the input feature map A first passes through a convolutional layer to extract features. The module then replicates these features into three components: Query, Key, and Value. Each component is reshaped before the positional attention calculation begins. After reshaping, the Query feature map B is transposed to size HW×C, while the Key feature map C is reshaped to size C×HW. The attention scores between these two maps are then computed, reducing the feature map size to HW×HW. The module multiplies this encoded positional attention by the Value feature map *D*, and finally reshapes the result back to C×H×W. The computational process is as follows:

$$K = \mathrm{reshape}(D \cdot \mathrm{Softmax}(B^{T} \cdot C)) \tag{11}$$

where $B \in \mathbb{R}^{C\times HW}$, $C \in \mathbb{R}^{C\times HW}$, $D \in \mathbb{R}^{C\times HW}$, $K \in \mathbb{R}^{C\times H\times W}$.

After obtaining the feature map K modulated by the positional attention mechanism, the module multiplies it by the weighting coefficient α. It then fuses this result with the original input feature map A through addition:

$$E = \alpha K + A \tag{12}$$

where α is set to 0.5.

Similarly, within the Channel Attention Module, the input feature map A is first reshaped to C×HW. The module then multiplies A by its transpose A^T, producing a matrix of size C×C. After that, the feature is normalised using the Softmax function to generate attention scores between channels. These attention scores are then multiplied by A and finally reshaped back to C×H×W. The computational process is as follows:

$$U = \mathrm{reshape}(\mathrm{Softmax}(A \cdot A^{T}) \cdot A) \tag{13}$$

where $A \in \mathbb{R}^{C\times HW}$ and $U \in \mathbb{R}^{C\times H\times W}$.

After obtaining the feature map U modulated by the channel attention mechanism, the module multiplies it by the weighting coefficient β and fuses it with the input feature map A through summation:

$$E = \beta U + A \tag{14}$$

where β is set to 0.2.

The outputs from the two attention modules are fused through summation, enriching the spatial contextual information and strengthening the discriminative power of the channel features.

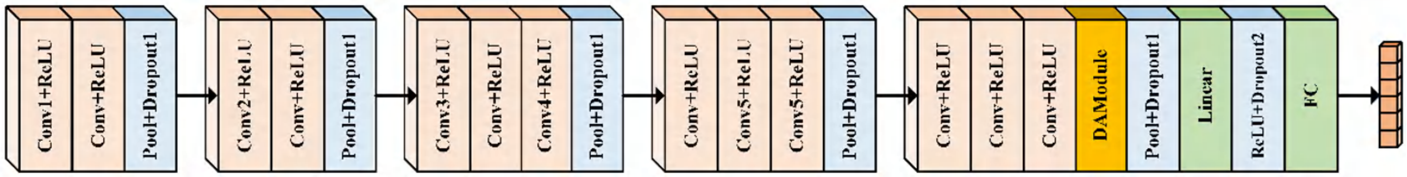


Figure 5. Backbone architecture of the PCG feature vector extraction module. PCG, phonocardiogram; Conv, convolutional layer; ReLU, Linear rectification function; Pool, pooling layer; Dropout, dropout layer; DA Module, Dual Relation-Aware Attention Network Module; Linear, linear layer; FC, fully connected layer.

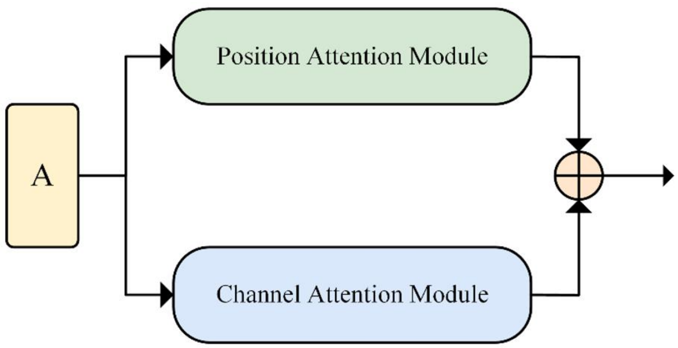


Figure 6. DA Module structural diagram. DA Module, Dual Relation-Aware Attention Network Module.

2.4.2 ECG feature vector extraction module

This study uses ResNet18 as the baseline model for fine-tuning and modification. The network architecture and parameters are shown in **Table 2** and **Figure 8**.

Here, SC represents the residual connections, DSConv represents depthwise separable convolutions, and MidConv denotes the convolutional layers where the input and output dimensions are identical.

To reduce the number of parameters and computational load, this study replaces the convolutional layers within residual blocks with depthwise separable convolutions. Depthwise separable convolutions decompose standard convolution operations into two lighter, decoupled steps: depthwise convolution and pointwise convolution. The depthwise convolution uses fixed-size kernels to independently convolve each channel and stack the resulting feature maps. The subsequent pointwise convolution employs a 1×1 kernel to aggregate information from the depthwise stage, enabling linear combinations across channels. To further enhance inter-channel information fusion, a MidConv layer is introduced to strengthen feature interactions among different channels.

Additionally, this study incorporates the Selective Kernel Module (SK Module) to improve classification accuracy [21]. The architecture of the SK Module is shown in **Figure 9**.

The SK Module includes three components: Split, Fuse, and Select.

In the Split section, two convolutional kernels of different sizes are first applied to the input feature map X, producing Ũ and Û, respectively. These two outputs are then added together to generate U. The computational process is as follows:

$$\tilde{U} = \tilde{F}_{3\times3}(X) \quad (15)$$

$$\hat{U} = \tilde{F}_{5\times5}(X) \quad (16)$$

$$U = \tilde{U} + \hat{U} \quad (17)$$

where $X, \tilde{U}, \hat{U} \in \mathbb{R}^{H\times W\times C}$.

The Fuse component uses global average pooling to extract key information from the overall context:

$$S = F_{gp}(U) \quad (18)$$

where $S \in \mathbb{R}^{1\times1\times C}$.

Subsequently, a fully connected layer is used to compress the key information S into a compact representation Z:

$$Z = \mathrm{ReLU}(\mathrm{BatchNorm}(F_{fc}(S))) \quad (19)$$

where $Z \in \mathbb{R}^{1\times1\times d}$ and $d<C$.

The Select component then processes Z usingtwo fully connected layers of size $d\times C$. The resulting outputs are concatenated and normalized via the softmax function to generate the weights $W \in \mathbb{R}^{2\times1\times C}$:

$$W_a = \tilde{F}_{fc}(Z) \quad (20)$$

$$W_b = \hat{F}_{fc}(Z) \quad (21)$$

$$W = \mathrm{softmax}(\mathrm{Concat}(W_a, W_b)) \quad (22)$$

where $W_a, W_b \in \mathbb{R}^{1\times1\times C}$.

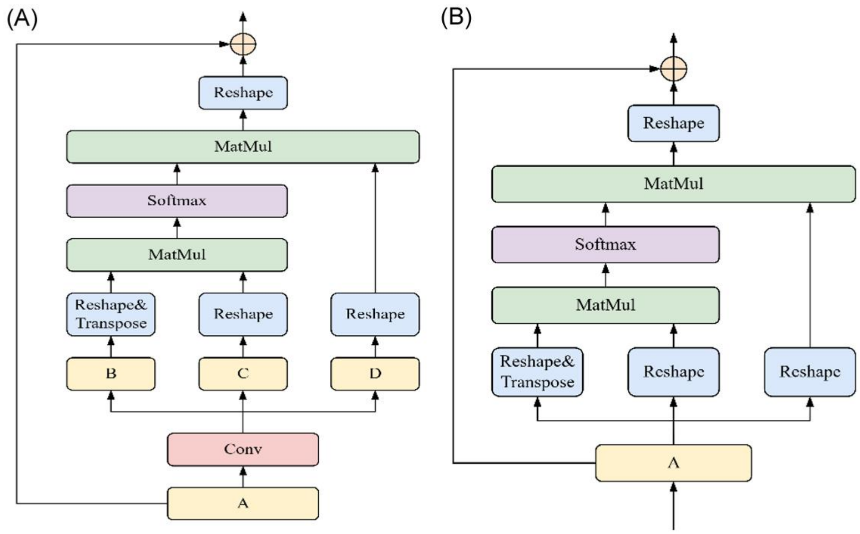


Figure 7. Submodule architecture of the DA Module. (A) Architecture of the position attention module; (B) Architecture of the channel attention module. DA Module, Dual Relation-Aware Attention Network Module.

Table 2. Parameters of the ECG feature vector extraction module

Layer	Structure
Conv	7×7, Cin=3, Cout=64, S=2
DSConv1	3×3, Cin=Cout=64, S=1
DSConv2	3×3, Cin=Cout=128, S=1
DSConv3	3×3, Cin=Cout=256, S=1
DSConv4	3×3, Cin=Cout=512, S=1
MidConv	3×3, Cin=Cout
Dropout1	p=0.4
Dropout2	p=0.7
Pool	AdaptiveAvgPool (1, 1)
SK Module	Kernels=[1, 3, 5, 7], Reduction=8, L=32
Linear	In=512, Out=64
Fully connected	In=64, Out=2

Note: ECG, electrocardiogram; Conv, convolutional layer; Cin, the number of input channels for the convolutional layer; Cout, number of output channels; S, stride; DSConv, depthwise separable convolution; MidConv, intermediate convolution; Dropout, dropout layer; Dropout, dropout layer; p, the dropout probability for the Dropout layer; AdaptiveAvgPool, adaptive average pooling; (1, 1), pooling the feature map to a spatial size of (1, 1); SK Module, Selective Kernel Module; Kernels, the convolutional kernel sizes used in the SK Module; Reduction, the compression ratio; L, the minimum dimension.

Finally, the first dimension is multiplied by weight W and applied to Ũ, while the second dimension is multiplied by weight W and applied to Û. This process produces the feature map V, which is modulated by the attention mechanism with selectable kernels:

$$V = W[0] \odot \tilde{U} + W[1] \odot \hat{U} \tag{23}$$

where $Z \in \mathbb{R}^{H \times W \times C}$.

2.5 Feature fusion and recognition

This study employed a feature vector extraction module to obtain 64-dimensional features from each modality. These features were then concatenated to form a 128-dimensional feature vector. Finally, classification was performed using a Support Vector Machine (SVM) model.

3 EXPERIMENTAL SETUP

3.1 Experimental data

To validate the effectiveness of the proposed method, data from the publicly available Physio+Net/Computing in Cardiology Challenge 2016 dataset were used. The *training*-a subset of this dataset contains 409 samples. After excluding samples without paired ECG recordings and those officially annotated as unclassifiable due to excessive noise, 388 samples remained. Each sample includes synchronously recorded PCG and ECG signals and is labeled as either normal or abnormal. The sampling frequency is 2,000 Hz. Statistical analysis indicates that the sample durations in the *training*-a set range from 9 to 36 seconds. The 388 samples were divided into training and test sets at a ratio of 9:1. To prevent data leakage, segments derived from the same recording were assigned exclusively to either the training set or the test set. To better reflect real-world conditions, record-level classification results were used instead of segment-level results. Specifically, record-level predictions were obtained by averaging the segment-level output probabilities.

3.2 Experimental setup

The model was trained on a system equipped with a single NVIDIA RTX 4060 GPU (16 GB VRAM). The implementation was based on the PyTorch deep learning framework (version 2.5.1). During training, the Adam optimizer was employed for 60 epochs with a learning rate of 1×10^{-5}. The batch size was set to 32, and the cross-entropy loss function was used.

3.3 Evaluation criteria

This study uses accuracy (Acc), sensitivity (Sen), specificity (Spe), and F1 score to evaluate model performance. The definitions of these four metrics are as follows:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{24}$$

$$Sensitivity = \frac{TP}{TP + FN} \tag{25}$$

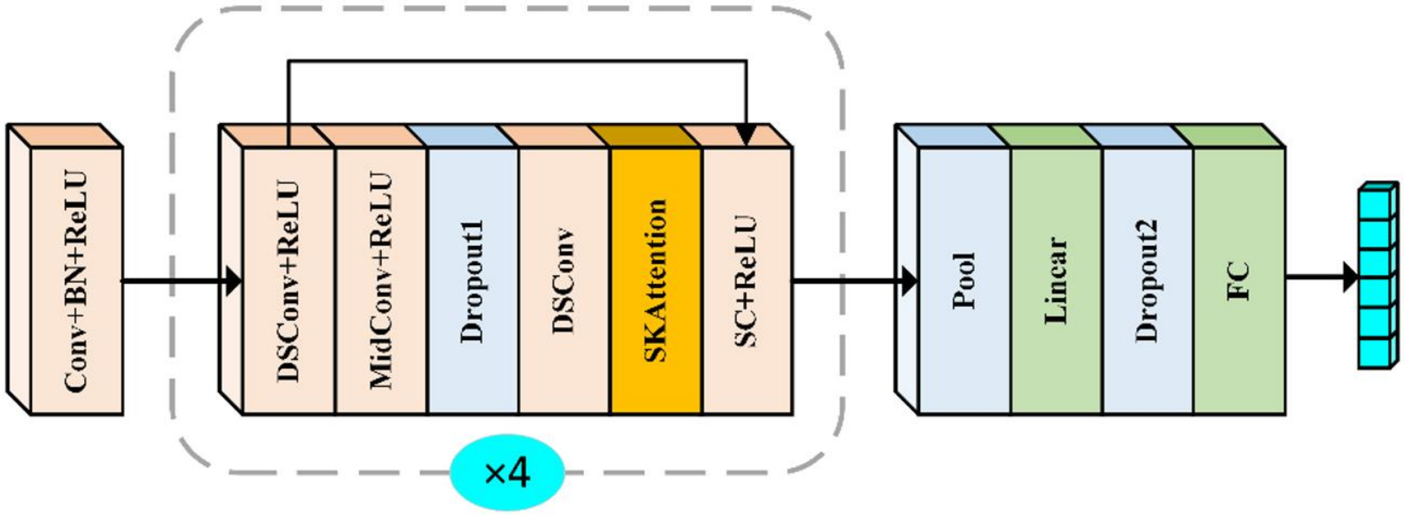


Figure 8. Backbone architecture of the ECG feature vector extraction module. ECG, electrocardiogram; Conv, convolutional layer; BN, batch normalization; ReLU, Linear rectification function; MidConv, intermediate convolution; Dropout, dropout layer; DSConv, depthwise separable convolutions; SK Attention, Selective Kernel Module; SC, Skip Connection; Linear, linear layer; Pool, pooling layer; FC, fully connected layer.

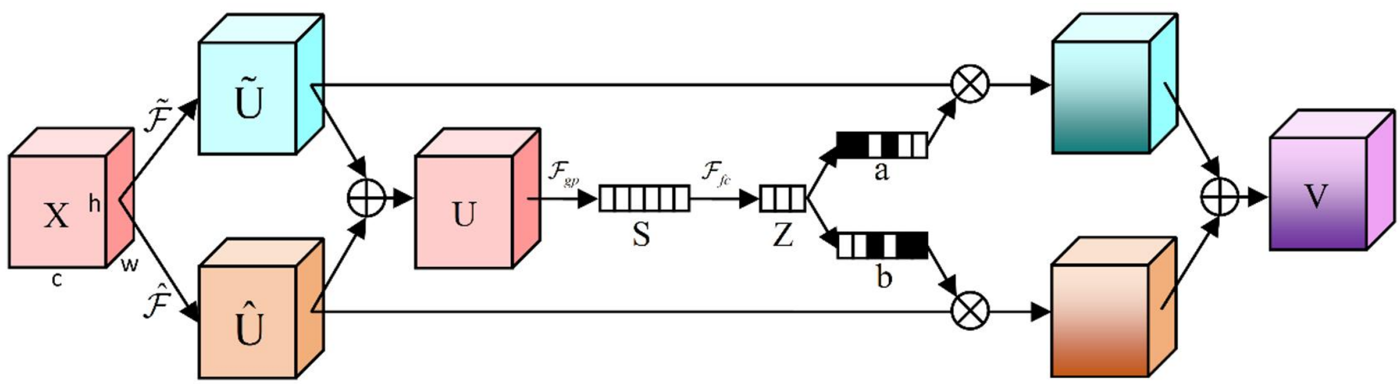


Figure 9. Structure of the SK Module. SK Module, Selective Kernel Module.

$$Specificity = \frac{TN}{TN + FP} \tag{26}$$

$$F1 = 2 \times \frac{Sen \times Pre}{Sen + Pre} \tag{27}$$

where *TP*, *FP*, *TN*, and *FN* represent the number of true positives, false positives, true negatives, and false negatives in the confusion matrix, respectively.

In binary classification, sensitivity measures the proportion of actual positive samples correctly predicted as positive. Higher sensitivity indicates a stronger ability to identify positive samples, resulting in fewer missed detections. Specificity measures the proportion of actual negative samples correctly predicted as negative. Higher specificity indicates a stronger ability to exclude negative examples, resulting in fewer false positives. Precision measures the proportion of true positives among all samples predicted as positive, reflecting the accuracy of the model's predictions. The F1 score is the harmonic mean of sensitivity and precision, providing a balanced metric for overall model performance.

4 RESULTS AND DISCUSSION

As shown in **Table 3**, the improved PCG model achieved higher Acc, Sen, Spe, and F1 score. The PCG model with the Axial Attention Module achieved the highest Spe. However, its overall Acc remained lower than that of the model with the DA Module, which reached 85.3%. As shown in **Table 4**, the improved ECG model incorporating the SK Module achieved the highest Acc, Sen, and F1 scores. **Table 5** presents a comparative assessment of dual-modality versus single-modality results. The PCG and ECG dual-modality network outperformed the single-modality approach across all four metrics, with particularly strong complementary improvement in specificity. While time–frequency images effectively detect abnor-

Table 3. Evaluation metrics for the baseline single-modality PCG model and models with different modules

	Acc	Sen	Spe	F1
Baseline VGG16	0.827	0.890	0.681	0.878
Benchmark	0.840	0.901	0.698	0.888
Benchmark+CBAM	0.822	0.879	0.690	0.874
Benchmark+AA	0.802	0.790	0.828	0.848
Benchmark+DA	0.853	0.901	0.741	0.896

Note: PCG, phonocardiogram; Acc, accuracy; Sen, sensitivity; Spe, specificity; F1, F1 score; VGG16, Visual Geometry Group 16-layer network; CBAM, Convolutional Block Attention Module; AA, Axial Attention; DA, Dual Relation-Aware Attention Network.

Table 4. Evaluation metrics for the baseline single-modality ECG model and models with different modules

	Acc	Sen	Spe	F1
Baseline ResNet18	0.820	0.820	0.819	0.864
Benchmark	0.840	0.853	0.810	0.882
Benchmark+CBAM	0.796	0.846	0.681	0.853
Benchmark+AA	0.874	0.879	0.862	0.907
Benchmark+SK	0.876	0.930	0.750	0.913

Note: ECG, electrocardiogram; Acc, accuracy; Sen, sensitivity; Spe, specificity; F1, F1 score; ResNet18, Residual Network 18-layer; CBAM, Convolutional Block Attention Module; AA, Axial Attention; SK, Selective Kernel.

Table 5. Evaluation metrics for single-modal and dual-modal classification

	Acc	Sen	Spe	F1
PCG-Only	0.853	0.901	0.741	0.896
ECG-Only	0.876	0.930	0.750	0.913
PCG+ECG	0.954	0.974	0.905	0.967

Note: Acc, accuracy; Sen, sensitivity; Spe, specificity; F1, F1 score; PCG, phonocardiogram; ECG, electrocardiogram.

mal cardiac function, the dual-modality approach increases the model's confidence in negative predictions and reduces false alarms. The "Benchmark" in **Tables 3** and **4** refers to the backbone network without any additional modules introduced.

Table 6 compares different classifiers. The decision tree model achieved the highest specificity, but its overall accuracy was relatively low. The SVM achieved the highest Acc, Sen, and F1 score.

Table 7 compares metrics across studies. This research achieved an Acc of 95.4% and a Sen of 97.4%, outperforming other studies. However, there is still room to improve Spe and the F1 score.

5 CONCLUSION

This study proposed a cardiac function state recognition model based on dual-modality time–frequency representations of PCG and ECG signals. First, both PCG and ECG signals underwent preprocessing for noise reduction and segmentation. Next, the one-dimensional time-series signals were converted into time–frequency plots to better capture pathological information. For these two-dimensional time–frequency plots, distinct automatic feature extraction neural networks were designed. Accuracy and sensitivity were improved by refining the baseline model and integrating attention modules. Since minimizing missed diagnoses is crucial in disease screening, this study prioritized high sensitivity. Finally, the 64-dimensional feature vectors from each modality were concatenated and input into an SVM, achieving a classification accuracy of 95.4% and a sensitivity of 97.4%.

Table 6. Evaluation metrics for the ablation experiments of the replacement classifier

	Acc	Sen	Spe	F1
KNN	0.915	0.960	0.810	0.941
DT	0.879	0.864	0.914	0.909
GBDT	0.943	0.971	0.879	0.960
XGBoost	0.951	0.974	0.897	0.965
SVM	0.954	0.974	0.905	0.967

Note: Acc, accuracy; Sen, sensitivity; Spe, specificity; F1, F1 score; KNN, K-Nearest Neighbor; DT, Decision Tree; GBDT, Gradient Boosting Decision Tree; XGBoost, eXtreme Gradient Boosting; SVM, Support Vector Machine.

Table 7. Evaluation metrics for comparative experiments

	Acc	Sen	Spe	F1
Li P et al. [9]	0.873	0.903	0.845	0.874
Zhang et al. [12]	0.946	0.949	0.940	0.974
Li J et al. [11]	0.864	0.850	0.931	0.911
Hettiarachchi et al. [22]	0.904	0.947	0.750	0.939
Chakir et al. [8]	0.925	0.923	0.929	0.941
Approach in this study	0.954	0.974	0.905	0.967

Note: Acc, accuracy; Sen, sensitivity; Spe, specificity; F1, F1 score.

Compared with state-of-the-art methods, the approach developed in this study achieved higher Acc and Sen on public datasets. In future work, this research will focus on refining the methodology and models to improve their generalization capabilities, allowing them to be applied in scenarios with missing modalities.

DECLARATIONS

Author contributions

Mingzhi Zhang was responsible for data processing, time–frequency conversion, model construction, and manuscript writing. Piding Li made important contributions to data analysis.

Funding

This research received no external funding.

Data availability

This study utilised only publicly available datasets.

Ethics approval and consent to participate

Not applicable.

Consent for publication

The authors agree to the publication of this manuscript in both print and electronic formats in ZENTIME and grant the publisher permission to make necessary editorial and formatting revisions.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

The authors confirm that this manuscript is original, unpublished, and not under consideration elsewhere. All authors have approved the manuscript and confirm adherence to ethical guidelines, including disclosure of potential conflicts of interest.

REFERENCES

[1] World Health Organization. WHO reveals leading causes of death and disability worldwide: 2000-2019 [Internet]. Geneva (Switzerland): WHO; 2020 [updated 2020 Dec 9; cited 2025 Nov 20]. Available from: https://www.who.int/news/item/09-12-2020-who-reveals-leading-causes-of-death-and-disability-worldwide-2000-2019

[2] National Center for Cardiovascular Diseases, The Writing Committee of the Report on Cardiovascular Health and Diseases in China. Report on cardiovascular health and diseases in China 2023: An updated summary. Chin Circ J. 2024 Jul 27;39(7):625-660. https://doi.org/10.3969/j.issn.1000-3614.2024.07.001

[3] Karhade J, Dash S, Ghosh SK, Dash DK, Tripathy RK. Time–frequency-domain deep learning framework for the automated detection of heart valve disorders using PCG signals. IEEE Trans Instrum Meas. 2022 Mar 29;71:1-11. https://doi.org/10.1109/TIM.2022.3163156

[4] Houssein EH, Ibrahim IE, Neggaz N, Hassaballah M, Wazery YM. An efficient ECG arrhythmia classification method based on manta ray foraging optimization. Expert Syst Appl. 2021 Nov 1;181:115131. https://doi.org/10.1016/j.eswa.2021.115131

[5] Liu C, Springer D, Li Q, Moody B, Juan RA, Chorro FJ, et al. An open access database for the evaluation of heart sound algorithms. Physiol Meas. 2016 Dec;37(12):2181. https://doi.org/10.1088/0967-3334/37/12/2181

[6] Han H, Huang X, Chang H, Fan H, Chen P, Chen J. Review of self-supervised learning methods in field of ECG. J Front Comput Sci Technol. 2024;18(7):1683-1704. https://doi.org/10.3778/j.issn.1673-9418.2310043

[7] Cosoli G, Poli A, Scalise L, Spinsante S. Measurement of multimodal physiological signals for stimulation detection by wearable devices. Measurement. 2021 Nov;184:109966. https://doi.org/10.1016/j.measurement.2021.109966

[8] Chakir F, Jilbab A, Nacir C, Hammouch A. Recognition of cardiac abnormalities from synchronized ECG and PCG signals. Phys Eng Sci Med. 2020 Jun;43(2):673-677. https://doi.org/10.1007/s13246-020-00875-2

[9] Li P, Hu Y, Liu ZP. Prediction of cardiovascular diseases by integrating multi-modal features with machine learning methods. Biomed Signal Process Control. 2021 Apr;66:102474. https://doi.org/10.1016/j.bspc.2021.102474

[10] Sun C, Liu X, Liu C, Wang X, Liu Y, Zhao S, et al. Enhanced CAD detection using novel multi-modal learning: Integration of ECG, PCG, and coupling signals. Basel. 2024 Oct 30;11(11):1093. https://doi.org/10.3390/bioengineering11111093

[11] Li J, Ke L, Du Q, Chen X, Ding X. Multi-modal cardiac function signals classification algorithm based on improved D-S evidence theory. Biomed Signal Process Control. 2022 Jan;71(Part A):103078. https://doi.org/10.1016/j.bspc.2021.103078

[12] Zhang H, Zhang P, Lin F, Chao L, Wang Z, Ma F, et al. Co-learning-assisted progressive dense fusion network for cardiovascular disease detection using ECG and PCG signals. Expert Syst Appl. 2024 Mar 15;238(Part F):122144. https://doi.org/10.1016/j.eswa.2023.122144

[13] Zhu J, Liu H, Liu X, Chen C, Shu M. Cardiovascular disease detection based on deep learning and multi-modal data fusion. Biomed Signal Process Control. 2025 Jan;99:106882. https://doi.org/10.1016/j.bspc.2024.106882

[14] Liu X, You L, Lv C, Chen M, Wei L, Zheng Y, et al. Integrating ECG and PCG signals through a Dual-Modal ViT for coronary artery disease detection. IEEE Journal of Biomedical and Health Informatics 2025. https://doi.org/10.1109/JBHI.2025.3589257

[15] Xu W, Yu K, Ye J, Li H, Chen J, Yin F, et al. Automatic pediatric congenital heart disease classification based on heart sound signal. Artif Intell Med. 2022 Apr;126:102257. https://doi.org/10.1016/j.artmed.2022.102257

[16] Lu G, Tang T, Qi J, Wang Y, Zhao Y. Heart sound denoising based on CEEMDAN and wavelet analysis with adaptive dual thresholds. J Nanjing Univ Posts Telecommun (Nat Sci). 2025 Aug;45(4):36-47. https://doi.org/10.14132/j.cnki.1673-5439.2025.04.005

[17] Kabir MA, Shahnaz C. Denoising of ECG signals based on noise reduction algorithms in EMD and wavelet domains. Biomed Signal Process Control. 2012 Sep;7(5):481-489. https://doi.org/10.1016/j.bspc.2011.11.003

[18] Suryady Z, Rahman AWA, Oktarina D. Human emotion classification based on phonocardiography signals (PCG). AIP Conf Proc. 2024 Feb 7;2729(1):020002. https://doi.org/10.1063/5.0168132

[19] Daubechies I, Lu J, Wu HT. Synchrosqueezed wavelet transforms: An empirical mode decomposition-like tool. Appl

Comput Harmon A. 2011 Mar;30(2):243-261. https://doi.org/10.1016/j.acha.2010.08.002

[20] Fu J, Liu J, Tian H, Li Y, Bao Y, Fang Z, et al. Dual attention network for scene segmentation. Proceedings of the IEEE/CVF conference on computer vision and pattern recognition; 2019. https://doi.org/10.1109/cvpr.2019.00326

[21] Li X, Wang W, Hu X, Yang J. Selective Kernel Networks. 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR); 2019 June 15-20; Long Beach, CA, USA. IEEE. p. 510-519. https://doi.org/10.1109/CVPR.2019.00060

[22] Hettiarachchi R, Haputhanthri U, Herath K, Kariyawasam H, Munasinghe S, Wickramasinghe K, et al. A novel transfer learning-based approach for screening pre-existing heart diseases using synchronized ECG signals and heart sounds. 2021 IEEE International Symposium on Circuits and Systems (ISCAS); 2021 May 22-28; Daegu, Korea. IEEE; 2021. p. 1-5. https://doi.org/10.1109/ISCAS51556.2021.9401093

Progress in
Medical Devices

REVIEW ARTICLE

Research progress on energy-based tissue fusion technologies and related medical devices

Junjie Shen, Zhongxin Hu, Chengli Song, Lin Mao

Shanghai Institute for Minimally Invasive Therapy, School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Lin Mao.

Address correspondence to: Lin Mao, Shanghai Institute for Minimally Invasive Therapy, School of Health Science and Engineering, University of Shanghai for Science and Technology, No. 516 Jungong Road, Yangpu District, Shanghai 200093, China. Tel: +86-21-55572159.
E-mail: linmao@usst.edu.cn.

Received January 24, 2026; Accepted March 25, 2026; Published June 24, 2026

DOI: 10.61189/748101ldqptn

Abstract

Traditional methods of tissue closure, such as sutures and staples, have long been the gold standard in surgery. However, they have major drawbacks, such as the body's reaction to foreign materials and the difficulty of the technical aspects of minimally invasive procedures. Energy-based tissue fusion (EBTF) technology is a revolutionary alternative that uses energy to produce autologous tissue sealing. This review aims to provide a comprehensive analysis of the biophysical principles, technological evolution, and clinical applications of current EBTF technologies and related devices. The fundamental mechanisms of EBTF technologies are investigated, with a focus on collagen denaturation and cross-linking induced by different energy modalities such as radiofrequency (RF) current, ultrasound, and laser. Three representative systems are critically evaluated: the impedance-controlled bipolar system (LigaSure™), the ultrasonic coagulating shears (Harmonic™), and the hybrid ultrasonic-bipolar device (Thunderbeat™). Their performance is compared in terms of vessel sealing efficacy, thermal spread, operative time, and complication rates across various surgical specialties. The clinical evidence indicates that the primary advantage of RF device lies in safety, whereas the ultrasonic devices offer reduced lateral thermal damage, and the hybrid device demonstrate superior versatility and procedural speed. The review concludes by identifying future trends, including the integration of artificial intelligence and robotic platforms, which promise to further enhance the safety and precision of surgical energy devices.

Keywords: Energy-based tissue fusion, Medical devices, Vessel sealing system, Electrosurgery, Minimally invasive surgery

Highlights

- Mechanistic comparison of radiofrequency, ultrasonic, and laser energy modalities for achieving collagen denaturation in tissue fusion.
- Critical evaluation of three leading device platforms (LigaSure™, Harmonic™, and Thunderbeat™) across surgical specialties and performance metrics.
- Future integration of artificial intelligence and robotic systems to enhance precision and safety in energy-based surgical devices.

1 INTRODUCTION

The history of surgical operations has gone hand in hand with the evolution of tissue closure and hemostasis methods, which have become one of the main engines of minimally invasive surgery development [1, 2]. Over the past century, surgical practitioners have embraced traditional tissue closure techniques such as sutures, staples, and clips as the most reliable

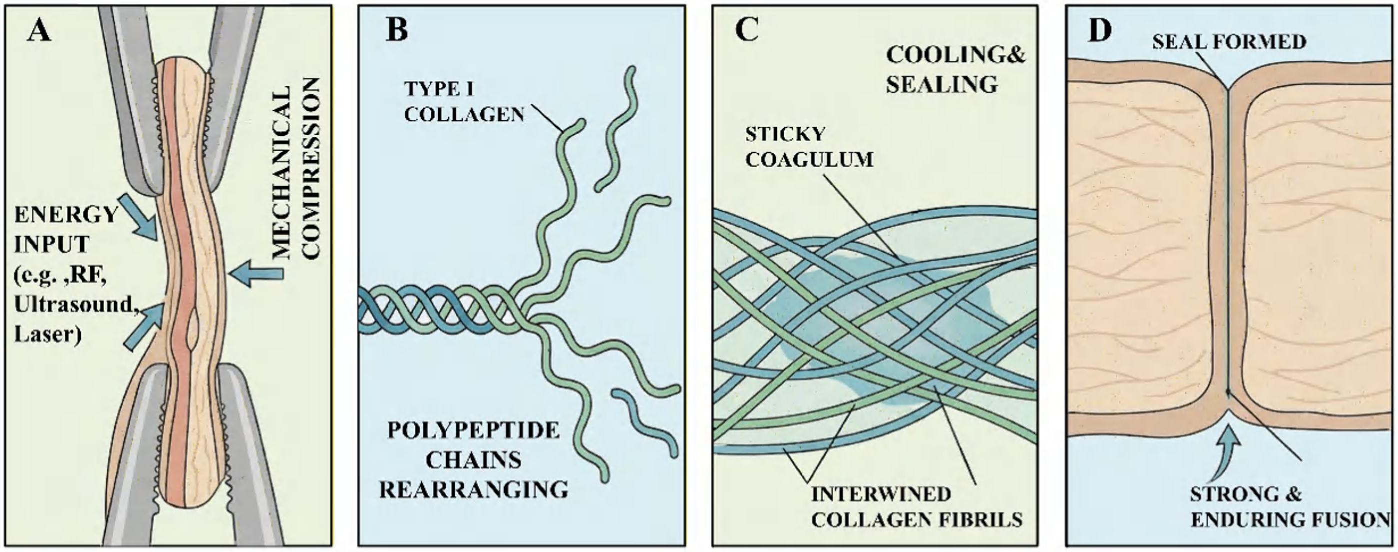


Figure 1. Energy-based tissue fusion mechanism. (A) Energy application and mechanical compression; (B) Collagen denaturation; (C) Cooling and sealing; (D) Tissue fusion. RF, radiofrequency.

methods of tissue closure in the operating room [3-5]. However, these methods involve forcibly holding tissue together by applying mechanical tension and hence have some inherent limitations. Since these are foreign materials, they are either permanently or semi-permanently lodged in the body, which may lead to a foreign body granuloma, tissue adhesion, or even infection [6, 7]. Besides, intracorporeal suturing during minimally invasive and robotic-assisted surgeries, which is among the most technically demanding skills for many surgeons, can be time-consuming and, on top of that, have a very steep learning curve. Tissue adhesive could serve as an alternative to suturing for tissue closure; though their use in certain anastomoses is still very rare [8, 9]. Issues such as failure to maintain proper strength in a wet environment, inability to conform to a moving tissue surface, and difficulty in applying accurately make them unreliable as the main closure method [10]. These limitations speak to the obvious demand for a more effective and consistently safer anastomotic technology.

Driven by the progress in biomedical engineering, energy-based tissue fusion (EBTF) technology has been massively transformed during the last three decades and has thus impacted surgical procedures in a way hardly ever seen before. At its core, this technology combines precisely measured energy input, e.g., radiofrequency (RF), ultrasound, or laser, and simultaneous mechanical compression resulting in tissue sealing and cutting being achieved at the same time (as shown in **Figure 1**). When energy is further applied, the temperature of tissue rises to a critical level, which is generally between 60 °C and 90 °C. At this level, the main structural protein, type I collagen, is denatured: the triple-helix structure opens, allowing the polypeptide chains to re-arrange [11]. Then, due to cooling and under pressure, these new collagen fibrils become intertwined with coagulated extracellular proteins, thus a strong and enduring seal is formed, exhibiting considerable tensile strength [12-14]. This mechanism of collagen denaturation and subsequent cellular and tissue cooling into a "sticky coagulum" may account for the success of EBTF [15]. Unlike conventional electrocautery, which relies on thermal destruction of tissue to achieve hemostasis, EBTF technology can form a genuine molecular fusion. Su et al. were the first to use Raman spectroscopy to study the characteristics of in vitro tissue fusion. Through spectral analysis of fused porcine blood vessels and small intestines, they demonstrated for the first time at the molecular level that collagen denaturation and crosslink reorganization occurred in the fused tissue [16]. The discovery has provided a vital theoretical basis for the subsequent development of EBTF technology. Nowadays, with the modernization of medical science continuously increasing, the clinical applications of EBTF have expanded on a large scale. Especially, in minimally invasive and robotic-assisted surgeries, where it is highly technical and time-consuming to suture, EBTF can help make the procedure faster and more efficient.

This review aims to provide a comprehensive overview of the current state of EBTF technology. First, the different biophysical mechanisms of the various energy modalities and how each of them affects tissue remodeling are described. Subsequently, the design philosophy and clinical performance of several market-leading devices are evaluated, stressing their strengths and limitations within surgical practice. At the end, a perspective is offered on the future trajectory of EBTF technology and related medical devices, focusing on the potential integration of intelligent sensing algorithms and robotic surgical systems.

2 FORM OF ENERGY

Current clinical EBTF technologies predominantly rely on three principal energy modalities: RF energy, ultrasonic energy, and laser energy. Even though EBTF technologies differ a lot as far as how the energy is delivered, they all come together biologically to one common result: using heat to change and rearrange structural proteins in tissues. To get more insight into how each individual modality works, it is essential to first master the mechanisms and evolution of different energy forms.

2.1 RF energy

RF energy refers to alternating electrical currents within the frequency range of 300 kHz to 300 GHz, and is generally regarded as one of the most mature and widely used forms of energy in contemporary electrosurgery. In surgical tissue fusion, current frequency generally varies between 300 kHz and 500 kHz. The surgical use of RF energy primarily revolves around a triad of capabilities: a clean cut, coagulation to stop bleeding, and the sealing of tissues for anastomosis. The high frequency of current notably provides numerous significant benefits: more versatility in use, energy delivery with much higher precision within a shorter period, and a lower probability of neuromuscular stimulation compared with lower-frequency currents.

The core method of RF-driven tissue fusion is based on Joule heating principles. Basically, tissue electrical impedance causes resistance to electrical current flow, which then results in rapid agitation of ions within the cellular and extracellular milieu. The ionic friction thereby directly converts electrical energy into thermal energy, which locally heats the tissue. The effects that consequently occur include both: a non-thermal aspect of electroporation at the level of cell membranes and, more prominently, a deep thermal effect [17]. It is the orderly application of heat that makes the fusion possible.

The history of using RF energy for tissue anastomosis dates back to the early 1990s, when the Paton Welding Institute in Ukraine was the first to conduct pioneering research uniting the concept of tissue welding with controlled energy. In 1998, a major breakthrough came when Kennedy et al. demonstrated, for the first time, the application of RF energy to create an arteriovenous anastomosis, thus proving the possibility of such an approach in vascular tissue [18]. Over the next ten years, the technology went through rapid clinical adoption and development. It was Lamberton et al. who conducted a landmark study that shifted perceptions in 2008, when they compared four laparoscopic vessel ligation devices [19]. Among the devices tested, the RF energy-based system provided the best results in terms of burst pressures and sealing times; therefore, it was confirmed as the most effective for vessel sealing. The technology was also subjected to extensive testing on large animal models. For example, Pan et al. used a porcine model in 2020 to compare RF tissue fusion with traditional suture methods for intestinal anastomosis [20]. Histological examination showed that the gaps in the muscularis layer at the anastomotic site in the RF group were replaced by newly formed collagen fibers. This demonstrates that the tissue was not only sealed but also actively supported during the healing process, thus confirming the method to be a practical and safe alternative.

2.2 Ultrasound

Ultrasonic energy represents another cornerstone modality in EBTF, which is set apart by its special mode of mechanical, rather than electrical, energy conversion. In surgical applications, ultrasonic devices work at very high frequencies, usually between 20 kHz and 60 kHz, well above the limit of human hearing. Contrary to RF energy, which depends on electrical current and heat generated by resistance, ultrasonic emission makes use of piezoelectric transducers in the instrument handle to convert electrical energy into very high-frequency longitudinal mechanical vibrations at the working part of the instrument such as a blade or a jaw [21, 22].

Initially, mechanical friction at a microscopic level is the principal effect when the vibrating blade touches the tissue under the application of a compressive force. Such friction produces a small-scale heat spot exactly at the tissue-blade interface as well as inside the compressed tissue area. The temperature usually rises more slowly and locally in the case of the ultrasonic method than in RF, and the peak temperature is generally only between 80 °C and 100 °C. This thermal effect is a secondary consequence of the primary mechanical energy delivery. The heat causes the denaturation of tissue proteins, mainly collagen and elastin, while the simultaneous, long-lasting pressure and vibration help release water vapor from the tissue. When tissue proteins denature, they create a thick coagulum. The persistent ultrasonic vibration under pressure breaks and reforms the protein bonds, resulting in their realignment and intermixing across the apposed tissue surfaces. They solidify into a single, welded seal on cooling and release of pressure. Very importantly, the significantly lower peak temperatures and the effect of ultrasonic energy in removing water often lead to less lateral thermal damage and less smoke production when compared to RF energy [23].

Ultrasonic energy has shifted from being merely a conceptual instrument to a clinically permitted mode, backed up by an ever-increasing body of research that shows its effectiveness and safety in various surgical fields. In 2002, Foschi et al. used transmission and scanning electron microscopy to study vascular tissue after ultrasonic anastomosis [24]. Their ultrastructural study indicated that the thermal effect caused by ultrasonic energy helps the breakdown and later reformation of tissue proteins, therefore giving a mechanistic explanation of successful tissue fusion. Going from experimentation to clinical testing, Sista et al. performed a comparison study of 211 consecutive patients undergoing right hemicolectomy due to colon cancer [25]. Patients were divided into two groups: one group using a novel ultrasonic energy handpiece (108 patients) and the other group treated with conventional hemostasis methods (103 patients). The data showed that the use of ultrasonic energy notably reduced surgical time, the amount of blood lost during surgery, and postoperative complications. Moreover, Tolone et al. assessed the application of the ultrasonic scalpel for endocrine surgery by comparing total thyroidectomy procedures carried out with an ultrasonic scalpel and conventional suture ligation [26]. Their data demonstrated that ultrasonic energy is a safe, effective, and time-saving alternative. It can facilitate the operations without causing an increase in drainage volume, complications, or length of hospital stay.

2.3 Laser

Laser energy holds a special place in tissue fusion as a separate modality, mainly because of the pure photonic nature of laser. The process of laser generation can be described in simple terms as follows. Electrons in an atomic medium that are excited to a higher energy state by laser light will come down to their ground state and, in doing so, emit photons in a highly synchronized manner. This phenomenon results in the production of light that is monochromatic, coherent, and highly directional [27]. From the surgical point of view, this means that energy can be delivered with extremely high spatial precision, thus the predetermined tissue plane for interaction can be targeted and collateral damage can be minimized. One of the main inherent benefits of laser energy is that, being a form of light, it does not introduce any foreign material to the wound, thus eliminating the main source of infection or chronic inflammatory reactions associated with suture or staple lines.

Laser tissue fusion (LTF) is mainly mediated by the photothermal effect. Thus, when a laser beam at a certain wavelength is absorbed by water or hemoglobin, the energy released is initially converted into heat. This locally generated heat at the spot causes a phase change of the tissue structural proteins. Collagen is one of the proteins in the extracellular matrix that is commonly affected by heat. The triple-helical structure of collagen is mainly maintained by hydrogen bonds. The breaking of these hydrogen bonds occurs at a certain temperature, and this causes the collagen fibrils to lose their structure—going from their highly ordered state into a chaotic, random coil conformation. Importantly, while this process occurs, the tissue's overall architecture is preserved through mechanical apposition. Following cessation of laser irradiation and during a strict cooling phase, collagen from both tissue surfaces, which is now unfolded and quite mobile, will find each other and, by undergoing recrystallization, form bonds together. Here the polypeptide chains intertwine to form new intermolecular bonds thus producing a uniform, cross-linked protein matrix that closes the wound [28]. It is like a biological "weld" without any gaps, the strength of which comes from the reforged collagen network rather than from the use of foreign materials or mere coagulation necrosis.

The use of laser energy in tissue anastomosis has transitioned through various milestone stages. Firstly, the pioneering work by Jain et al. in 1979 laid down the basis for the feasibility of laser vascular repair [29]. After that, some experiments helped to define the laser welding process parameters more clearly. As an illustration, Bremmer et al. showed that the seal strength depends to a great extent on fine control of laser power, irradiation time, and simultaneous application of pressure in a rabbit artery model [30]. Also, some research articles further detailed the cellular heating mechanisms. For example, Li et al. pointed out that the optimum temperature range for the skin weld is approximately 55–65 °C, above which the mechanical properties become compromised because of significant protein denaturation [31]. Beyond immediate sealing performance, there is still evidence supporting favorable healing outcomes. According to Ghosh et al., laser-sealed mouse skin wounds regained function much earlier than those closed with sutures, suggesting that the laser might support physiological wound healing [32]. All these accomplishments together show the development of LTF from a new idea to a highly technical process. Currently, the emphasis is on broadening its application and integrating in vivo monitoring to achieve better control over the procedure.

3 EBTF MEDICAL DEVICES

The evolution of surgical energy instruments has been one of the paradigm shifts in modern medicine. It signifies a trend from simple anatomical dissection to sophisticated tissue management. As a matter of fact, hemostatic methods over the last 100 years have evolved from suturing and monopolar cautery to the present-day advanced and intelligent tissue fusion platforms. This major change is mostly due to the increasing demand for surgical safety, efficiency, and minimally invasive methods. In this section, we have clinically reviewed several tissue fusion systems that are prevalently used in hospitals and have compared them to each other based on publicly available data from publications indexed in the Science Citation Index Expanded, peer-reviewed biomedical engineering literature, and clinical comparative studies indexed in PubMed (as shown in **Table 1**). We have focused on explaining the devices' working principle, preclinical performance metrics, and clinical results in various types of surgery. This multidisciplinary work presents a fair evaluation of the available technologies and points to future innovations required to develop energy-based surgical devices.

3.1 LigaSure™

From an engineering viewpoint, the fusion of tissues by means of RF energy is based on two main electrode configurations (as shown in **Figure 2**). The monopolar configuration uses only

Table 1. Summary of key literature comparing three EBTF medical devices across various procedures

Study	Device evaluated	Surgical site/ procedure	Sample size/study type	Key outcomes and clinical findings
Amirkazem et al. [33]	LigaSure™ vs. Silk Suture	Spleen (traumatic splenectomy)	Randomized clinical trial	Drastically shortened operative time (12 vs. 21 min) and reduced intraoperative blood loss (80 vs. 280 mL).
Akhtar et al. [34]	LigaSure™ vs. Milligan-Morgan	Hemorrhoids (hemorrhoidectomy)	70 patients	Significantly less postoperative pain (VAS at 6–24 h), reduced blood loss, and shorter operative time compared to conventional technique.
Vidal et al. [35]	LigaSure™ Small Jaw vs. LigaSure™ Precise	Thyroid (total thyroidectomy)	2,000 patients	Small Jaw reduced operative time (40 vs. 65 min), incision length (4 vs. 7 cm), and drainage needs, with no increase in complications.
Constant et al. [36]	LigaSure™	Kidney (donor nephrectomy)	124 patients	Lowest estimated blood loss (90±53 mL); immediate graft function observed in all transplanted kidneys.
Revelli et al. [37]	Harmonic Scalpel	Thyroid surgery	Meta-analysis of 21 RCTs	Significantly shorter operative time (-25 to -26 min), reduced blood loss (-30 mL), and lower postoperative drainage volume.
Dong et al. [38]	Harmonic Scalpel vs. Electrocautery	Breast (axillary dissection)	23 patients (retrospective case-control study)	Reduced intraoperative blood loss (33 vs. 90 mL), lower drainage volume (177 vs. 272 mL), and shorter drainage duration.
Waraich et al. [39]	Harmonic Scalpel vs. Electrocautery	Esophagus (esophagectomy)	142 patients (retrospective study)	Fewer blood transfusions, less median blood loss (500 vs. 700 mL), and lower respiratory complications (13.6% vs. 17.3%).
Abo-hashem et al. [40]	Harmonic Scalpel vs. Bipolar	Hemorrhoids (hemorrhoidectomy)	RCT	Significant drop in postoperative pain, lower urinary retention (9.4% vs. 34.4%), and faster return to work.
Kumar et al. [41]	Harmonic Scalpel	Vascular surgery (SEPS)	Observational study	Allowed accurate cutting with minimal thermal injury; average operative time of 40 min with low complication rate.
Seehofer et al. [42]	Thunderbeat™ vs. Harmonic/LigaSure™	Animal model	Animal study	Significantly faster cutting speed; achieved average burst pressure of 734 mmHg for large vessels, outperforming comparators.
Liberman et al. [43]	Thunderbeat™	Thoracic (ex vivo pulmonary artery)	Ex vivo model	Highest average burst pressure (875 mmHg) with zero sealing failures; other bipolar systems exhibited complete failures.
Milsom et al. [44]	Thunderbeat™	Colorectal (laparoscopic)	Prospective study	Safely ligated major vessels, such as the IMA, without adverse events; functioned as a single tool for dissection and hemostasis.
Kuipers et al. [45]	Thunderbeat™ vs. Electrocautery	Head and neck (neck dissection)	RCT	Shortened median operative time by approximately 50 min and reduced blood loss (210 vs. 431 mL).
Suzuki et al. [46]	Thunderbeat™ vs. traditional methods/ LigaSure™	Head and neck (neck dissection)	Retrospective study	Reduced operative time compared with traditional methods (66 vs. 96 min); lower complication rate (3.4%) compared with the LigaSure group (20%).

Note: EBTF, energy-based tissue fusion; VAS, visual analog scale; RCT, randomized controlled trial; SEPS, subfascial endoscopic perforator surgery; IMA, inferior mesenteric artery.

one active electrode at the surgical site and the current then flows from the patient's body to a remote dispersive electrode. This is good for superficial cutting and coagulation, but its wider current path makes it less accurate for tissue fusion. On the other hand, the bipolar configuration has both electrodes incorporated in the instrument jaws. The electric current flows only through the area of the tissue held between them, thus the thermal and electric effects are restricted to a small volume [47]. This concentrated energy release allows the fusion to be very precise, with minimal damage to surrounding tissues, so bipolar systems are suitable for sealing tissues.

To ensure safe surgical seals, many contemporary systems integrate real-time tissue impedance monitoring with automated energy termination. The LigaSure™ system (Medtronic) is a representative example of a closed-loop technology based on tissue impedance monitoring and automated energy termination. Representative LigaSure™ instruments, including the LigaSure Impact™ and LigaSure Atlas™, are shown in **Figure 3**. As a leader in advanced bipolar vessel sealing, LigaSure™'s main technological feature is its generator platform (ForceTriad™), which contains the exclusive TissueFect™ sensing system [49]. As technical reports and experimental

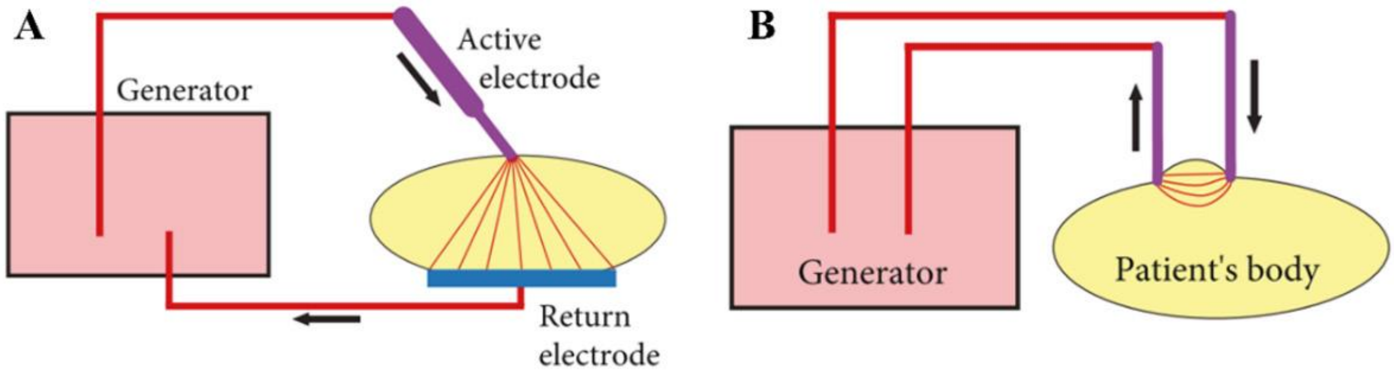


Figure 2. Two main electrode configurations. (A) Monopolar circuit; (B) Bipolar circuit. The figure is reproduced from [48].

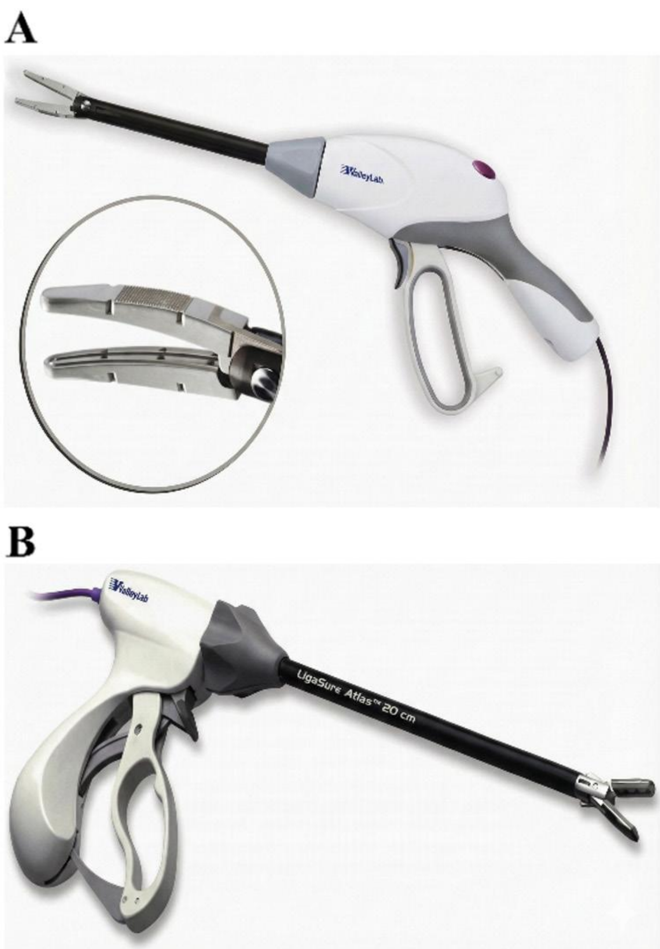


Figure 3. LigaSure™ instruments. (A) LigaSure Impact; (B) LigaSure Atlas. The figure is reproduced from [52].

studies show, the TissueFect™ system measures tissue impedance at a frequency of more than 3,000 measurements per second [50]. The change in tissue impedance induced by the release of intracellular water through vaporization and collagen denaturation follows a typical nonlinear pattern during the sealing cycle. The device continually adjusts voltage and current based on these impedance changes, thus ensuring that the tissue remains within a thermal range that is ideal for fusion while avoiding drying or burning [51]. An algorithm detects an impedance level or changes indicative of seal completion and stops energy delivery automatically. Determining the operative endpoint is fully automated, so subjective surgeon judgment is replaced by objective signal-based control, which in turn leads to improvements in procedural consistency, safety, and seal reliability across different tissue types.

Numerous randomized controlled trials (RCTs) and retrospective studies within different surgical specialties have confirmed the clinical effectiveness of the LigaSure™ system. Present research highlights three main areas of focus where current evidence is concentrated: more efficient hemostasis superior to traditional methods, device-specific optimization advantages, and safety in complex surgeries.

LigaSure™, compared with traditional suturing or ligation methods, can significantly shorten operative time and minimize the amount of bleeding during the operation. In a clinical RCT of traumatic splenectomy, the use of LigaSure™ reduced the average duration of the surgery significantly (12 vs. 21 min, $P<0.05$) and also considerably reduced the intraoperative blood loss (80 vs. 280 mL, $P<0.05$) as compared with the conventional silk suture ligation [33]. Likewise, in a case of hemorrhoidectomy for grade III/IV hemorrhoids, the device achieved better outcomes than the traditional Milligan-Morgan method as it was not only associated with lower blood loss and shorter operative time but also postoperative pain scores, measured 6–24 hours postoperatively, were significantly lower ($P=0.001$) [34].

Also, anatomically site-specific models of LigaSure™ have further improved surgical outcomes. According to a retrospective review of 2,000 total thyroidectomy patients, the performance of different generations of devices was compared, and the LigaSure™ Small Jaw—specifically created for use in tight spaces—was found to be better than the standard LigaSure Precise™ [35]. The Small Jaw was capable of significantly reducing the operative time (40 vs. 65 min, $P=0.002$), incision size, and the amount of drainage needed without increasing the

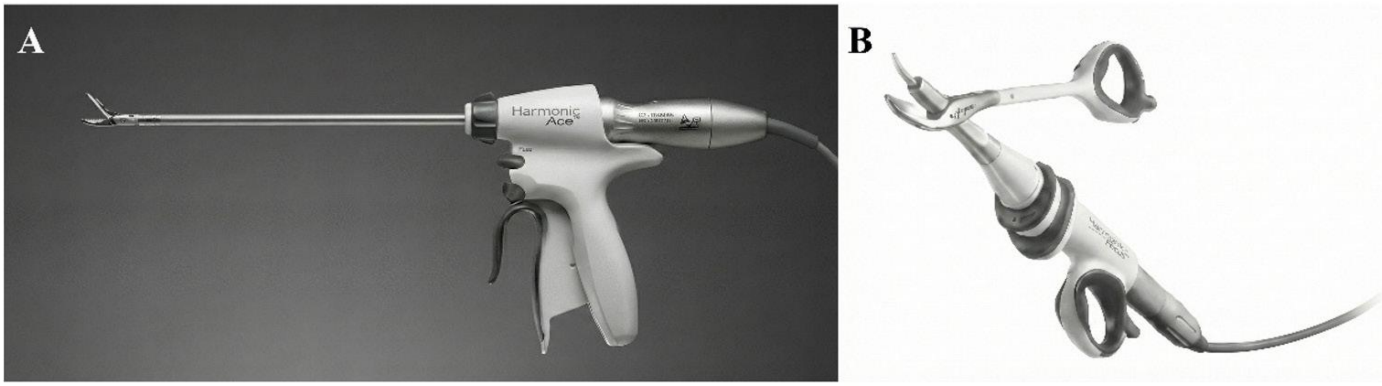


Figure 4. HARMONIC™ instruments. (A) Harmonic Ace for laparoscopic use; (B) Harmonic Focus for open use. The figure is reproduced from [54].

risk of complications, thus confirming the clinical significance of anatomic device design.

Moreover, in transplant surgery, where the utmost safety precautions are taken, the system has demonstrated reliable vessel sealing functions. In this study of 124 laparoscopic living donor nephrectomies, the device was useful in keeping blood loss to a minimum during the separation of the main vascular branches in the nephrectomy procedure (90±53 mL) [36]. Most importantly, all the transplanted kidneys showed immediate function after reperfusion, thus proving that the device had a good safety record in terms of thermal spread and vessel integrity.

Overall, the research results obtained demonstrate that the LigaSure™ system is a highly versatile and effective advancement in surgical energy devices. It has become an alternative to conventional ligation methods to some extent by continually reducing operative time, blood loss, and postoperative pain, while ensuring high levels of safety in both routine and advanced procedures.

3.2 HARMONIC™

When working in limited anatomical areas or close to vital structures, it is crucial to use surgical energy very precisely. HARMONIC™ products are widely recognized as the leader in the ultrasonic energy field (as shown in **Figure 4**). The device's principle of operation is radically different from that of bipolar RF devices. Through the piezoelectric effect, a transducer in the handpiece converts electrical energy into mechanical vibration, which causes the blade to oscillate longitudinally at a frequency of 55,500 Hz. Such high-frequency vibration leads to the intensification of molecular friction within tissues. In this way, heat for protein denaturation is produced [49]. Since the heat is produced by internal friction and not by the passage of external electrical current, the device's temperature is usually lower, and it is significantly lower than that of electrosurgical devices. Moreover, the greatest characteristic of HARMONIC™ is that it causes minimal lateral thermal damage. A lot of comparative studies involving both infrared (IR) thermography and histopathology have demonstrated that the lateral thermal spread resulting from the use of Harmonic is usually only about 1–2 mm, which is much less than that of monopolar cautery and some bipolar devices [53]. This feature makes it an excellent tool for delicate dissection in areas with a high concentration of nerves and blood vessels.

It is noteworthy, however, that despite minimal lateral thermal spread, the blade itself can accumulate heat after prolonged activation, necessitating a certain cooling period. Another drawback of the ultrasonic scalpel is that the instrument produces a very visible and substantial surgical smoke (aerosol). Unlike smoke produced by electrocautery, this kind of aerosol consists of larger particle diameters (0.35–6 μm) and has higher levels of moisture and lipid components. Chemical analysis using mass spectrometry (Rapid Evaporative Ionization Mass Spectrometry, REIMS) shows that aerosol is mainly lipids such as diglycerides and triglycerides, as opposed to electrocautery smoke that is rich in glycerophospholipids. Regarding biological risks, even if the chemical toxicity of ultrasonic smoke is lower than that of electrocautery smoke, there is a theoretical risk that cellular materials can be carried because of the larger particle size [55, 56].

The HARMONIC™ systems show the traits of versatility and clinical benefits that can be observed across several surgical specialties. The currently available studies point to three major areas of these effects: greater hemostatic effectiveness in oncological and glandular surgeries, an improved post-surgical quality of life, and highly precise vascular interventions.

In particular, during major resections and glandular surgeries, the instrument has been shown to consistently improve fluid management and operative efficiency compared with traditional methods. A meta-analysis of 21 RCTs in thyroid surgery brought out such effectiveness, revealing that the Harmonic Scalpel (HS) and Harmonic Focus notably reduced surgical duration (by 25–26 min, $P<0.0001$) and reduced blood loss during surgery (by ~30 mL, $P<0.01$) as compared to conventional

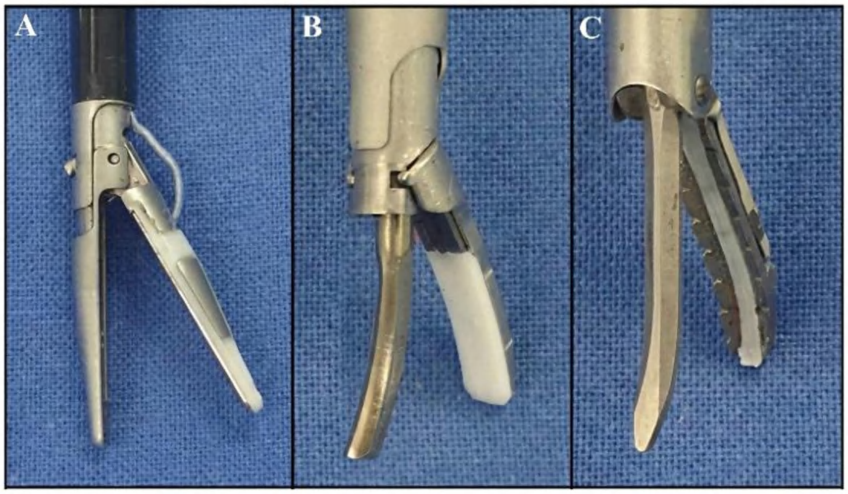


Figure 5. Detailed view of three instruments. (A) LigaSure™; (B) HARMONIC™; (C) Thunderbeat™. The figure is reproduced from [42].

hemostasis [37]. These gains have also been supported in breast cancer surgery, where the use of the HS in axillary dissection led to a significantly smaller amount of blood loss during surgery (33 vs. 90 mL, P<0.001) and less drainage volume and duration as compared to electrocautery [38]. Similarly, in esophageal cancer surgery, the device was superior to traditional electrocautery in reducing the need for blood transfusion (median 0 vs. 2 units, P=0.003), reducing blood loss, and lowering the rate of postoperative respiratory complications (13.6% vs. 17.3%) [39].

Moreover, the HS can also significantly improve patient recovery and quality of life long after an operation. In a proctological surgery study, it was found that the HS had markedly better postoperative pain outcomes than bipolar electrocautery [40]. Additionally, it resulted in less frequent use of narcotics and non-steroidal anti-inflammatory drugs, and there was a noticeable decrease in the occurrence of urinary retention (9.4% vs. 34.4%, P<0.05). Therefore, due to earlier discharge and return to work, after only two weeks, 75% of the patients in the HS group had returned to work, compared with 45% in the control group, signifying the socioeconomic benefits of this device.

Furthermore, the fundamental principle of the device's operation enhances precision in complex vascular interventions. In subfascial endoscopic perforator vein surgery for chronic venous insufficiency, the HS has been demonstrated as an effective alternative to metal clips or electrocautery. By coagulating and cutting simultaneously to produce only a minimal lateral thermal injury, surgeons have been able to achieve precise exposure of perforator veins [41]. The high level of precision here brought about the performance of fast operations (average 40 min) with minimal risk of complications.

In conclusion, HARMONIC™ is a unique system that not only improves surgical efficiency but also results in better patient outcomes. The system minimizes tissue harm and controls bleeding which benefits the patient in two ways: on the one hand, it lowers the physiological impact of surgery. As demonstrated by reduced blood loss and drainage in tumor cases. On the other hand, it accelerates patients' functional recovery, as shown in proctological and vascular procedures.

3.3 Thunderbeat™

Thunderbeat™ is a hybrid energy system that combines ultrasonic and bipolar RF technologies in one device. Such a design addresses the complementary limitations of each modality: ultrasonic instruments are fast in cutting but have relatively limited hemostasis, whereas bipolar devices, although slower in dissection, provide strong coagulation [49]. As shown in **Figure 5**, the innovative jaw design of this instrument, consisting of a central ultrasonic probe enclosed by bipolar electrodes, enables the device to operate in two main modes: 'Seal & Cut' for simultaneous cutting and sealing, and 'Seal' for dedicated coagulation. Preclinical studies in porcine models showed that Thunderbeat™ was able to achieve significantly shorter dissection times than Harmonic Ace, LigaSure V, and EnSeal when working with vessels of various sizes (2–7 mm, P<0.01). Moreover, this device was given a top versatility rating, thus recognizing its ability to streamline surgical procedures through less frequent instrument changes [57].

The combination of ultrasonic friction and electrical resistance heating, however, accounts for a more intricate thermal footprint. In a direct in vivo porcine model study comparing Thunderbeat™ with EnSeal and LigaSure across multiple tissue types (liver, mesentery, muscle, and spleen), Thunderbeat™ produced significantly higher peak operating temperatures than the other two devices in a consistent manner. It led to the smallest total thermal injury zone in liver tissue, yet its collateral thermal spread in the spleen, mesentery, and muscle was significantly higher than that of EnSeal and LigaSure, which proved to be less thermally intensive and caused negligible collateral damage [58]. Therefore, when dissecting near critical neural or vascular structures, surgeons should take into account the possibility that Thunderbeat™ may have a propensity for greater lateral thermal spread.

There are three key surgical parameters that the existing literature mainly focuses on: the impact of integration on operative speed, the consistency of burst pressure for large vessels, and the efficiency of a single-instrument workflow. Through preclinical studies, the biophysical characteristics of this hybrid technology have been thoroughly defined on a theoretical level, and its dual action of ultrasonic cutting and bipolar coagulation has been demonstrated.

In an animal study, Seehofer et al. found that the device performed cutting operations at a faster rate than single-modality ultrasonic or bipolar devices, yet it was still able to maintain an average burst pressure of 734 mmHg for vessels of 5–7 mm diameter [42]. The sealing capacity was taken to the limit in a model of the lung artery ex vivo by Liberman et al. [43]. Their findings showed that the device was able to close pulmonary branches with an average burst pressure of 875 mmHg, thus demonstrating its potential applicability in the high-pressure vasculature of the chest.

These physical characteristics, when taken together, suggest that the device supports a multipurpose operation workflow, especially during complicated laparoscopic surgeries. Milsom et al. confirmed this aspect of the device in their prospective study on laparoscopic colorectal surgery [44]. They found that the device was able to handle challenging vascular pedicles such as the inferior mesenteric artery and could also be used as the main dissector. Hence, the hybrid instrument concept is supported by this case where the use of different instruments can be avoided by combining dissection and hemostasis into one single operation.

Versatility of these devices translates most advantageously to head and neck surgical settings in terms of documented operative time. In this domain, the highest level of evidence exists. Kuipers et al., comparing the device with traditional electrocautery, measured a median reduction in operative time of almost 50 minutes besides a decrease in blood loss (210 mL vs. 431 mL) [45]. To add to that, Suzuki et al. in a retrospective study looked at the device versus two other methods: the traditional approach and the LigaSure™ system [46]. While noticing shorter surgery duration as well, this research pinpointed the fact that the Thunderbeat™ cohort encountered fewer complications (3.4%) than the LigaSure™ group (20%). Taken together, these results imply that the main clinical benefit of the device is the reduction of operative time without compromising safety.

In essence, the Thunderbeat™ system is an effective alternative in the energy device market. Marrying ultrasonic speed with bipolar hemostasis, it delivers a device that is not only highly efficient but also very flexible, thus meeting a wide range of surgical situations such as, for example, procedures that require both thorough dissection and safe handling of bigger blood vessels, thereby assisting in attaining surgical efficiency in different disciplines.

3.4 Laser-based tissue fusion systems

Unlike RF and ultrasonic devices that rely on mechanically compressive jawed instruments, LTF or laser tissue soldering (LTS) primarily utilizes non-contact fiber-optic delivery systems combined with precise optical and thermal sensors. The fundamental engineering challenge in laser-based equipment is temperature regulation. The success of laser fusion depends entirely on maintaining the tissue surface temperature within a very narrow therapeutic window (typically 60 °C to 65 °C for optimal collagen denaturation). Overheating causes irreversible thermal necrosis and tissue vaporization, while underheating leads to insufficient protein cross-linking and weak tissue bonds.

To address this, modern clinical laser fusion equipment heavily relies on closed-loop temperature control systems. A typical temperature-controlled laser soldering system incorporates a laser generator (e.g., carbon dioxide, neodymium-doped yttrium aluminum garnet, or near-IR diode lasers), a flexible delivery optical fiber, and a real-time IR thermometry circuit or thermal imaging feedback circuit [59]. During the procedure, the IR sensor continuously monitors the thermal footprint at the tissue interface. A control algorithm processes this data in real-time to dynamically modulate the laser power output, ensuring that the tissue temperature stays within the target range without manual intervention.

Such systems have made a smooth transition to clinical applications. For example, during a pilot clinical trial of laparoscopic cholecystectomy, the temperature-controlled laser soldering system was used to close human skin incisions. The system was regularly successful in producing sealed wounds. Besides, the net time for closure was significantly reduced compared with traditional suturing methods, thus confirming the clinical feasibility of temperature-controlled laser devices [60].

Furthermore, the latest frontier in laser fusion equipment involves the integration of these temperature-controlled optical systems with surgical robotic platforms, known as robot-assisted LTS systems [61]. By mounting the laser applicator and IR sensors onto robotic arms, robot-assisted LTS systems eliminate the inconsistencies caused by human hand tremors and variable scanning speeds. This automated, high-precision control over both the spatial delivery of laser energy and the thermal parameters yields highly reproducible tissue bonds, paving the way for advanced, semi-automated wound closure in minimally invasive surgeries.

4 FUTURE PERSPECTIVES: ARTIFICIAL INTELLIGENCE (AI) AND ROBOTIC INTEGRATION

As EBTF technologies mature, the next frontier in surgical innovation will be the transition from passive energy delivery tools to active, intelligent surgical platforms. This shift is enabled by perfectly combining AI, machine learning, and advanced robotic systems.

4.1 AI-driven energy modulation and real-time tissue identification

At present, the majority of EBTF devices use simple impedance feedback to stop energy delivery. Nevertheless, the use of machine learning algorithms can bring about highly dynamic energy modulation. Through the constant analysis of multi-modal sensor inputs, e.g., tissue impedance, thermal dissipation, and hydration levels, AI models are able to forecast the best sealing endpoint for heterogeneous tissues instantly, thus achieving maximum hemostasis while minimizing iatrogenic thermal injury.

Furthermore, a groundbreaking application of AI in EBTF is the real-time identification of tissues through the analysis of surgical smoke. When EBTF devices cut or coagulate tissue, they generate aerosols rich in cellular metabolites and lipids. Technologies such as REIMS—often referred to as the "iKnife" concept—can capture this smoke. By coupling REIMS with machine learning classifiers (such as principal component analysis and neural networks), the system can analyze the mass spectrometric profile in milliseconds [62]. This allows the EBTF instrument to act as a smart diagnostic probe, instantly distinguishing between healthy tissue, critical structures (like nerves and blood vessels), and malignant tumor margins during oncological surgeries.

4.2 Integration with robotic surgical systems

The physical execution of tissue fusion is also being revolutionized by robotic platforms, such as the da Vinci Surgical System. Integrating advanced EBTF instruments—like fully articulated robotic vessel sealers—into robotic arms overcomes the physical limitations of human surgeons. Robotic integration offers three-dimensional high-definition visualization, eliminates physiological hand tremors, and provides a 360-degree range of motion. This allows for unparalleled precision when applying energy and mechanical compression in confined, complex anatomical spaces, such as the deep pelvis or the pancreatic head [63].

Looking ahead, the fusion of AI-based computer vision and robotic EBTF devices will allow for semi-autonomous surgical maneuvers. It is anticipated that future robotic systems will be able to locate vessels automatically, measure the precise pressure needed, and deliver the specific energy dosage for each patient, thereby ensuring consistent surgical results and lowering the mental burden on surgeons.

5 CONCLUSION

In short, tissue closure technology has seen a drastic change from traditional mechanical anastomosis to EBTF, which is nowadays regarded as a significant paradigm shift in surgical practice. The article has deeply discussed the mechanism of EBTF, which uses controlled thermal energy to cause collagen denaturation and molecular rearrangement, thus avoiding the major problems of sutures and staplers still commonly used today, such as foreign body reactions and technical difficulties in minimally invasive surgeries. It does not really matter if the tissue is heated through Joule heating by RF energy, through the mechanical friction of ultrasonic vibrations, or through the photothermal effect of laser energy; biologically, all these methods target the same outcome: a strong, self-healing tissue seal. Taken to clinical practice, these concepts have led to the development of revolutionary medical devices. A good example of this is the LigaSure™ system which uses a real-time impedance feedback technology to continuously and accurately seal tissues. The HARMONIC™ system relies on ultrasonic vibrations to precisely cut tissue while effectively limiting lateral thermal damage. The Thunderbeat™ system is a hybrid of the two techniques that offers the speed of ultrasound combined with the safety of bipolar sealing. The large amount of clinical data presented in this paper demonstrates that these techniques are not only effective in achieving hemostasis but also, through their associated medical devices, greatly facilitate the surgeon's work, reduce the duration of the operation, and minimize blood loss. On top of that, they have such a wide range of application in basically every specialty that they can be used in procedures ranging from simple vessel ligation to digestive, endocrine, and urological surgeries.

Looking forward, the development of energy-based surgical devices will be focused on more than just the power generation aspect. These devices will be more intelligent, precise, and integrated with robotic technologies [64]. Energy devices will become seamlessly integrated with robotic systems as the use of robotic platforms for minimally invasive surgery increases. Although advanced energy-based devices are associated with higher costs at the initial stage, their capability of lessening the duration of the operation and the rate of complications is translated into a favorable long-term cost-benefit scenario for the healthcare system [65]. Furthermore, the use of AI and computer vision is expected to completely change the concept of surgical safety. Intelligent surgical systems of the future will most probably use machine learning algorithms to identify the nature of tissues during surgery in real-time [66]. As a result, next-generation surgical devices might be capable of automatically identifying nerves, blood vessels, or other tissues. In this way, they can dynamically adjust energy output to maximize hemostasis while minimizing iatrogenic thermal injury. Ultimately, these innovations are intended to be able to strike the perfect combination of anastomotic strength, minimal tissue injury, and optimal healing results.

DECLARATIONS

Author contributions

Junjie Shen and Zhongxin Hu were responsible for writing the original draft and conducting the investigation. Lin Mao and Chengli Song were responsible for writing, reviewing, and editing the manuscript, as well as supervising the project.

Funding

This research received no external funding.

Data availability

No new data were generated or analyzed in this review. All data cited in this article are from previously published studies, which have been appropriately referenced in the text.

Ethics approval and consent to participate

Not applicable.

Consent for publication

All authors acknowledge and agree to submit this article to the journal Progress in Medical Devices.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.

REFERENCES

[1] Banga A, Saeed Muhammad S, Misra S, Ying X, Patino C, Rathore Sawai S, et al. Sealing the deal: A meta-analysis of suture-only and hybrid closure techniques for femoral access hemostasis after percutaneous cardiovascular procedures. J Am Coll Cardiol. 2025 Apr 1;85(12_Supplement):854-854. https://doi.org/10.1016/S0735-1097(25)01338-5

[2] Chen Y, Chai M, Xuan C, Lin J, Yang H, Li C, et al. Tuning the properties of surgical polymeric materials for improved soft-tissue wound closure and healing. Prog Mater Sci. 2024 Jun 1;143:101249. https://doi.org/10.1016/j.pmatsci.2024.101249

[3] Crumplin MKH. Surgical suture needles. Br J Surg. 2023 Mar 2;110(5):548-552. https://doi.org/10.1093/bjs/znad053

[4] Yang N, Venezuela J, Allavena R, Lau C, Dargusch M. Zinc-based subcuticular absorbable staples: An in vivo and in vitro study. Acta Biomater. 2023 Sep 1;167:593-607. https://doi.org/10.1016/j.actbio.2023.06.030

[5] Albéniz E, Álvarez MA, Espinós JC, Nogales O, Guarner C, Alonso P, et al. Clip closure after resection of large colorectal lesions with substantial risk of bleeding. Gastroenterology. 2019 Nov 1;157(5):1213-1221. https://doi.org/10.1053/j.gastro.2019.07.037

[6] Pagán AJ, Ramakrishnan L. The formation and function of granulomas. Annu Rev Immunol. 2018 Feb 5;36:639-665. https://doi.org/10.1146/annurev-immunol-032712-100022

[7] Gong Y, Li C, Zhu B, Li F, Feng L, Zhao Y, et al. Engineering injectable bioadhesives with sealing and anti-fouling properties for postoperative anti-adhesion. Chem Eng J. 2024 Nov 15;500:156747. https://doi.org/10.1016/j.cej.2024.156747

[8] Kelmansky R, McAlvin BJ, Nyska A, Dohlman JC, Chiang HH, Hashimoto M, et al. Strong tissue glue with tunable elasticity. Acta Biomater. 2017 Apr 15;53:93-99. https://doi.org/10.1016/j.actbio.2017.02.009

[9] Peng W, Lai Y, Fan J, Dai C, Shen J, Liu P. Simple component self-gelling powders as multifunctional tissue adhesives for sutureless wound healing. Chem Eng J. 2024 Sep 1;495:153514. https://doi.org/10.1016/j.cej.2024.153514

[10] Zheng Y, Baidya A, Annabi N. Molecular design of an ultra-strong tissue adhesive hydrogel with tunable multifunctionality. Bioact Mater. 2023 Nov 1;29:214-229. https://doi.org/10.1016/j.bioactmat.2023.06.007

[11] Xu F, Lu T, Seffen KA. Biothermomechanical behavior of skin tissue. Acta Mech Sin. 2008 Feb 1;24(1):1-23. https://doi.org/10.1007/s10409-007-0128-8

[12] Leo L, Bridelli MG, Polverini E. Reversible processes in collagen dehydration: A molecular dynamics study. Arch Biochem Biophys. 2021 Dec 15;714:109079. https://doi.org/10.1016/j.abb.2021.109079

[13] Mogilner IG, Ruderman G, Grigera JR. Collagen stability, hydration and native state. J Mol Graph Model. 2002 Dec 1; 21(3):209-213. https://doi.org/10.1016/S1093-3263(02)00145-6

[14] Wright NT, Humphrey JD. Denaturation of collagen via heating: An irreversible rate process. Annu Rev Biomed Eng. 2002 Jul 16;4:109-128. https://doi.org/10.1146/annurev.bioeng.4.101001.131546

[15] Kramer EA, Rentschler ME. Energy-based tissue fusion for sutureless closure: Applications, mechanisms, and potential for functional recovery. Annu Rev Biomed Eng. 2018 Jun 4;20:1-20. https://doi.org/10.1146/annurev-bioeng-071516-044702

[16] Su L, Cloyd KL, Arya S, Hedegaard MA, Steele JA, Elson DS, et al. Raman spectroscopic evidence of tissue restructuring in heat-induced tissue fusion. J Biophotonics. 2013 Nov 15;7(9):713-723. https://doi.org/10.1002/jbio.201300099

[17] Zhou D, Yang G, Xu J, Ling B, Wang S. Non-thermal effect of radio frequency treatments verified by the multi-scale structure and in-vitro digestibility of sweet potato starch. Innov Food Sci Emerg Technol. 2023 Jul 1;87:103412. https://doi.org/10.1016/j.ifset.2023.103412

[18] Kennedy JS, Stranahan PL, Taylor KD, Chandler JG. High-burst-strength, feedback-controlled bipolar vessel sealing. Surg Endosc. 1998 Jun 1;12(6):876-878. https://doi.org/10.1007/s004649900733

[19] Lamberton GR, Hsi RS, Jin DH, Lindler TU, Jellison FC, Baldwin DD. Prospective comparison of four laparoscopic vessel ligation devices. J Endourol. 2008 Oct 1;22(10):2307-2312. https://doi.org/10.1089/end.2008.9715

[20] Pan H, Leung KKC, Ng EKW. Tissue fusion technology versus suture and staple in porcine bowel anastomosis: An in vivo study. Braz J Med Biol Res. 2020 Apr 17;53(5):e9305. https://doi.org/10.1590/1414-431X20209305

[21] Cardoso GS, Martins YA, Lopez RFV, Carneiro AAO, Lopes JH, Pavan TZ. Medical and biological applications of Langevin-type ultrasonic transducers: A narrative review. Ultrasonics. 2026 Feb 1;158:107825. https://doi.org/10.1016/j.ultras.2025.107825

[22] Saber Jahromi M, Abedini R. Design optimization and performance characteristics of a hemostatic ultrasonic scalpel. Int J

Interact Des Manuf. 2025 Sep 1;19(9):6319-6330. https://doi.org/10.1007/s12008-024-02216-z

[23] Yang X, Cao J, Yan Y, Liu F, Li T, Han L, et al. Comparison of the safety of electrotome, Harmonic scalpel, and LigaSure for management of thyroid surgery. Head Neck. 2017 Jun 1;39(6):1078-1085. https://doi.org/10.1002/hed.24701

[24] Foschi D, Cellerino P, Corsi F, Taidelli T, Morandi E, Rizzi A, et al. The mechanisms of blood vessel closure in humans by the application of ultrasonic energy. Surg Endosc Other Interv Tech. 2002 May 1;16(5):814-819. https://doi.org/10.1007/s00464-001-9074-x

[25] Sista F, Abruzzese V, Schietroma M, Cecilia EM, Mattei A, Amicucci G. New harmonic scalpel versus conventional hemostasis in right colon surgery: A prospective randomized controlled clinical trial. Dig Surg. 2013 Jan 1;30(4-6):355-361. https://doi.org/10.1159/000354864

[26] Tolone S, Bondanese M, Ruggiero R, Gili S, Pirozzi R, Parisi S, et al. Outcomes of sutureless total thyroidectomy in elderly. Int J Surg. 2016 Sep 1;33 Suppl 1:S16-S19. https://doi.org/10.1016/j.ijsu.2016.05.056

[27] Wright VC, Riopelle MA. Laser physics for surgeons: An introduction to carbon dioxide laser surgery. Acta Obstet Gynecol Scand. 1984 Jan 1;63(S125):5-15. https://doi.org/10.3109/00016348409157012

[28] Bass LS, Moazami N, Pocsidio J, Oz MC, Logerfo P, Treat MR. Changes in type I collagen following laser welding. Lasers Surg Med. 1992 Jan 1;12(5):500-505. https://doi.org/10.1002/lsm.1900120508

[29] Jain KK, Gorisch W. Repair of small blood vessels with the neodymium-YAG laser: A preliminary report. Surgery. 1979 Jun;85(6):684-688.

[30] Bremmer J, van Doormaal TPC, Verweij BH, van der Zwan A, Tulleken CAF, Verdaasdonk R. Vessel wall perforation mechanism of the excimer laser-assisted non-occlusive anastomosis technique. Lasers Med Sci. 2016 May 25;31(6):1169-1175. https://doi.org/10.1007/s10103-016-1950-7

[31] Li C, Wang K. Effect of welding temperature and protein denaturation on strength of laser biological tissues welding. Opt Laser Technol. 2021 Jun 1;138:106862. https://doi.org/10.1016/j.optlastec.2020.106862

[32] Ghosh D, Salinas CM, Pallod S, Roberts J, Makin IRS, Yaron JR, et al. Temporal evaluation of efficacy and quality of tissue repair upon laser-activated sealing. Bioeng Transl Med. 2022 Sep 28;8(2):e10412. https://doi.org/10.1002/btm2.10412

[33] Amirkazem VS, Malihe K. Randomized clinical trial of ligasure™ versus conventional splenectomy for injured spleen in blunt abdominal trauma. Int J Surg. 2017 Feb 1;38:48-51. https://doi.org/10.1016/j.ijsu.2016.12.036

[34] Akhtar F, Rauf S, Noor A, Ayubi A, Sakhizada F. 314 Comparison of hemorrhoidectomy by LigaSure and conventional Milligan Morgan's technique. Br J Surg. 2023 Aug 21;110(Suppleme nt_6):znad241.478. https://doi.org/10.1093/bjs/znad241.478

[35] Vidal O, Saavedra-Perez D, Valentini M, Astudillo E, Fernández-Cruz L, García-Valdecasas JC. Surgical outcomes of total thyroidectomy using the LigaSure™ Small Jaw versus LigaSure Precise™: A retrospective study of 2000 consecutive patients. Int J Surg. 2017 Jan 1;37:8-12. https://doi.org/10.1016/j.ijsu.2016.08.241

[36] Constant DL, Florman SS, Mendez F, Thomas R, Slakey DP. Use of the LigaSure vessel sealing device in laparoscopic living-donor nephrectomy. Transplantation. 2004 Dec 15;78(11):1661-1664. https://doi.org/10.1097/01.tp.0000144379.29943.ed

[37] Revelli L, Damiani G, Bianchi CBNA, Vanella S, Ricciardi W, Raffaelli M, et al. Complications in thyroid surgery. Harmonic scalpel, harmonic focus versus conventional hemostasis: A meta-analysis. Int J Surg. 2016 Apr 1;28:S22-S32. https://doi.org/10.1016/j.ijsu.2015.12.050

[38] Dong W. 87P Therapeutic effect of harmonic scalpel versus electrocautery in axillary dissection of breast cancer. Ann Oncol. 2016 Dec 1;27:25-26. https://doi.org/10.1016/S0923-7534(21)00247-7

[39] Waraich N, Ahmed J, Rashid F, Mulvey D, Leeder P, Iftikhar SY. Is harmonic scalpel an effective tool for oesophagectomy? Int J Surg. 2009 Jan 1;7(4):330-333. https://doi.org/10.1016/j.ijsu.2009.03.001

[40] Abo-hashem AA, Sarhan A, Aly AM. Harmonic Scalpel® compared with bipolar electro-cautery hemorrhoidectomy: A randomized controlled trial. Int J Surg. 2010 Jan 1;8(3):243-247. https://doi.org/10.1016/j.ijsu.2010.01.010

[41] Kumar A, Agarwal PN, Kumar Garg P. Evaluation of subfascial endoscopic perforator vein surgery (SEPS) using harmonic scalpel in varicose veins: An observational study. Int J Surg. 2009 Jan 1;7(3):253-256. https://doi.org/10.1016/j.ijsu.2009.04.005

[42] Seehofer D, Mogl M, Boas-Knoop S, Unger J, Schirmeier A, Chopra S, et al. Safety and efficacy of new integrated bipolar and ultrasonic scissors compared to conventional laparoscopic 5-mm sealing and cutting instruments. Surg Endosc. 2012 Mar 27;26(9):2541-2549. https://doi.org/10.1007/s00464-012-2229-0

[43] Liberman M, Khereba M, Goudie E, Kazakov J, Thiffault V, Lafontaine E, et al. Pilot study of pulmonary arterial branch sealing using energy devices in an ex vivo model. J Thorac Cardiovasc Surg. 2014 Aug 16;148(6):3219-3223. https://doi.org/10.1016/j.jtcvs.2014.05.089

[44] Milsom JW, Trencheva K, Sonoda T, Nandakumar G, Shukla PJ, Lee S. A prospective trial evaluating the clinical performance of a novel surgical energy device in laparoscopic colon surgery. Surg Endosc. 2015 Aug 27;29(5):1161-1166. https://doi.org/10.1007/s00464-014-3783-4

[45] Kuipers NC, de Kleijn BJ, Wedman J, van der Laan B, Plaat BEC, Halmos GB. A comparison of the Thunderbeat and standard electrocautery devices in head and neck surgery: A prospective randomized controlled trial. Eur Arch Otorhinolaryngol. 2021 Mar 19;278(12):4987-4996. https://doi.org/10.1007/s00405-021-06739-z

[46] Suzuki K, Shimizu M, Sakagami T, Yagi M, Fujisawa T, Sawada S, et al. A comparison of short-term outcomes of neck dissection for head and neck cancers using Thunderbeat, LigaSure or treatment without an energy-based device: A case controlled study. Int J Surg. 2018 Sep 21;58:60-64. https://doi.org/10.1016/j.ijsu.2018.09.009

[47] Marynskyy GS, Tkachenko VA, Bysko VO, Podpryatov SE, Podpriatov SS, Grabovskyi SD, et al. High-frequency equipment for live tissue welding. Paton Weld J. 2023 Jan 4;(1):23-30. https://doi.org/10.37434/tpwj2023.01.04

[48] Taheri A, Mansoori P, Sandoval LF, Feldman SR, Pearce D, Williford PM. Electrosurgery: part II. Technology, applications, and safety of electrosurgical devices. J Am Acad Dermatol. 2014 Mar 13;70(4):607.e1-607.e12. https://doi.org/10.1016/j.jaad.2013.09.055

[49] Sran H, Sebastian J, Hossain MA. Electrosurgical devices: Are we closer to finding the ideal appliance? A critical review of current evidence for the use of electrosurgical devices in general surgery. Expert Rev Med Devices. 2016 Feb 1;13(2):203-215. https://doi.org/10.1586/17434440.2016.1134312

[50] Sutton PA, Awad S, Perkins AC, Lobo DN. Comparison of lateral thermal spread using monopolar and bipolar diathermy, the Harmonic Scalpel™ and the Ligasure™. Br J Surg. 2010 Jan 27;97(3):428-433. https://doi.org/10.1002/bjs.6901

[51] Palazzo FF, Francis DL, Clifton MA. Randomized clinical trial of Ligasure™ versus open haemorrhoidectomy. Br J Surg. 2002 Feb 22;89(2):154-157. https://doi.org/10.1046/j.1365-2168.2002.01993.x

[52] Schuchert MJ, Abbas G, Landreneau JP, Luketich JD, Landreneau RJ. Use of energy-based coagulative fusion technology and lung sealants during anatomic lung resection. J Thorac Cardiovasc Surg. 2012 Sep 1;144(3):S48-S51. https://doi.org/10.1016/j.jtcvs.2012.06.001

[53] Manoli E, Mason S, Ford L, Adebesin A, Bodai Z, Darzi A, et al. Validation of ultrasonic harmonic scalpel for real-time tissue identification using rapid evaporative ionization mass spectrometry. Anal Chem. 2021 Apr 13;93(14):5906-5916. https://doi.org/10.1021/acs.analchem.1c00270

[54] Nanashima A, Abo T, Arai J, Takagi K, Matsumoto H, Takeshita H, et al. Usefulness of vessel-sealing devices combined with crush clamping method for hepatectomy: A retrospective cohort study. Int J Surg. 2013 Nov 1;11(9):891-897. https://doi.org/10.1016/j.ijsu.2013.07.012

[55] Krones CJ, Conze J, Hoelzl F, Stumpf M, Klinge U, Möller M, et al. Chemical composition of surgical smoke produced by electrocautery, harmonic scalpel and argon beaming - a short study. Eur Surg. 2007 Apr 1;39(2):118-121. https://doi.org/10.1007/s10353-006-0305-1

[56] Barrett WL, Garber SM. Surgical smoke: A review of the literature. Surg Endosc Other Interv Tech. 2003 Jun 1;17(6):979-987. https://doi.org/10.1007/s00464-002-8584-5

[57] Milsom J, Trencheva K, Monette S, Pavoor R, Shukla P, Ma J, et al. Evaluation of the safety, efficacy, and versatility of a new surgical energy device (THUNDERBEAT) in comparison with Harmonic ACE, LigaSure V, and EnSeal devices in a porcine model. J Laparoendosc Adv Surg Tech. 2012 May 1;22(4):378-386. https://doi.org/10.1089/lap.2011.0420

[58] Dunay MP, Lipcsey Z, Arany-Tóth A, Németh T, Solymosi N, Venczel L, et al. Experimental assessment of three electrosurgical tissue-sealing devices in a porcine model. Acta Vet Hung. 2020 Sep 30;68(3):318-322. https://doi.org/10.1556/004.2020.00051

[59] Dizman B, Ruhi MK. Development of an infrared array sensor-integrated laser system for precision and efficacy in medical applications. Lasers Med Sci. 2025 Jun 5;40(1):254. https://doi.org/10.1007/s10103-025-04510-y

[60] David Simhon MD, Ilan G, Gregory S, Tamar V, Israel N, Roberto M, et al. Temperature-controlled laser-soldering system and its clinical application for bonding skin incisions. J Biomed Opt. 2015 Dec 1;20(12):128002. https://doi.org/10.1117/1.JBO.20.12.128002

[61] Basov S, Milstein A, Sulimani E, Platkov M, Peretz E, Rattunde M, et al. Robot-assisted laser tissue soldering system. Biomed Opt Express. 2018 Nov 1;9(11):5635-5644. https://doi.org/10.1364/BOE.9.005635

[62] Balog J, Sasi-Szabó L, Kinross J, Lewis MR, Muirhead LJ, Veselkov K, et al. Intraoperative tissue identification using rapid evaporative ionization mass spectrometry. Sci Transl Med. 2013 Jul 17;5(194):194ra193. https://doi.org/10.1126/scitranslmed.3005623

[63] Abaza R, Henderson SJ, Martinez O. Robotic vessel sealer device for lymphocele prevention after pelvic lymphadenectomy: Results of a randomized trial. J Laparoendosc Adv Surg Tech. 2022 May 31;32(7):721-726. https://doi.org/10.1089/lap.2021.0531

[64] Saeidi H, Opfermann JD, Kam M, Wei S, Leonard S, Hsieh MH, et al. Autonomous robotic laparoscopic surgery for intestinal anastomosis. Sci Robot. 2022 Jan 26;7(62):eabj2908. https://doi.org/10.1126/scirobotics.abj2908

[65] Vettoretto N, Foglia E, Gerardi C, Lettieri E, Nocco U, Botteri E, et al. High-energy devices in different surgical settings: Lessons learnt from a full health technology assessment report developed by SICE (Società Italiana di Chirurgia Endoscopica). Surg Endosc. 2023 Apr 1;37(4):2548-2565. https://doi.org/10.1007/s00464-022-09734-5

[66] Hashimoto DA, Rosman G, Rus D, Meireles OR. Artificial intelligence in surgery: Promises and perils. Ann Surg. 2018 Jul 1;268(1):70-76. https://doi.org/10.1097/sla.0000000000002693

Progress in
Medical Devices

RESEARCH ARTICLE

A multi-frequency power amplifier for detecting tiny metal in the human body

Yuming Liu, Piding Li

School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Piding Li.

Address correspondence to: Piding Li, School of Health Science and Engineering, University of Shanghai for Science and Technology, No. 516 Jungong Road, Yangpu District, Shanghai 200093, China. E-mail: lpdbyusst@163.com.

Received March 11, 2026; Accepted April 20, 2026; Published June 25, 2026

DOI: 10.61189/744920nwaoek

Abstract

Objective: Tiny metallic foreign bodies may remain in the human body after accidental ingestion, surgery, or ballistic injury, potentially causing inflammation, tissue damage, and other complications. Although X-ray and CT are widely used for detection and localization, intraoperative motion and workflow constraints may reduce localization accuracy and real-time retrieval efficiency. This study aims to develop a portable multi-frequency electromagnetic excitation circuit to assist the detection and localization of tiny metallic foreign bodies in the human body. **Methods:** A multi-frequency electromagnetic excitation circuit was designed for balanced-coil eddy-current sensing. The proposed transmitter combines Selective Harmonic Elimination Pulse-Width Modulation (SHE-PWM) with a full-bridge Class-D power amplifier to generate synchronous multi-frequency excitation currents at 50 kHz, 150 kHz, 350 kHz, and 850 kHz. The use of multiple excitation frequencies provides complementary depth sensitivity, in which low-frequency excitation improves penetration depth for deeply embedded targets, while high-frequency excitation enhances the response and spatial resolution of small or superficial objects. Circuit simulations and hardware measurements were conducted to evaluate the time-domain and frequency-domain characteristics of the proposed circuit. **Results:** Simulation and experimental results showed good agreement with theoretical predictions. The proposed circuit successfully generated synchronous multi-frequency excitation currents with controllable spectral components. The results confirmed that the combination of SHE-PWM and a full-bridge Class-D power amplifier can provide spectrally controllable and energy-efficient excitation suitable for balanced-coil eddy-current sensing. **Conclusions:** The proposed multi-frequency electromagnetic excitation circuit provides a feasible supplementary solution for tiny metallic foreign-body detection and localization. Its low-cost, portable, and energy-efficient characteristics make it potentially suitable for bedside and intraoperative electromagnetic assistance, especially in scenarios where conventional imaging methods are limited by workflow constraints or real-time localization requirements.

Keywords: SHE-PWM, Class-D power amplifier, Eddy current testing, Balanced metal coil, Detecting tiny metal in body

Highlights

- A multi-frequency electromagnetic excitation scheme is proposed for the detection and localization of tiny metallic foreign bodies inside the human body.
- By integrating SHE-PWM with a full-bridge Class-D power amplifier, the transmitter achieves energy-efficient, spectrally controllable, and synchronous multi-frequency excitation.

- Optimized switching angles enhance designated frequency components and suppress non-target harmonics, improving both excitation efficiency and frequency selectivity.
- Low-frequency excitation increases penetration depth for deeply embedded objects, while high-frequency excitation enhances the response and resolution of small or superficial targets.
- Circuit simulations and hardware measurements show good agreement with theoretical analysis, validating the feasibility and effectiveness of the proposed method.

1 INTRODUCTION

In modern medical practice, inadvertent introduction and retention of metallic foreign bodies in the human body are not uncommon. Representative scenarios include accidental ingestion of metal fragments, traumatic retention of shrapnel or steel pellets, and intraoperative loss of small metallic items such as suture needles, hemostatic clips, guidewires, or miniature instrument components [1, 2]. Once retained, such foreign bodies may trigger localized inflammatory reactions, tissue injury, elevated infection risk, and even secondary harm during subsequent examinations or interventions. Accordingly, clinical management typically requires a prompt closed-loop workflow encompassing screening, localization, and retrieval. At present, X-ray radiography and computed tomography (CT) remain the mainstream modalities for metal localization, as they can directly depict the morphology of the foreign body and its spatial relationship with surrounding tissues. However, during perioperative or emergency management, physiological motions—including muscle contraction, respiration, and gastrointestinal peristalsis—can cause foreign-body migration; moreover, the use of X-ray or CT is often constrained by the size and limited maneuverability of these large-scale systems, which may delay screening and removal. Therefore, the development of a low-cost, portable, and real-time auxiliary localization technique capable of rapid confirmation and dynamic tracking beyond conventional imaging has clear engineering value and clinical significance.

Compared with conventional industrial metal detection, human-body metallic foreign-body detection is more challenging. Targets may be deeply embedded across different tissue layers with variable stand-off distance from the sensing coil. In addition, foreign bodies exhibit diverse morphologies (needle-like, plate-like, or particulate), and dynamic orientations driven by respiration, muscle contraction, and peristalsis. Meanwhile, biological tissues are conductive and dielectric with substantial inter-subject variability, and the presence of metallic surgical instruments further complicates the electromagnetic background and increases interference. Under such conditions, a single excitation frequency is often insufficient to satisfy both sensitivity and stability requirements. Hence, multi-frequency eddy-current techniques are desirable to improve detection efficiency: lower frequencies enhance effective penetration depth and facilitate detectability of deeply embedded objects, whereas higher frequencies strengthen near-surface responses and improve resolution for small or superficial targets, thereby improving reliability under complex backgrounds and posture variations [3, 4].

Eddy-current-based detection provides a feasible route for rapid localization of metallic foreign bodies in the human body. An excitation coil generates an alternating magnetic field that induces eddy currents in the metallic target (or magnetization responses in ferromagnetic objects) [5-7]. The resulting magnetic-field perturbation is then sensed by a receiving coil, enabling the determination of target presence and relative position. Similar to industrial in-line rejection systems, this approach relies on high-sensitivity acquisition and robust decision-making to detect weak electromagnetic signals.

The overall architecture of an eddy-current detection system is illustrated in **Figure 1**, typically comprising a power transmitter, a receiving and signal-conditioning front end, an analog-to-digital converter (ADC), and a digital processing unit (e.g., a microcontroller or field-programmable gate array [FPGA]). Among these components, the power transmitter is indispensable for improving the signal-to-noise ratio (SNR) and overall sensitivity: by delivering high-current excitation, it generates a sufficiently strong alternating electromagnetic field, thereby increasing the magnitude of target-induced field perturbations and the induced voltage at the receiver [8, 9]. In multi-frequency eddy-current detection, critical information is encoded in the frequency-dependent responses associated with target material properties, size, and depth. Therefore, the transmitter must provide not only high-power output but also synchronous, stable, and controllable multi-frequency excitation, ensuring phase coherence, amplitude consistency, and spectral purity across frequencies to meet the requirements of demodulation and feature extraction, and to avoid inter-frequency coupling and amplitude/phase drift that would otherwise degrade localization or identification accuracy. To this end, this work designs a synchronous multi-frequency transmitter employing a selective harmonic elimination (SHE) strategy to synthesize the excitation current waveform, enhancing the desired frequency components while suppressing non-target harmonics, thereby providing a stable transmission foundation for high-sensitivity detection under multi-frequency excitation.

2 EDDY CURRENT PRINCIPLE

Figure 2 schematically illustrates the operating principle of eddy-current–based detection. The output of the power amplifier drives the transmit coil. To enhance the coupling (sensing) coefficient between coils in the metal-detection probe, the

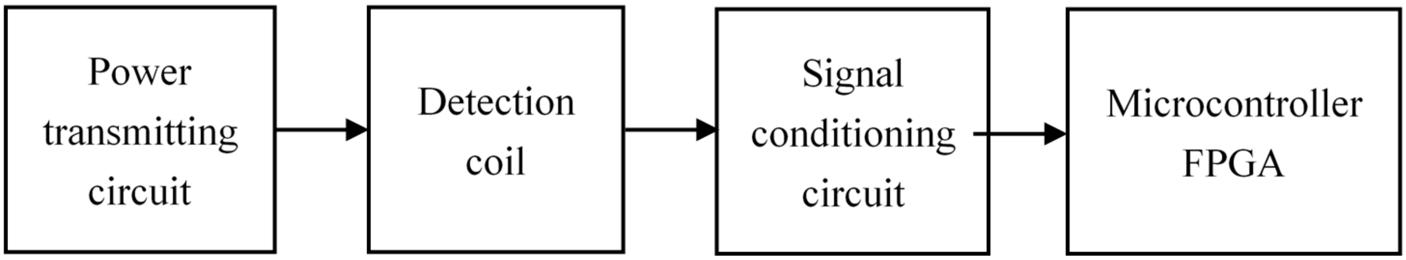


Figure 1. Overall architecture of an eddy-current detection system. FPGA, field-programmable gate array.

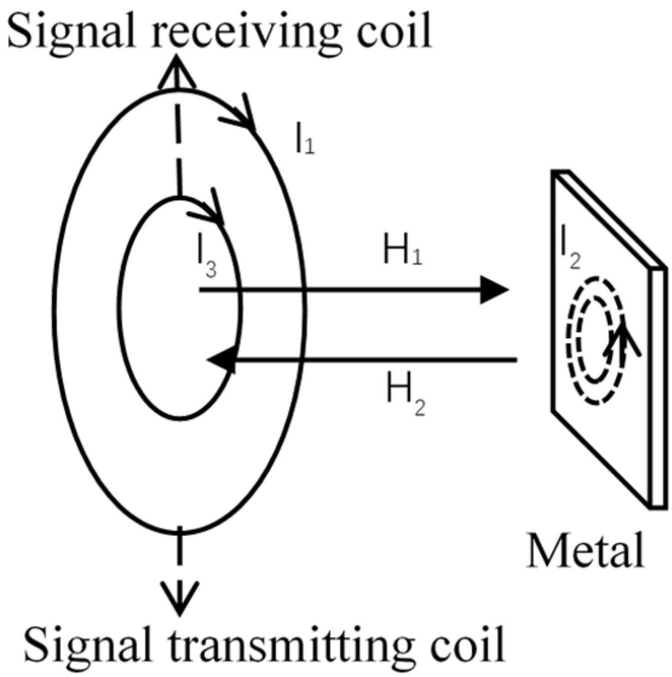


Figure 2. Operating principle of eddy-current-based detection.

transmit and receive coils are arranged in a coaxial coplanar configuration, with the transmit coil placed externally and the receive coil internally. The alternating magnetic field generated by the detection coil induces a rotational electric field in the conductive target, thereby driving closed-loop eddy-current distributions. According to Lenz's law, these eddy currents generate a secondary magnetic field that couples back to the primary excitation field, thereby altering the magnetic flux linkage and the equivalent impedance of the detection coil [10, 11]. The resulting impedance variation comprises both a resistive component, associated with Joule losses induced by eddy currents, and an inductive component, associated with magnetic-flux redistribution (with explicit frequency dependence) [12, 13]. This impedance response is jointly governed by the target's electrical conductivity and magnetic permeability, as well as its geometry, spatial position, and the excitation frequency (via the skin-depth effect). By measuring and demodulating the amplitude and phase of the coil voltage or current (or equivalently the in-phase and quadrature [I/Q] components), one can detect the presence of metallic targets and further enable parameter characterization, localization, and identification.

3 SELECTION OF POWER AMPLIFIER TYPE

In the above-described electromagnetic detection system, the transmitting coil requires excitation currents of sufficiently high magnitude, making the choice of power amplifier configuration critical to overall system performance. For single-frequency eddy-current detection systems, a resonant driving scheme can be employed, in which the transmit coil and matching network form a resonant circuit capable of delivering power efficiently at a single frequency. However, for multi-frequency excitations, simultaneous resonance across multiple frequency components is impractical due to the inherently narrowband characteristics of resonant amplifiers.

In the context of multi-frequency excitations, power amplifier performance directly affects the quality of the modulated signal, the harmonic composition of synthesized waveforms, and the overall energy efficiency of the detection system. Previous studies have shown that Class-AB power amplifiers are preferred in analog and/or audio applications owing to their favorable linearity and low distortion levels [14]. Nevertheless, their application in high-power broadband or multi-frequency eddy current setups is limited. This limitation arises from suboptimal power conversion efficiency, significant thermal dissipation, restricted operational bandwidth, and insufficient current-driving capability when directly powering excitation coils. Moreover, the high thermal output of Class-AB amplifiers reduces system efficiency and longevity and increases costs owing to the need for additional thermal management.

To address these drawbacks of conventional power amplifiers, switch-mode power amplifiers, particularly Class-D types, are gaining popularity. In a Class-D power amplifier, the output transistors operate as electronic switches, alternating between fully ON and fully OFF states at high frequency. This switching minimizes switching and conduction power losses, resulting in remarkably high amplifier efficiency [15, 16]. **Figure 3** illustrates the circuit diagram of a full-bridge (H-bridge) Class-D amplifier, a configuration widely used in modern high-efficiency designs. The combination of a Class-D power amplifier with multi-frequency selective harmonic elimination pulse-width modulation (SHE-PWM) enables direct generation of complex multi-frequency current signals from a direct current (DC) power source, without requiring linear amplification of each frequency component. This approach over-

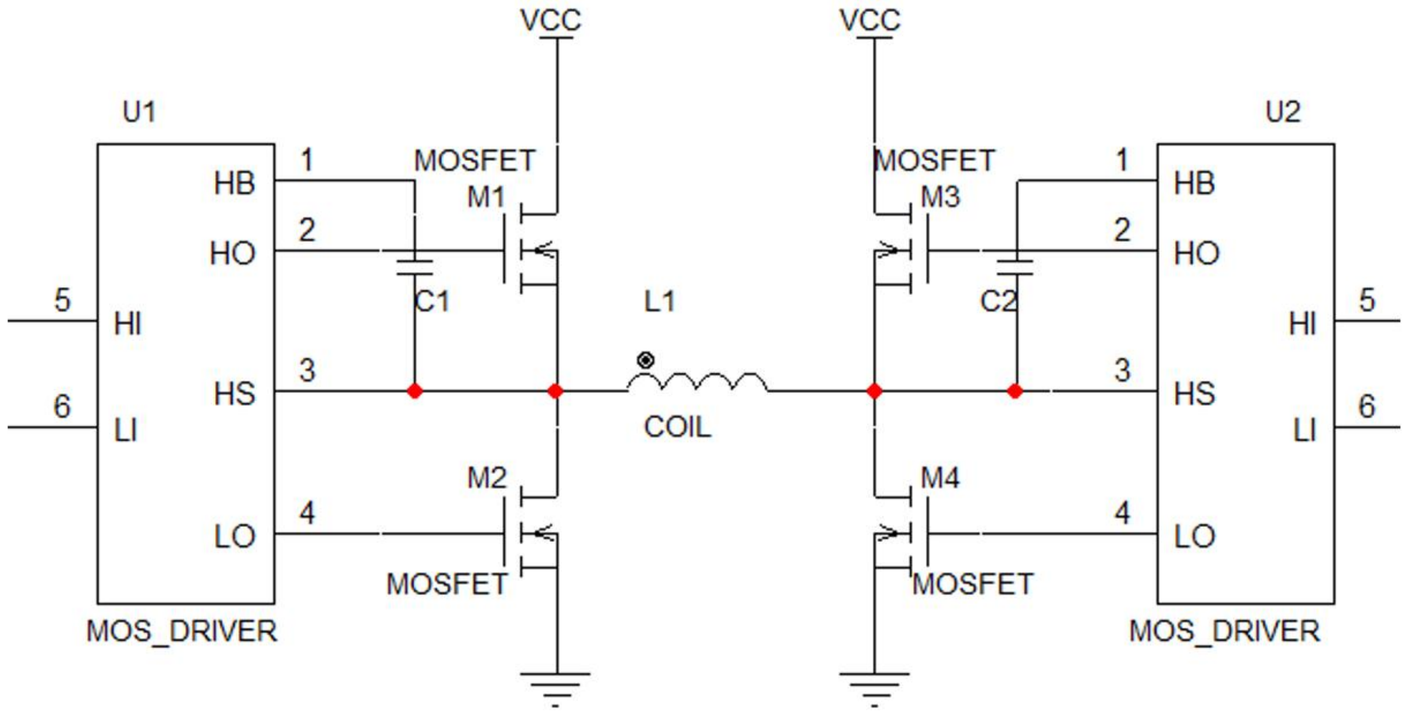


Figure 3. Class-D power amplifier (H-bridge).

Table 1. Class-AB power amplifiers and Class-D power amplifiers

Characteristic	Class-AB power amplifiers	Class-D power amplifiers
Operating principle	Linear amplification	Switching amplification
Efficiency	Typically 30–60%	Typically 85–95%
Suitable frequency range	Best suited to low-to-mid frequencies (<100 kHz).	Well suited to mid-to-high frequencies (kHz to MHz)
Output-current capability	Limited	Very high
Thermal loss	High	Very low

comes the efficiency bottleneck of conventional amplifiers while maintaining high spectral purity of the output current signals and their harmonics. It allows real-time generation of broadband excitation signals with both high efficiency and high spectral purity, making it suitable for advanced multi-frequency eddy-current detection systems. **Table 1** summarizes the performance differences between Class-D and Class-AB power amplifiers.

4 PRINCIPLE OF SELECTIVE HARMONIC ELIMINATION

4.1 Fourier series representation of inductor current

To achieve high SNR, low drift, and repeatability in eddy-current–based localization of metallic foreign bodies in the human body, the transmitter must not only provide sufficient excitation strength but also deliver a controllable and stable excitation spectrum. This requirement is particularly critical under synchronous multi-frequency excitation, where inter-frequency amplitude and phase consistency directly determines the stability of I/Q-demodulated features and the achievable localization accuracy. Although Class-D power amplifiers offer high efficiency and are suitable for portable implementations, their switching outputs inherently contain rich harmonic components. If such harmonics are not properly controlled, the tissue–coil system may exhibit additional background responses, baseline drift, and inter-frequency crosstalk, thereby reducing detectability of tiny targets and increasing the risk of false positives. To address this issue, this study adopts a SHE strategy, optimizing switching angles to synthesize the desired excitation at target frequencies while suppressing critical non-target harmonics. This approach enhances the robustness of multi-frequency detection and improves the reliability of target localization.

In the SHE strategy, the switching angles are designed to constrain the output harmonic spectrum. As a result, the excitation current achieves the desired amplitude at the designated target frequencies while minimizing specific non-target harmonic components.

The periodic pulse-width modulation (PWM) voltage generated by the power electronic converter can be represented using a Fourier series, which forms the basis for harmonic angle optimization, as shown in Equation (1) [17]:

$$V_N(t) = \frac{a_0}{2} + \sum_{n=1}^{N}\left(a_n \cos\left(\frac{2\pi nt}{T}\right) + b_n \sin\left(\frac{2\pi nt}{T}\right)\right) \tag{1}$$

where a_0 indicates the DC term, a_n the cosine coefficients, and b_n the sine coefficients. The Fourier series representation of the PWM waveform is inherently complex due to multiple switching intervals and asymmetric pulse durations.

To cosine the analytical solution and facilitate harmonic management, the PWM waveforms can be assumed to possess quarter-wave symmetry. Under this assumption, the waveform exhibits half-period sign inversion symmetry, which mathematically eliminates certain Fourier components, specifically the DC term, even harmonics, and sine coefficients of odd harmonics. This assumption greatly simplifies the nonlinear equations used to determine the switching angles for the SHE-PWM strategy, enabling effective synthesis of the multi-frequency excitation waveform.

In the full-bridge Class-D power amplifier under the symmetrical conditions, the voltage waveform applied to the load can be described in a compact Fourier series form. The resulting series, presented in Equation (2), provides the mathematical basis for determining the targeted harmonic components and their corresponding switching angles in the SHE-PWM approach:

$$V_N = \sum_{n=1}^{N}\left(b_n \sin\left(\frac{2\pi nt}{T}\right)\right) \tag{2}$$

Under the quarter-wave symmetry assumption, the cosine Fourier coefficients related to the harmonic components of the output voltage can be further derived. These coefficients quantify the amplitudes of the retained harmonics and relate the desired output spectrum to the switching angles of the inverter. The general form of the cosine series coefficients is expressed in Equation (3):

$$b_n = \frac{4V_0}{n\pi}\left(\sum_{i=1}^{N}((-1)^k \cos(n\alpha_i))\right)$$
$$k = \begin{cases} 1 & \textit{Switching angle } a_i \textit{ at rising edge} \\ 0 & \textit{Switching angle } a_i \textit{ at falling edge} \end{cases} \tag{3}$$

Here, n takes only odd integers due to the use of quarter-wave symmetry, which restricts harmonics to odd orders. α_i denotes the switching angles of the PWM waveform, i.e., the time at which the output voltage transitions between levels within a half-wave cycle. V_0 represents the drain voltage of the MOSFET switch, which, in the context of the full-bridge Class-D amplifier represented in **Figure 2**, corresponds to the supply voltage V_{CC}.

In the above Class-D power amplifier system investigated in this study, the excitation coil can be approximated as an ideal inductive load. By this approach, the terminal voltage and current satisfy the constitutive relation of an ideal inductor, which facilitates waveform synthesis. In this case, the instantaneous voltage across the coil is given by:

$$u(t) = L\frac{di(t)}{dt} \tag{4}$$

where *L* denotes the inductance of the excitation coil. Therefore, the instantaneous current through the excitation coil can be modeled as the time integral of the applied pulsed voltage:

$$i(t) = \frac{1}{L}\int u(t)\,dt \tag{5}$$

This model captures the accumulative property of the inductor current, reflecting the increase in magnetic energy of the coil as a function of the applied voltage and duration of the excitation signal. The output voltage of the Class-D power amplifier takes only discrete values, usually $+V_0$ and $-V_0$, forming a series of rectangular pulses in the time domain. Accordingly, the current through the inductive load represents the integral of these voltage pulses. **Figure 4** illustrates the voltage waveform and the corresponding current waveform.

Let $\omega=2\pi/T$ denote the basic angular frequency of the periodic excitation signal. By substituting Equations (2), (3), and (5) into the previous expressions, the instantaneous current in the excitation coil can be represented as a Fourier series expansion, as in Equation (6). The coefficients b_n, representing the amplitude of the harmonic components in the coil current, are described in Equation (7).

$$G(t) = \sum_{n=1}^{N} b_n sin(n\omega t) \tag{6}$$

$$b_n = -\frac{4\pi V_0}{n^2 \pi\omega L}\sum_{n=1}^{N}(-1)^k \sin(n\alpha_i) \tag{7}$$

4.2 Frequency selection of inductor current

In the previous section, the Fourier series representation of the inductor current was derived from the PWM voltage waveform of the inverter. Consequently, various current waveforms can be synthesized by appropriately selecting the cosine series coefficients b_n. In this case, the proper selection of these coefficients is critical for defining the quality and spectral content of the excitation currents.

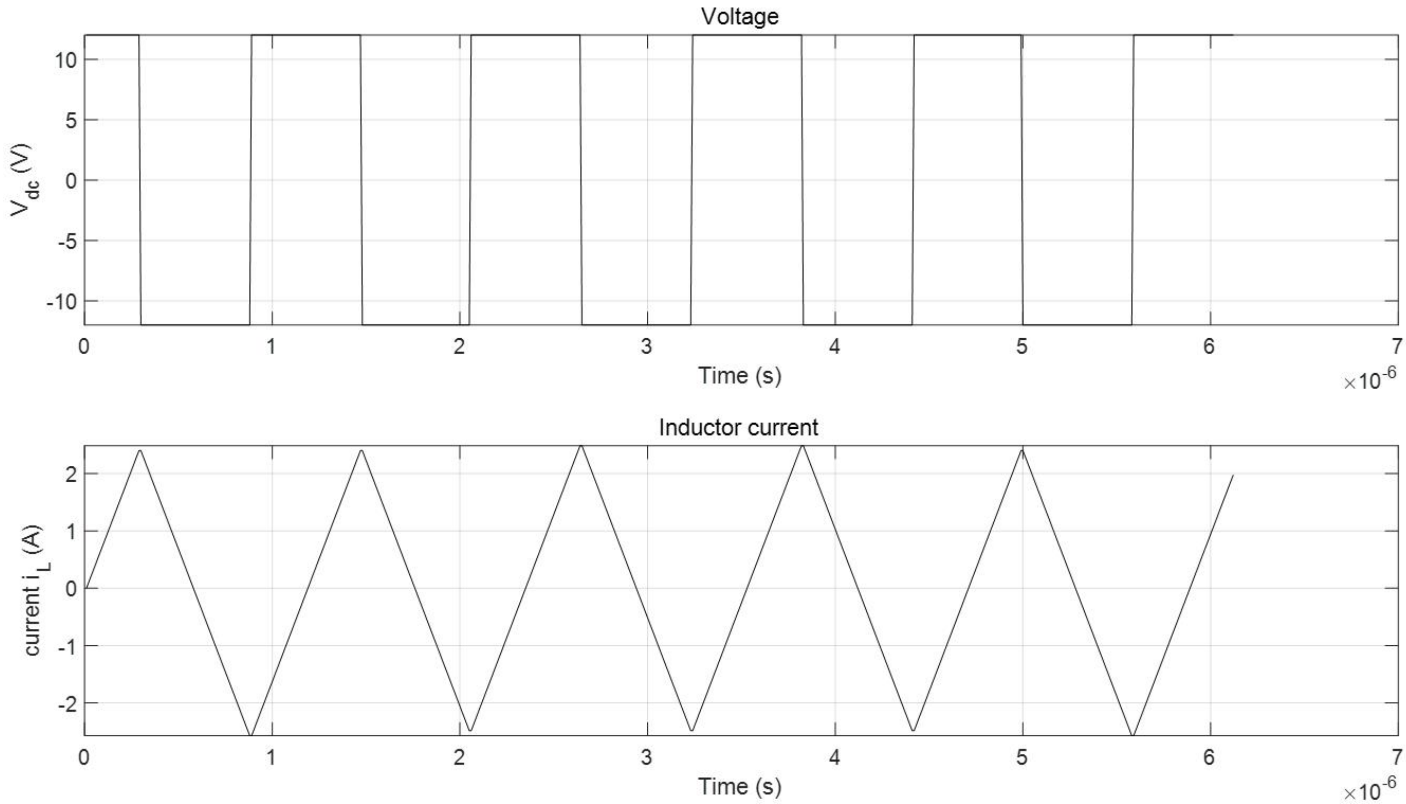


Figure 4. Schematic of inductor voltage and current waveforms.

To obtain a physically realizable and spectrally controlled current shape, the desired shape must first be defined in the time domain. Morphologically, each switching interval should exhibit a high level of linearity to allow an accurate approximation using piecewise linear segments. In addition, to simplify the Fourier series coefficients and maintain a manageable mathematical form, the target waveform is designed with quarter-wave symmetry, consisting of the fundamental and odd harmonics. This design preserves symmetry across each half cycle, eliminates redundant even-order harmonics, while retaining only the harmonic components required for multi-frequency excitation.

In eddy-current-based assistance for localization of metallic foreign bodies within the human body, multi-frequency excitation is commonly employed to obtain complementary electromagnetic response features, so as to accommodate variations in both target materials (ferromagnetic and non-ferromagnetic) and burial depth or orientation. The theoretical basis is that the eddy-current distribution and associated secondary magnetic field are jointly determined by the electrical conductivity σ, magnetic permeability μ, and the excitation angular frequency ω, exhibiting pronounced frequency dependence. According to the skin effect, the characteristic skin depth is given by:

$$\delta = \sqrt{\frac{2}{\omega\mu\sigma}} \tag{8}$$

As the excitation frequency increases, the skin depth decreases: high-frequency excitation concentrates eddy currents near the target surface, thereby enhancing sensitivity to superficial and small-sized foreign bodies; in contrast, low-frequency excitation provides a greater effective electromagnetic penetration depth, thereby improving detectability of deeply embedded targets. Moreover, materials with higher μ and σ (e.g., ferromagnetic metals or high-conductivity metals) tend to exhibit stronger skin effects and more pronounced amplitude/phase variations at the same frequency, facilitating material discrimination and localization.

Based on this principle and leveraging the spectrally controllable synthesis enabled by SHE, this study selects 50 kHz, 150 kHz, 350 kHz, and 850 kHz as the multi-frequency excitation tones. The corresponding current waveform are designed to provide complementary amplitude–phase responses for subsequent signal processing. By using synchronous multi-frequency excitation, an amplitude–phase (I/Q) feature space can be constructed, allowing joint evaluation of responses across frequencies. This approach improves robustness against tissue-background effects, variations in electromagnetic coupling, and motion-induced amplitude/phase perturbations, thereby enhancing both the detectability and localization resolution for tiny metallic foreign bodies in the human body. The mathematical expression of the synthesized inductor current can is given

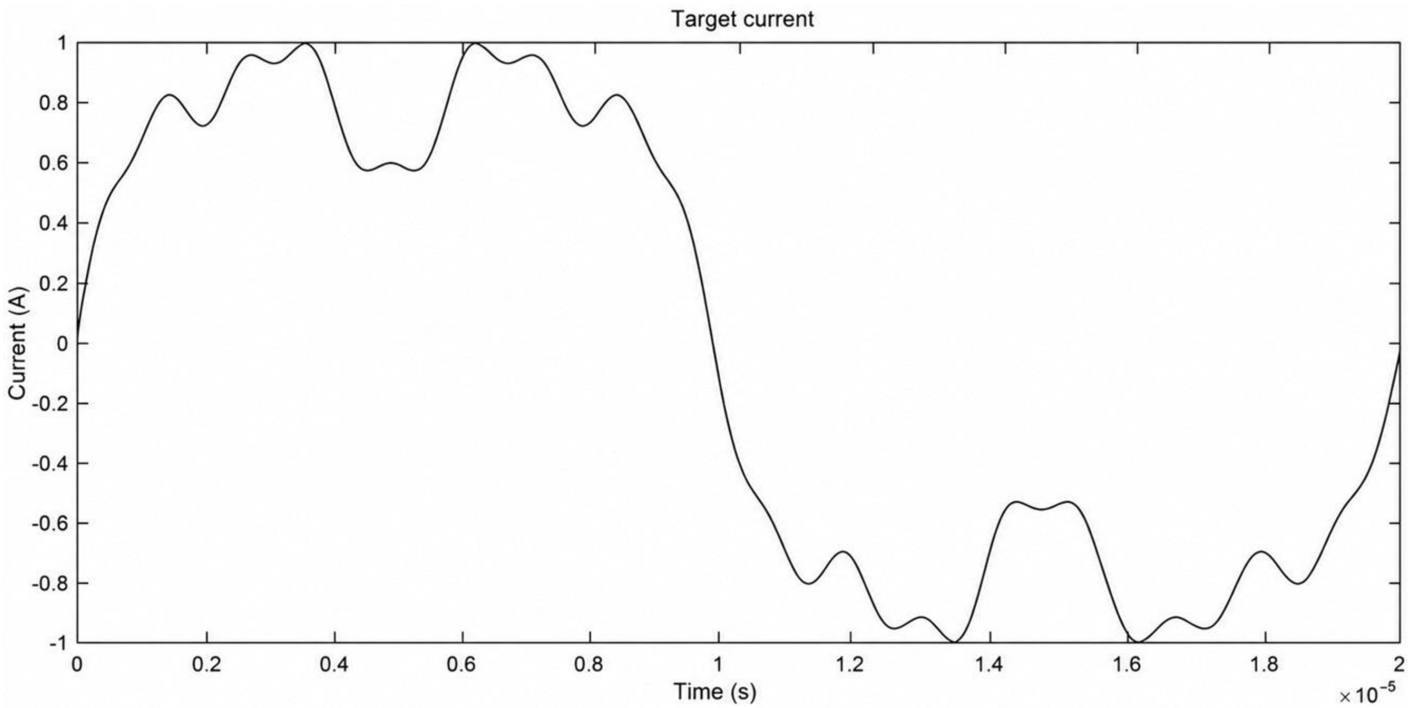


Figure 5. Target inductor current waveform for multi-frequency excitation.

by Equation (9), and the corresponding time-domain current waveforms are illustrated in **Figure 5**.

$$f(t) = \sin(\omega t) + \frac{1}{3}\sin(3\omega t) + \frac{1}{7}\sin(7\omega t) + \frac{1}{17}\sin(17\omega t) \quad (9)$$
$$\omega = 2\pi * 50 * 10^3$$

4.3 Optimization of the spectral coefficients of inductor current

The primary objective of the SHE-PWM is to modulate the low-order harmonics of the output waveform through optimized switching angles α_i. This approach allows elimination of unwanted harmonics from high-power converters while preserving the required spectral characteristics of the excitation signal. A major challenge in implementing SHE-PWM lies in solving a system of nonlinear transcendental equations that map the switching angles α_i to the amplitudes of the harmonics. The complexity of this system depends on the waveform characteristics and the number of discrete voltage levels involved.

Several methods have been proposed to address this challenge, including numerical solutions, resultant theory methods, and optimization-based techniques. Numerical methods usually face difficulties due to the nonlinear expressions inherent in standard SHE-PWM models, especially when high-order harmonics or two-level inverter topologies are used. Resultant theory approaches encounter increasing computational complexity as the polynomial order rises with the number of harmonics to be eliminated.

With advances in computational power, optimization-based strategies have become the preferred approach. In this framework, the SHE-PWM equations are reformulated as an optimization problem subject to constraints. The objective function is defined as the residual magnitude of the undesired harmonics, with the goal of minimizing this value. Constraints enforce the elimination of targeted harmonics and the symmetry of the waveform.

In this context, the switching angles α_i can be determined via an optimization algorithm. The procedure begins by identifying the local extrema of the target current waveform. These extrema serve as initial estimates for the switching angles, providing a valid starting point for subsequent iterations. In system employing a single H-bridge configuration, the slope of the inductor current remains constant in magnitude during each switching interval, allowing the previous target current waveform to be approximated by line segments with equal slope magnitude.

Figure 6 illustrates the optimization procedure for one-quarter of the waveform duration ($\theta=\pi/2\approx1.57$). The green curve represents the desired target current in the first quarter of the waveform, while the red polylines show piecewise-linear approximation used in the optimization. The remaining three quarters of the waveform are derived through symmetry. The initial switching angles derived from this approximation method are provided in **Table 2**.

However, when a suitable initial guess for the switching angles exists, gradient-based algorithms can converge substantially

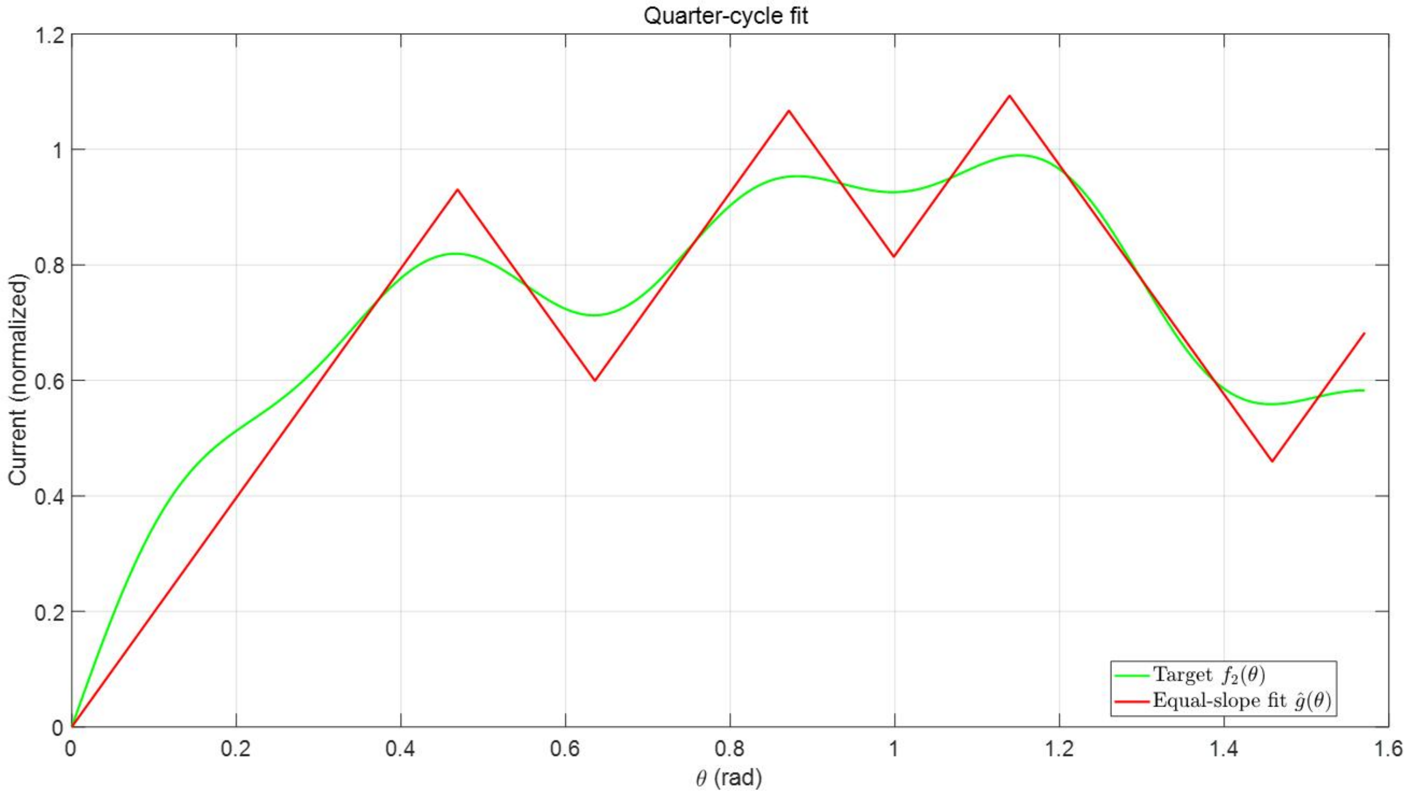


Figure 6. First-quarter current waveform fitting for SHE-PWM optimization.

Table 2. Initial switching angle

Switching angle	α_1	α_2	α_3	α_4	α_5	α_6
θ (rad)	0.47	0.64	0.87	1.00	1.14	1.46
Time (us)	1.48	2.02	2.80	3.18	3.66	4.64

faster than stochastic or population-based algorithms [18]. In this study, a gradient-based algorithm was used to determine the amplitudes of the harmonic components, as formulated in Equation (10):

$$\min_{0\le\alpha_1\le\alpha_2\le\alpha_3\le\alpha_4\le\alpha_5\le\alpha_6\le\frac{\pi}{2},m_p} \sum_{p\in\phi} \mu_p\left(\frac{p\omega L}{V_0}b_p - m_p\right)^2 \quad \text{subject to} \left|\frac{p\omega L}{V_0}b_p\right| \le \varepsilon_p \quad p\in\phi \tag{10}$$

Here, the modulation index is defined as:

$$m_p = \frac{V_p}{V_0} = \frac{p\omega L}{V_0}b_p \tag{11}$$

where V_p represents the amplitude of the p^{th} harmonic of the output voltage, and V_0 denotes the DC bus voltage of the Class-D amplifier. The weighting factor μ_p scales the contribution of each harmonic to the objective function. Φ indicates the set of desired harmonics to be accurately synthesized, while Ψ represents the set of undesired harmonics to be reduced. The parameter ε_p defines the permissible level of each non-target harmonic, acting as a soft constraint to prevent excessive distortion.

Minimizing the residual error between the calculated Fourier coefficient b_p and the modulation index m_p, while respecting the constraints on non-target harmonics, enables iterative refinement of the switching angles to accurately reproduce the desired harmonics and suppress undesired components.

Using the initial switching angles derived from the piecewise-linear approximation of the target current waveform g(θ) and the coefficient definition from Equation (7), the harmonic components can be analytically derived by Equation (12):

$$b_p = -\frac{4V_0}{p^2\pi\omega L}\left(2\sum_{q=1}^{6}(-1)^q\sin(p\alpha_q) - \sin\left(\frac{p\pi}{2}\right)\right) \tag{12}$$

At each switching angle α_i, two successive voltage transitions occur: from $+V_0$ to 0 and then from 0 to $-V_0$, or vice versa depending on the switch polarity, which must be accounted for in the Fourier series computation.

When θ=π/2, only a single downward transition ($+V_0$ to 0) occurs, corresponding to the end of quarter cycle. Substituting

Table 3. Initial values and optimized values

	ε_5	ε_9	ε_{11}	ε_{13}	ε_{15}	Switching angle (rad)						m_p			
						α_1	α_2	α_3	α_4	α_5	α_6	50 k	150 k	350 k	850 k
Initial	0.10	0.18	0.22	0.26	0.30	0.47	0.64	0.87	1.00	1.14	1.46	0.53	0.43	0.46	0.74
Optimize	0.10	0.18	0.22	0.26	0.30	0.46	0.63	0.86	0.97	1.12	1.47	0.49	0.49	0.49	0.67

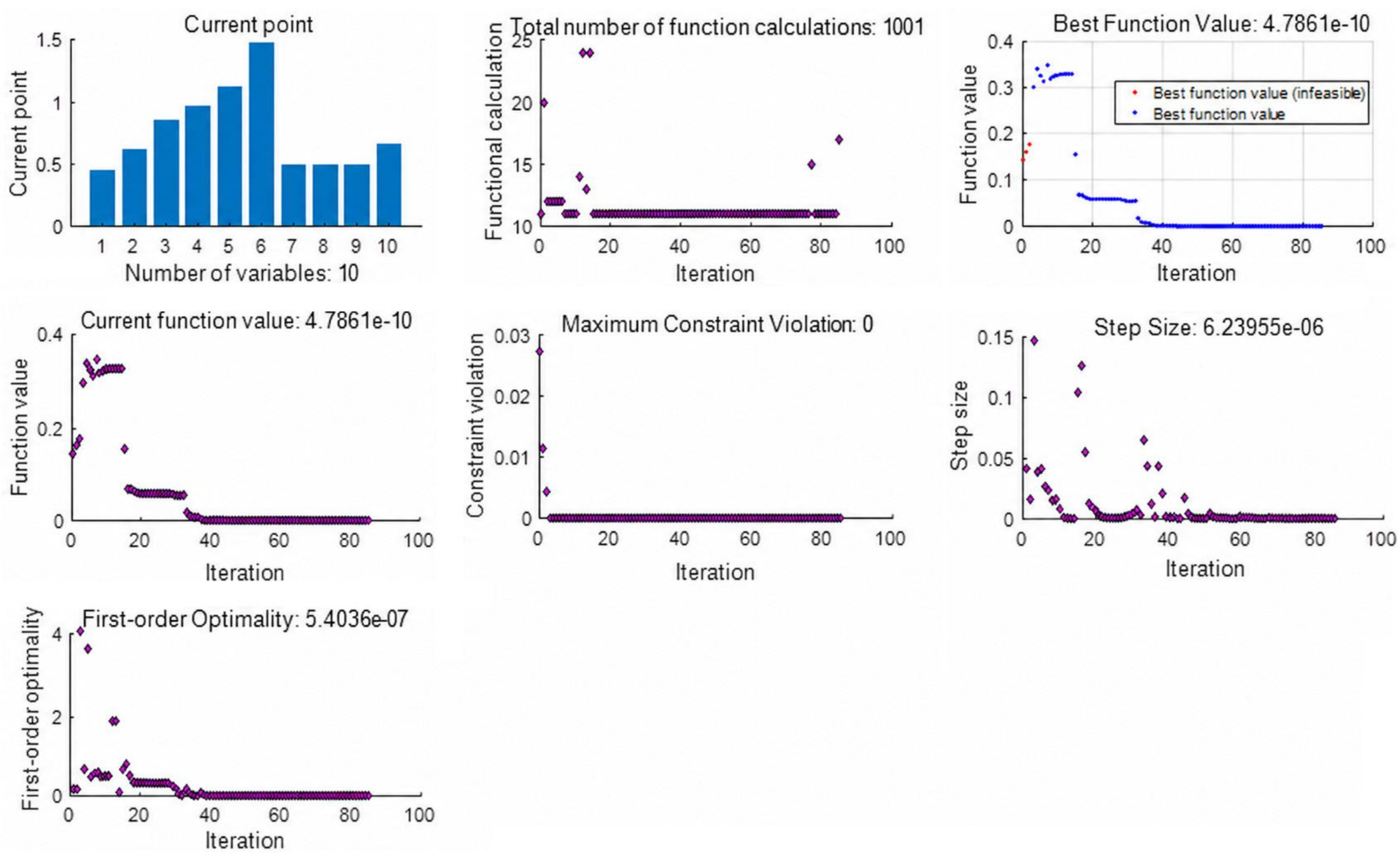


Figure 7. Optimization of switching angles and harmonic residuals using the *fmincon* function.

Equation (12) into the residual minimization problem of Equation (10) yields Equation (13), which explicitly relates the residual harmonic amplitudes to the switching angles, enabling selective synthesis of harmonics in the Class-D power amplifier system.

$$\min_{0\le\alpha_1\le\cdots\le\alpha_6\le\frac{\pi}{2},m_p} \sum_{p=1,3,7,17} \mu_p\left(-\frac{8}{p\pi}\sum_{q=1}^{6}(-1)^q\sin(p\alpha_q)-\frac{4}{p\pi}\sin\left(\frac{p\pi}{2}\right)+m_p\right)^2$$
$$\text{subject to}\left|-\frac{8}{p\pi}\sum_{q=1}^{6}(-1)^q\sin(p\alpha_q)-\frac{4}{p\pi}\sin\left(\frac{p\pi}{2}\right)\right|\le\varepsilon_p \quad (13)$$
$$p\in\{5,9,11,13,15\}$$

The initial parameters for the optimization algorithm are shown in the first column of **Table 3**. In this case, the weighting factors μ_p were set to unity (μ_p=1). Once the initial switching angles and weights are determined, the constrained optimization problem is solved using the *fmincon* function in MATLAB. The objective function quantifies the deviation between predicted and target harmonic magnitudes, while the constraints enforce the permissible bounds for non-target harmonics.

The *fmincon* function (constrained function minimization) is the primary procedure within the MATLAB Optimization Toolbox used to solve constrained continuous optimization problems where the objective and constraint functions are continuously differentiable. This function can handle continuous decision-variable problems subject to linear and/or nonlinear constraints. **Figure 7** illustrates the optimization results for the switching angles as formulated in Equation (12).

The Maximum Constraint Violation metric decreases drastically from approximately 3×10^{-2} in the early iterations to nearly zero by the 10–15th iteration. This trend indicates that the barrier term and line-search procedures within the interior-point method effectively guide the iterates from infeasible to feasible regions. Once feasibility is achieved, the iterates remain strictly feasible throughout the rest of the optimization procedure, demonstrating that all constraints are satisfied.

Simultaneously, the Best Function Value decreases steadily from approximately 3.5×10^{-1} to the order of 10^{-4} over 30-40 iterations, eventually stabilizing at a minimal of 4.7861×10^{-10}. The convergence behavior can be understood by examining the plots of feasible and infeasible solutions. In the initial iterations, several infeasible points (red stars in **Figure 7**) approach the feasible region but remain constrained by the limits. However, the best feasible objective (blue stars) steadily improves until it converges to the best infeasible objective at the global minima point. This observation clarifies a common misconception that the best feasible point must be optimized before the infeasible solution. Ultimately, the evolution of the Best Function Value reflects the transition from satisfying the feasibility constraints to the final optimization of the switching angles within the feasible region, thereby minimizing the residuals of the harmonic components.

The Step Size parameter displays several backtracking and enlargement pulses during the first TEN iterations. After this phase, the step size stabilizes at approximately 10^{-4}, indicating the onset of the local refinement phase of the optimization. This pattern of behavior is consistent with the Interior-Point Algorithm's barrier parameter update and centralization mechanism, wherein the solver fine-tunes the solution in the region of the optimal point once feasibility has been established.

The First-Order Optimality value reduces from an initial order of $O(1)$ to approximately 5.4×10^{-7}, remaining below the default optimality tolerance of approximately 10^{-6} used in the *fmincon* solver. Along with the step size reduction and stabilization of the objective function, these results indicate that the algorithm has approached a stationary point satisfying the Karush-Kuhn-Tucker (KKT) conditions, where the first-order necessary conditions for constrained optimality are approximately satisfied.

The number of function evaluations was 1,001, requiring fewer than 100 iterations. In the case of a ten-dimensional optimization problem with six switching angles and four modulation indices involving nonlinear equality and inequality constraints, this moderate computational effort reflects good differentiability of the objective and constraint functions along with minimal backtracking during the line search.

The current point bar chart shows that the six optimized switching angles converge to the range of 0.5–1.5 radians, remaining strictly within their prescribed bounds (0 to $\pi/2$). This implies that the solution is largely driven by harmonic constraints rather than boundary saturation. Similarly, the four modulation indices converge to the range of 0.5–0.8, well within he physically acceptable limits. Together with the observation that the Maximum Constraint Violation remains at zero, this confirms feasibility and realizability of the solution.

In general, the optimization algorithm exhibits the conventional behavior of an interior-point method: (1) rapid approach toward feasibility, (2) orderly and monotonic decrease of the objective function, (3) gradual reduction of the step size accompanied by improved first-order optimality, and (4) attainment of a strictly feasible solution satisfying the KKT conditions approximated to the desired precision. The final objective function value of 4.7861×10^{-10} suggests that the residual harmonics are minimized to numerical precision, confirming accurate representation of both target and non-target harmonics within the prescribed thresholds.

From an engineering perspective, the optimized switching angles and modulation indices jointly achieve the desired spectral characteristics without boundary saturation, ensuring robust tolerance to timing errors and dead-time quantization. **Figure 8** compares the voltage-current waveforms before and after optimization. The reduction in harmonic distortion validates the effectiveness of the proposed method in generating high-purity multi-frequency excitation currents suitable for high-sensitivity localization of tiny metallic foreign bodies in the human body.

5 HARDWARE CIRCUIT DESIGN AND SIMULATION

5.1 Device selection

As illustrated in **Figure 8**, the output voltage waveform exhibits a maximum switching frequency of approximately 3–4 MHz. Accordingly, the gate driver and MOSFET devices must operate reliably at such high switching speeds. In this study, the isolated gate driver IC 2EDF7275K was employed to drive the power stage. This gate driver provides galvanic isolation between the power and control circuits, mitigating the propagation of high-current transients from the power stage to sensitive control inputs, such as those of the FPGA, which could otherwise lead to permanent damage. The 2EDF7275K IC supports switching frequencies of up to 10 MHz, easily accommodating the PWM switching pulses described previously. Moreover, it allows flexible gate-level adjustments, facilitating circuit tuning and debugging.

For the power switching components, BSC057N08NS3 MOSFETs were chosen. These devices have a drain-source voltage of 80 V, a typical on-resistance $R_{DS}(on)$ of 5.7 mΩ, along with fast transition times ($t_r\approx14$ ns, $t_f\approx9$ ns), minimizing transition losses and reducing waveform distortion during high-frequency Class-D operation. According to the Safe Operating Area (SOA) diagram shown in **Figure 9**, the MOSFETs can be safely operated under the chosen conditions of 12 V input and a 2 A drain current. These specifications provide sufficient margin to ensure reliable performance during continuous high-frequency switching, enabling stable excitation currents in the coil necessary for high-sensitivity eddy-current detection of tiny metallic foreign bodies.

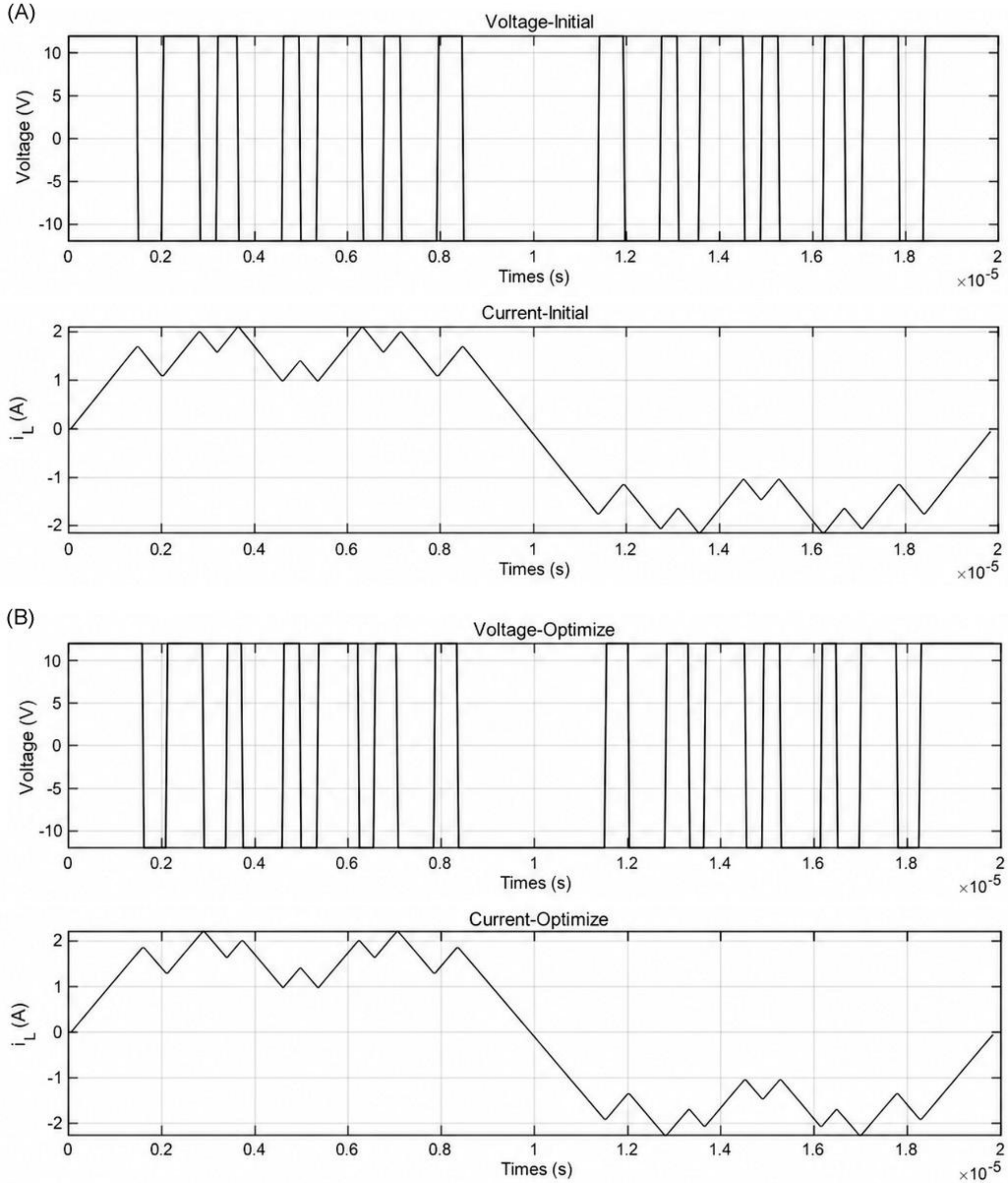


Figure 8. Initial and optimized voltage-current waveforms of the excitation coil. (A) Initial voltage-current waveforms prior to SHE-PWM optimization; (B) Voltage-current waveforms after optimization, showing improved harmonic purity and adherence to the target multi-frequency excitation profile.

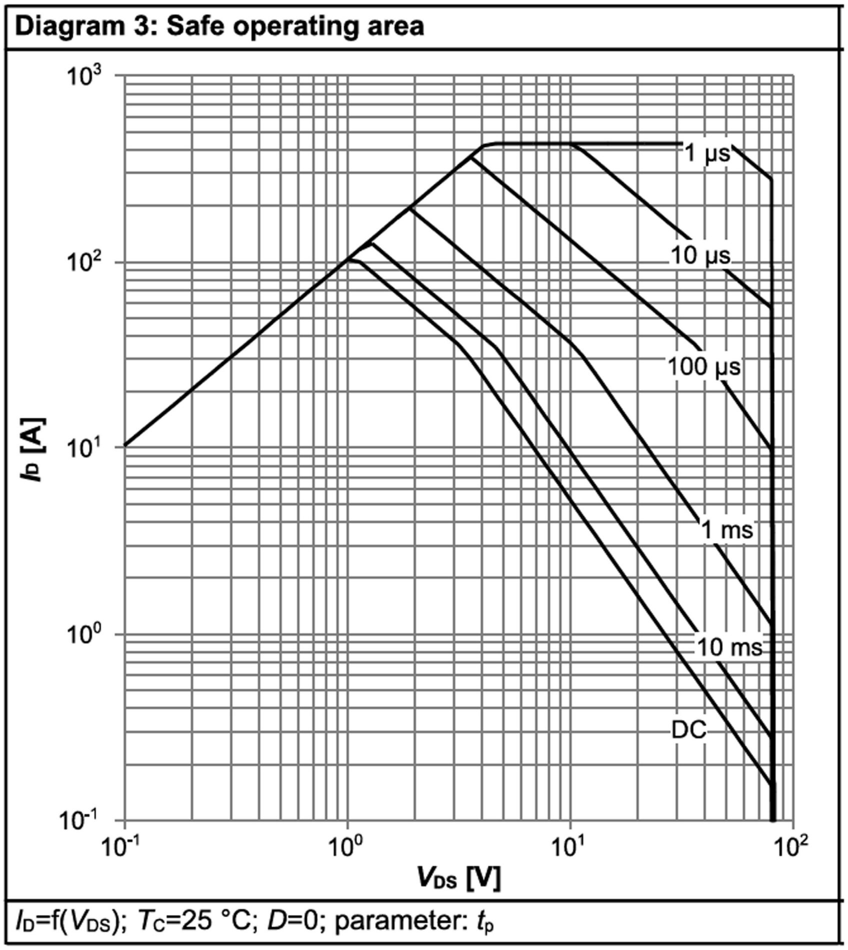


Figure 9. Safe operating area.

5.2 Hardware circuit simulation

Figure 10 shows the voltage and current waveforms and their corresponding spectra, obtained through PSpice simulations. PSpice is a widely used general-purpose circuit simulator in the field of electronic design automation (EDA), mainly used to evaluate and develop analog and mixed-signal (analog-digital) electronic circuits. It allows designers to predict circuit performance before physical implementation.

Simulation outcomes indicate slight oscillations in the output voltage waveform during switching transitions. These slight oscillations, however, do not affect circuit functionality. In the time-domain current waveform of the excitation coil, the inductor current closely follows the target waveform, confirming the correct operation of the Class-D power amplifier and the effectiveness of the SHE-PWM control strategy.

Frequency-domain analysis further confirms these observations. Distinct spectral peaks appear at 50 kHz, 150 kHz, 350 kHz, and 850 kHz, accurately matching the four predefined excitation frequencies. Simultaneously, the amplitudes of non-target harmonics remain below the specified thresholds, indicating that the optimized switching angles and modulation indices successfully achieve precise harmonic synthesis. These results confirm that the excitation current exhibits a well-defined multi-frequency pattern with high spectral purity, which is essential for robust and accurate eddy-current detection of tiny metallic foreign bodies in the human body.

5.3 Hardware circuit design

The power supply module of the hardware platform is designed to provide multiple regulated voltage rails to satisfy the requirements of various circuit sub-components. As illustrated in **Figure 11**, the initial 12 V DC input is first stepped down to 8 V to provide the gate-drive voltage for the MOSFETs in the H-bridge power stage. To ensure galvanic isolation between the control and power planes, the 8 V line follows is further stepped down using an isolated DC-DC converter to provide 5 V for the gate-driver ICs. Subsequently, a linear voltage regulator generates the 3.3–5 V control-side voltages needed by the driver circuitry.

Figure 12 illustrates the schematic diagram of the H-bridge module. For clarity, only half of the bridge is illustrated, as the other half is symmetrical. The COL+ node connects to the first terminal of the excitation coil, while the second terminal is connected to COL−. The input network pins IN+ and IN− (pins 3 and 2 of the second 2EDF7275K isolated gate-driver IC) are configured to guarantee complementary switching between the two half-bridges and to maintain the desired differential phase relationship of the output voltage.

In the gate-drive circuitry, B340A Schottky Rectangular Diodes (D4 & D5) are paralleled with resistors R9 & R13. This configuration facilitates rapid turn-off transitions of the MOSFETs, as the current is directed through the low-impedance path of the B340A Rectangular Diodes, reducing gate-charge removal time [19-21]. During turn-on, the gate charge current flows from the driver's output (OUTA and/or OUTB) to the MOSFET gate. In this phase, the B340A diodes are reverse-biased and non-conductive, and the resistive network controls the gate charging rate, resulting in a "slow turn-on, fast turn-off" behavior. This asymmetrical switching mitigates parasitic inductance-induced voltage overshoot and ringing while maintaining efficient switching transitions and minimizing electromagnetic interference in the H-bridge circuit.

Figure 13 illustrates the experimentally recorded H-bridge output voltage. In this diagram, the yellow curve corresponds to the voltage at the COL+ terminal of the excitation coil,

Figure 10. Voltage, current, and current spectrum of the excitation coil from hardware circuit simulation. (A) Time-domain voltage waveform of the H-bridge output; (B) Time-domain current waveform through the excitation coil; (C) Frequency-domain spectrum of the coil current showing multi-frequency excitation peaks at 50 kHz, 150 kHz, 350 kHz, and 850 kHz.

whereas the pink curve corresponds to the COL− terminal. A clear voltage sag appears across the coil during specific intervals, attributed to the charge-discharge behavior of the bootstrap capacitor (C24) used in the gate-drive circuit of the high-side MOSFETs. When the low-side MOSFET conducts for a short duration, the bootstrap capacitor is not fully charged. Consequently, during the subsequent long conduction interval of the high-side MOSFET, the capacitor discharges over an

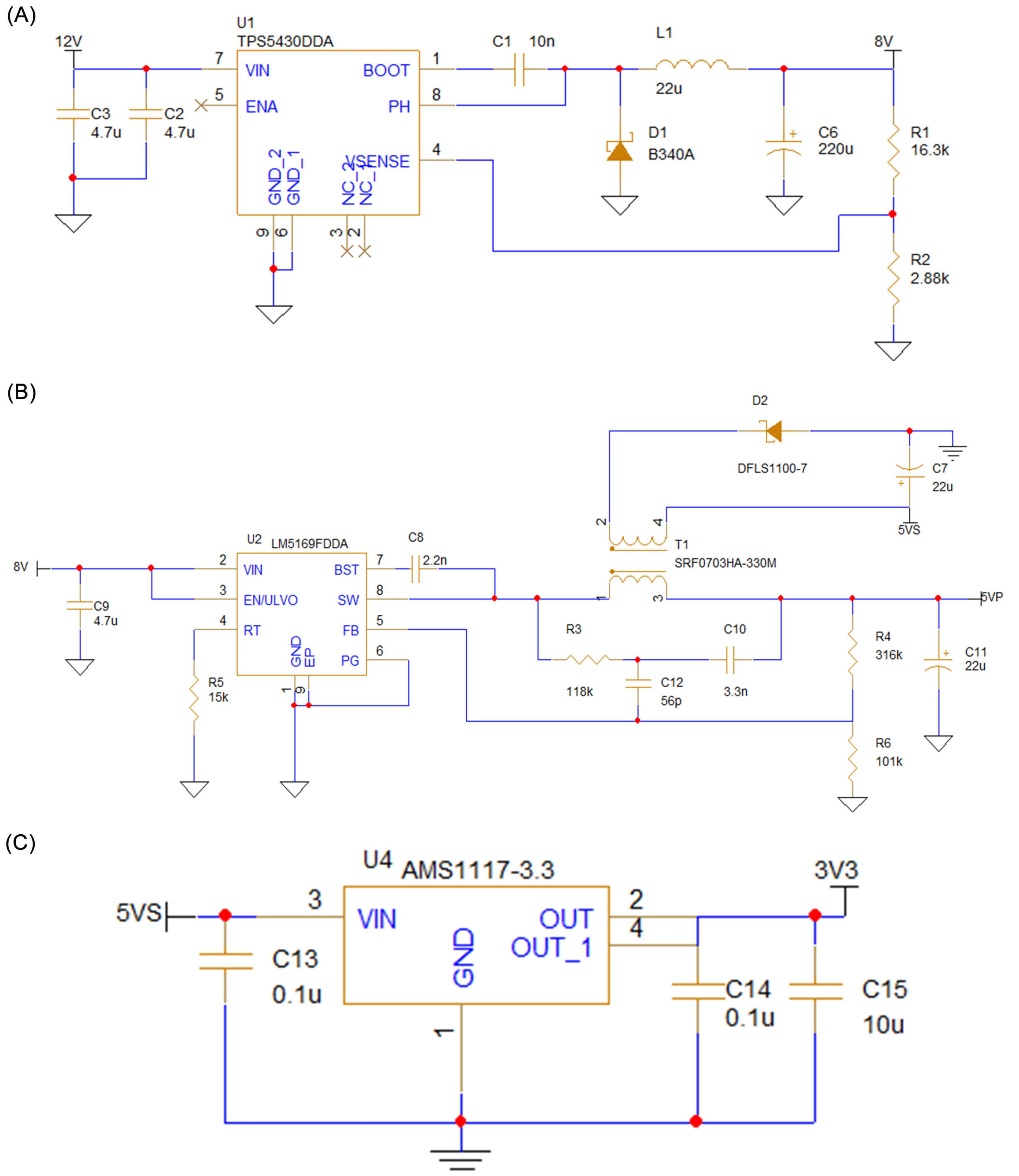


Figure 11. Power supply module of the hardware platform. (A) Step-down conversion from 12 V to 8 V; (B) Isolated DC-DC conversion from 8 V to 5 V; (C) Linear regulation from 5 V to 3.3 V.

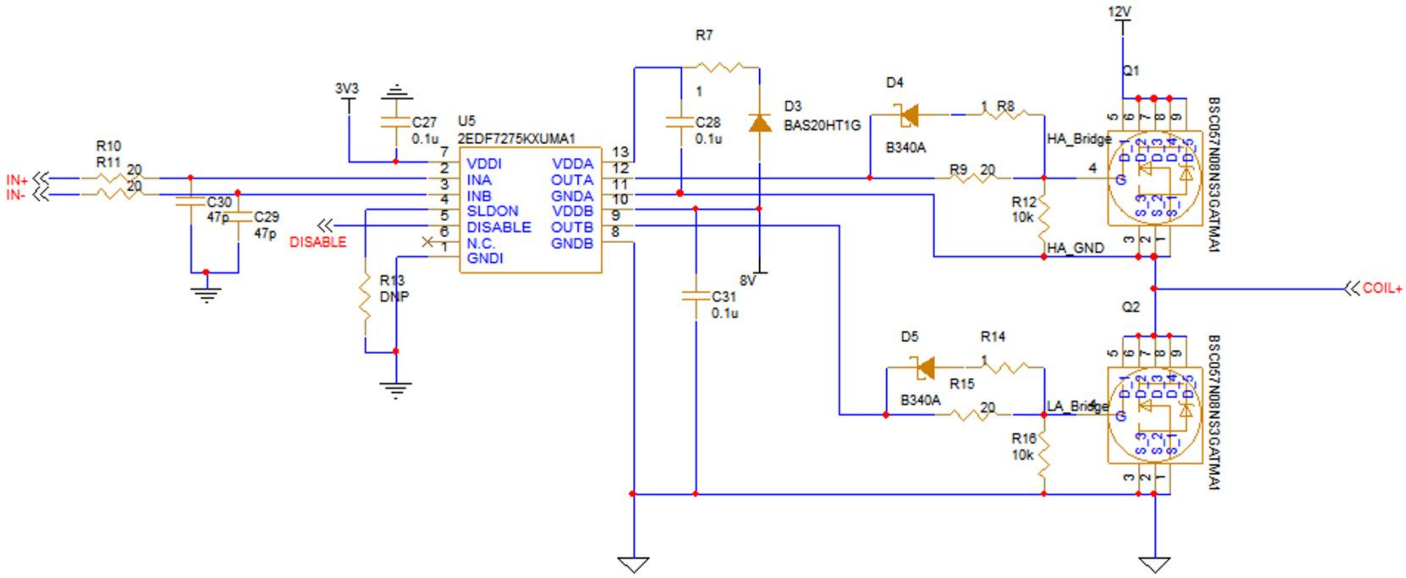


Figure 12. Schematic diagram of the H-bridge (half).

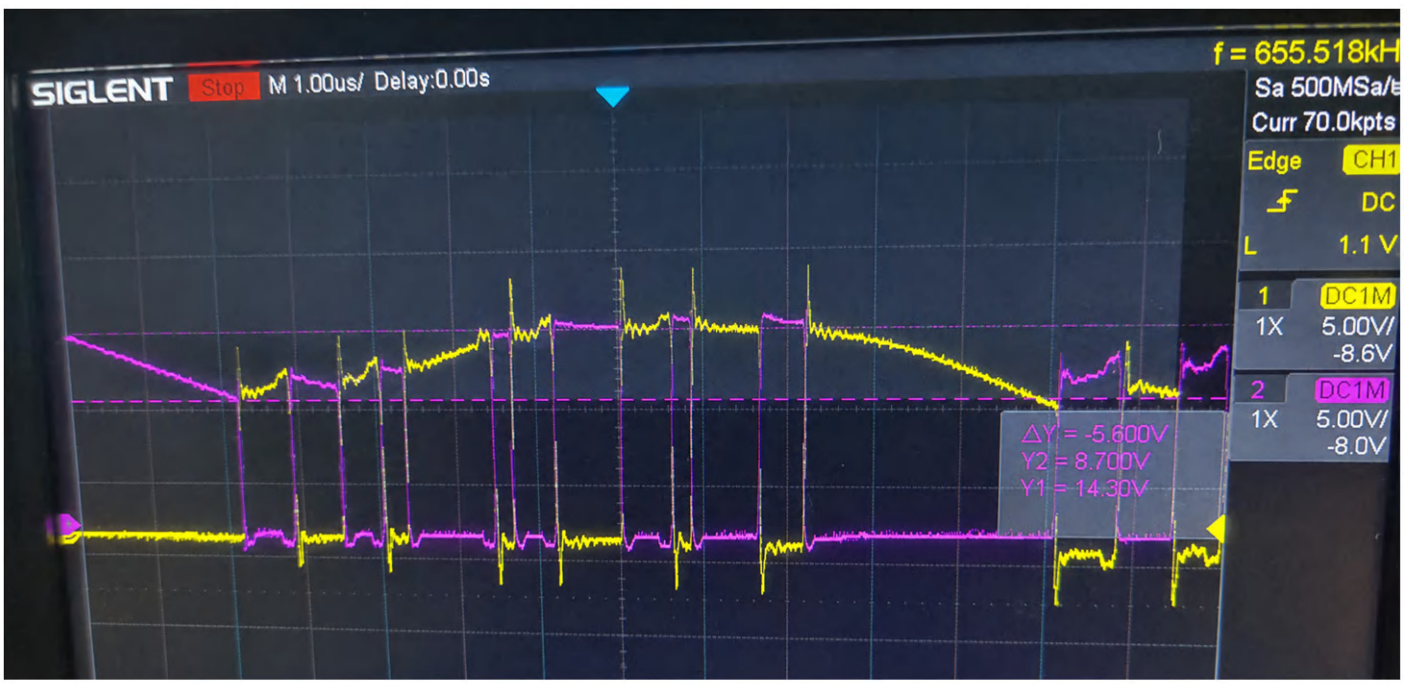


Figure 13. Experimentally recorded H-bridge output voltage.

extended period, resulting in reduced gate-drive voltage and the observed voltage sag across the coil.

6 CONCLUSION

This study addresses the clinical challenge of assisted screening and localization of tiny metallic foreign bodies in the human body (e.g., suture needles, hemostatic clips, and shrapnel). Leveraging the SHE approach, we developed a transmitter driving scheme capable of generating synchronous multi-frequency excitation currents, providing an energy-efficient and spectrally controllable excitation-circuit foundation for eddy-current-based detection systems.

By optimizing the switching angles, the transmitter achieves precise current components at the designated target frequencies while effectively suppressing critical non-target harmonics. This, in turn, mitigates unwanted background responses, baseline drift, and inter-frequency crosstalk in the tissue–coil coupling system, thereby improving the stability and repeatability of multi-frequency excitation. Furthermore, circuit-level simulations and hardware implementation validate the engineering feasibility of the proposed scheme. These results provide a robust hardware foundation for future development of clinical-oriented multi-frequency eddy-current detection techniques, enabling feature-based discrimination and accurate localization of tiny metallic foreign bodies.

DECLARATIONS

Author contributions

Yuming Liu performed data acquisition, processing, analysis and model building and wrote the manuscript. Piding Li made substantial contributions to data analysis and manuscript preparation.

Funding

This research received no external funding.

Data availability

This study utilized only publicly available datasets.

Ethics approval and consent to participate

Not applicable.

Consent for publication

We agree to the publication of this manuscript in both print and electronic formats in ZENTIME, and grant the publisher permission to make necessary editorial and formatting revisions to the manuscript.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.

REFERENCES

[1] Bozorgmehr R, Bahadorinia M, Pouyanfar S, Ahmadinejad M, Bahri MH, Bagherpour JZ. A rare case of abdominal foreign bodies; laparoscopic removal of a sewing needle. Ann Med Surg (Lond). 2022 Oct;82:104747. https://doi.org/10.1016/j.amsu.2022.104747

[2] Hall Barbosa C. Localization of firearm projectiles in the human body using a superconducting quantum interference device magnetometer: A theoretical study. Rev Sci Instrum. 2004 Jun 1;75(6):2098-2106. https://doi.org/10.1063/1.1753679

[3] AbdAlla AN, Faraj MA, Samsuri F, Rifai D, Ali K, Al-Douri Y. Challenges in improving the performance of eddy current testing: Review. Meas Control. 2019 Jan;52(1-2):46-64. https://doi.org/10.1177/0020294018801382

[4] Grochowalski JM, Chady T. Rapid identification of material defects based on pulsed multifrequency eddy current testing and the k-nearest neighbor method. Materials. 2023 Oct 11;16(20):6650. https://doi.org/10.3390/ma16206650

[5] Svatos J, Vedral J, Pospisil T. Advanced instrumentation for polyharmonic metal detectors. IEEE Trans Magn. 2016 May;52(5):1-4. https://doi.org/10.1109/TMAG.2015.2507780

[6] Yamazaki S, Nakane H, Tanaka A. Basic analysis of a metal detector. IEEE Trans Instrum Meas. 2002 Aug;51(4):810-814. https://doi.org/10.1109/TIM.2002.803397

[7] Hu W, Liu Z. Study of metal detection based on chaotic theory. In: 2010 8th World Congress on Intelligent Control and Automation; 2010 Jul 7-9; Jinan, China. IEEE; 2010. p 2309-2314. https://doi.org/10.1109/WCICA.2010.5554170

[8] García-Martín J, Gómez-Gil J, Vázquez-Sánchez E. Non-destructive techniques based on eddy current testing. Sensors. 2011;11(3):2525-2565. https://doi.org/10.3390/s110302525

[9] Bernieri A, Betta G, Ferrigno L, Laracca M. Crack depth estimation by using a multi-frequency ECT method. IEEE Trans Instrum Meas. 2023 Mar;62(3):544-552. https://doi.org/10.1109/TIM.2012.2232471

[10] Machado MA. Eddy currents probe design for NDT applications: A review. Sensors. 2024 Sep 7;24(17):5819. https://doi.org/10.3390/s24175819

[11] Sophian A, Tian G, Fan M. Pulsed eddy current non-destructive testing and evaluation: A review. Chin J Mech Eng. 2017 May;30:500-514. https://doi.org/10.1007/s10033-017-0122-4

[12] Blitz J, Peat TS. The application of multi-frequency eddy currents to testing ferromagnetic metals. NDT E Int. 1981 Feb;4(1):15-17. https://doi.org/10.1016/0308-9126(81)90058-4

[13] Antoun CA, Würsch C, Köchli C, Perriard Y. Balanced metal detector based on optimized frequencies and spatial phase profile responses to differentiate metal rods. IEEE Magn Lett. 2017;8:1-5. https://doi.org/10.1109/LMAG.2017.2674619

[14] Zhang B, Yin G, Zhang Y, Lin J. Research on highly sensitive electromagnetic method based on balance coil for underground metal detection. IOP Conf Ser: Earth Environ Sci. 2021;660(1):012004. https://doi.org/10.1088/1755-1315/660/1/012004

[15] Duraij M, Xiao Y, Zsurzsan G, Zhang Z. A comparative study on Class AB and Class D amplifier topologies for high temperature power line communication circuits. In: 2021 IEEE International Symposium on Power Line Communications and its Applications (ISPLC); 2021 Oct 26-27; Aachen, Germany. IEEE; 2021. p. 102-107. https://doi.org/10.1109/ISPLC52837.2021.9628537

[16] Mei S, Hu Y, Xu H, Wen H. The class D audio power amplifier: A review. Electronics. 2022 Oct 9;11(19):3244. https://doi.org/10.3390/electronics11193244

[17] Dahidah MSA, Konstantinou G, Agelidis VG. A review of multilevel selective harmonic elimination PWM: Formulations, solving algorithms, implementation and applications. IEEE Trans Power Electron. 2015 Aug;30(8):4091-4106. https://doi.org/10.1109/TPEL.2014.2355226

[18] Boyd S, Vandenberghe L. Convex optimization. Cambridge (UK): Cambridge University Press; 2004. 727 p. https://doi.org/10.1017/CBO9780511804441

[19] Rujas A, López VM, Mir L, Nieva T. Gate driver for high power SiC modules: Design considerations, development and experimental validation. IET Power Electron. 2018;11(6):977-983. https://doi.org/10.1049/iet-pel.2017.0535

[20] Jiang Y, Feng C, Yang Z, Zhao X, Li H. A new active gate driver for MOSFET to suppress turn-off spike and oscillation. Chin J Electr Eng. 2018 Jun;4(2):43-48. https://doi.org/10.23919/CJEE.2018.8409349

[21] Aboadla EH, Khan S, Abdul Kadir K, Md Yusof Z, Habaebi MH, Habib S, et al. Suppressing voltage spikes of MOSFET in H-bridge inverter circuit. Electronics. 2021 Feb 5;10(4):390. https://doi.org/10.3390/electronics10040390

Progress in
Medical Devices

Zentime

REVIEW ARTICLE

Deep learning for prostate intervention: Recent advances in non-rigid magnetic resonance imaging–transrectal ultrasound image registration

Peiyu Chen, Xudong Guo

School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.

Corresponding author: Xudong Guo.

Address correspondence to: Xudong Guo, School of Health Science and Engineering, University of Shanghai for Science and Technology, No. 516 Jungong Road, Yangpu District, Shanghai 200093, China.
E-mail: guoxd@usst.edu.cn.

Received December 29, 2025; Accepted March 6, 2026; Published June 26, 2026

DOI: 10.61189/692164snwggk

Abstract

The treatment of prostate cancer (PCa) is shifting towards the use of highly accurate image-guided procedures in order to achieve better oncologic results. A common strategy consists of using both pre-operative multiparametric magnetic resonance imaging and intra-operative transrectal ultrasound during a prostate biopsy or focal ablation procedure, thus offering enhanced localisation information through high spatial resolution and dynamic response, respectively. However, reliable non-rigid registration is still technically challenging owing to differences in cross-modal imaging physics as well as large deformations between the two modalities caused by rectal probe compression; this paper reviews how deep learning has evolved, focusing on convolutional neural networks, generative models, including generative adversarial networks and diffusion models, and transformer-based architectures. We discuss the extent to which they utilise biomechanical priors to inform the solution of registration problems, against standardized challenges such as μ-RegPro. State-of-the-art approaches achieve sub-millimetre target registration errors with real-time inference times for intra-operative deployment. Addressing outstanding challenges related to interpretation and generalization, this review provides an outlook of the road map to develop "physics-aware" smart interventional systems. All these developments represent important steps toward a fully automated, precise, and minimally invasive PCa management pipeline.

Keywords: Deep learning, Image registration, Non-rigid, Transrectal ultrasound, Magnetic resonance imaging

Highlights

- This review systematically reviews the evolution of deep learning-based non-rigid prostate magnetic resonance imaging–transrectal ultrasound registration.
- This review analyzes dominant paradigms: hybrid convolutional neural networks, generative adversarial networks/diffusion models, and transformers.
- This review explores integrating anatomical priors and physical constraints to address label scarcity.
- This review critically evaluates the generalization gap between state-of-the-art benchmarks and clinical workflows.
- This review proposes future directions in physics-aware artificial intelligence and intelligent robotic interventions.

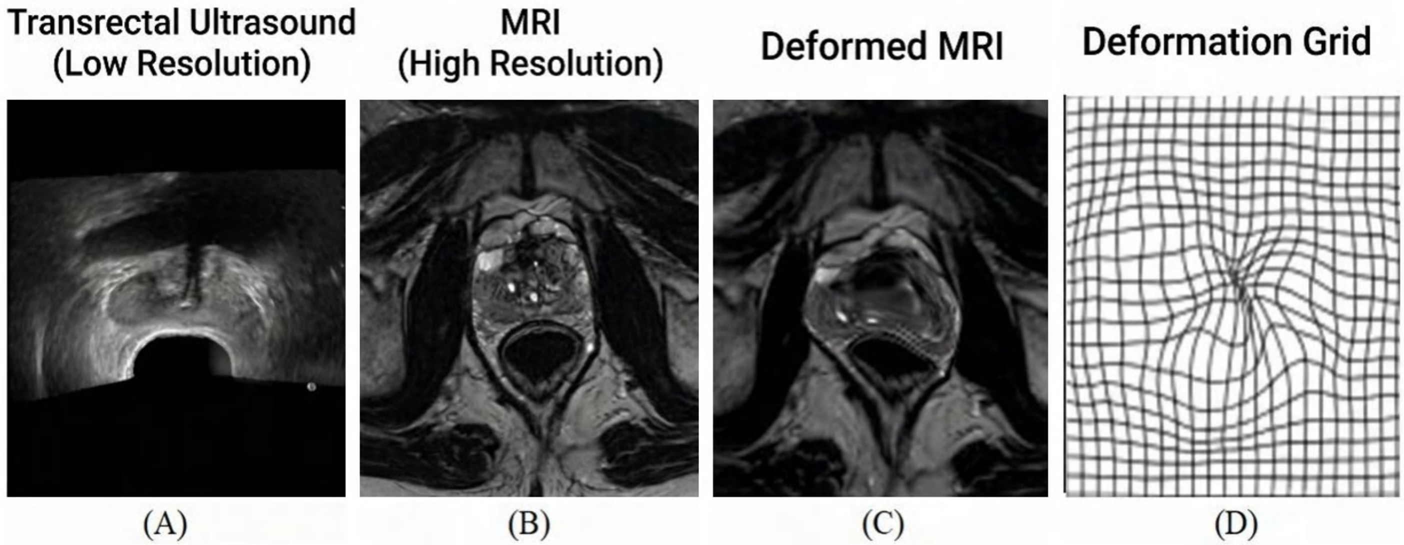


Figure 1. Visualization of core challenges in cross-modal registration. (A) Intraoperative TRUS: Target ultrasound image characterized by low contrast, speckle noise, and a limited FOV, highlighting the significant modality gap; (B) Preoperative MRI: High-resolution T2-weighted image showing clear anatomical details; (C) Warped MRI: MRI image after non-rigid spatial transformation, aligned to match the prostate anatomy under probe compression; (D) Deformation Grid: Visualization of the dense displacement field; grid distortion reflects the non-linear spatial mapping. MRI, magnetic resonance imaging; TRUS, transrectal ultrasound; FOV, field of view; T2, transverse relaxation time 2.

1 INTRODUCTION

Prostate cancer (PCa) remains one of the most prevalent malignancies among men worldwide, ranking as a leading cause of both cancer incidence and mortality [1]. Based on Global Cancer Observatory statistics for the year 2022, more than one and a half million new cases of cancer were diagnosed worldwide, leading to about 400,000 deaths [2]. Early diagnosis and accurate treatment are crucial to improve the five-year survival rate and quality of life in PCa patients. Currently, clinical diagnosis relies heavily on systematic biopsy guided by transrectal ultrasound (TRUS). However, conventional TRUS-guided biopsy is hindered by two intrinsic limitations [3, 4]. First, as illustrated in **Figure 1A**, TRUS is inherently plagued by inferior image quality characteristics such as a low signal-to-noise ratio, speckle noise, and poor soft-tissue contrast, which obscure tumour borders and hamper the differentiation of malignant lesions from benign hyperplasia. Second, the process is highly dependent on the operator, and the accurate placement of a needle depends on the clinician's experience and hand steadiness. All these factors lead to a high false-negative rate, with a risk of missing clinically significant prostate cancer (csPCa), resulting in a delay in care. In contrast, multiparametric magnetic resonance imaging (MRI) offers superior soft-tissue resolution and multi-sequence imaging capabilities (**Figure 1B**), which enable accurate detection, localization, and risk stratification of csPCa. However, MRI is expensive and time-consuming, and it cannot guide the procedure in real time. Consequently, preoperative MRI-TRUS image fusion-targeted biopsy has been proposed. This technology aligns high-precision preoperative three-dimensional (3D) MRI with real-time intraoperative two-dimensional (2D)/3D TRUS, "projecting" MRI-identified suspicious lesions onto the TRUS space to provide clinicians with precise navigation, thereby significantly enhancing the diagnostic yield and detection rate of biopsy [5, 6].

In such a clinical workflow, image registration is the key technology to realize accurate navigation. Image registration refers to the task of searching for a spatial transformation that can map images of either the same subject or different subjects into each other. MRI-TRUS registration is a classic cross-modal problem fraught with severe technical hurdles. First, the disparate imaging physics of MRI and TRUS result in vast differences in intensity distributions, texture features, and tissue contrast. This makes conventional intensity-based similarity measures—such as normalized mutual information, mean squared error, and modality-independent neighborhood descriptor—struggle to establish effective spatial correspondences [7-10]. Second, the prostate is a compliant, deformable organ; intraoperative probe pressure, changes in patient positioning, and respiratory motion cause significant non-rigid deformation [11]. This physical shift renders preoperative anatomy inconsistent with intraoperative observations. Registration algorithms must therefore apply complex spatial transformations to the MRI to fit the TRUS geometry (**Figure 1C**). This spatial mapping is highly non-linear and non-uniform, as visualized by the dense deformation grid in **Figure 1D**. Moreover, the low signal-to-noise ratio of TRUS, acoustic artifacts, as well as the difference in spatial resolution and field of view among modalities add extra hurdles for obtaining an exact voxel-wise correspondence. Conventional non-rigid registration approach-

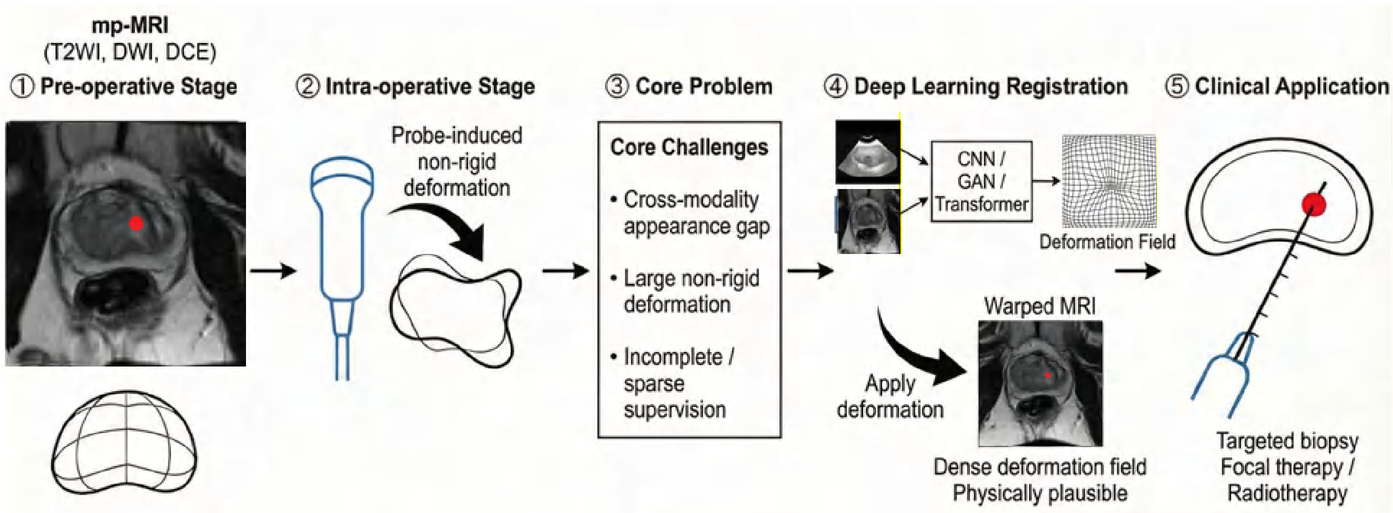


Figure 2. Overview of the DL-based prostate MRI-TRUS registration clinical workflow and technical framework. The process comprises five stages: ① Preoperative acquisition of mp-MRI to localize tumors; ② Intraoperative TRUS imaging, during which non-rigid deformation occurs due to probe pressure; ③ Identification of core registration challenges; ④ DL registration module (using CNN, GAN, or Transformer backbones) to predict dense deformation fields; ⑤ Final clinical application, providing precision navigation for targeted biopsy or focal therapy. DL, deep learning; mp-MRI, multiparametric magnetic resonance imaging; MRI, magnetic resonance imaging; TRUS, transrectal ultrasound; CNN, convolutional neural network; GAN, generative adversarial network; T2WI, T2-weighted imaging; DWI, diffusion-weighted imaging; DCE, dynamic contrast-enhanced imaging.

es are typically classified into spline-based methods (e.g., B-spline Free-Form Deformation and Thin-Plate Splines) and physics-based models [12, 13]. However, both paradigms have significant disadvantages. Spline-based approaches are often unable to model very local and non-uniform deformation as they assume a global smoothness constraint. Physics-based models are biomechanically plausible, but often require tedious iterative optimisation of complicated energy functions, making them prone to local minima and unsuitable for the real-time requirements of clinical interventions [14, 15].

To address such issues, deep learning (DL), in particular convolutional neural network (CNN)-based architectures, has enabled a new era of learning-based medical image (MI) registration, which does not need to optimize for each specific pair: DL methods consider registration as an end-to-end (E2E) regression problem. By training deep neural networks with a large dataset, this network learns a complicated non-linear mapping from an input image pair to the output deformation field [16]. After learning, the network is able to predict displacement fields in less than a millisecond and therefore enables online registration for clinicians with better reliability. In recent years, researchers have also widely used DL in prostate MRI-TRUS registration with customised designs according to the particularities of this task. Among these, we can mention the use of CNNs and weakly-supervised approaches to extract features at different levels, adopting generative models, including generative adversarial networks (GANs) and diffusion models, to address the cross-modal discrepancy issue and large deformation generation, and incorporating transformers for modeling long-range context relations. These developments are moving prostate-directed biopsy away from "experience-dependent" cognitive fusion and closer to fully automatic, precise, and intelligent "software fusion".

The aim of this review is to provide a systematic summary and critical discussion on the status quo of DL-based non-rigid prostate MRI-TRUS registration. First, we describe the key technical paradigms and representative work in the domain. Second, we summarize the datasets and evaluation metrics for performance assessment and quantitative comparison. Third, we examine how to integrate DL technologies into a clinical surgical workflow and introduce the open problems for their clinical deployment. Last but not least, we present our perspectives regarding the directions of future work and clinical translation so as to serve as a useful guide to researchers and clinicians alike. We summarize the overall DL-based prostate MRI-TRUS registration clinical workflow and technical pipeline presented throughout this manuscript in **Figure 2**.

2 DL PARADIGMS FOR PROSTATE MRI-TRUS REGISTRATION

Over the past few years, DL has revolutionized prostate MRI-TRUS image registration and generated multiple types of approaches that tackle the problem from different angles. To better understand how each technology has evolved over time, in this subsection we divide the existing work into three categories according to what they are innovating: first, we present CNN-based approaches, which are the most established para-

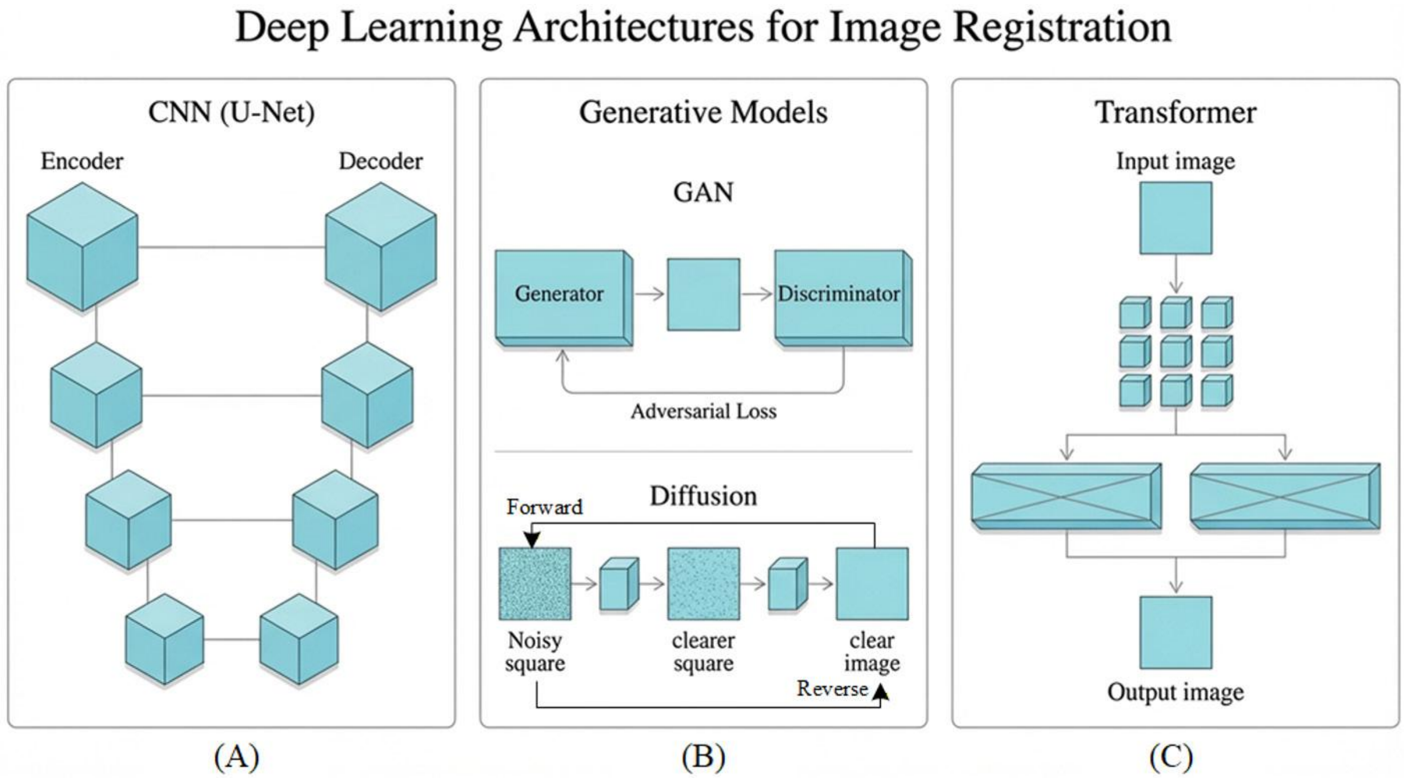


Figure 3. Schematic overview of deep learning architectures for prostate MRI-TRUS registration. (A) CNN Architecture: Represented by the U-Net backbone, utilizing encoder-decoder paths and skip connections for end-to-end regression; (B) Generative Models: Illustrating two distinct paradigms: GANs (top), which use adversarial training between a generator and discriminator to bridge modality gaps; and diffusion models (bottom), which employ an iterative denoising process (from noisy to clear) to generate high-fidelity deformation fields or images; (C) Transformer Architecture: Demonstrating the tokenization of input images into patches and the use of self-attention mechanisms to capture global anatomical dependencies. MRI, magnetic resonance imaging; TRUS, transrectal ultrasound; CNN, convolutional neural network; GAN, generative adversarial network.

digm, including advances in real-time architecture optimization and improved feature representation by incorporating geometry and physics priors (e.g., segmentation contours, biomechanical models). The second category is the generative model-based methods, including GANs that use adversarial learning to close the modality gap, and the more recent diffusion models, which are known for having strong distribution modeling capabilities. Lastly, we include attention mechanism and transformer-based models, where we explain how such an approach solves the local receptive field problem that comes with the convolution operation by adding a self-attention module or as a hybrid model. This architecture can therefore capture global, long-range dependencies as well as prominent anatomical features to perform this classification task. An illustration of the classification architecture used is shown in **Figure 3**.

2.1 CNN-based methods

As one of the first technically mature DL-driven image registration methods, CNN-based methods are designed based on a simple assumption that strong local feature representation capability and nonlinear transformation capacity can be utilized by modeling the image registration problem as an E2E regressor. Pioneering works, such as the framework that applied CNNs as similarity metrics for driving 3D ultrasound registration, achieved an error below 5 mm in 81% of the samples, proving that deep features have greater potential than conventional intensity-based measures or even manual alignment [17].

Later, popular approaches (e.g., VoxelMorph) established the trend of utilizing encoder-decoder networks for direct prediction of displacement fields [16]. Based on these works, researchers first worked on optimizing registration workflows and basic architectures for addressing the challenges of real-time performance and large deformations intrinsic to prostate interventions. For example, Dupuy et al. proposed an online 2D/3D registration approach based on CNNs, incorporating the registration result of the previous frame as a trajectory prior and substantially improving temporal consistency [11]. To address large deformation, Guo et al. proposed the multi-stage registration architecture with a "coarse-to-fine" approach and a synthesized error-scaled data generation method to address the missing coverage of large-deformation labels in training, reducing the surface registration error to 3.57 mm [18]. In a subsequent study, Guo et al. proposed a deep learning-based ultrasound frame-to-volume registration pipeline for registering 2D ultrasound frames to 3D volumes [19]. With the integration of a frame-to-frame registration network and a slice-correction network, and a similarity filtering mechanism, the pipeline was able to achieve an approximate target navigation error of 1.93 mm at a clinical-grade frame rate of 5–14 frames per second, demonstrating both high accuracy and fast inference speed suitable for intraoperative use in an MRI-ultrasound fusion-guided biopsy system.

Besides improving the network structure and inference process, another important line of work is exploring more detailed anatomical features and prior knowledge in order to facilitate better inter-modality correspondence learning with extra semantic content provided by inputs or additional geometrical regularization imposed explicitly. One common approach is to enrich input semantics via segmentation probability maps. Zeng et al. proposed a hybrid supervision method that first uses two independent fully convolutional networks to produce prostate probability maps on MRI and TRUS, respectively, which are further fed to the registration network as auxiliary channels. By combining Dice loss and surface distance loss, they obtained a mean target registration error (TRE) value of 2.53 mm using this technique in 36 patients [20]. As research deepened, to elevate feature expression from a frequency-domain perspective, Jiang et al. proposed a method based on joint learning and a multi-level wavelet feature pyramid. Through a cyclic mutual-assistance mechanism between the segmentation and registration networks, they achieved superior accuracy across 642 public cases [21]. Additionally, to endow networks with physical interpretability, incorporating physical model constraints has become an important direction. Fu et al. proposed a biomechanically constrained registration method based on 3D point-cloud matching as a proxy task. They used deformation fields obtained from finite element analysis (FEA) as supervision, enabling them to encode these biomechanical constraints in the network and achieve high-precision MR-TRUS registration [22].

2.2 Generative model-based methods (adversarial networks and diffusion models)

A GAN is composed of two competitive networks called a generator and a discriminator. Given an image registration task, the generator usually predicts the deformation field or generates the warped image, whereas the discriminator discriminates the warped image from the ground-truth target image, compelling the Generator to generate more realistic and plausible results. Yan et al. first proposed such an idea for multimodal MI registration using the framework of adversarial image registration network. Using a Wasserstein GAN approach, the generator directly estimates rigid transformation parameters, while the discriminator is used to evaluate the alignment quality, achieving fast (<100 ms) and accurate (TRE 3.48 mm) registration for 763 cases, effectively validating the viability of adversarial learning [23].

To cope with such anatomical complexity, later works focused on maintaining prominent regional information as well as adding some physics-based restrictions. For example, Lian et al. presented an region of interest-guided deformable image-to-image translator, which converts one modality into another while enforcing consistent structural prior knowledge in order to generate plausible TRUS-like images, simplifying later registration challenges and preserving geometric truthfulness [24]. Likewise, Feng et al. proposed a salient region matching approach; while not a pure GAN structure, they integrate cross-modal spatial attention into a generative framework. They use a salient region matching loss to regularize the shape and intensity consistency between the region of interest of the prostate, and they obtained a Dice coefficient of 85.9%, thus improving the robustness of a purely automatic registration [25]. In addition, to address the problem of potential unphysical deformations in pure data-driven generation, we use a variational autoencoder as a component of our biomechanics-informed generative registration framework architecture, which encodes biomechanical priors (computed via FEM) onto a latent manifold. In registration, the network searches within such a manifold for an optimal deformation field fulfilling both image information and biomechanical constraints, effectively suppressing unreasonable local distortion [26].

Although successful for cross-modal registration with GANs, due to its unstable training process as well as the mode collapse issue, it is not powerful enough to model the complicated distribution of large deformations. Recently, diffusion probabilistic models have been introduced. Based on their good distribution coverage and stable training performance, they have been brought into the registration domain to break through such a bottleneck. Yao et al. conspicuously introduced diffusion models to the weakly supervised setting. They designed a label-aware diffusion model, which can correct the initial bias by aligning label centroids, followed by producing a high-quality deformation field through a feature-guided module with label supervision. It is highly robust for large deformations, substantially lowering the TRE to 0.940 mm, with an exceptionally small Jacobian determinant fold-over rate (0.134), and is much more precise than classical unsupervised or supervised registration techniques such as VoxelMorph [27].

2.3 Transformer-based registration methods

Although CNN-based registration approaches are highly efficient with good local texture extraction ability, they are based on a strong inductive bias—i.e., the set of structural assumptions that the model makes for generalization to new data. For CNNs, this bias mainly consists of locality and translation invariance, which means the model assumes the relevant fea-

Table 1. Characteristics of representative public datasets for prostate MRI-TRUS registration

Dataset	Scale	Raw resolution	Preprocessing and characteristics
μ-RegPro [32]	76 cases	MRI: 120×128×128 voxels TRUS: 81×118×88 voxels	Resampled to 0.8 mm^3 isotropic resolution
TCIA [33]	Prostate subset	T2WI: 0.5 mm/plane, 3.6 mm slice DWI: 2.0 mm/plane, 3.6 mm slice DCE: 1.5 mm/plane, 4.0 mm slice	Raw DICOM; Heterogeneous resolutions; Significant slice thickness

Note: MRI, magnetic resonance imaging; TRUS, transrectal ultrasound; TCIA, The Cancer Imaging Archive; T2WI, T2-weighted imaging; DWI, diffusion-weighted imaging; DCE, dynamic contrast-enhanced imaging; DICOM, Digital Imaging and Communications in Medicine.

ture is spatially clustered and its relative position remains unchanged. While this facilitates feature learning, it nevertheless results in the problem of the local receptive field, which brings great difficulties in modeling large-scale non-rigid deformation.

This limitation is especially important for prostate MRI-TRUS registration. In a clinical setting, inserting and manipulating the transrectal probe applies considerable pressure to the gland, inducing large-scale global deformations and non-uniform displacements of the whole organ, which have long-range spatial correlations, where the pressure on the rear end directly determines the shape and location of the front tip. Since the prostate is a soft tissue, all these global variations are distributed across the whole field of view. It is difficult for CNNs to learn such inherent correlations among far-away anatomical structures because of their nature as local window-based operations. In order to address this limitation, equipping them with an attention mechanism, i.e., transformers for global context extraction, is now a mainstream direction of development.

Originally having great success in natural language processing, transformers use their key component, self-attention, which is able to compute association weights between any two positions in an input sequence adaptively. Since vision transformers for unsupervised volumetric medical image registration was introduced into MI registration in 2021, the field quickly moved on, exploring shifts from pure global attention to hierarchical feature representations, such as hierarchical vision transformer [28, 29]. Nevertheless, the pure transformer network usually sacrifices matching precision for minute structures because there is no local texture information. Therefore, the “hybrid architectures” that combine the local feature extraction ability of CNN and the global screening ability of attention mechanisms are becoming mainstream choices.

For instance, Mahmoudi et al. proposed 3D-Wavelet-Deep-Separable-Attention-PMorph, which is a hybrid of Swin Transformer and CNN with an encoder-decoder architecture [30]. They introduced their main contribution as the Wavelet-Deep-Separable-Attention, which relies on wavelet transform to achieve multi-scale space-frequency fusion that overcomes the shortcomings of conventional transformer in representing frequency-domain information. Assuming standardized μ-RegPro dataset and pre-aligned conditions, this approach reached a TRE of less than one millimetre [30]. Parallel to this is another technical route focusing on clinical robustness and salient feature mining. Yu et al. proposed a weakly-supervised framework integrating an Attention-enhanced U-shaped CNN and a Residual-Enhanced Registration Network [31]. This method utilizes a sparse set of expert-annotated anatomical landmarks to compute loss functions, alleviating the annotation burden. In this work, we not only evaluated our algorithms with the publicly available μ-RegPro dataset, but also provided a patient cohort including Prostate Imaging Reporting and Data System scores as an additional test set [32]. Despite the fact that, due to its complexity, the mean TRE increased up to 8.63 mm compared to the public dataset, the study showed that weakly supervised DL can be used to assist biopsy navigation in realistic clinical workflow scenarios [31]. The above two recent works show that either learning the multi-scale frequency perception by transformer blocks or promoting salient feature representation with attention mechanisms, enhancing the model’s perception of global topology and important anatomical structures, is the critical path to solving the challenging non-rigid prostate registration problem.

3 DATASETS AND EVALUATION METRICS

Consistent datasets and unified evaluation criteria are essential to objectively evaluate and compare different DL-based registration approaches. In this part, we list typical public resources and basic evaluation criteria used by the prostate MRI-TRUS registration community.

3.1 Public datasets

The size, variety and annotation quality of a dataset directly determine how well a DL model will perform and to what extent it can be generalized. Today, most studies on prostate MRI-TRUS registration are based on a few publicly available datasets and internal private datasets, as patient privacy strictly restricts the collection and distribution of high-quality, large-scale labelled datasets. A comparison between the two most representative public datasets and their peculiarities is provided in **Table 1**.

Datasets used in recent works are based on the following two approaches to algorithm verification. μ-RegPro is a very well-controlled testbed [32]. In this dataset, MRI and TRUS volumes are resampled to the same 0.8 mm^3 isotropic resolution. Although this strict standardization removes the impact of resolution disparity and improves the repeatability of registration outcomes, it can lose some of the finer anatomy that is present in raw data.

On the other hand, The Cancer Imaging Archive maintains the raw clinical state of the imaging data [33]. It comprises diagnostic-grade Digital Imaging and Communications in Medicine sequences with heterogeneous resolutions, with good in-plane resolution (e.g., 0.5 mm, T2-weighted imaging) but large slice thicknesses (up to 4.0 mm). The Cancer Imaging Archive requires stronger preprocessing and adaptation capabilities for the registration models but provides better ecological validity in that it reflects the natural heterogeneity and artifacts of clinical acquisitions.

3.2 Evaluation metrics

To evaluate registration performance in a quantitative manner, several widely used evaluation criteria are employed to measure registration quality based on different aspects, such as volume intersection ratio, surface point-to-surface distance, and inlier correspondences between intrinsic features.

TRE: TRE, which has been regarded as a "gold standard" to assess registration accuracy, quantifies the Euclidean distance between one fixed set of predefined, known corresponding points (landmarks) in the source and target images, which usually correspond to important anatomical landmarks, such as the urethra, prostate calcifications, or cysts. The TRE is defined as:

$$TRE = sqrt\left(\frac{1}{N} * \sum_{i=1}^{N} (\| \mathrm{p}_i^m - T(\mathrm{p}_i^s) \|^2)\right) \tag{1}$$

where p_i^m and p_i^s denote the coordinates of the i-th corresponding landmark in the target and source images, respectively, T represents the learned transformation, and N is the total number of points. A lower TRE indicates higher precision. In clinical practice, a TRE of less than 2.0–3.0 mm is generally considered acceptable for targeted interventions.

Dice similarity coefficient (DSC): DSC measures the volume overlap between two binary masks (e.g., organ segmentations). It has values ranging from 0 to 1, where 1 represents an exact match. This is used to measure the similarity between the warped source mask and the target mask in a registration problem:

$$DSC = 2 * \frac{|\mathrm{A} \cap \mathrm{B}|}{(|\mathrm{A}| + |\mathrm{B}|)} \tag{2}$$

where A and B represent the source and target masks. DSC is a highly effective metric for assessing the overall global alignment of an organ.

Hausdorff distance (HD): HD measures the maximum mismatch between two surfaces:

$$h(\mathrm{A},\mathrm{B}) = \min_{\{a\in A\}} \min_{\{b\in B\}} \| \mathrm{a} - \mathrm{b} \| \tag{3}$$

$$HD = \max\left(h(\mathrm{A},\mathrm{B}), h(\mathrm{B},\mathrm{A})\right) \tag{4}$$

To reduce the impact of noise and outlier points, the 95th percentile HD is usually reported in the literature instead. It picks the maximal distance after removing the top 5 percent of the largest distances and serves as a strong indicator of local misalignment at the boundary.

Mean surface distance (MSD): MSD measures the mean distance between every pair of corresponding points on each surface. Compared with HD, MSD is not as sensitive to local outliers and thus better reflects the overall surface matching quality.

Table 2 outlines the performance of state-of-the-art (SOTA) methods, categorized by their respective evaluation datasets. While direct numerical benchmarking is hindered by varying normalization protocols and dataset scales, these metrics offer a comprehensive trajectory of the field's technical maturation and its ongoing shift toward clinical integration.

While the different origins of data, as well as the lack of standardization for data preprocessing, prevent us from making an absolute quantitative comparison among them, horizontally, the quantitative metrics summarized in **Table 2** highlight a clear technical evolution, as well as the present-day bottlenecks in translating these advances into clinical practice. We also note, however, that performance can only ever be as good as the dataset on which it has been tested, with "best-case scenario" performance typically being obtained when testing on already aligned or curated datasets. Performance remains very limited on raw clinical data.

The first is that developments in model architecture have directly led to a decade's worth of orders-of-magnitude improvements in registration accuracy. Initial CNN and GAN-based approaches (e.g., Zhu et al.; Yan et al.), which were restricted to the local receptive field and simple adversarial strategies, have TREs in a range of approximately 3.0–5.0 mm most of the time [17, 23]. By comparison, works in 2025 using transformers or diffusion models manage to lower the TRE below one millimeter (<1.0 mm) [30]. Note that those SOTA results mainly report on very standard benchmarks, where spatial pre-alignment reduces the initial pose variance. This indicates that, although modeling of complex deformations is important, the quality of the input data is one of the most important factors that determines the degree of reported success.

Table 2. Summary of performance metrics for representative DL-based prostate registration methods

Study	Method	TRE (mm)	DSC	HD95 (mm)	MSD (mm)	Dataset sources
Yan et al. 2018 [23]	Adversarial (AIR-net)	3.480	-	-	-	Private
Zhu et al. 2019 [17]	CNN-based similarity	<5.000 (in 81% cases)	-	-	-	Private
Zeng et al. 2020 [20]	Label-driven weakly supervised	2.530±1.390	0.910±0.020	4.410	0.880	Private
Fu et al. 2021 [22]	Biomechanically constrained	1.570±0.770	0.940±0.020	2.960±1.000	0.900±0.230	Private
Guo et al. 2023 [19]	FVReg (CNN)	1.930	-	-	-	Private
Lian et al. 2025 [24]	Deformation-aware GAN	2.240±1.020	0.876	-	-	Private
Feng et al. 2025 [25]	Salient region match	4.650±1.760	0.859±0.035	-	-	μ-RegPro
Mahmoudi et al. 2025 [30]	3D-WDA-PMorph	0.850	0.941	-	-	TCIA
Jiang et al. 2025 [21]	Joint learning (MWFP)	-	0.811±0.018	-	-	TCIA
Yu et al. 2025 [31]	Weakly supervised (RERN)	7.860 (Pub); 8.630 (Clin)	0.748 (Pub); 0.730 (Clin)	10.180 (Pub); 11.180 (Clin)	-	μ-RegPro
Yao et al. 2025 [27]	Label-aware diffusion	0.940	0.880	-	-	μ-RegPro
Hao et al. 2026 [26]	BGProReg (Phy-GAN)	3.720±1.050	0.924±0.027	3.000±0.700	-	μ-RegPro and TCIA

Note: TRE, target registration error; DSC, Dice similarity coefficient; HD95, 95th percentile Hausdorff distance; MSD, mean surface distance; DL, deep learning; AIR-net, adversarial image registration network; CNN, convolutional neural network; FVReg, frame-to-volume registration; GAN, generative adversarial network; MWFP, multi-level wavelet feature pyramid; RERN, residual-enhanced registration network; BGProReg, biomechanics-informed generative registration framework; Phy-GAN, physics-informed generative adversarial network; 3D-WDA-PMorph, three-dimensional wavelet-deep-separable-attention PMorph; TCIA, The Cancer Imaging Archive; Pub, public dataset; Clin, clinical dataset.

Second, physics-based constraints are important for achieving plausible deformation, although purely data-driven approaches typically achieve lower TRE: the absence of anatomical constraints leads to non-physical topology changes, as shown in works by Fu et al. and Hao et al. [22, 26]. While the TREs (approximately 1.57–3.72 mm) are still slightly higher than those from recent pure DL methods, the use of biomechanical constraints or FEM priors can prevent topological errors, e.g., Jacobian determinant folding. It stresses that forcing a physically realistic deformation field is at the core of ensuring safety for clinical deployment, even with low errors.

Lastly, there is still a large performance gap between controlled experiments and the true clinical setting. As shown in the work of Yu et al., the same algorithm performed significantly worse on more heterogeneous clinical data than it did on very well-controlled public datasets, with an increase in TRE from 7.86 mm to 8.63 mm [31]. This indicates that the good performance obtained by SOTA methods when operating on preprocessed data could be an overestimate of their performance in realistic settings, and future efforts should not focus solely on improving performance on a reference dataset but should also attempt to take into account domain shift, multimodal interference, e.g., artifacts, and multicenter pathological variation to improve the clinical robustness of such models.

4 CLINICAL APPLICATIONS AND CHALLENGES

DL-based MRI-TRUS registration is not only a matter of developing algorithms, but it is also useful when incorporated into clinical diagnosis and treatment processes. It could improve the accuracy and safety of prostate biopsy and other interventions: these tools are changing the paradigm of treatment. Here we highlight some fundamental applications to clinical practice, and discuss various factors that make their clinical adoption difficult.

4.1 Clinical application scenarios

At present, MRI-TRUS fusion-targeted biopsy is the most mature and popular application scenario for fusion technology. Conventional systematic biopsy is usually restricted by the poor specificity of ultrasound, which results in “blind sampling” and a high omission rate of csPCa [34-36]. Fusion-guided systems instead map high-risk regions identified on MRI in real time, e.g., Prostate Imaging Reporting and Data System 4/5 lesions, onto intra-operative TRUS images to guide a precise trajectory of the needle, thereby improving the diagnostic yield of csPCa with a minimal number of cores and related complications [37]. For example, the PROST-Net model implemented on the PROST robotic platform to perform transperineal biopsy with real-time prostate segmentation and tracking, thus minimizing errors due to subjective human factors [38]. Wang et al. developed an E2E pipeline including a preprocessing module, the DeepReg registration network, and a live planning module, which generates both trajectory suggestions and “collision warnings” between the organs at risk closest to each other, showing that it is possible to have completely intelligent surgical assistance [39].

Precise registration technology is a key support for achieving focal therapy, such as high-intensity focused ultrasound or cryoablation [40, 41]. For both treatments, the accurate delivery of energy to the preoperative MRI-targeted plan and real-time MRI-TRUS registration facilitate accurate execution of the treatment plan as well as an adaptive response to tissue deformation due to thermal ablation or ice ball growth [40, 42]. Additionally, for minimally invasive procedures such as brachytherapy, registering the pre-operative MRI and intra-operative ultrasound improves targeting accuracy while sparing organs at risk to decrease morbidity [43]. In order to tackle registration challenges due to the presence of metallic needle artifacts on ultrasound, researchers have suggested weakly supervised approaches for propagating the MRI-based prostate contours onto the treatment-planning ultrasound image, enabling efficient dose allocation and analysis [44].

4.2 Technical and clinical translation challenges

Despite these promising results in isolated experiments, there are still some challenges that need to be addressed before applying a DL model as an evidence-based medical tool in real-life conditions [45].

4.2.1 Generalization and data bottlenecks

However, deep models, especially complicated CNNs or transformers, are prone to highly fitting the particular distribution in the training set, which results in a severe performance drop when facing the heterogeneity of data from different centers, as the data are often unstandardized across device manufacturers or imaging protocols [46]. Additionally, due to strict patient privacy rules, e.g., General Data Protection Regulation and Health Insurance Portability and Accountability Act, and the prohibitive cost of expert annotation, most research is based on small-scale and single-center datasets. Large-scale standardized multicenter data are one of the major limitations for demonstrating generalizability and reaching the maximum achievable performance.

4.2.2 Technical robustness and interpretability in complex scenarios

Although registration error is reported at a millimeter scale, the robustness of algorithms to inevitable "clinical noise", e.g., bowel air-gas artifacts, non-uniform probe pressure, or extreme tissue distortion, remains to be rigorously stress-tested [47]. Furthermore, due to the "black-box" nature of DL, the decision-making process is opaque: if a clinician cannot interpret the registration logic, e.g., which landmarks have been matched, they will not necessarily believe in the outcome of the system. Therefore, designing explainable artificial intelligence (AI) to explain how the model reaches its conclusions can be necessary before building confidence between engineers and clinicians [48].

4.2.3 Workflow integration and regulatory hurdles

A successful algorithm will need to seamlessly integrate into existing hospital information systems, i.e., picture archiving and communication systems/radiology information systems, and match the user interface/user experience preferences of surgical teams. As medical devices, such systems have to be approved by regulatory authorities such as the Food and Drug Administration or National Medical Products Administration, and there must be an built-in fault detection and error handling within the system itself. When the model's confidence is low or an anomaly has been detected, it should warn and hand over control back to the doctor, adhering to the "human-in-the-loop" paradigm as a baseline for surgical safety.

5 FUTURE DIRECTIONS AND PERSPECTIVES

To solve the problems mentioned above, high-precision prostate MRI-TRUS registration technology is developing toward robustness and deep clinical integration. With leading-edge technical trends as well as pressing clinical needs, we envision that future studies will focus on the following five main aspects:

The first is the development of a new generation of hybrid architectures as well as the exploration of more generative frameworks in future registration models. The key point of future registration models should be how to combine global information with local information from different scales. On one hand, transformer-based networks can effectively capture long-range dependencies, which have obvious advantages when dealing with the estimation of large-scale non-rigid deformation. On the other hand, it is believed that hybrid "transformer+CNN"-based networks would be responsible for keeping the global topological structure consistent while enforcing local pixel-wise correspondence. Furthermore, recently emerging generative networks, especially diffusion models, have shown remarkable performance in distribution modeling. Applying them to the task of cross-modal image synthesis or utilizing them as priors for deformation field generation can outperform conventional similarity measures, contributing new mathematical tools to the field of unsupervised registration.

Second, future studies should focus on learning from limited resources. In order to avoid requiring massive amounts of expert supervision, few-shot learning and domain adaptation would be another major focus in future studies. A meta-learning or metric-learning framework could help improve the rapid adaptation capability when facing data from novel clinical centers or imaging devices. Furthermore, exploring efficient transfer learning schemes, where the anatomical representation learned in well-populated domains, e.g., mono-modal MRI registration, is distilled to data-scarce MRI-TRUS tasks, is one such possible route toward alleviating the annotation bottleneck problem.

Third, future studies should focus on deep mining of multi-modal radiomics. The existing methods mainly use single-channel intensity or a binary mask. The paradigm is changing toward multi-parameter and multidimensional information fusion. Fusing multi-sequence MRI information, e.g., T2-weighted imaging, diffusion-weighted imaging, and dynamic contrast-enhanced imaging, and possibly including the metabolic functional signature of prostate-specific membrane antigen positron emission tomography/computed tomography, may be useful. Researchers may build high-dimensional radiomics feature spaces. These types of multi-modal constraints will help disambiguate the registration problem where there is ambiguity due to, for example, low contrast between tissues and/or similar grey values across modalities, thus improving the stability of soft tissue matching.

Fourth, future studies should focus on the incorporation of physical laws and biomechanical constraints. Purely data-driven models tend to suffer from non-physical deformations, e.g., tissue folding. The integration of physics priors is critical for guaranteeing a unique and safe solution. Future work will focus more on explicitly incorporating diffeomorphic mapping, volume preservation, and biomechanical consistency constraints in terms of losses or networks. This guarantees that the learned deformation fields are also physically invertible, respecting the hyperelastic behavior of prostate tissues. FEA-based strategies using FEA results to guide latent space learning would be important baselines in building "physics-aware" neural networks.

Finally, future studies should focus on "perception-planning-execution" closed-loop intelligent system building. The final value of registration algorithms should be their translation into clinical practice. The mainstream direction is deeply integrating AI registration with surgical robotic control and real-time intraoperative imaging systems. The development of smart navigators that are able to track deformations in real time and dynamically correct trajectories, enabling a fully automatic closed loop between imaging diagnosis/pre-operative planning and robotically accurate intervention, is the ultimate way toward bringing PCa management to the age of automation, personalization, and non-invasive precision. Importantly, in order for such a shift to occur, further work is needed to reduce the "black-box" aspect of DL and focus on explainable AI techniques. Providing interpretable models with associated measures of confidence or maps showing important anatomical features would allow clinicians to understand how the model made its prediction, thus enabling the clinical interpretability and trust required for high-risk intervention adoption.

6 CONCLUSION

Non-rigid MRI-TRUS registration is the key technology to achieve precision-targeted biopsy or focal therapy of PCa. Although it plays a very important role, accurate alignment has long been a challenge in the field of MI computing due to different cross-modal imaging physics, complex soft-tissue deformations, and the ubiquitous absence of voxel-level annotations. DL has brought about a revolution in tackling such a non-linear mapping issue. In this survey, we have described clearly how the field has evolved: starting with fully supervised CNN E2E regressors, progressing to mixed CNN architectures using both weak supervision and physics-based constraints; from the use of GANs for bridging modality gaps to high-precision generative registration with diffusion models; and, finally, toward joint modeling of local fine structure and global topology through hybrid architectures and attention mechanisms. To date, it has been shown by the following works that DL techniques are making significant progress in improving registration accuracy, acceleration, and convergence.

Currently, SOTA DL models have achieved sub-millimeter-level accuracy (TRE<1.0 mm) on standard public datasets, with short (<0.5 s) inference times, making them suitable for intraoperative real-time navigation. All these advances provide solid anatomical guidance for accurate biopsy, ablative therapy, and brachytherapy. However, we have to admit that all the above results are built upon idealized experiments; there exists a large "generalization gap" from lab benchmarks to real-world clinical scenarios. The stability and safety of such models against multicenter heterogeneous data, pathologically complex cases, and intraoperative artifacts require additional validation in large-scale, prospective clinical studies.

In future work, we expect to shift our attention away from small improvements in accuracy and toward improving utility for clinical practice by focusing on out-of-sample performance across institutions, developing low-resource training techniques, and strongly incorporating physics and biomechanical prior knowledge into neural network architecture design for plausible and safe deformation predictions. The final goal is to build highly automated, explainable, and effortlessly deployed E2E intelligent interventional systems through the synergistic coupling between advanced algorithms and surgical robots. These tools are expected to bring the field of PCa care into an age of automation, precision, and intelligent minimally invasive interventions.

DECLARATIONS

Author contributions

Peiyu Chen conducted the data analysis and drafted the initial manuscript. Xudong Guo supervised the entire study and provided critical contributions to the writing and revision of the paper.

Funding

This work was supported by the National Natural Science Foundation of China (62373253), the National Key Research and Development Program of China (2023YFF1105002), the Open Project Fund of the State Key Laboratory of Cardiovascular Diseases (2024SKL-TJ012), and the Huzhou Key Laboratory of Precise Diagnosis and Treatment of Urinary Tumor Fund Project (HZMNZL202401).

Data availability

The datasets analyzed in this review are available in the public repositories cited within the article. No new data were created during this study.

Ethics approval and consent to participate

Not applicable.

Consent for publication

Not applicable.

Competing interests

The authors declare that they have no competing interests.

Acknowledgements

Not applicable.

REFERENCES

[1] Hassan MR, Islam MF, Uddin MZ, Ghoshal G, Hassan MM, Huda S, et al. Prostate cancer classification from ultrasound and MRI images using deep learning based Explainable Artificial Intelligence. Future Generation Computer Systems. 2022;127:462-472. https://doi.org/10.1016/j.future.2021.09.030

[2] Bray F, Laversanne M, Sung H, Ferlay J, Siegel RL, Soerjomataram I, et al. Global cancer statistics 2022: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2024 May-Jun;74(3):229-263. https://doi.org/10.3322/caac.21834

[3] Avolio PP, Lughezzani G, Anidjar M, Hassan T, Rompré-Brodeur A, Buffi NM, et al. The diagnostic accuracy of micro-ultrasound for prostate cancer diagnosis: A review. World J Urol. 2023 Nov;41(11):3267-3276. https://doi.org/10.1007/s00345-023-04521-w

[4] Bektaş M, Chia CM, Burchell GL, Daams F, Bonjer HJ, van der Peet DL. Artificial intelligence-aided ultrasound imaging in hepatopancreatobiliary surgery: Where are we now? Surg Endosc. 2024 Sep;38(9):4869-4879. https://doi.org/10.1007/s00464-024-11130-0

[5] Erbin A, Caglar U, Turkay R. Evaluating the effectiveness and safety of robotic-assisted MRI/TRUS fusion transperineal prostate biopsy systems: A narrative review based on current literature. Eurasian J Med. 2023 Dec;55(Suppl 1):S125-S130. https://doi.org/10.5152/eurasianjmed.2023.23370

[6] Inoue T, Shin T. Current magnetic resonance imaging-based diagnostic strategies for prostate cancer. Int J Urol. 2023 Dec;30(12):1078-1086. https://doi.org/10.1111/iju.15281

[7] Studholme C, Hill DLG, Hawkes DJ. An overlap invariant entropy measure of 3D medical image alignment. Pattern Recognition. 1999;32(1):71-86. https://doi.org/10.1016/s0031-3203(98)00091-0

[8] Wang Z, Bovik AC. Mean squared error: Love it or leave it? A new look at signal fidelity measures. IEEE Signal Processing Magazine. 2009;26(1):98-117. https://doi.org/10.1109/MSP.2008.930649

[9] Heinrich MP, Jenkinson M, Bhushan M, Matin T, Gleeson FV, Brady SM, et al. MIND: Modality independent neighbourhood descriptor for multi-modal deformable registration. Med Image Anal. 2012 Oct;16(7):1423-1435. https://doi.org/10.1016/j.media.2012.05.008

[10] Shakeri S, Le W, Ménard C, Kadoury S. Deformable MRI to transrectal ultrasound registration for prostate interventions with shape-based deep variational auto-encoders. 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI). 2021; 174-178. https://doi.org/10.1109/ISBI48211.2021.9434101

[11] Dupuy T, Beitone C, Troccaz J, Voros S. 2D/3D deep registration for real-time prostate biopsy navigation. Medical Imaging 2021: Image-Guided Procedures, Robotic Interventions, and Modeling. 2021;58. https://doi.org/10.1117/12.2579874

[12] Mitra J, Martí R, Oliver A, Lladó X, Ghose S, Vilanova JC, et al. Prostate multimodality image registration based on B-splines and quadrature local energy. Int J Comput Assist Radiol Surg. 2012 May;7(3):445-454. https://doi.org/10.1007/s11548-011-0635-8

[13] Zettinig O, Shah A, Hennersperger C, Eiber M, Kroll C, Kübler H, et al. Multimodal image-guided prostate fusion biopsy based on automatic deformable registration. Int J Comput Assist Radiol Surg. 2015 Dec;10(12):1997-2007. https://doi.org/10.1007/s11548-015-1233-y

[14] Hu Y, Gibson E, Ahmed HU, Moore CM, Emberton M, Barratt DC. Population-based prediction of subject-specific prostate deformation for MR-to-ultrasound image registration. Med Image Anal. 2015 Dec;26(1):332-344. https://doi.org/10.1016/j.media.2015.10.006

[15] Yang X, Rossi P, Mao H, Jani AB, Ogunleye T, Curran WJ, et al. A MR-TRUS registration method for ultrasound-guided prostate interventions. Proc SPIE Int Soc Opt Eng. 2015 Feb;9415:94151Y. https://doi.org/10.1117/12.2077825

[16] Balakrishnan G, Zhao A, Sabuncu MR, Guttag J, Dalca AV. VoxelMorph: A learning framework for deformable medical image registration. IEEE Trans Med Imaging. 2019 Aug;38(8):1788-1800. https://doi.org/10.1109/TMI.2019.2897538

[17] Zhu N, Najafi M, Han B, Hancock S, Hristov D. Feasibility of image registration for ultrasound-guided prostate radiotherapy based on similarity measurement by a convolutional neural network. Technol Cancer Res Treat. 2019 Jan 1;18:1533033818821964. https://doi.org/10.1177/1533033818821964

[18] Guo H, Kruger M, Xu S, Wood BJ, Yan P. Deep adaptive registration of multi-modal prostate images. Comput Med Imaging Graph. 2020 Sep;84:101769. https://doi.org/10.1016/j.compmedimag.2020.101769

[19] Guo H, Xu X, Song X, Xu S, Chao H, Myers J, et al. Ultrasound frame-to-volume registration via deep learning for interventional guidance. IEEE Trans Ultrason Ferroelectr Freq Control. 2023 Sep;70(9):1016-1025. https://doi.org/10.1109/tuffc.2022.3229903

[20] Zeng Q, Fu Y, Tian Z, Lei Y, Zhang Y, Wang T, et al. Label-driven magnetic resonance imaging (MRI)-transrectal ultrasound (TRUS) registration using weakly supervised learning for MRI-guided prostate radiotherapy. Phys Med Biol. 2020 Jun 26;65(13):135002. https://doi.org/10.1088/1361-6560/ab8cd6

[21] Jiang H, Hu J, Qian X, Zheng Y, Zhou Z, Dai Y. Registration method of MRI-TRUS images based on joint learning and multi-level wavelet feature pyramid. Computer Engineering. 2025;51(10):270-283. https://doi.org/10.19678/j.issn.1000-3428.0069093

[22] Fu Y, Lei Y, Wang T, Patel P, Jani AB, Mao H, et al. Biomechanically constrained non-rigid MR-TRUS prostate registration using deep learning based 3D point cloud matching. Med Image Anal. 2021 Jan;67:101845. https://doi.org/10.1016/j.media.2020.101845

[23] Yan P, Xu S, Rastinehad AR, Wood BJ. Adversarial image registration with application for MR and TRUS image fusion. Machine Learning in Medical Imaging. 2018;197-204. https://doi.org/10.1007/978-3-030-00919-9_23

[24] Lian J, Pang Y, Chen X, Huang Y, Lian B, Yap P. Deformation-aware MR-TRUS image translation for prostate cancer brachytherapy. Res Sq. 2025 Nov 18;rs.3.rs-7958252. https://doi.org/10.21203/rs.3.rs-7958252/v1

[25] Feng Z, Ni D, Wang Y. Salient region matching for fully automated MR-TRUS registration. 2025 IEEE 22nd International Symposium on Biomedical Imaging (ISBI). 2025;1-5. https://doi.org/10.1109/isbi60581.2025.10981019

[26] Hao Y, Li S, Zhao J, Zhang C. BGProReg: A biomechanically generative framework for prostate MRI-TRUS deformable image registration. Information Fusion. 2026;127:103911. https://doi.org/10.1016/j.inffus.2025.103911

[27] Yao Z, Chen J, Wen T. A label-aware diffusion model for weakly supervised deformable registration of multimodal MRI–TRUS in prostate cancer. Int J Comput Assist Radiol Surg. 2025 Nov 6. https://doi.org/10.1007/s11548-025-03538-3

[28] Chen J, He Y, Frey EC, Li Y, Du Y. ViT-V-Net: Vision transformer for unsupervised volumetric medical image registration. arXiv. 2021. https://doi.org/10.48550/arXiv.2104.06468

[29] Ghahremani M, Khateri M, Jian B, Wiestler B, Adeli E, Wachinger C. H-ViT: A Hierarchical vision transformer for deformable image registration. 2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR). 2024;11513-11523. https://doi.org/10.1109/CVPR52733.2024.01094

[30] Mahmoudi H, Ramadan H, Riffi J, Tairi H. 3D-WDA-PMorph: Efficient 3D MRI/TRUS prostate registration using transformer-CNN network and wavelet-3D-depthwise-attention. J Imaging Inform Med. 2026 Apr;39(2):1474-1493. https://doi.org/10.1007/s10278-025-01615-2

[31] Yu J, Xiao B, Wang J, Mi H, Wang E, Huang N, et al. Weakly supervised deep learning for multimodal MRI-TRUS registration: Toward assisting prostate biopsy guidance. Digit Health. 2025 Sep 8;11:20552076251375870. https://doi.org/10.1177/20552076251375870

[32] Baum ZMC, Saeed SU, Min Z, Hu Y, Barratt DC. MR to ultrasound registration for prostate challenge - Dataset. 2023. https://doi.org/10.5281/zenodo.8004388

[33] Natarajan S, Priester A, Margolis D, Huang J, Marks L. Prostate MRI and ultrasound with pathology and coordinates of tracked biopsy (Prostate-MRI-US-Biopsy) [Data set]. The Cancer Imaging Archive (TCIA). 2020.

[34] Benelli A, Vaccaro C, Guzzo S, Nedbal C, Varca V, Gregori A. The role of MRI/TRUS fusion biopsy in the diagnosis of clinically significant prostate cancer. Ther Adv Urol. 2020 May 18;12:1756287220916613. https://doi.org/10.1177/1756287220916613

[35] Liu W, Kadier A, Shen D, He Y, Dong S, Zhu K, et al. Combined MRI-TRUS fusion targeted and systematic biopsy versus systematic biopsy alone for the detection of prostate cancer: protocol for a prospective single-centre trial. BMJ Open. 2024 Mar 1;14(3):e080593. https://doi.org/10.1136/bmjopen-2023-080593

[36] Wu Y, Lu Q, Liu Z, Wu J, He X, Yang Y, et al. Advances in imaging and artificial intelligence for precision diagnosis and biopsy guidance in prostate cancer. Front Oncol. 2025 Oct 13;15:1614891. https://doi.org/10.3389/fonc.2025.1614891

[37] Soerensen SJC, Fan RE, Seetharaman A, Chen L, Shao W, Bhattacharya I, et al. Deep learning improves speed and accuracy of prostate gland segmentations on magnetic resonance imaging for targeted biopsy. J Urol. 2021 Sep;206(3):604-612. https://doi.org/10.1097/ju.0000000000001783

[38] Palladino L, Maris B, Antonelli A, Fiorini P. PROST-Net: A deep learning approach to support real-time fusion in prostate biopsy. IEEE Trans Med Robot Bionics. 2022;4(2):323-326. https://doi.org/10.1109/tmrb.2022.3145667

[39] Wang B, Yu L, Pan Z, Ni H, Lin Z, Fan Y. MRI/TRUS-guided auxiliary system for transperineal prostate biopsy based on deep learning. 2023 6th International Conference on Information Communication and Signal Processing (ICICSP). 2023;914-921. https://doi.org/10.1109/ICICSP59554.2023.10390795

[40] Alabousi M, Ghai S. Magnetic resonance imaging-guided ultrasound ablation for prostate cancer – A contemporary review of performance. Front Oncol. 2023 Jan 4;12:1069518. https://doi.org/10.3389/fonc.2022.1069518

[41] Karwacki J, Kiełbasa J, Szczepaniak Z, Zagórski K, Kobylański M, Gurwin A, et al. Current status of cryoablation in prostate cancer management. Clin Med Insights Oncol. 2025 Jul 3;19:11795549251350830. https://doi.org/10.1177/11795549251350830

[42] Beek E, Hata N, Tuncali K, Moreira P. Image-guided adaptive cryotherapy for prostate cancer treatment. Ann Biomed Eng. 2025 Nov;53(11):3237-3246. https://doi.org/10.1007/s10439-025-03833-9

[43] van Luijtelaar A, Fütterer JJ, Bomers JGR. Minimally invasive magnetic resonance image-guided prostate interventions. Br J Radiol. 2022 Mar 1;95(1131):20210698. https://doi.org/10.1259/bjr.20210698

[44] Chen Y, Xing L, Yu L, Liu W, Pooya Fahimian B, Niedermayr T, et al. MR to ultrasound image registration with segmentation-based learning for HDR prostate brachytherapy. Med Phys. 2021 Jun;48(6):3074-3083. https://doi.org/10.1002/mp.14901

[45] Yang Y, Zhang H, Gichoya JW, Katabi D, Ghassemi M. The limits of fair medical imaging AI in real-world generalization. Nat Med. 2024 Oct;30(10):2838-2848. https://doi.org/10.1038/s41591-024-03113-4

[46] Kushol R, Wilman AH, Kalra S, Yang Y. DSMRI: Domain shift analyzer for multi-center MRI datasets. Diagnostics (Basel, Switzerland). 2023 Sep 14;13(18):2947. https://doi.org/10.3390/diagnostics13182947

[47] Karnik VV, Fenster A, Bax J, Cool DW, Gardi L, Gyacskov I, et al. Assessment of image registration accuracy in three-dimensional transrectal ultrasound guided prostate biopsy. Med Phys. 2010 Feb;37(2):802-813. https://doi.org/10.1118/1.3298010

[48] Chaddad A, Peng J, Xu J, Bouridane A. Survey of explainable AI techniques in healthcare. Sensors (Basel, Switzerland). 2023 Jan 5;23(2):634. https://doi.org/10.3390/s23020634