

MEDICAL ARTIFICIAL INTELLIGENCE

ISSN: 2957-5524

Published by ZENTIME PUBLISHING CORPORATION

Volume 1 Number 1





A survey on the application of deep learning in knee joint cartilage ultrasound image segmentation

Jintao Duan¹, Miao Zhou², Yuxiang Wang¹, Fangfang Chen¹, Liangqing Lin³, Qinghua Wu³, Haipo Cui¹

¹School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. ²Jiangsu Cancer Hospital, Nanjing 213164, Jiangsu Province, China. ³The First Hospital of Putian, Putian 351100, Fujian Province, China.

Corresponding author: Haipo Cui.

Declaration of conflict of interest: None.

Received September 27, 2024; Accepted January 22, 2025; Published March 31, 2025

Highlights

- A systematic review of deep learning (DL) approaches for femoral cartilage segmentation in knee joint ultrasound images.
- Evaluation of popular DL models (U-Net, U-Net++, Siam U-Net, Mask R-CNNs) using various metrics.
- Discussion on dataset scarcity, data preprocessing methods, and future directions for DL-based ultrasound segmentation.

Abstract

Objectives: The femoral cartilage in the knee joint is prone to degenerative changes and injuries, often requiring Magnetic Resonance Imaging as the diagnostic gold standard. However, due to the high cost and limited availability of Magnetic Resonance Imaging, ultrasound is explored as a viable alternative. This paper presents a comprehensive review of current deep learning (DL) strategies for knee femoral cartilage segmentation in ultrasound images, focusing on commonly used datasets, data preprocessing techniques, and state-of-the-art DL models. **Methods:** We systematically reviewed the literature from major medical and engineering databases, summarizing key contributions to knee femoral cartilage segmentation. We focused on (1) the scarcity of large-scale public ultrasound datasets and its impact on model training, (2) popular DL architectures (e.g., U-Net variants, Mask Region-based Convolutional Neural Network), and (3) evaluation techniques, particularly the Dice Similarity Coefficient. We also examined image preprocessing and data augmentation strategies aimed at mitigating data insufficiency. **Results:** Our review shows that U-Net and its variants (e.g., Siam U-Net, U-Net++) commonly achieve competitive Dice Similarity Coefficient values (around 0.70–0.80) for knee cartilage segmentation, despite the limitations in training data. Advanced networks like Mask Region-based Convolutional Neural Network, when combined with robust image preprocessing and transfer learning (e.g., using COCO/ImageNet pretrained weights), can improve Dice Similarity Coefficient by over 20%. However, the absence of standardized public datasets limits direct comparisons between studies and affects reproducibility. **Conclusion:** DL holds significant potential for accurate and cost-effective femoral cartilage segmentation in knee joint ultrasound images, offering a feasible alternative or complement to Magnetic Resonance Imaging-based assessments. However, challenges remain due to the lack of large-scale, open-access ultrasound datasets and inconsistent evaluation protocols. Future work should focus on establishing public benchmarks, refining novel network architectures, and enhancing real-time clinical deployment to foster wider adoption and greater clinical impact.

Keywords: Deep learning, knee joint cartilage, ultrasound image segmentation

Introduction

The knee joint, one of the body's largest and most complex structures, endures substantial stress during daily activities. Femoral cartilage

damage, commonly detected during arthroscopic examinations, contributes to patient discomfort, functional impairment, and decreased quality of life [1]. As populations age and lifestyles evolve, the prevalence of knee joint disor-



Table 1. Literature on knee femoral cartilage segmentation using deep learning and ultrasound imaging

Methods	Evaluation metrics	Authors
U-Net	DSC ROI DSCUB	Antico et al.(2020) [23]
U-Net++	DSC A/VPD HD	du Toit et al. (2022) [24]
Siam-U-Net	DSC CE	Dunnhofer et al. (2020) [25]
Mask R-CNNs	DSC, Jaccard	Kompella et al. (2019) [22]

Note: DSC, Dice Similarity Coefficient; CNNs, Convolutional neural networks; ROI, region of interest; Mask R-CNNs, Mask Region-based Convolutional Neural Networks.

ders has risen, including degenerative knee osteoarthritis, cartilage injuries, chondromalacia, and subchondral bone fractures [2]. Accurate diagnosis and timely treatment are crucial for pain relief, mitigating disease progression, and restoring knee function. Although magnetic resonance imaging (MRI) is the current diagnostic gold standard, its high cost, lengthy scan times, and limited availability reduce its practicality. Consequently, Schmitz proposed ultrasound as an effective alternative, citing its lower cost, real-time imaging capability, suitability for use in operating rooms, and value for surgical guidance [3-5]. Ultrasound-guided arthroscopic surgeries, as demonstrated by Wu, highlight its growing role in clinical applications [6].

Medical image segmentation plays a vital role in clinical tasks such as diagnosis, treatment planning, and disease monitoring, with ultrasound imaging being particularly favored for its non-invasive, real-time, and cost-effective features. However, ultrasound image segmentation presents inherent challenges due to artifacts, intensity inhomogeneities, operator-dependent probe placement, soft tissue deformation, and motion artifacts related to knee joint structures [7, 8]. These issues complicate segmentation [9]. To address these challenges, researchers have proposed various solutions, ranging from traditional digital image processing to machine learning. For example, Srensen et al. developed an automated bone segmentation method using dynamic programming, while Faisal et al. applied the Local Statistical Level Set Method for cartilage segmentation, achieving an average Dice Similarity Coefficient (DSC) of 0.91 [10-12]. Among the traditional (i.e., non-deep learning) segmentation approaches, Srensen et al. and Faisal et al. both demonstrated how explicit, model-based methods can be applied to medical image segmentation [10, 11]. Dice et al. employed a dynamic programming technique, while Faisal et al. adopted the Local Statistical Level Set Method for knee cartilage segmentation, achieving an average DSC of 0.91 [12]. Although their studies targeted different anatomical structures (bone vs. cartilage), both exemplify the effectiveness of traditional algo-

rithms in enhancing segmentation accuracy. Despite these advances, traditional methods remain time-consuming and often require manual parameter adjustments for accurate feature delineation.

Deep learning (DL) has revolutionized computer vision, with algorithms that excel across a range of medical imaging applications. Key milestones include the development of convolutional layers and architectures such as Alex-Net, U-Net, region-based convolutional neural networks (R-CNNs), and Transformer models [13-17]. DL techniques have shown impressive success in ultrasound image segmentation across various anatomical sites, such as the breast, thyroid, distal humerus cartilage, and left atrium [18-21]. However, knee ultrasound segmentation continues to present unique challenges. Recent approaches, including Mask region-based convolutional neural networks (Mask R-CNNs) introduced by Gayatri Kompella and Maria Antico, show promise for high-quality, fully automated segmentation, although research remains limited in this area [22].

This paper reviews the application of DL to femoral cartilage segmentation in knee ultrasound images. As shown in **Table 1**, we highlight four representative articles covering segmentation methods, publication years, and evaluation metrics in this field. In addition to the content in **Table 1**, Section 3 discusses additional DL architectures and various modifications applied to different knee structures, providing a more comprehensive overview. The organization of this paper is as follows: Section 2 introduces data preprocessing techniques for knee ultrasound images; Section 3 covers commonly used DL architectures, including those in **Table 1** and other relevant studies, along with evaluation metrics. Finally, Sections 4 and 5 present the discussion and conclusions, respectively.

Data and preprocessing

Dataset

In DL model training, data is a crucial compo-

Table 2. Medical image dataset

Datasets	Imaging modality	Anatomical region	Number of images
MRNet [30]	MRI	Knee	1,370
FastMRI [31]	MRI	Knee and cranial nerve	11,500
TN-SCUI 2020 [32]	Ultrasonic	Thyroid nodule	4,554
CAMUS [33]	Ultrasonic	Left ventricle and left atrial endocardium	1,000

Note: MRI, magnetic resonance imaging.

nent. It has been shown that extensive training on large datasets enhances the ability of models to learn accurate weights and parameters, thereby significantly improving performance. However, unlike other domains, medical image segmentation often faces challenges related to data scarcity and reliance on non-public datasets. These issues can lead to suboptimal model performance and problems such as overfitting, particularly when the available data is limited [26]. Creating large, publicly accessible datasets in the medical domain is particularly difficult due to patient privacy concerns and unauthorized access, leading most datasets to remain private [27]. Additionally, image annotation often requires manual labeling by experienced physicians. It is not feasible for a single doctor to annotate large datasets, so the involvement of multiple medical staff is necessary, which may result in inconsistencies in annotation standards due to individual differences among doctors. Moreover, the high financial cost of acquiring medical image data necessitates the involvement of volunteers and patients, and, due to safety concerns, some images cannot be collected in large quantities, further exacerbating data scarcity. The lack of sufficient data can lead to significant positive and negative sample biases due to the unique characteristics of medical data, ultimately hindering effective model training.

Models typically perform better when large public datasets are available. Several studies have successfully utilized public datasets, such as the Automatic Cardiac Diagnosis Challenge and the Multimodal Brain Tumor Segmentation Challenge, to achieve notable results and improve model performance [28, 29]. These datasets include a diverse range of imaging data, such as ultrasound, MRI, and computed tomography scans, and have contributed significantly to advancements in DL within the medical field. Given the current lack of a public dataset for knee ultrasound images, **Table 2** presents several public ultrasound databases, as well as datasets containing knee images, which can serve as a starting point for future research.

Data preprocessing

High-quality original medical images can be used for model training and yield reasonable results. However, implementing appropriate preprocessing steps can further enhance the utility of these images. While preprocessing cannot directly address the issue of limited data volume, it can alleviate model-fitting and performance challenges associated with data scarcity. Image enhancement techniques, such as adjusting brightness and contrast and applying blurring, combined with data augmentation methods (e.g., random flips and rotations), help expand the dataset, thus diversifying the training set and mitigating data limitations.

In medical data, data imbalance is a common challenge. Techniques like oversampling and undersampling have proven effective in reducing this issue. Ultrasound images, in particular, often suffer from speckle noise, leading to suboptimal signal-to-noise ratios. Kompella demonstrated that applying Gaussian filtering to remove high-frequency noise, combined with fine-tuning blurring parameters, reduces noise while preserving the clarity of anatomical structures, such as femoral cartilage [22]. Similarly, Antico et al. employed Matrix Laboratory (MATLAB) to segment 3D ultrasound images of the knee joint into 2D slices [23]. By adjusting pixel dimensions and filling in black pixels, they achieved favorable training results. Tang used thin-plate spline interpolation to model biological deformations and expand the dataset [34]. This interpolation method, which smoothly approximates sparse data points, supports biological variability, increasing the dataset and enhancing the model's generalization and performance, ultimately improving image model training.

Methods and Materials

Convolutional neural networks (CNNs)

CNNs are widely used in computer vision and DL due to their ability to extract features through convolution. This makes them particu-

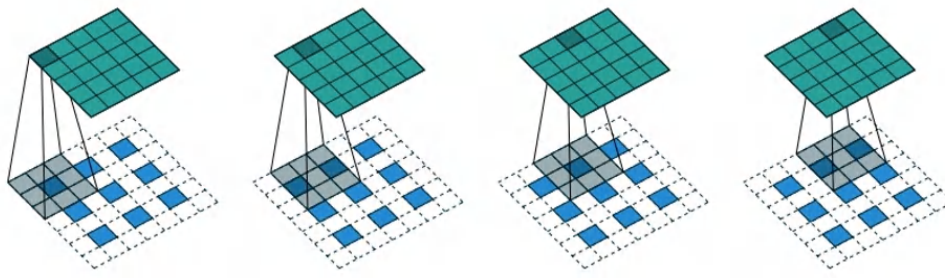


Figure 1. The fundamental architecture of CNNs. This figure is cited from [35]. CNNs, convolutional neural networks.

larily suited for tasks such as image classification, object detection, and image segmentation (**Figure 1**). CNNs are structured in a hierarchical layer configuration, with each layer processing the output from the previous one. The first layer, the input layer, directly receives the images. Subsequent layers filter the input data to identify local features, using filters (or kernels) that are adjusted through backpropagation to optimize their values. The size of these kernels is determined by the model designer.

In CNNs, the area of an input image influenced by a neuron is called its receptive field. Larger kernels expand this field, enabling the model to capture more contextual information. This capability enhances the model's ability to distinguish between targets and backgrounds, improving its reasoning. However, enlarging the receptive field can increase computational demands and reduce local detail resolution.

Activation layers are typically added to introduce non-linearity, and pooling layers are used to reduce the dimensionality of the output, thus lowering the computational load. Pooling can be categorized into max pooling, which highlights features and preserves finer details, and average pooling, which may reduce overfitting. Connected layers are then used to identify higher-level features. During training, kernel weights and neural connections are iteratively optimized through backpropagation until an optimal setup is achieved.

This canonical CNNs framework has paved the way for architectures such as LeNet and GoogLeNet [36, 37]. Depending on the input image type, CNNs can be 2D, 2.5D, or 3D. The 2.5D configuration, which captures spatial information more comprehensively than 2D but requires fewer computational resources than 3D, has been shown to be effective. Prasoon et al. demonstrated its utility by creating a 2.5D representation of a knee joint by aggregating three orthogonal 2D MRI image slices (XY, YZ,

and XZ) [38]. They successfully segmented knee cartilage using this 2.5D approach.

U-Net

Following the advent of CNNs, U-Net was introduced as a groundbreaking neural network architecture, as shown in **Figure 2** [15]. Prior to the innovative proposal of U-Net by Ronneberger et al., efforts to segment medical images had been made but largely produced suboptimal results [15]. The introduction of U-Net, however, significantly advanced medical image segmentation, culminating in a victory at the 2015 ISBI Cell Tracking Challenge. U-Net represents a pivotal development in this field, with numerous algorithms subsequently built upon this model, achieving notable successes.

As shown in **Figure 2**, the U-Net architecture consists of two main components: an encoder (contracting pathway) and a decoder (expanding pathway). The encoder, similar to a conventional CNNs, uses layers of down-sampling and convolution to extract features from images. The decoder, on the other hand, employs up-sampling and skip connections to the contracting pathway, which enhances the resolution of features extracted by the encoder and transmits them to the final convolution layer, resulting in a fully segmented image. The name "U-Net" comes from the distinctive U-shaped architecture formed by the encoder-decoder interaction.

Antico et al. utilized a five-layer U-Net model to process 3D knee ultrasound images by segmenting them into 2D slices along the sagittal plane [23]. These slices were then used for model training, enabling successful segmentation of the distal femoral cartilage. They applied the Dice coefficient as the loss function, ReLU for the activation function, and Adam as the optimizer, which contributed to promising results [39, 40].

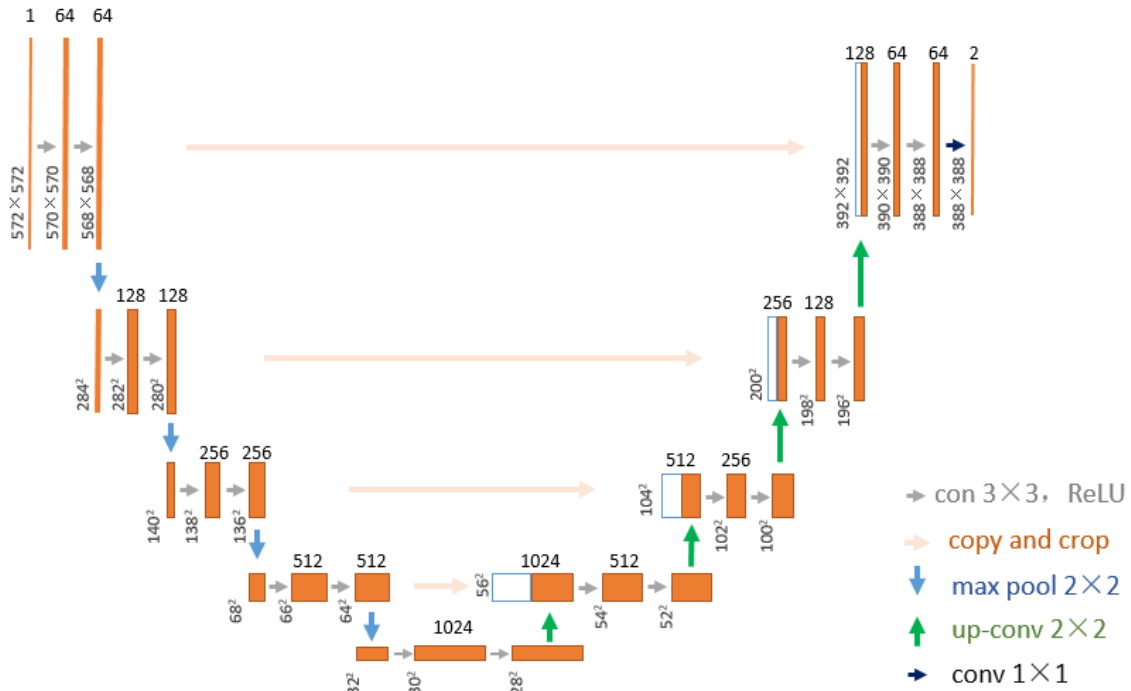


Figure 2. The fundamental architecture of U-Net. This figure is cited from [15].

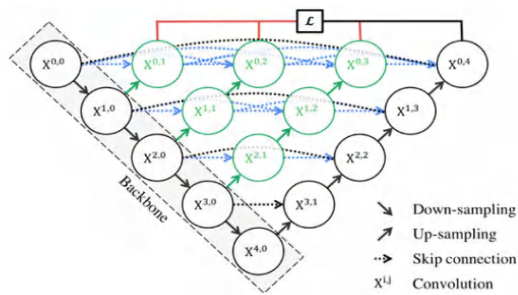


Figure 3. The fundamental architecture of U-Net++. This figure is cited from [41].

U-Net++

Building on the U-Net framework, U-Net++ represents an evolution in neural network architectures for image segmentation, as illustrated in Figure 3 [41]. This model features an interconnected network structure, enhancing its ability to capture features and understand contextual information. U-Net++ incorporates a branching system that combines features from both the encoder and decoder levels, improving feature transmission and information integration. These branching connections enhance the model's ability to capture context and preserve fine details, leading to more accurate semantic segmentation. Compared to the original U-Net, these architectural improvements make U-Net++ more effective at extracting contextual knowledge, optimizing parameter usage, and enhancing robustness.

When applied to imaging tasks, the 3D U-Net++ model benefits from the unique data characteristics of 3D images compared to 2D images. Through convolution and pooling operations across dimensions, the 3D U-Net++ model excels at capturing and utilizing the inherent information in 3D image data. In fields requiring the analysis of volumetric shapes and structures, such as medical image segmentation, the 3D U-Net++ model is particularly effective at delineating contours and configurations. Its ability to extract features in three dimensions allows it to retain more information during convolution and pooling than its 2D counterpart. However, it comes with the drawback of reduced efficiency compared to the 2D version, due to the larger amount of data in 3D medical imaging. On the other hand, acquiring image data for 2D U-Net models is simpler and has a broader range of applications.

Addressing the limitations of 2D ultrasound imaging, which heavily relies on probe placement and operator skill, and exploring bedside alternatives that compensate for MRI's constraints, du Toit et al. employed a U-Net++ model with a ResNet-50 backbone instead of conventional U-Net models [24, 42]. They partitioned 3D ultrasound images into 2D slices and applied data augmentation techniques, including random assortments, translations, rotations, and scalings, to train the model. After reconstructing the 2D ultrasound slices, the predictions were reinserted into their respective slots within the 3D volume, linking the boundary points

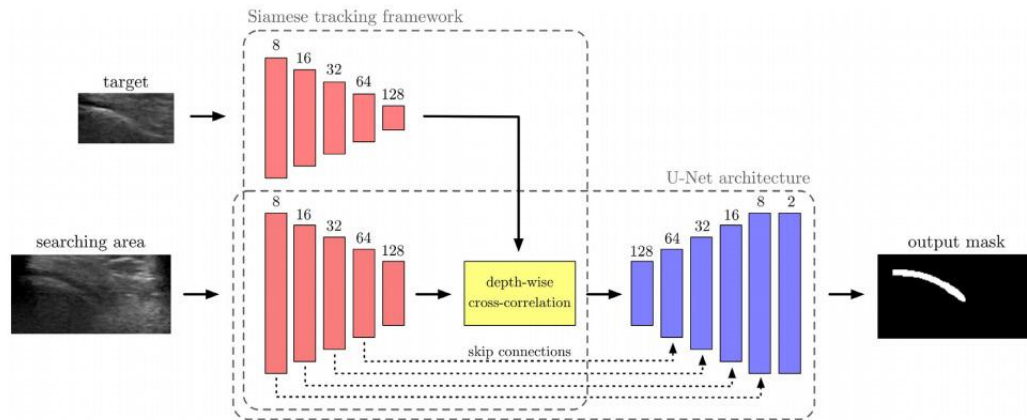


Figure 4. The fundamental architecture of Siam U-Net. This figure is cited from [25].

of adjacent slices. A smoothing filter was then applied, and the volume's extremities were sealed, resulting in a complete 3D surface. Despite these advancements, U-Net demonstrated a slight edge over U-Net++ in all evaluated metrics.

Siam U-Net

The Siam U-Net model is an enhanced version of the original U-Net architecture, specifically designed to improve performance in knee cartilage image segmentation. Dunnhofer et al. successfully implemented this model for segmenting ultrasound images of knee cartilage, achieving significant results [25]. The model's innovation lies in integrating the U-Net with the Siamese framework, as shown in **Figure 4** [43, 44]. It accepts two cropped images as inputs: one for preliminary segmentation and the other for finer segmentation. This dual-level segmentation approach allows the model to process larger images while preserving small details.

The images are processed through an encoder consisting of five computational units, each containing a 3x3 convolutional layer and a 2x2 max pooling operation. The encoder further refines each convolutional layer using batch normalization, ReLU activation, and a dropout layer. After processing, a depth-level cross-correlation operation is performed on the feature maps. Zero-padding is then applied to these cross-correlated feature maps to match the dimensions of the search area embedding. This enables the comparison of the target cartilage image with a specific slice region, resulting in a richer similarity map than standard cross-correlation operations.

The output binary mask is generated through a decoder branch sequence, consisting of four blocks: bilinear upsampling of the previous fea-

ture map, followed by a 2x2 convolution; connection with the feature representations from each encoder block in the search area (referred to as skip connections); and two 3x3 convolutional layers. Finally, a 1x1 convolutional layer is used to generate the output segmentation, consisting of two channels: one for predicting the foreground object (cartilage) and the other for predicting the background pixels.

The incorporation of depth-level cross-correlation and skip connections enables more accurate cartilage segmentation within the search area. Compared to the original U-Net, the Siam U-Net features smaller computational block sizes, reducing computational load and increasing processing speed. Additionally, the inclusion of a dropout layer improves the model's generalization capabilities.

Mask R-CNNs

Kompella and colleagues utilized a Mask R-CNNs for cartilage image segmentation [22]. In their study, 3D images of the cartilage were sliced sagittally into 256 2D slices, each with dimensions of 272x510 pixels, which were then used for training, as shown in **Figure 5**.

Mask R-CNNs is an extension of the R-CNNs model, designed for object detection and segmentation tasks. It excels at both detecting object locations and generating corresponding segmentation masks. The Mask R-CNNs architecture consists of four components: a backbone network, a Region Proposal Network (RPN), an object detection network, and a pixel-level mask generation network. The backbone network, such as ResNet or VGG, is responsible for extracting image features [45]. By utilizing this backbone, Mask R-CNNs extracts high-level features from the input images to support subsequent object detection and seg-

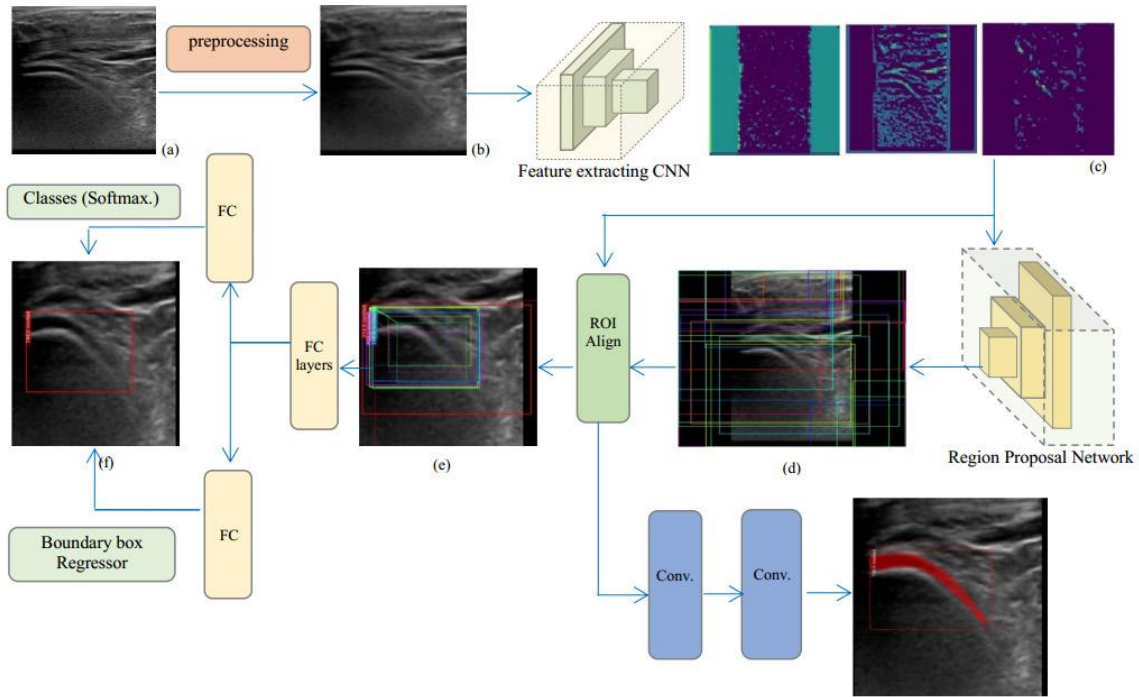


Figure 5. The fundamental architecture of Mask R-CNNs. (a) The original US image is fed into a feature-extracting CNNs; (b) The RPN generates candidate ROIs; (c) Each ROI is precisely aligned via ROI Align, avoiding misalignments from standard pooling; (d) FC layers predict object classes (Softmax) and refine bounding boxes (regressor); (e) A separate branch outputs instance-level segmentation masks; (f) Final predicted bounding box obtained from the bounding box regressor, along with the object probability obtained from the softmax classifier, after passing through the fully convolutional network. Through this multi-branch approach, Mask R-CNNs concurrently performs object detection and segmentation, enabling automatic knee cartilage analysis in ultrasound images. This figure is cited from [22]. US, ultrasound; CNNs, convolutional neural networks; RPN, Region Proposal Network; ROIs, regions of interest; FC, Fully connected; Mask R-CNNs, Mask Region-based Convolutional Neural Networks.

mentation tasks.

The RPN component generates object regions within the image. It is a convolutional network that captures features at various positions and scales using a sliding window approach, determining whether an object is present at each location. The RPN also offers boundary box regression adjustments for each region. The object recognition network then takes these regions, classifies them, and refines the bounding boxes for precise object delineation. This network typically consists of fully connected layers or a combination of connected layers.

The unique feature of Mask R-CNNs is its pixel-level mask generation network, which receives candidate regions from the RPN, extracts corresponding features, and generates pixel-level masks for the object using a Fully Convolutional Network [46]. This allows the network to represent each object instance with a detailed mask. Mask R-CNNs thus performs simultaneous object detection and pixel-level segmentation, providing both accurate object localization and detailed segmentation results. The backbone network extracts high-level se-

mantic features, the RPN generates candidate regions, the object detection network classifies and refines the bounding boxes, and the mask generation network produces pixel-level masks for each object.

Evaluation metrics

Model performance in segmentation is typically evaluated using a variety of metrics. Relying on a single metric to assess a model's performance is unreliable; hence, using a combination of metrics provides a more comprehensive evaluation. Among these, the DSC is commonly regarded as the most reliable metric for evaluating segmentation models, measuring the overlap between two binary masks [12]. Segmentation accuracy is quantitatively assessed by comparing the model's output with manual annotations, considered the gold standard. Precision is typically evaluated through overlap measurements or surface distance. The DSC, which measures the overlap between two regions, is employed in this study to compare the proposed method's segmentation results with the ground truth. As shown in Equation 1, $R1$ represents the ground truth area, while $R2$ is

Table 3. Comparison of Mean and Min/Max Dice Similarity Coefficients (DSC) Achieved by U-Net and Intra-Observer Expert

Results	Mean DSC		Min/Max DSC	
	U-Net	Intra-observer	U-Net	Intra-observer
1	0.65	0	0/0.95	
2	0.68	0.77	0/0.94	0/0.95
3	0.69	0.73	0/0.95	0/0.92
4	0.68	0.40	0/0.94	0/0.92
5	0.71	0.70	0/0.94	0/0.92
Overall	0.68	0.64	-	-

Note: DSC, Dice Similarity Coefficient.

Table 4. Comparison of mean-DSC achieved by 2D Transverse and 3D Reconstruction Segmentation Performance Using U-Net and U-Net++

Results	U-Net Mean-DSC	U-Net++ Mean-DSC
2D Transverse	87.10±4.60	83.80±2.90
3D Reconstruction	73.10±3.90	74.30±1.70
P-value	0.063	0.064

Note: DSC, Dice Similarity Coefficient.

the segmented area. A higher DSC value indicates better segmentation accuracy. Given the diverse evaluation metrics used in the reviewed studies, the DSC was selected as the primary evaluation criterion for this study.

$$DSC = 2 * \frac{|R1 \cap R2|}{|R1| + |R2|} \quad (1)$$

Results

Performance of U-Net on 2D ultrasound data

The U-Net architecture has become a pivotal model in medical image segmentation, serving as the foundation for numerous derivatives, while maintaining strong performance on its own. Antico et al. applied the U-Net model to segment femoral cartilage images from knee ultrasound scans, obtaining promising results [23]. A comparative analysis was conducted between the DSC obtained from the U-Net model and those from expert manual annotations. **Table 3** presents the calculated average DSC, along with the maximum and minimum DSC values. As shown in the table, the U-Net model achieved an average DSC between 0.65 and 0.71, with an overall mean of 0.68. In comparison, expert annotations across test sets (Test Sets 1 to 5) ranged from 0 to 0.77, with an average of 0.64. Zero DSC values from experts, as seen in Test 1 of **Table 3**, indicate an anomalous case. The data suggests that the U-Net model performs well in image segmentation, with its DSC values approximately 1-2% higher

than those of the experts for both complete images and selected regions. This implies that the U-Net's segmentation performance is either comparable to or slightly better than that of human experts. However, both the U-Net and expert segmentations are less accurate in images with blurred boundaries or poorly defined cartilage. To improve performance, specialized strategies may be needed to address such challenging cases.

Performance of U-Net++ on 2D/3D ultrasound data

The U-Net++ model is an enhanced version of the original U-Net, featuring a more complex network architecture with additional layer hierarchies and branching connections. This modification enhances the model's ability to capture more detailed features and integrate contextual information. However, its segmentation performance does not always surpass that of the original U-Net model, and under certain conditions, it may produce inferior results.

du Toit et al. applied both U-Net and U-Net++ models to 3D knee ultrasound images, which were sliced into 2D images for training [24]. The 2D segmented images were then reassembled into their original 3D structure, connecting the boundaries of adjacent slices. A smoothing filter was applied, and the volume extremes were sealed to reconstruct a continuous 3D surface. In their study, du Toit et al. assessed the mean DSC between the 2D slices and the

Table 5. Comparison of mean DSC among Siam-U-Net, Temporal and Spatio-Temporal Tracking, U-Net, and Léger's U-Net

Results	Mean DSC		Mean DSC		
	Temporal tracking	Spatio-temporal	Siam-U-Net	U-Net	Léger's U-Net [47]
1	0.74±0.15	0.73±0.16	0.74±0.15	0.61±0.24	0.60±0.23
2	0.69±0.20	0.71±0.16	0.69±0.20	0.62±0.22	0.65±0.23
3	0.69±0.16	0.70±0.15	0.69±0.16	0.68±0.18	0.68±0.21
4	0.69±0.17	0.68±0.18	0.69±0.17	0.62±0.23	0.64±0.22
5	0.73±0.14	0.73±0.14	0.73±0.14	0.66±0.21	0.67±0.20
6	0.69±0.15	0.68±0.16	0.69±0.15	0.63±0.23	0.62±0.28
Overall	0.68	0.71±0.16	0.70±0.16	0.64±0.22	0.64±0.23

Note: DSC, Dice Similarity Coefficient.

Table 6. Comparison of segmentation results with and without preprocessing, using ImageNet, COCO, and the Level Set Method

Results	Pre-processing	Mean-DSC
Pretrain with ImageNet [48]	No	0.22
	Yes	0.64
Pretrain with COCO [49]	No	0.67
	Yes	0.80
Level Set [50]	No	0.52
	Yes	0.54

Note: DSC, Dice Similarity Coefficient.

reconstructed 3D images, and evaluated the P-value to analyze the variance in segmentation performance, as shown in **Table 4**. The results showed that the segmentation quality of the 2D images was significantly better than that of the 3D reconstructions. This discrepancy was due to the reconstruction method, which produced rectangular surfaces that intersected three distinct regions, as manually segmented. Additionally, the smoothing filter simplified the 2D predictions, lowering the DSC. However, it also removed outliers, making the segmentation results for both 2D and 3D images more comparable for both U-Net models.

Further performance tests with knee osteoarthritis images revealed no significant differences in segmentation outcomes between U-Net and U-Net++, aligning with results from healthy knee images.

Performance of Siam-U-Net on temporal and spatio-temporal sequences of 2D ultrasound data

Siam-U-Net is an innovative variant of the U-Net neural network, integrating a Siamese tracking framework. Dunnhofer et al. collected a dataset comprising 35 3D+time series knee

ultrasound images [25]. Manual segmentation of 2D ultrasound images was performed by an orthopedic surgeon, generating temporal and spatio-temporal 2D+time series from the 3D+time series data. A total of 18,278 slices within the sagittal planes of the ultrasound volume were annotated. The dataset was split into six subsets, with one used for testing and five for training. The model was applied to segment femoral cartilage in 2D ultrasound images. The average DSC for both temporal and spatio-temporal tracking were compared, alongside the performance of the original U-Net and the U-Net variant proposed by Léger et al., as shown in **Table 5** [47]. The results indicate that Siam-U-Net outperforms these comparative models in terms of performance metrics.

Performance of Mask R-CNNs on 2D ultrasound data

Mask R-CNNs, an advanced extension of R-CNNs, simultaneously detects object locations and generates their masks. Using a 4D ultrasound probe (5-13 MHz frequency range), Kompella et al. acquired static 3D ultrasound images of knee joints from healthy volunteers [22]. These images were manually segmented by an experienced physician and split into 256

2D slices in the coronal plane, with 85% allocated to the training set and 15% to the test set. Preprocessing was done according to COCO and ImageNet datasets before training. The efficacy of Mask R-CNNs was evaluated using the Mean DSC, comparing segmentation accuracy among networks pretrained on COCO and ImageNet datasets to those with no pretraining. Additionally, Mask R-CNNs' performance was compared with the conventional level-set segmentation method, which requires manual initialization. As shown in **Table 6**, networks pretrained on the COCO2016 dataset exhibited a 20% increase in DSC over those without pretraining. Pretraining with the ImageNet dataset resulted in a 190% improvement in performance, surpassing the level-set technique's accuracy.

Performance comparison with different models

Comparing DSC scores (U-Net: 0.68, Intra-observer: 0.64), U-Net slightly outperforms intra-observer segmentation. Furthermore, in recognizing 2D (87.1, 73.1) transverse and 3D reconstructed images of femoral cartilage, U-Net outperforms U-Net++ (83.8, 73.1). This is due to U-Net++'s increased structural complexity, stemming from additional pathways and skip connections that enable it to incorporate more image information. Despite these improvements, optimal results are not achieved in femoral cartilage image recognition. The integration of the U-Net model with a Siamese network structure in Siam-U-Net allows for the use of two cropped images as inputs, one for preliminary segmentation and the other for precise segmentation. This multi-level segmentation architecture improves the model's ability to preserve fine details while handling larger images, resulting in a slight increase in DSC from U-Net (0.64) to Siam-U-Net (0.70). Preprocessing also plays a significant role in performance, as demonstrated by DSC scores in Mask R-CNNs. Preprocessing the COCO dataset increased the DSC score from 0.67 to 0.80, with similar enhancements observed in other datasets. These improvements were due to the extraction of relevant features, noise reduction, and the use of data augmentation techniques that enriched the data and improved the model's generalization capability.

Discussion

Summary

The literature reviewed in this article demonstrates the partial application of DL for the

segmentation of femoral cartilage in knee ultrasound images. The article begins with an introduction to common data types used in DL, followed by an overview of publicly available datasets and preprocessing methods for data management. Various methods suitable for femoral cartilage segmentation in knee ultrasound images are discussed, with a focus on the U-Net architecture and its variants, which have been predominantly used. Other models, such as Mask R-CNNs, have also shown promising results. During the review, it was noted that some studies employed less common evaluation metrics, and many relied on non-public datasets. The performance of various models is highlighted, with the DSC being the most frequently used evaluation metric.

Challenges

Despite these advancements, challenges remain in the DL segmentation of femoral cartilage in knee ultrasound images. First, there is a lack of large-scale, publicly accessible datasets specifically for knee femoral cartilage ultrasound imaging. While data augmentation and the fusion of supervised and unsupervised learning can help alleviate this issue, the effectiveness of the models is still constrained by their inherent limitations. Thus, the creation of a standardized, publicly available dataset is crucial. Second, early research on knee femoral cartilage ultrasound segmentation often failed to incorporate the latest neural network models and architectures, leading to performance gaps compared to newer approaches. Ongoing research is expected to address these gaps. Additionally, there is growing interest in intelligent medical devices for medical image processing. However, the significant computational power required for processing medical images presents challenges in ensuring both accuracy and real-time performance when deploying models on medical devices. The model size is also a critical factor. These challenges may be gradually overcome with advancements in the miniaturization of DL models for medical image segmentation and the development of intelligent medical devices.

Future research directions

Future research in automated femoral cartilage detection, powered by DL, aims to enhance real-time operational efficiency, accuracy, and user-friendliness of the model. By integrating the model with ultrasound images of femoral cartilage obtained via the probe, immediate identification could be achieved, thus improving the efficiency of healthcare providers and strength-

ening the reliability of clinical decision-making. Another key focus is the development of a comprehensive public dataset for femoral cartilage. As the dataset grows, model detection accuracy is expected to improve. Simultaneously, given the current data limitations, there is a need for progress in identifying various distinct morphological forms of femoral cartilage. Therefore, a major research direction is to enhance the model's ability to generalize and identify these unique cartilage morphologies.

Conclusion

This research reviews the latest advancements in DL for knee ultrasound image segmentation and compiles publicly accessible medical datasets and methods employed in femoral cartilage segmentation. The study also includes an analysis of several evaluation metrics. DL has the potential to significantly enhance the diagnostic efficiency of physicians. For patients in medically underserved areas, online availability of trained models could provide substantial support, improving medical services. This survey aims to offer valuable insights and encourage further research on femoral cartilage segmentation in knee ultrasound images.

Author contributions: Jintao Duan: Conceptualization of the review; Literature search and screening; Writing—original draft preparation; Data interpretation; Final manuscript revision. Miao Zhou: Assisted in literature collection; Contributed to data organization; Reviewed and edited the manuscript for language and structure. Yuxiang Wang: Data curation; Figure/table preparation; Provided feedback on manuscript sections related to deep learning architectures. Fangfang Chen: Literature analysis; Drafted parts of the methodology and discussion; Verified references for accuracy. Liangqing Lin: Provided medical guidance on knee joint clinical aspects; Ensured the clinical relevance of the review; Critically reviewed the manuscript for important intellectual content. Qinghua Wu: Offered medical guidance regarding ultrasound imaging protocols; Gave expert opinions on clinical feasibility; Revised the manuscript sections involving ultrasound data. Haipo Cui (Corresponding Author): Overall supervision; Funding acquisition; Methodology design in deep learning perspective; Final approval and submission of the manuscript.

References

[1] Curl WW, Krome J, Gordon ES, et al. Cartilage injuries: a review of 31,516 knee arthroscopies. *Arthroscopy* 1997;13(4):456-

460.

- [2] Cui A, Li H, Wang D, et al. Global, regional prevalence, incidence and risk factors of knee osteoarthritis in population-based studies. *EClinicalMedicine* 2020;29-30:100587.
- [3] Schmitz RJ, Wang HM, Polprasert DR, et al. Evaluation of knee cartilage thickness: a comparison between ultrasound and magnetic resonance imaging methods. *Knee* 2017;24(2):217-223.
- [4] Antico M, Sasazawa F, Wu L, et al. Ultrasound guidance in minimally invasive robotic procedures. *Med Image Anal* 2019;54:149-167.
- [5] Lueders DR, Smith J, Sellon JL. Ultrasound-guided knee procedures. *Phys Med Rehabil Clin N Am* 2016;27(3):631-648.
- [6] Wu L, Jaiprakash A, Pandey AK, et al. 29 - Robotic and Image-Guided Knee Arthroscopy. In: Abedin-Nasab MH, editors. *Handbook of Robotic and Image-Guided Surgery*. Elsevier, 2020. p. 493-514.
- [7] Vlad VM. Pitfalls in Musculoskeletal Ultrasound. In: El Miedany Y, editors. *Musculoskeletal Ultrasonography in Rheumatic Diseases*. Cham. Springer International Publishing, 2015. p. 21-55.
- [8] Shrimali V, Anand RS, Kumar V. Current trends in segmentation of medical ultrasound B-mode images: a review. *IETE Tech Rev* 2009;26:8-17.
- [9] Noble JA, Boukerroui D. Ultrasound image segmentation: a survey. *IEEE Trans Med Imaging* 2006;25(8):987-1010.
- [10] Srensen TA, Srensen TJ. A method of establishing groups of equal amplitude in plant sociology based on similarity of species and its application to analyses of the vegetation on Danish commons. *Biol Skr Dan Vid Sel* 1948;5:1-34
- [11] Faisal A, Ng SC, Goh SL, et al. Knee cartilage segmentation and thickness computation from ultrasound images. *Med Biol Eng Comput* 2018;56(4):657-669.
- [12] Dice LR. Measures of the amount of ecologic association between species. *Ecology* 1945;26:297-302.
- [13] Fukushima K. Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biol Cybern* 1980;36:193-202.
- [14] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM* 2017;60:84-90.
- [15] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. *Medical Image Computing and Computer-Assisted Intervention – MICCAI* 2015 2015;234-241.

- [16] Girshick R, Donahue J, Darrell T, et al. Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation. 2014 IEEE Conference on Computer Vision and Pattern Recognition 2014;580-587.
- [17] Vaswani A, Shazeer N, Parmar N, et al. Attention is all you need. Proceedings of the 31st International Conference on Neural Information Processing Systems 2017;6000–6010.
- [18] Chen G, Liu Y, Qian J, et al. DSEU-net: A novel deep supervision SEU-net for medical ultrasound image segmentation. Expert Syst Appl 2023;223:119939.
- [19] Chen Y, Zhang X, Li D, et al. Automatic segmentation of thyroid with the assistance of the devised boundary improvement based on multicomponent small dataset. Appl Intell (Dordr) 2023;1-16.
- [20] Zhao W, Su X, Guo Y, et al. Deep learning based ultrasonic visualization of distal humeral cartilage for image-guided therapy: a pilot validation study. Quant Imaging Med Surg 2023;13(8):5306-5320.
- [21] Degel MA, Navab N, Albarqouni S. Domain and Geometry Agnostic CNNs for Left Atrium Segmentation in 3D Ultrasound. Medical Image Computing and Computer Assisted Intervention – MICCAI 2018 2018;630-637.
- [22] Kompella G, Antico M, Sasazawa F, et al. Segmentation of Femoral Cartilage from Knee Ultrasound Images Using Mask R-CNN. Annu Int Conf IEEE Eng Med Biol Soc 2019;2019:966-969.
- [23] Antico M, Sasazawa F, Dunnhofer M, et al. Deep learning-based femoral cartilage automatic segmentation in ultrasound imaging for guidance in robotic knee arthroscopy. Ultrasound Med Biol 2020;46(2):422-435.
- [24] du Toit C, Orlando N, Papernick S, et al. Automatic femoral articular cartilage segmentation using deep learning in three-dimensional ultrasound images of the knee. Osteoarthritis Cartilage 2022;4(3):100290.
- [25] Dunnhofer M, Antico M, Sasazawa F, et al. Siam-U-Net: encoder-decoder siamese network for knee cartilage tracking in ultrasound images. Med Image Anal 2020;60:101631.
- [26] Hesamian MH, Jia W, He X, et al. Deep learning techniques for medical image segmentation: achievements and challenges. J Digit Imaging 2019;32(4):582-596.
- [27] Razzak MI, Naz S, Zaib A. Deep Learning for Medical Image Processing: Overview, Challenges and the Future. In: Dey N, Ashour AS, Borra S, editors. Classification in BioApps: Automation of Decision Making. Cham Springer International Publishing, 2018. p. 323-350.
- [28] Bernard O, Lalande A, Zotti C, et al. Deep learning techniques for automatic MRI cardiac multi-structures segmentation and diagnosis: is the problem solved? IEEE Trans Med Imaging 2018;37(11):2514-2525.
- [29] Menze BH, Jakab A, Bauer S, et al. The multimodal brain tumor image segmentation benchmark (BRATS). IEEE Trans Med Imaging 2015;34(10):1993-2024.
- [30] Bien N, Rajpurkar P, Ball RL, et al. Deep-learning-assisted diagnosis for knee magnetic resonance imaging: development and retrospective validation of MRNet. PLoS Med 2018;15(11):e1002699.
- [31] Zbontar J, Knoll F, Sriram A, et al. fastMRI: An Open Dataset and Benchmarks for Accelerated MRI. ArXiv 2018;abs/1811.08839.
- [32] Shusharina N, Heinrich MP, Huang R. Segmentation, Classification, and Registration of Multi-modality Medical Imaging Data: MICCAI 2020 Challenges, ABCs 2020, L2R 2020, TN-SCUI 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4–8, 2020, Proceedings. 2021;XIX,156.
- [33] Leclerc S, Smistad E, Pedrosa J, et al. Deep learning for segmentation using an open large-scale dataset in 2D echocardiography. IEEE Trans Med Imaging 2019;38(9):2198-2210.
- [34] Tang Z, Duan J, Sun Y, et al. A combined deformable model and medical transformer algorithm for medical image segmentation. Med Biol Eng Comput 2023;61(1):129-137.
- [35] Dumoulin V, Visin F. A guide to convolution arithmetic for deep learning. ArXiv 2016;abs/1603.07285.
- [36] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. Proceedings of the IEEE 1998;86:2278-2324.
- [37] Szegedy C, Wei L, Yangqing J, et al. Going deeper with convolutions. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2015;1-9.
- [38] Prasoon A, Petersen K, Igel C, et al. Deep Feature Learning for Knee Cartilage Segmentation Using a Triplanar Convolutional Neural Network. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013 2013;246-253.
- [39] Glorot X, Bordes A, Bengio Y. Deep Sparse Rectifier Neural Networks. International Conference on Artificial Intelligence and Statistics 2011.
- [40] Kingma DP, Ba J. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980 2014.
- [41] Zhou Z, Rahman Siddiquee MM, Tajbakhsh N,

- et al. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support 2018;3-11.
- [42] He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016;770-778.
- [43] Bertinetto L, Valmadre J, Henriques JF, et al. Fully-Convolutional Siamese Networks for Object Tracking. Computer Vision – ECCV 2016 Workshops 2016;850-865.
- [44] Bertinetto L, Henriques JF, Valmadre J, et al. Learning feed-forward one-shot learners. Proceedings of the 30th International Conference on Neural Information Processing Systems 2016;523–531.
- [45] Simonyan K, Zisserman A. Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 2014.
- [46] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2015;3431-3440.
- [47] Léger J, Brion E, Javard U, et al. Contour Propagation in CT Scans with Convolutional Neural Networks. Advanced Concepts for Intelligent Vision Systems 2018;380-391.
- [48] Deng J, Dong W, Socher R, et al. ImageNet: A large-scale hierarchical image database. 2009 IEEE Conference on Computer Vision and Pattern Recognition 2009;248-255.
- [49] Lin TY, Maire M, Belongie S, et al. Microsoft COCO: Common Objects in Context. Computer Vision – ECCV 2014 2014;740-755.
- [50] Osher S, Sethian JA. Fronts propagating with curvature-dependent speed: Algorithms based on Hamilton-Jacobi formulations. J Comput Phys 1988;79:12-49.



Application of traditional methods and deep learning in breast ultrasound image segmentation

Fangfang Chen¹, Miao Zhou², Jintao Duan¹, Yuxiang Wang¹, Liangqing Lin⁴, Wenhui Guo³, Yongchu Hu⁵

¹School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. ²Jiangsu Cancer Hospital, Nanjing 213164, Jiangsu Province, China. ³School of Anesthesiology, Naval Medical University, Shanghai 200433, China. ⁴The First Hospital of Putian, Putian 351100, Fujian Province, China. ⁵The Department of Anesthesiology, Long March Hospital, Shanghai 200003, China.

Corresponding authors: Wenhui Guo and Yongchu Hu.

Declaration of conflict of interest: None.

Received September 27, 2024; Accepted January 22, 2025; Published April 1, 2025

Highlights

- **Comparison of Methods:** This article compares traditional techniques, such as thresholding and edge detection, with deep learning-based methods for breast ultrasound segmentation.
- **U-Net's Effectiveness:** U-Net is identified as the benchmark for medical image segmentation due to its efficiency and ability to preserve details.
- **Advantages of Deep Learning:** Deep learning models, like CNNs and FCNs, improve segmentation accuracy by learning directly from raw data, while also mitigating noise.

Abstract

Breast ultrasound image segmentation is vital in medical imaging, enabling precise delineation of tissues and lesions, which contributes to the diagnosis and treatment of breast diseases. This article reviews both traditional methods and recent advancements in deep learning techniques for breast ultrasound image segmentation. The discussion begins by highlighting the significance of image segmentation in breast disease diagnosis and its background within medical imaging. Traditional segmentation methods, such as thresholding, edge detection, and region growing, are examined, with an analysis of their applications and limitations in breast ultrasound segmentation. Subsequently, the focus shifts to deep learning approaches, including classic models like Convolutional Neural Networks, Fully Convolutional Networks, and U-Net, along with their improved algorithms. These methods learn hierarchical features directly from raw data, reducing reliance on manual preprocessing. U-Net, in particular, is highlighted as the benchmark for medical image segmentation due to its efficient data usage and ability to preserve fine-grained details. Comparative analysis demonstrates the advantages of deep learning in enhancing segmentation accuracy, reducing noise, and handling complex texture structures. The article concludes by summarizing current achievements and challenges in the field, while offering an outlook on the future developments aimed at advancing breast ultrasound image segmentation for improved diagnosis and treatment of breast diseases.

Keywords: Breast ultrasound image segmentation, medical imaging, deep learning

Introduction

Breast diseases encompass a range of conditions affecting breast tissue, including mastitis, breast hyperplasia, fibroadenoma, and breast cancer, with female being the primary demographic affected [1]. In 2023, the National Health Commission of the People's Republic of China underscored the significance of breast

cancer as a leading malignancy threatening women's health, emphasizing the need for prioritizing the prevention and management of breast diseases.

Early diagnosis is crucial for improving treatment outcomes and survival rates for breast cancer patients. Breast ultrasound plays a key role in early screening and diagnosis by

Address correspondence to: Wenhui Guo, School of Anesthesiology, Naval Medical University, No.25 Zhongyuan Road, Shanghai 200093, China. E-mail: wendyguo17@outlook.com. **Yongchu Hu**, The Department of Anesthesiology, Long March Hospital, No.415 Fengyang Road, Shanghai 200003, China. E-mail address: Adsfoxcn@sina.com.cn.





Figure 1. Breast ultrasound image segmentation based on threshold value. This figure is cited from [4].

accurately segmenting breast tissue and other regions in ultrasound images, thereby providing high-resolution, detailed tissue information. This segmentation process is vital in the diagnosis and treatment of breast diseases. However, challenges arise due to the complex nature of breast tissue, including noise, low contrast, and image blur, which complicate the differentiation between the breast area and surrounding tissues. Moreover, the interpretation of ultrasound images heavily relies on the experience and skills of clinicians, leading to variability in diagnostic accuracy.

Accurate breast ultrasound image segmentation enables rapid and accurate identification of disease types and locations. Recent advancements in image processing and computer vision have significantly enhanced the segmentation efficiency. While traditional image segmentation techniques, such as thresholding, edge detection, and region growing, have been extensively applied in breast ultrasound image segmentation, they often struggle to handle the complex textures and indistinct boundaries in ultrasound images, resulting in inaccurate or inconsistent segmentation. To overcome these limitations, there has been a growing focus on deep learning-based methods for breast ultrasound image segmentation. Deep learning techniques, particularly convolutional neural networks (CNNs), can automatically learn deep features from large datasets, demonstrating remarkable performance improvements in image segmentation tasks [2]. These approaches not only enhance the differentiation of breast tissues and lesions but also reduce the impact of individual and device-related image variations, providing doctors with more reliable diagnostic tools.

Traditional breast ultrasound segmentation methods

Threshold based segmentation method

Threshold-based segmentation methods are commonly categorized into two types: binary segmentation and multi-level segmentation, depending on the number of resulting segments [3]. Binary segmentation is the most used threshold segmentation method, dividing the image into two regions: foreground (e.g., white) and background (e.g., black), as illustrated in **Figure 1** [4]. This technique is generally applied in simpler image segmentation tasks, such as object detection and image binarization.

In contrast, multi-level segmentation is employed for more complex images, where the image is divided into multiple regions, each representing a different category. This method uses several thresholds to allocate pixels to different categories based on their grayscale values, creating a multi-valued image suitable for complex image segmentation tasks. However, in practical applications, using multi-level thresholds for image segmentation may encounter difficulties, such as parameter selection, over-segmentation, high computational complexity, and sensitivity to noise, necessitating a careful consideration of the trade-offs based on specific circumstances when selecting an appropriate segmentation method [3]. Breast ultrasound delineation involves accurately and automatically segmenting breast tissue from surrounding tissues in ultrasound images. This process integrates principles from multiple fields, including image processing and machine learning. By comparing the pixel grayscale values of breast ultrasound images with predefined thresholds, the image can be segmented into breast tissue and background regions. The selection of thresholds significantly influences the results and is most effective for relatively simple image scenarios.

In the 1990s, with advancements in digital image processing and computer vision, the utilization of computer algorithms for medical images analysis started to be explored. Automatic segmentation of breast ultrasound images has

emerged as a significant research direction, attracting increasing attention. In 2001, Horsch et al. proposed a computationally efficient algorithm for segmenting breast masses in ultrasound images [5]. This algorithm maximizes utility functions defined on segmentation boundaries using grayscale value thresholds in preprocessed images. The performance of the segmentation algorithm was evaluated in a database containing 400 cases using two methods. The first method compared the computer-delineated boundaries with the manually outlined ones, and the second method compared the performance of computer-aided diagnostic techniques on edges delineated by the computer versus those delineated manually [5]. The results demonstrated that the automatic classifier developed by Horsch et al. achieved Area under the ROC Curve (AUC) values of 0.91 and 0.87 for distinguishing between benign and malignant breast lesions in fully automated segmentation and manual segmentation, respectively.

Furthermore, when integrated with their automatic classifier, the performance of the algorithm in distinguishing breast lesions was comparable to manual segmentation [5]. In another study, Cui et al. designed a method for automatically segmenting breast masses in ultrasound images. This method automatically estimated an initial contour based on a manually identified point near the mass center [6]. It then employed a two-stage Active Contour (AC) approach to iteratively refine the initial contour. The segmentation results were subject to self-inspection and correction, with performance evaluation comparing computer-based segmentation to those performed by experienced radiologists. The results demonstrated that their Computer-Aided Diagnosis (CAD) system performed similarly to manual segmentation by experienced professionals [6].

Recently, researchers have made new advancements. Zebari et al. proposed an improved segmentation model based on thresholding and trainable techniques for extracting regions of interest [7]. Furthermore, they developed a hybrid segmentation method for breast region and chest muscle boundary segmentation in breast X-ray images, integrating thresholding and machine learning techniques. In this approach, wavelet transform bands were removed to highlight the breast region and optimize the estimated breast boundary. This novel thresholding technique contributes to determining the initial breast boundary and partially overcomes the limitations of manual image segmentation.

Method based on edge detection

Edge detection is a commonly used method for image segmentation that involves detecting edge information within an image [8]. Edges typically correspond to boundaries between objects or transitions between regions, making them crucial for segmenting an image into distinct areas or objects [9]. This method segments the image by detecting contours and boundaries, effectively outlining the contours of the breast region in ultrasound images. The edge detection-based segmentation approach generally follows these steps: first, edge detection is performed; next, edge tracking, edge linking, or region merging is conducted based on the detected edge information to achieve image segmentation. This approach is suitable for breast tissues with clearly defined edges but may be less effective for complex or poorly defined breast tissue structures.

To address these challenges, researchers have been developing enhanced techniques. In 2019, Rampun et al. proposed a modified version of the Holistically-Nested Edge Detection network method to accurately segment the pectoralis major muscle from breast X-ray images, showing that the modified Holistically-Nested Edge Detection network model could segment the muscle more effectively with higher accuracy and robustness compared to traditional methods [10]. In 2022, Daoud et al. proposed an edge-based approach to improve the localization of Region of Interest (RoI) in breast ultrasound images, overcoming the limitations of single-object detection models in diverse scenarios [11]. This method integrates results from multiple deep learning-based object detection models to process ultrasound images and obtain their respective RoI localizations. Subsequently, edge detection algorithms, such as the Canny edge detector, are utilized to extract edge information from the images. Based on the edge information, the RoIs generated by the different models are filtered and merged to achieve more precise RoI localization. Further details on the application of edge detection in breast-related research can be found in **Table 1**.

Region growing algorithm

The region growing algorithm is a widely used image segmentation technique based on seed points, which partitions images into regions with similar characteristics. Starting with one or several seed points, it progressively clusters adjacent pixels into a unified region via incremental growth or pixel amalgamation, particularly suitable for segmenting homogeneous

Table 1. Literature on the application of edge detection in breast diagnosis

Authors	Methods	Evaluation metrics
Chakraborty et al. [12] (2016)	A new automated edge detection technique integrates Sobel edge detection with an active contour model.	Accuracy=92.5 Sensitivity=93 Specificity=85
Guo et al. [13] (2019)	A strong image segmentation method based on interval uncertainties	PSNR (The Canny filter outperforms other methods for low-variance salt-and-pepper noise.)
Anjum et al. [14] (2021)	HOG, Canny	Recall=0.94
Gupta et al. [15] (2022)	Canny	Accuracy=88.88 Sensitivity=94.44 Specificity=83.33
Triwibowo et al. [16] (2023)	Canny, SVM	Accuracy=95

Note: HOG, Histograms of oriented gradient; SVM, Support Vector Machine.

backgrounds and contiguous breast regions. Its fundamental principle involves expanding the region from the seed point by adding neighboring pixels that share similar characteristics, typically based on grayscale or color value differences, until a predefined stopping criterion is met [17]. Pixels whose differences fall below a specified threshold are grouped into the same region. Despite demonstrating resilience to noise and the ability to manage local image features without a prerequisite on the number of regions to be segmented, this algorithm still has limitations. Its performance is sensitive to the initial seed point selection, and the process of setting appropriate parameters can be complex. To address these challenges, researchers introduced enhancements in seed point selection and region propagation in the early 21st century. For example, Xia et al. introduced a best-first approach to enhance seed accuracy, which simplifies the bipartite matching process to unipartite matching by utilizing epipolar and continuity constraints [18]. Additionally, they implemented a dynamically adaptive window, replacing the traditionally used larger one, thus improving both the temporal and spatial efficiency of the search process.

However, in practical applications, the region growing algorithm is infrequently employed in isolation. Instead, it is commonly combined with other image processing techniques to support breast cancer diagnosis. Nonetheless, it offers effective segmentation, feature extraction, and quantitative analysis in breast ultrasound image processing, aiding in the diagnosis and treatment of breast diseases. Consequently, research into enhancing the region growing algorithm remains an ongoing endeavor. Recently, Wang et al. introduced a segmentation method for breast ultrasound images, leveraging adaptive region growing with variable level set techniques [19]. This approach seeks to address challenges associated with breast ultrasound

images, including high noise levels, blurred boundaries, and substandard imaging quality. Notably, the results demonstrated diagnostic values exceeding 0.99 for both Jaccard and Dice indices in benign tumor segmentation, highlighting the method's potential to improve segmentation accuracy in breast ultrasound images.

Application of deep learning in breast ultrasound segmentation

Deep learning, a subset of artificial intelligence, is a pivotal machine learning technique that primarily employs neural network models to capture and represent intricate data patterns [19, 20]. Its defining feature is the ability to perform feature learning and representation learning via multi-layer neural networks, which facilitates the automatic extraction of relevant features from data, enabling the modeling and resolution of complex problems. Common neural network structures utilized in deep learning include Multi-Layer Perceptrons (MLP), CNNs, and Recurrent Neural Networks [21]. These structures refine their weight parameters via the backpropagation algorithm, consistently enhancing the learning capabilities and overall performance.

Deep learning has led to remarkable achievements in various fields, including computer vision, natural language processing, and speech recognition. In image processing, it has proven effective in tasks such as image classification, object detection, and image segmentation. In medical imaging, deep learning is widely employed in tasks like breast ultrasound image segmentation, tumor detection, and disease diagnosis, significantly enhancing both accuracy and efficiency. Traditional image segmentation methods often require manual design of feature extractors and segmentation rules, limiting their performance and making them vulnerable

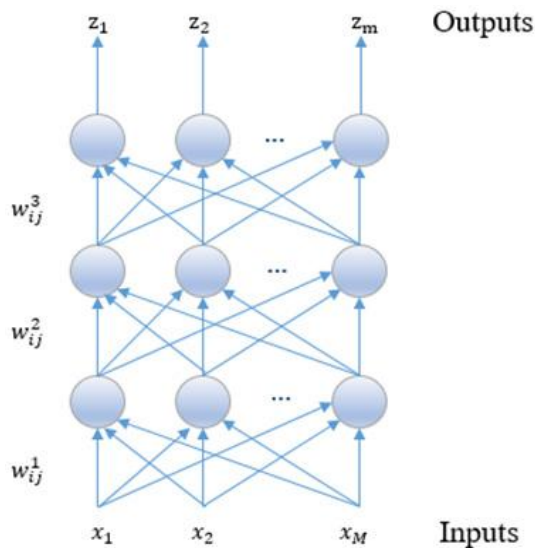


Figure 2. Multi-layer perceptron.

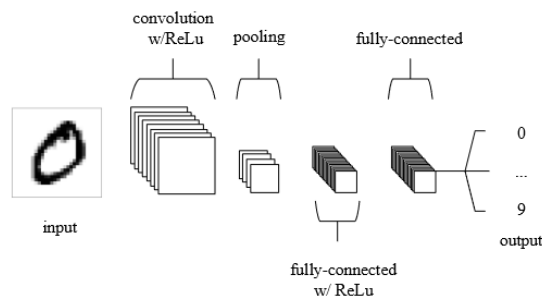


Figure 3. A simple CNN architecture. This figure is cited from [25]. CNNs, convolutional neural networks.

to noise and other interferences. In contrast, deep learning techniques, through multi-layer neural networks, can automatically learn features from breast ultrasound images and offer robust capabilities for nonlinear fitting, making it more capable of handling complex image structures and variations [2]. In fact, deep learning has superseded conventional detection methods. For instance, Freitas et al. demonstrated the use of a novel neural network approach to replace the Sobel filter, utilizing it to an automatic method for calculating depth edges in breast contours [22].

Deep learning common neural networks

MLP

MLP is one of the most fundamental and widely utilized structures within artificial neural networks [23]. It consists of an input layer, one or several hidden layers, and an output layer, forming a fundamental architecture, as depicted in **Figure 2**.

The input layer functions as the gateway for the MLP, receiving feature vectors as input data. The hidden layers, or the intermediate layer, which may contain one or multiple layers of neurons, process this input. Each neuron receives the output from the preceding layer as its input and applies a nonlinear transformation via an activation function. The output layer, located at the final stage, produces the MLP's results. The number of neurons within the output layer depends on the specific task: a single neuron is sufficient for binary classification, whereas multiple neurons are required for multi-class classification, with sigmoid or softmax activation functions typically used for classification.

Despite its simplicity and effectiveness, the MLP has limitations when handling complex tasks such as image processing. Its expressive power is constrained, making more advanced architectures like CNNs better suited for such tasks.

CNNs

CNNs emerge as one of the most prominent architectures in artificial neural networks, widely used for tasks such as image recognition and computer vision. As a type of feedforward neural network, CNNs are particularly effective at automatically extracting features from images and processing data with a grid-like structure. Key components of CNNs include convolutional layers, pooling layers, and fully connected layers [24]. A simplified structure of a CNNs is depicted in **Figure 3** [25]. The convolutional layers perform feature extraction by convolving the input image with specific convolutional kernels, resulting in a series of feature maps. Pooling layers reduce the dimensionality of these feature maps, thus decreasing computational complexity and improving the network's robustness. Fully connected layers then map the feature maps from the pooling layers to their corresponding labels, completing the classification process.

In addition to the previously mentioned components, CNNs also incorporate activation functions, loss functions, and the backpropagation algorithm. **Figure 4** [25] depicts a representative CNNs structure, where activation functions (e.g., ReLUs) are applied between the convolutional and pooling layers, followed by one or more fully connected layers also utilizing ReLUs activations to enhance learning and feature mapping.

Fully Convolutional Network (FCNs)

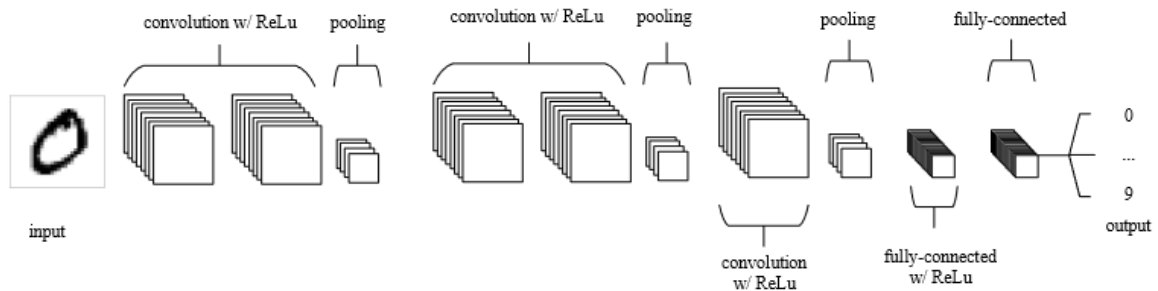


Figure 4. Common CNN structures. This figure is cited from [25]. CNNs, convolutional neural networks.

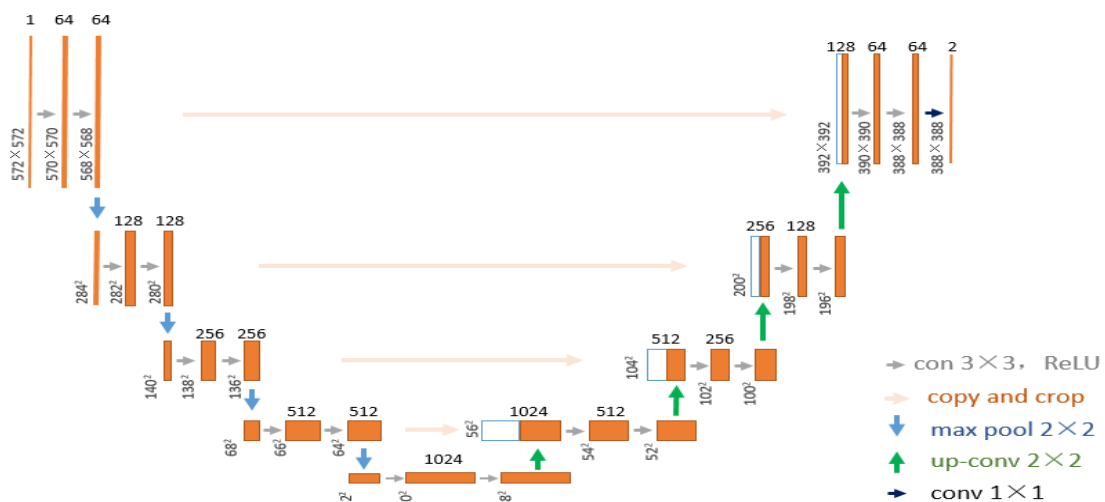


Figure 5. U-net architecture. This figure is cited from [28].

FCNs, are specialized type of neural network introduced by Jonathan Long, Evan Shelhamer, and Trevor Darrell in 2015 [26]. Their innovation lies in applying FCNs to the complex task of image semantic segmentation. The key distinction of FCNs is the replacement of traditional CNN's fully connected layers with convolutional layers. This modification enables FCNs to process input images of arbitrary sizes. In standard CNNs, convolutional layers are responsible for extracting pertinent image features, whereas fully connected layers subsequently map these features into classification or regression outputs. Conversely, FCNs accomplish the intricate transformation from feature maps to output maps solely through convolutional layers, allowing them to handle images of diverse sizes and generate output images with corresponding dimensions. The development of FCNs was motivated by the need to address a critical limitation of traditional CNNs, namely their inability to process input images of varying sizes [26, 27].

Deep learning algorithm applied to image segmentation

U-Net

U-Net, a specialized form of CNNs, is specifically tailored for image segmentation tasks. Its architecture, illustrated in **Figure 5**, operates at a low resolution of 32×32 pixels. Initially proposed by Ronneberger et al. in 2015, U-Net addresses the challenge of spatial information loss during the inter-layer transmission within CNNs, a limitation that can potentially compromise image segmentation accuracy [28]. Unlike traditional CNNs, U-Net incorporates skip connections to preserve information, enabling decoder layers to directly access both shallow and deep features from the encoder. Consequently, it mitigates issues such as information loss and vanishing gradients, thereby enhancing the spatial accuracy and semantic coherence of segmentation results. Owing to these advantages, U-Net is well-suited for tasks like breast ultrasound image segmentation.

The introduction of U-Net established a solid foundation for its subsequent utilization in breast ultrasound image segmentation. In 2018, Almajalid et al. implemented the U-Net architecture to segment breast ultrasound

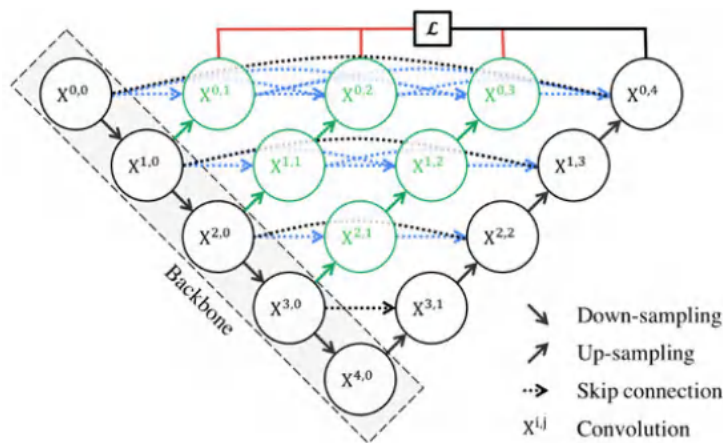


Figure 6. U-net++ architecture. This figure is cited from [36].

images [29]. Their approach began with pre-processing techniques, including histogram equalization to enhance contrast and anisotropic diffusion filtering to reduce speckle noise. Subsequently, to expand the training dataset, they employed data augmentation techniques, such as rotation and elastic deformation. Following this, noise regions were excised from the segmentation outcomes. Employing double cross-validation, the U-Net model was trained and tested on a dataset encompassing 221 breast ultrasound images, achieving an average DICE score of 82.52%, demonstrating its robustness and accuracy for tumor image segmentation. Similarly, Yap et al. proposed a patch-based LeNet and U-Net model, along with a transfer learning approach utilizing a pre-trained FCNS-AlexNet [30]. Their study compared these methods to four advanced lesion detection algorithms: radial gradient index, multifractal filtering, rule-based region ranking, and deformable part models. Additionally, they conducted a comparative analysis on datasets obtained from two distinct ultrasound systems, improving the overall performance of deep learning-based techniques for lesion detection.

Multiple automatic methods have been proposed for detecting and segmenting breast lesions. Nevertheless, challenges persist due to ultrasound artifacts and the complex shapes and locations of lesions, which hinder accurate segmentation. To address these issues, Amiri et al. introduced a two-stage approach, where U-Net was initially employed for lesion detection, followed by a second U-Net for segmenting the identified lesion areas [31]. They also developed a test-time augmentation technique to evaluate the effectiveness of the detection phase, which resulted in a 1.8% improvement in the average Dice score. In 2021, Guo et al. proposed an enhanced U-Net with an extended training method, which preserved the textural

nuances and edge characteristics of breast tumors in output maps [32].

Notably, training U-Net with gray-scale probability labels accelerated the process compared to conventional labeling methods, leading to clearer tumor boundaries and improving the safety and precision of robot-assisted breast surgeries. In 2022, Zhao et al. proposed a breast tumor ultrasound image segmentation method based on the U-Net framework, combining residual blocks and attention mechanisms

[33]. Meanwhile, Yan and Liu et al. developed an attention-enhanced U-Net model with mixed dilated convolutions (AE U-net with HDC) for tumor segmentation in breast ultrasound images [34]. Recently, Pramanik et al. presented a novel network structure, dual branch U-Net (DBU-Net), which improves upon the classical U-Net architecture [35]. DBU-Net incorporates a dual-branch structure to enhance segmentation accuracy and robustness by handling feature information at different scales. One branch captures global and contextual information, while the other branch focuses on local and detailed information. This dual-branch approach allows DBU-Net to better capture tumor structures in breast ultrasound images, demonstrating significant improvements in segmentation performance over traditional U-Net. These advancements highlight the potential clinical applicability of DBU-Net in breast tumor image segmentation.

UNet++

UNet++, an extended version of the U-Net algorithm, was proposed in 2018 by a research team led by Zhou from Indonesia. Its network structure, shown in **Figure 6**, aims to enhance the performance of image segmentation tasks [36]. The primary improvement of UNet++ lies in its feature propagation and reuse capabilities, which are achieved by introducing multiple down-sampling and up-sampling pathways [36, 37]. Unlike the original U-Net architecture, which has a single down-sampling and up-sampling path, UNet++ introduces deeper feature fusion pathways by incorporating multiple nested, dense skip connections. These skip connections enable the network to propagate and reuse feature information across different levels, significantly improving segmentation accuracy.

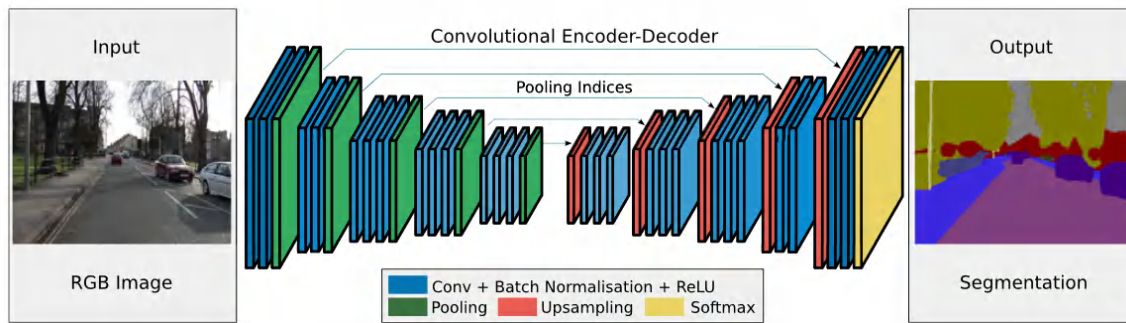


Figure 7. SegNet architecture. This figure is cited from [42].

UNet++ has found wide applications in medical image segmentation due to its numerous advantages. Wu et al. utilized UNet++ as the base structure and substituted its standard convolutions with dilated convolutions [38]. This modification increased the receptive field, allowing the model to incorporate more global contextual information, thereby enhancing organ segmentation accuracy.

Currently, the utilization of UNet++ in breast-related tasks primarily focuses on segmenting breast cancer cell nuclei. In 2020, Wang et al. proposed an improved breast cancer cell nucleus segmentation method using UNet++, with Inception-ResNet-V2 as the backbone [39]. The performance of this enhanced UNet++ was compared with the original UNet++ architecture using a breast cancer cell nucleus segmentation database, demonstrating the superiority of the improved version. The following year, Dinh et al. introduced a novel approach combining UNet++ with pre-trained EfficientNet as the backbone architecture for breast tumor cell nucleus segmentation, along with multi-organ transfer learning methods [40]. This new approach outperformed other techniques in segmenting breast tumor cell nuclei, significantly increasing accuracy. More recently, You et al. proposed a novel EfficientUNet++ model for breast tumor segmentation [41]. This model integrates ResNet18 as the encoder, a channel attention mechanism, and deep supervision to address the issue of gradient vanishing and enhance feature extraction. The channel attention module further improves the model's accuracy, particularly in handling tumor boundaries.

SegNet

SegNet is a CNN architecture designed for semantic image segmentation tasks. Initially proposed in 2017 by a team from the Computer Laboratory at the University of Cambridge, including Badrinarayanan, its network structure is shown in **Figure 7** [42]. SegNet's design focuses on pixel-level image segmentation,

incorporating an encoder for feature extraction and a decoder for feature restoration and segmentation.

Compared to U-Net, it offers satisfactory segmentation performance while requiring lower computational and storage resources, making it suitable for resource-constrained environments. In contrast, U-Net is specifically optimized for medical image segmentation, considering factors such as high resolution and complex structures, which contribute to its superior performance in medical imaging tasks [43].

U-Net and SegNet were both applied to lesion segmentation in breast ultrasound images by Vianna et al. in 2021, with comparative analysis conducted using paired sample t-tests [44]. The results demonstrated U-Net's superior performance in their task settings. Among the configurations tested, U-Net with Dice as the loss function and ReLU as the activation unit achieved the best results, with higher Dice similarity coefficients and shorter training times. This study concluded that, despite SegNet's efficiency, U-Net outperformed it in dataset segmentation tasks. In the same year, Vianna et al. presented a performance analysis of SegNet for automatic breast ultrasound image segmentation [45]. The results showed that SegNet achieved accuracy similar to other models, although performance varied depending on the specific configurations used.

While the U-Net algorithm remains the primary choice for breast ultrasound image segmentation, SegNet offers advantages such as high computational efficiency and low memory usage, making it a promising candidate for other medical image processing tasks. SegNet has also been successfully applied to retinal vessel segmentation in fundus images, where accurate segmentation of vessel networks is crucial for the early detection and treatment of conditions like diabetic retinopathy and hypertensive retinopathy [46]. Moreover, SegNet has

Table 2. New techniques for breast diagnosis

Authors	Methods	Results
Shu et al. [56] (2020)	CNN (DenseNet169 + max pooling)	Mass: AUC=0.87 ACC=0.796 Calcification: AUC=0.805 ACC=0.707
Ekici et al. [57] (2020)	CNN	ACC=98.95
Heidari et al. [58] (2019)	CNN, SVM	Accuracy=92 Sensitivity=89 Specificity=95 AUC=0.8451
Nyayapathi et al. [59] (2019)	DSM	AUC=0.8451
Wang et al. [60] (2020)	ResNet-50 MS MI	AUC=0.862
Zhao et al. [33] (2022)	Improved residual U-Net	Dice=92.10 IoU=85.10 MAD=4.99
Ning et al. [61] (2021)	SMU-Net	Dice=88.27 IoU=80.57 F1=88.31 Sensitivity=89.69 Precision=89.07
Gare et al. [62] (2022)	W-Net	ACC=0.852 MIoU=0.507
Chen et al. [63] (2022)	C-Net	Accuracy=97.51 Jaccard=72.41 Precision=82.70 Recall=82.42 Specificity=99.13
Chen et al. [64] (2022)	AAU-net	Jaccard=69.10 Precision=78.83 Recall=82.22 Specificity=98.82 Dice=78.14
Alruily et al. [65] (2023)	GAN, U-Net3+	Dice=95.49 Accuracy=95.67 Precision=95.59 Recall=95.68
Tekin et al. [66] (2023)	Tubule-U-Net	Dice=95.33 Recall=93.74 Specificity=90.02
He et al. [67] (2023)	HCTNet	Dice=97.23 Accuracy=97.41 Jaccard=94.63 Precision=95.59 Recall=97.14

Note: CNN, Convolutional Neural Network; SVM, Support Vector Machine; DSM, Dual scan mammoscope; MS, Multiscale; MI, Multi-Instance; SMU-Net, Saliency-guided morphology-aware U-Net; W-Net, Waveform-Net; C-Net, Cascaded Convolutional Neural Network; AAU-net, adaptive attention U-net; GAN, generative adversarial networks; U-Net3+, A Full-Scale Connected U-Net; HCTNet, hybrid CNN-transformer network; AUC, Area Under the Curve; ACC, Accuracy; Dice, Dice Similarity Coefficient; IoU, Intersection over Union; MAD, Mean Absolute Deviation; F1, F1-score; MIoU, Mean Intersection over Union.

demonstrated utility in lesion segmentation in liver and lung CT images, as well as in brain MRI images, particularly in applications related to brain imaging [47-49]. However, the applicability and performance of SegNet should be validated experimentally for each specific domain, depending on the task and dataset.

DeepLab

DeepLab is a series of image segmentation algorithms proposed by the Google team, utilizing deep convolutional neural networks and probabilistic graphical models, such as conditional random fields, for pixel-level image classification. Since its first introduction in 2014 by Chen

et al., the DeepLab series has evolved through multiple versions, including DeepLab v1, v2, v3, and v3+ [50, 51].

The DeepLab algorithms have been continuously optimized by researchers, improving the structure and parameters of deep convolutional neural networks and probabilistic models to enhance the accuracy of semantic image segmentation [52]. This progression has contributed significantly to the field of computer vision, offering valuable insights for understanding and analyzing breast ultrasound images. Shia et al. proposed a method for semantic segmentation of breast ultrasound images using the DeepLab v3+ model [53]. They analyzed a total of 684

breast tumor images and evaluated and found that DeepLab v3+ with a ResNet-50 decoder outperformed previous architectures in semantic segmentation tasks, particularly in identifying malignant features associated with BI-RADS malignant classifications in breast ultrasound (BUS) images. Additionally, Zhou et al. proposed a new deep learning-based breast region extraction method that integrates preprocessing techniques, including median filtering for noise suppression and contrast limited adaptive histogram equalization (CLAHE) for contrast enhancement, followed by semantic segmentation using the DeepLab v3+ model [54]. Moreover, this approach demonstrated state-of-the-art performance in extracting breast regions from mammography images. While DeepLab is less frequently applied in breast ultrasound segmentation, it has been widely used for segmenting human organs such as the brain [55]. Other literature on the application of deep learning in breast ultrasound image segmentation is presented in **Table 2**.

Discussion

Early research in breast disease diagnosis primarily focused on traditional image processing techniques for breast ultrasound image segmentation, such as thresholding, edge detection, and region growing. These methods typically rely on features such as grayscale intensity, texture, and shape for segmentation. For instance, tumor regions can be extracted using thresholding segmentation methods, which classify image pixels into distinct categories based on one or more thresholds; Edge detection methods, utilizing operators like Canny, Sobel, Roberts, and Laplacian, detect edges in the image to delineate the boundaries of the tumor region; Region growing methods, on the other hand, progressively merge adjacent pixels or regions into a unified area, starting from seed points and continuing until predefined stopping criteria are met.

However, as the incidence of breast diseases has risen globally, traditional ultrasound image segmentation methods have shown limitations like strong subjectivity and high sensitivity to noise, making them less effective for clinical applications. Consequently, breast ultrasound image segmentation has become a widely studied topic, garnering significant attention from numerous researchers.

In recent years, with the rapid advancement of deep learning technology, breast ultrasound image segmentation methods based on deep learning have garnered extensive attention and

research. These methods typically use deep learning models such as CNNs to automatically learn and extract features from large datasets, achieving more accurate and robust segmentation results. Notable deep learning models, such as U-Net, UNet++, SegNet, and DeepLab, have shown remarkable success in medical image segmentation. Among them, U-Net algorithm has demonstrated strong performance in breast ultrasound image segmentation, while UNet++, SegNet, and DeepLab algorithms, although less commonly used in breast imaging, have proven effective in other medical domains, showcasing substantial potential for broader use.

This study primarily focuses on algorithm improvements, evaluation of segmentation performance, and their clinical applications in breast disease diagnosis. Despite progress, challenges persist, such as poor image quality and the complex morphology of tumors, which continue to hinder segmentation accuracy. Furthermore, issues related to computational efficiency and real-time performance remain significant obstacles. As a result, no single algorithm has yet proven optimal across all scenarios. In practical applications, it is crucial to consider the specific characteristics of the images and the requirements of the task, selecting or combining algorithms accordingly for the best results.

Conclusion

This article reviews the various breast ultrasound image segmentation techniques, with a focus on traditional methods and the application of deep learning. Segmentation plays a crucial role in early breast cancer screening, tumor measurement, and surgical planning. While traditional techniques like thresholding and edge detection provide some value, they often struggle with noise and complex shapes. In contrast, deep learning methods offer improved accuracy, though challenges like data annotation and high computational demands remain. Future advancements in this field are expected to focus on multimodal image fusion, improvements in deep learning approaches, and the integration of ultrasound with MRI and X-ray for enhanced diagnostic accuracy and clinical utility.

Author contributions: Fangfang Chen: Responsible for the primary research work of the paper, including literature review, data analysis, and article writing. Miao Zhou: Responsible for part of the literature review and writing of specific chapters, particularly those related to relevant work and background. Jintao Duan:

Participated in data collection and organization, contributing to the discussion of technical details and methods. Yuxiang Wang: Participated in literature search, provided experimental data support, and created some of the figures and tables. Liangqing Lin: Participated in the revision and organization of the paper, mainly helping with formatting and language. Wenhui Guo, Yongchu Hu: Responsible for the overall conception and review of the paper, guiding the research direction, and finalizing and revising the paper.

References

- [1] Ilesanmi AE, Chaumrattanakul U, Makhanov SS. Methods for the segmentation and classification of breast ultrasound images: a review. *J Ultrasound* 2021;24(4):367-382.
- [2] Liu S, Wang Y, Yang X, et al. Deep learning in medical ultrasound analysis: a review. *Engineering* 2019;5(2):261-275.
- [3] Jia H, Ma J, Song W. Multilevel thresholding segmentation for color image using modified moth-flame optimization. *IEEE Access* 2019;7:44097-44134.
- [4] Rodrigues PS, Giraldi GA. Improving the non-extensive medical image segmentation based on Tsallis entropy. *Pattern Anal Applic* 2011;14(4):369-379.
- [5] Horsch K, Giger ML, Venta LA, et al. Automatic segmentation of breast lesions on ultrasound. *Med Phys* 2001;28(8):1652-1659.
- [6] Cui J, Sahiner B, Chan HP, et al. A new automated method for the segmentation and characterization of breast masses on ultrasound images. *Med Phys* 2009;36(5):1553-1565.
- [7] Zebari DA, Zeebaree DQ, Abdulazeez AM, et al. Improved threshold based and trainable fully automated segmentation for breast cancer boundary and pectoral muscle in mammogram images. *IEEE Access* 2020;8:203097-203116.
- [8] Ghassan MGA, Ahmed MA. Abushaala. Edge detection methods. 2015 2nd World Symposium on Web Applications and Networking (WSWAN) 2015;1-7.
- [9] Shao D, Xu C, Xiang Y, et al. Ultrasound image segmentation with multilevel threshold based on differential search algorithm. *IET Image Process* 2019;13(6):998-1005.
- [10] Rampun A, López-Linares K, Morrow PJ, et al. Breast pectoral muscle segmentation in mammograms using a modified holistically-nested edge detection network. *Med Image Anal* 2019;57:1-17.
- [11] Daoud MI, Al-Ali A, Alazrai R, et al. An Edge-Based Selection Method for Improving Regions-of-Interest Localizations Obtained Using Multiple Deep Learning Object-Detection Models in Breast Ultrasound Images. *Sensors (Basel)* 2022;22(18):6721.
- [12] Chakraborty S, Bhowmik MK, Ghosh AK, et al. Automated edge detection of breast masses on mammograms. 2016 IEEE Region 10 Conference (TENCON) 2016;1241-1245.
- [13] Guo G, Razmjoooy N. A new interval differential equation for edge detection and determining breast cancer regions in mammography images. *Syst Sci Control Eng* 2019;7(1):346-356.
- [14] Anjum R, Dipti RR, Rashid HO, et al. An Efficient Breast Cancer Analysis Technique by Using a Combination of HOG and Canny Edge Detection Techniques. 2021 5th International Conference on Trends in Electronics and Informatics (ICOEI) 2021;1290-1295.
- [15] Gupta KK, Vijay R, Pahadiya P. Detection of abnormality in breast thermograms using Canny edge detection algorithm for thermography images. *Int J Medical Eng Informatics* 2022;14:31-42.
- [16] Triwibowo DN, Dewa BP, Sumantri RBB, et al. Identification of Breast Tumors With Image Processing Using Canny Edge Detection. *J Adv Health Inf Res* 2023;1(1):28-34.
- [17] Angelina S, Suresh LP, Veni SHK. Image segmentation based on genetic algorithm for region growth and region merging. 2012 International Conference on Computing, Electronics and Electrical Technologies (ICCEET) 2012;970-974.
- [18] Xia X, Gai S. An Improved Dense Matching Algorithm for Face Based on Region Growing. *Computer Engineering and Networking* 2014;625-633.
- [19] Wang Y, Ye J, Wang T, et al. Breast Ultrasound Image Segmentation Algorithm Using Adaptive Region Growing and Variation Level Sets. 2022; 2022: 1752390.
- [20] LeCun Y, Bengio Y, Hinton G. Deep learning. *Nature* 2015;521(7553):436-444.
- [21] Deng L, Yu D. Deep Learning: Methods and Applications. *Found Trends Signal Process* 2014;7:197-387.
- [22] Freitas N, Silva D, Mavioso C, et al. Deep Edge Detection Methods for the Automatic Calculation of the Breast Contour. *Bioengineering (Basel)* 2023;10(4):401.
- [23] Taud H, Mas JF. Multilayer Perceptron (MLP). In: Camacho Olmedo MT, Paegelow M, Mas J-F, Escobar F, editors. *Geomatic Approaches for Modeling Land Change Scenarios*. Cham. Springer International Publishing, 2018. p.451-455.
- [24] Gu J, Wang Z, Kuen J, et al. Recent advances in convolutional neural networks. *Pattern Recognit* 2018;77:354-377.
- [25] O'Shea K, Nash R, "An Introduction to

- Convolutional Neural Networks." Available: <http://arxiv.org/abs/1511.08458>
- [26] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2015;3431-3440.
- [27] Wang J, Song L, Li Z, et al. End-to-End Object Detection with Fully Convolutional Network. 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR) 2021;15844-15853.
- [28] Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Medical Image Computing and Computer-Assisted Intervention - MICCAI 2015 2015;234-241.
- [29] Almajalid R, Shan J, Du Y, et al. Development of a Deep-Learning-Based Method for Breast Ultrasound Image Segmentation. 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA) 2018;1103-1108.
- [30] Yap MH, Pons G, Marti J, et al. Automated breast ultrasound lesions detection using convolutional neural networks. IEEE J Biomed Health Inform 2018;22(4):1218-1226.
- [31] Amiri M, Brooks R, Behboodi B, et al. Two-stage ultrasound image segmentation using U-Net and test time augmentation. Int J Comput Assist Radiol Surg 2020;15(6):981-988.
- [32] Guo Y, Duan X, Wang C, et al. Segmentation and recognition of breast ultrasound images based on an expanded U-Net. PLoS One 2021;16(6):e0253202.
- [33] Zhao T, Dai HJCI, Neuroscience. Breast Tumor Ultrasound Image Segmentation Method Based on Improved Residual U-Net Network. Comput Intell Neurosci 2022;2022:3905998.
- [34] Yan Y, Liu Y, Wu Y, et al. Accurate segmentation of breast tumors using AE U-net with HDC model in ultrasound images. Biomed Signal Process Control 2022;72:103299.
- [35] Pramanik P, Pramanik R, Schwenker F, et al. DBU-Net: Dual branch U-Net for tumor segmentation in breast ultrasound images. DBU-Net: Dual branch U-Net for tumor segmentation in breast ultrasound images. PLoS One 2023;18(11):e0293615.
- [36] Zhou Z, Siddiquee MMR, Tajbakhsh N, et al. Unet++: A nested u-net architecture for medical image segmentation. Deep Learn Med Image Anal Multimodal Learn Clin Decis Support (2018) 2018;11045:3-11.
- [37] Zhou Z, Siddiquee MMR, Tajbakhsh N, et al. Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. IEEE Trans Med Imaging 2020;39(6):1856-1867.
- [38] Wu S, Wang Z, Liu C, et al. Automatic Segmentation of Pelvic Organs After Hysterectomy by using Dilated Convolution U-Net++. 2019 IEEE 19th International Conference on Software Quality, Reliability and Security Companion (QRS-C) 2019;362-367.
- [39] Wu Z, Wen Z, Chen Y, et al. An Improved Cell nuclei Segmentation Method Based on U-Net++. 2023 International Conference on Machine Vision, Image Processing and Imaging Technology (MVIPT) 2023;78-84.
- [40] Dinh Tuan L, Kwon S-G, Lee S-H, et al. Breast Tumor Cell Nuclei Segmentation in Histopathology Images using EfficientUnet++ and Multi-organ Transfer Learning. J Korea Multimedia Soc 2021;24(8):1000-1011.
- [41] You G, Yang X, Lee X, et al. EfficientUNet: An efficient solution for breast tumour segmentation in ultrasound images. IET Image Process 2024;18(2):523-534.
- [42] Badrinarayanan V, Kendall A, Cipolla RJItopa, et al. Segnet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans Pattern Anal Mach Intell 2017;39(12):2481-2495.
- [43] Siddique N, Paheding S, Elkin C P, et al. U-Net and its variants for medical image segmentation: A review of theory and applications. Siddique N, Paheding S, Elkin CP, et al. U-Net and Its Variants for Medical Image Segmentation: A Review of Theory and Applications. IEEE Access 2021;9:82031-82057.
- [44] Vianna P, Farias R, de Albuquerque Pereira WC. U-Net and SegNet performances on lesion segmentation of breast ultrasonography images. Res Biomed Eng 2021;37(2):171-179.
- [45] Vianna PO, Farias R, Pereira WCA. Performance of the SegNet in the Segmentation of Breast Ultrasound Lesions. 2021 Global Medical Engineering Physics Exchanges/Pan American Health Care Exchanges (GMEPE/PAHCE) 2021;1-4.
- [46] Ryu J, Rehman MU, Nizami IF, et al. SegR-Net: A deep learning framework with multi-scale feature fusion for robust retinal vessel segmentation. Comput Biol Med 2023;163:107132.
- [47] Saood A, Hatem IJBMI. COVID-19 lung CT image segmentation using deep learning methods: U-Net versus SegNet. BMC Med Imaging 2021;21(1):19.
- [48] Almotairi S, Kareem G, Aouf M, et al. Liver tumor segmentation in CT scans using modified SegNet. Sensors (Basel) 2020;20(5):1516.
- [49] Alqazzaz S, Sun X, Yang X, et al. Automated brain tumor segmentation on multi-modal MR

- image using SegNet. *Comput Visual Media* 2019;5(2):209-219.
- [50] Chen LC, Papandreou G, Kokkinos I, et al. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. *IEEE Trans Pattern Anal Mach Intell* 2018;40(4):834-848.
- [51] Chen LC, Papandreou G, Schroff F, et al. Rethinking atrous convolution for semantic image segmentation. 2017;arXiv:1706.05587.
- [52] Chen L-C, Zhu Y, Papandreou G, et al. Encoder-Decoder with Atrous Separable Convolution for Semantic Image Segmentation. *Computer Vision - ECCV 2018*;833-851.
- [53] Shia WC, Hsu FR, Dai ST, et al. Semantic segmentation of the malignant breast imaging reporting and data system lexicon on breast ultrasound images by using DeepLab v3+. *Sensors (Basel)* 2022;22(14):5352.
- [54] Zhou K, Li W, Zhao D, et al. Deep learning-based breast region extraction of mammographic images combining pre-processing methods and semantic segmentation supported by Deeplab v3+. *Technol Health Care* 2022;30(S1):173-190.
- [55] Roy Choudhury A, Vanguri R, Jambawalikar SR, et al. Segmentation of Brain Tumors Using DeepLabv3+. *Brainlesion: Glioma, Multiple Sclerosis, Stroke and Traumatic Brain Injuries* 2019;154-167.
- [56] Shu X, Zhang L, Wang Z, et al. Deep neural networks with region-based pooling structures for mammographic image classification. *IEEE Trans Med Imaging* 2020;39(6):2246-2255.
- [57] Ekici S, Jawzal HJMh. Breast cancer diagnosis using thermography and convolutional neural networks. *Med Hypotheses* 2020;137:109542.
- [58] Heidari M, Mirniaharikandehei S, Liu W, et al. Development and assessment of a new global mammographic image feature analysis scheme to predict likelihood of malignant cases. *IEEE Trans Med Imaging* 2020;39(4):1235-1244.
- [59] Nyayapathi N, Lim R, Zhang H, et al. Dual scan mammoscope (DSM)—a new portable photoacoustic breast imaging system with scanning in craniocaudal plane. *IEEE Trans Biomed Eng* 2020;67(5):1321-1327.
- [60] Wang Z, Zhang L, Shu X, et al. An End-to-End Mammogram Diagnosis: A New Multi-Instance and Multiscale Method Based on Single-Image Feature. *IEEE Trans Cogn Dev Syst* 2021;13(3):535-545.
- [61] Ning Z, Zhong S, Feng Q, et al. SMU-Net: Saliency-guided morphology-aware U-Net for breast lesion segmentation in ultrasound image. *IEEE Trans Med Imaging* 2022;41(2):476-490.
- [62] Gare GR, Li J, Joshi R, et al. W-Net: Dense and diagnostic semantic segmentation of subcutaneous and breast tissue in ultrasound images by incorporating ultrasound RF waveform data. *Med Image Anal* 2022;76:102326.
- [63] Chen G, Dai Y, Zhang JJCM, et al. C-Net: Cascaded convolutional neural network with global guidance and refinement residuals for breast ultrasound images segmentation. *Comput Methods Programs Biomed* 2022;225:107086.
- [64] Chen G, Li L, Dai Y, et al. AAU-net: an adaptive attention U-net for breast lesions segmentation in ultrasound images. *IEEE Trans Med Imaging* 2023;42(5):1289-1300.
- [65] Alruily M, Said W, Mostafa AM, et al. Breast ultrasound images augmentation and segmentation using GAN with identity block and modified U-Net 3+. *Sensors (Basel)* 2023;23(20):8599.
- [66] Tekin E, Yazıcı Ç, Kusetogullari H, et al. Tubule-U-Net: a novel dataset and deep learning-based tubule segmentation framework in whole slide images of breast cancer. *Sci Rep* 2023;13(1):128.
- [67] He Q, Yang Q, Xie M, et al. HCTNet: A hybrid CNN-transformer network for breast ultrasound image segmentation. *Comput Biol Med* 2023;155:106629.



Application of U-Net and its variants in ultrasound image segmentation

Yuxiang Wang¹, Miao Zhou², Fangfang Chen¹, Jintao Duan¹, Liangqing Lin³, Qinghua Wu³, Wenhui Guo⁴, Haipo Cui¹

¹School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. ²Jiangsu Cancer Hospital, Nanjing 213164, China. ³Anesthesiology, The First Hospital of Putian, Putian 351100, China. ⁴School of Anesthesiology, Second Military Medical University/Naval Medical University, Shanghai 200433, China.

Corresponding author: Haipo Cui.

Declaration of conflict of interest: None.

Received September 27, 2024; Accepted January 22, 2025; Published April 1, 2025

Highlights

- This review provides a comprehensive overview of various U-Net architectures and their variants, including the original U-Net, U-Net++, Attention U-Net, and ResU-Net, along with a discussion on potential improvements to the U-Net architecture.
- The differences between these network architectures are analyzed in terms of training complexity and computational requirements.
- The review delves into the application of U-Net and its variants in ultrasound imaging, discussing both the advantages and limitations of each model in various ultrasound contexts. Relevant literature on the application of each network architecture in ultrasound is also summarized.

Abstract

Ultrasonography plays an important role in the fields of obstetrics, gynecology, cardiology, and hepatology, as well as ultrasound-guided nerve blocks, interventional therapy, and surgical navigation due to its non-invasive, real-time imaging and radiation-free characteristics. Recently, with the advancement of artificial intelligence, machine learning and deep learning algorithms have brought significant innovations to ultrasound imaging technology in the medical field. U-Net is widely recognized as one of the most commonly used deep learning models in medical image processing. This paper explores the application of the U-Net family of models in ultrasound imaging. The network architecture of the original U-Net, comprising encoder and decoder components, is first delineated. Next, classical variants, such as U-Net++, Attention U-Net, and ResU-Net, are introduced. The application of U-Net models in ultrasound and their segmentation performance are then reviewed, with Dice coefficients highlighted as the primary evaluation metric. Finally, the paper provides a comparative analysis of the advantages and disadvantages of the U-Net family of models.

Keywords: U-Net, U-Net++, Attention U-Net, ResU-Net, ultrasonography, segmentation performance

1 Introduction

Ultrasound, a widely utilized medical diagnostic tool, possesses real-time imaging capabilities, facilitating the real-time visualization of nerves, blood vessels, and muscles. This real-time imaging capability minimizes waiting times for both patients and medical practitioners. Despite its widespread use, ultrasound equipment has inherent limitations. The low contrast in

ultrasound images often results in similar echo characteristics across different tissues, posing challenges for doctors in distinguishing specific structures. For instance, during a lumbar plexus nerve block, the boundaries of the cloverleaf structure, which includes the quadratus lumborum, major lumbar muscle, and erector spinae muscle, may be unclear, leading to potential misidentification of abdominal cavity as transverse processes. Furthermore, the quality of



Table 1. Summary of the applicability, advantages, and disadvantages of each network structure

Model	Structural characteristics	Advantage	Disadvantage	Applicability
U-Net	Infrastructure	Aids in capturing features effectively.	May have limitations in dealing with multi-scale features.	Widely used in segmentation tasks for ultrasound images
U-Net++	Jump connections and deep supervision are introduced	Multi-level feature fusion contributes to comprehensive context understanding.	The network structure is relatively more complex.	More suitable for ultrasound images with multi-scale features
Attention U-Net	Attention mechanisms were introduced	The introduction of an attention mechanism improves the ability to capture key features.	More training samples are needed to learn the attention weights efficiently.	For tasks that require better attention to important structures
ResU-Net	Residual learning and U-Net were combined	Combines the advantages of residual learning and U-Net to provide deeper feature learning.	More training samples are needed to train the deep network efficiently.	For tasks requiring deeper feature learning and better generalization performance

ultrasound images is influenced by various positions and angles of the ultrasound probe [1]. This variability can result in images of differing quality and readability. As a result, diagnosing and testing diseases rely heavily on the subjective experience of the physician.

In recent years, the integration of artificial intelligence, particularly machine learning and deep learning, has gained prominence in the medical field [2]. Machine learning algorithms applied to ultrasound enable automated image analysis for tasks such as lesion detection, disease screening, segmentation, and feature extraction [3]. This automation improves the accuracy of disease detection and early diagnosis. Specifically, in breast tumor detection, ultrasound image segmentation plays a critical role. Deep learning models, particularly those based on the U-Net architecture, have proven effective for detecting and segmenting breast tumors [4-7].

The U-Net structure, recognized as one of the most widely employed deep learning models in image processing, excels in image segmentation and feature learning. Its proficiency in localizing and segmenting cells, tissues, or lesions in medical images makes it a powerful tool for precise segmentation tasks in medical images [8-11]. Ultrasound images, characterized by low contrast and high noise, present challenges for traditional image processing methods such as thresholding, edge detection, and mathematical morphology. The U-Net architecture addresses these challenges by employing jump connections, which effectively preserve high-resolution information and prevent information loss. This adaptability makes U-Net particularly well-suited for processing intricate medical ultrasound images. This paper focuses on the application of the U-Net family of models in ultrasound imaging.

II Methods and materials

Performance comparison with different methods

Each network structure has its unique characteristics, and no single model is universally applicable. In practical scenarios, experiments and evaluations should be tailored to the specific task, dataset features, and available computational resources. **Table 1** summarizes the applicability, advantages, and disadvantages of each network structure. In addition, the differences in training complexity and computational requirements across the various models are summarized in **Table 2**.

U-net

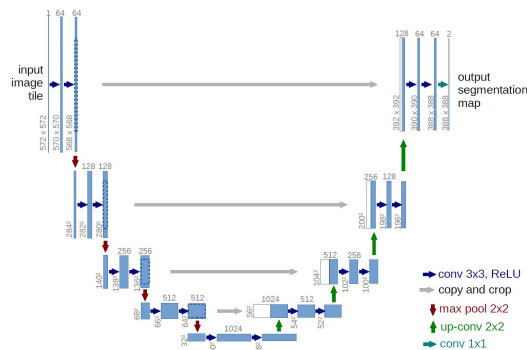
The U-Net architecture was proposed in 2015 by German researchers Olaf Ronneberger, Philipp Fischer, and Thomas Brox. The name “U-Net” comes from the network’s “U”-shaped structure, which is a distinctive feature of this architecture [12]. As depicted in **Figure 1**, the architecture comprises an encoder and a decoder. The encoder progressively extracts feature maps through multiple convolution and pooling operations, while the decoder enlarges the resolution through up-sampling. Skip connections link to the encoder and decoder, facilitating the fusion of feature maps, ultimately producing the segmentation result at the output layer.

Encoder

The encoder plays a crucial role in the U-Net architecture by extracting features from the input image through a sequence of convolutions, rectified linear unit (ReLU) activations, and max pooling. After each convolution, a feature map

Table 2. The differences in training complexity and computational requirements across the various models

Model	Training complexity	Computational requirements
U-Net	Due to its simple structure, U-Net requires fewer parameters during the training process, which makes the training process more efficient.	It has a relatively low computational load.
U-Net++	Due to the structural complexity, the number of parameters during training is higher than that of U-Net, necessitating greater computational resources for both training and tuning.	The computational requirements of U-Net++ have increased substantially due to the incorporation of additional skip and dense connections.
Attention U-Net	Although the attention mechanism is incorporated, the number of parameters in Attention U-Net does not increase as significantly as in U-Net++, while maintaining training efficiency.	The incorporation of the attention mechanism introduces additional computational and memory overhead.
ResU-Net	Despite the introduction of residual connectivity to accelerate training convergence, the structure of the residual network results in a higher training complexity compared to that of a regular U-Net.	Each convolutional layer incorporates residual connections, meaning that two paths (standard and residual) are utilized in each layer. This addition increases the computational burden, particularly when deep networks are involved, further exacerbating computational complexity.

**Figure 1. U-Net network framework.** This figure is cited from [12].

is produced. The ultimate downsampling layer comprises 1024 feature maps, each of size 28×28. The following sections explain the convolution operation in detail.

An input image of dimensions $H \times W \times C$ is processed, where H represents the height, W is the width, and C is the number of channels. A convolution operation is then performed, where each convolutional layer contains numerous convolutional kernels to execute the convolution operation on the input image. In the original U-Net structure, the convolution kernel is 3×3, with a stride of 1 and no padding. Upon completion of the convolution, a feature map is generated with the same number of convolution kernels, and the size of the feature map is as follows:

$$\text{Output Size} = \frac{\text{Input Size} - \text{Kernel Size} + 2 \times \text{Padding}}{\text{stride}} + 1$$

After the convolution, ReLU activation function is applied for feature extraction. ReLU is commonly used to enhance the network's ability to learn complex data patterns by introducing non-linearity [13]. It is computationally simple,

has a straightforward derivative (constant at 1 in most regions and 0 in others), and helps alleviate the vanishing gradient problem [14]. Its definition is as follows:

$$g(t) = \max(0, t)$$

In U-Net, ReLU improves the nonlinear representation of the network, enabling it to capture complex feature mappings and enhancing performance in medical image segmentation tasks. Each convolutional layer in the original U-Net undergoes two convolutions with convolutional kernels, and after each convolution, ReLU function is applied for feature extraction.

Following this, max pooling is typically employed in the pooling layer of U-Net. Max pooling involves sliding a window (of size $n \times n$) across the feature map with a stride of t and selecting the maximum value within the window at each position. The formula for calculating the maximum value is:

$$\text{Max_Value} = \max(X_{ij}, X_{i+1,j}, X_{i,j+1}, X_{i+1,j+1})$$

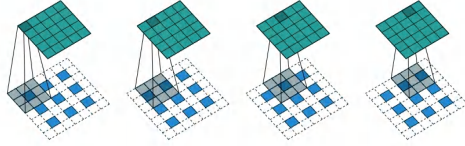
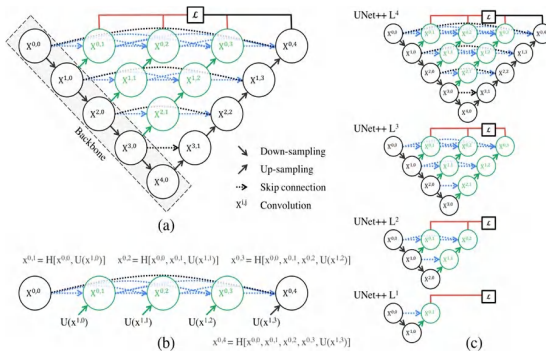
Where X_{ij} denotes the value at coordinate (i, j) of the input feature map.

Decoder

The decoder, functioning as a feature fusion network, aims to progressively restore the image dimensions and reconstruct the spatial details of the input image. This is achieved by integrating the feature maps from corresponding layers of the encoder with those of the decoder. Each stage of the decoder involves an upsampling operation, a convolution operation, and the amalgamation of feature maps.

Table 3. Summary of the development/improvements in the upsampling procedure

Presenters	Year	Achievement
Goodfellow et al. [16]	2014	GANs typically use transposed convolutional layers for upsampling.
Long et al. [17]	2015	Fractionally strided convolution was introduced for the first time.
Noh et al. [18]	2015	Applied to semantic segmentation tasks.

**Figure 2. Transposed convolution example.** This figure is cited from [15].**Figure 3. U-Net++ network framework.** (a) U-Net++ consists of an encoder and decoder that are connected through a series of nested dense convolutional blocks. The main idea behind U-Net++ is to bridge the semantic gap between the feature maps of the encoder and decoder prior to fusion. For example, the semantic gap between $(X^{0,0}, X^{1,3})$ is bridged using a dense convolution block with three convolution layers. In the graphical abstract, black indicates the original U-Net, green and blue show dense convolution blocks on the skip pathways, and red indicates deep supervision. Red, green, and blue components distinguish U-Net++ from U-Net; (b) Detailed analysis of the first skip pathway of U-Net++; (c) U-Net++ can be pruned at inference time, if trained with deep supervision. This figure is cited from [21].

The decoder begins with an upsampling operation, typically achieved through transposed convolution utilizing transposed convolution (also called inverse convolution). Unlike conventional convolution, transposed convolution accomplishes upsampling by inserting zeros between the outputs. A specific example of this process is illustrated in **Figure 2**. Suppose the input is a 5×5 matrix. Using a 3×3 convolution kernel with a stride of 2×2 and 1×1 zero-padding results in a 7×7 output matrix [15].

After up-sampling, the dimensions of the resulting feature maps are adjusted to match those of the feature maps from the corresponding encoder layer. **Table 3** summarizes the develop-

ment/improvements in the upsampling procedure.

Since transposed convolution enlarges the image rather than restoring it to its original size, feature concatenation is subsequently performed. This step is typically executed in a 'copy and crop' manner: the upsampled feature maps are copied to the corresponding encoder layer, and the encoder feature maps are cropped to match the dimensions. Typically, the cropping operation preserves the central region, ensuring identical dimensions in both height and width for both feature maps. Ultimately, the cropped feature maps are concatenated with the feature maps from the corresponding encoder level. This step facilitates the integration of low- and high-level features, which enhances the network's ability to recover image features.

Finally, a convolution operation is applied to the fused feature maps. This convolution in the decoder follows the same approach as in the encoder. After the first layer of the decoder completes its operations, this process is repeated for subsequent layers. Similar to the encoder, each decoder layer undergoes two convolution operations, with the image size doubling with each layer. However, it is noteworthy that in the final convolution of the last layer in the decoder, the original U-Net structure uses a 1×1 convolution kernel and does not apply an activation function.

Mainstream U-Net variants

U-Net++

U-Net++ was originally proposed by researchers from the University of Belgrade in Serbia, to address some limitations of the original U-Net architecture, such as information bottlenecks and feature loss in image segmentation tasks. U-Net++ introduces two key improvements over the traditional U-Net. First, it reconfigures the jump paths by incorporating jump connections. Unlike traditional U-Net networks, which solely fuses feature maps from encoders and decoders at the same level, U-Net++ introduces a dense convolutional block. This block facilitates jump connections, allowing the fusion of feature maps from encoders and decoders at diverse levels [19]. In this context, the formulation of the jump-connect path is delineated as

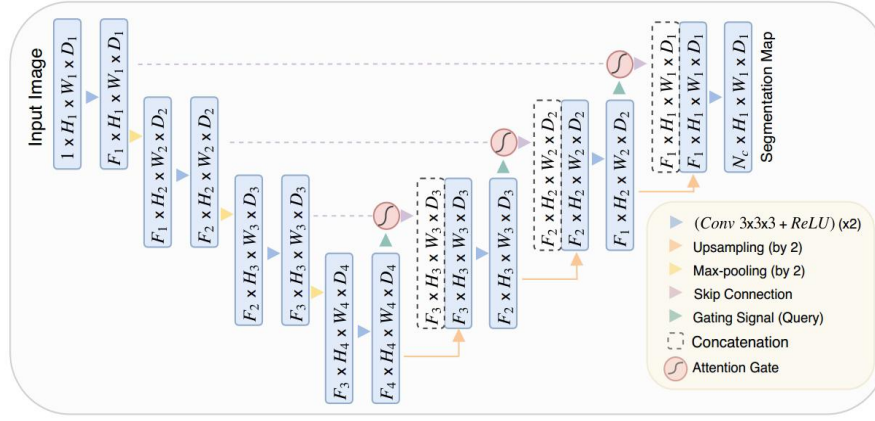


Figure 4. Attention U-Net network framework. This figure is cited from [24].

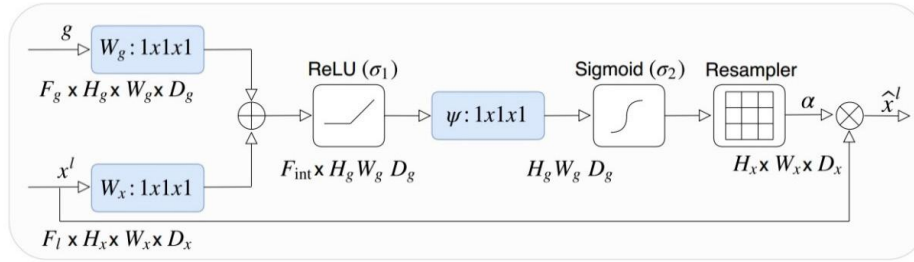


Figure 5. Schematic diagram of attention gate. This figure is cited from [24].

follows:

$$x^{i,j} = \begin{cases} R(n) & \text{If } n = 0 \\ R([x^{m,h}]_{h=0}^{m-1}, t(x^{m+1,n-1})), & \text{If } n > 0 \end{cases}$$

Second, deep supervision is incorporated in U-Net++, which involves integrating supervisory signals (loss functions) at various levels in the network [20]. These signals are introduced at each jump connection, enhancing the network's ability to learn features at various levels, thus refining segmentation performance. **Figure 3** illustrates a schematic representation of this network.

Attention U-Net

Attention U-Net introduces an attention mechanism that allows the network to automatically focus on specific areas of an image, mimicking human visual perception. When humans perceive stimuli, their attention is directed to particular areas based on their needs. This attention mechanism has been adopted in deep learning to enhance model performance by dynamically adjusting the focus on different parts of the input data, improving both performance and generalization [22]. Specifically, through the dynamic adjustment of attention weights, attention mechanisms enable the model to process different parts of input data in a more targeted manner.

Various types of attention mechanisms, such as channel attention and temporal attention, are tailored for various applications and tasks [23]. In U-Net architectures, the introduction of skip connections allows the concatenation of encoder features with corresponding decoder layers. However, this can lead to feature redundancy. To address this issue, Oktay et al. proposed the integration of an Attention Gate into the U-Net architecture, as depicted in **Figure 4** [24]. The integration of the Attention Gate within the skip connection dynamically adjusts the degree of attention on features, mitigating redundancy and enhancing the overall performance and generalization ability of the network.

In the design of attention gates, two inputs are initially involved, namely x and g . x denotes the feature map of the preceding layer in the encoder hierarchy, and g represents the global features, typically acquired by performing global pooling on the encoder output. Following this, x and g undergo a 1×1 convolution operation, followed by element-wise summation to derive the weighted, fused features. Subsequently, a non-linear transformation is introduced via the ReLU activation function, followed by another 1×1 convolution to enhance feature characterization. Attention weighting coefficients are then generated using the Sigmoid function, reflecting the importance of various features. Finally, the resulting coefficients are used to resample the features using trilinear interpolation and

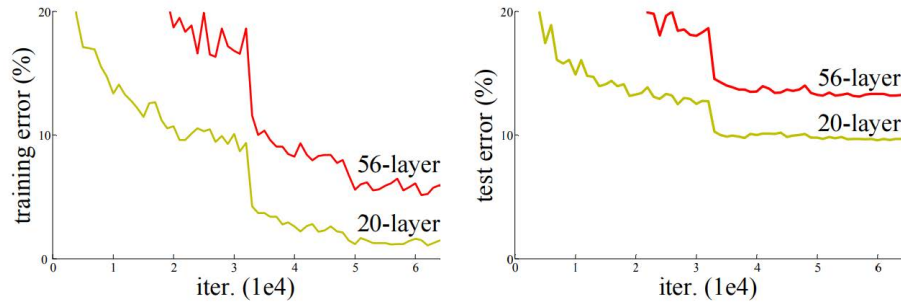


Figure 6. Training error and test error. This figure is cited from [25].

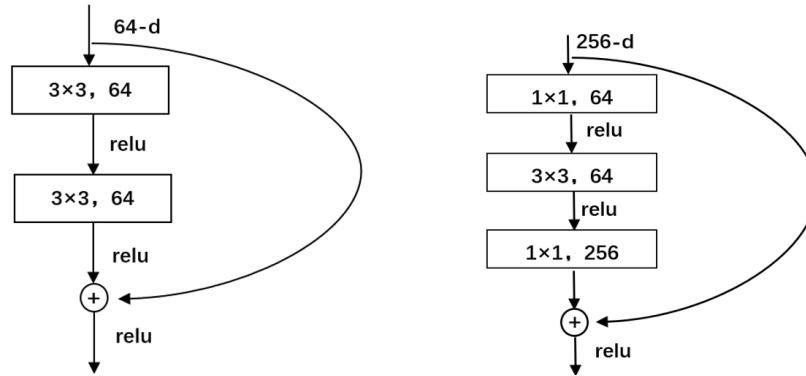


Figure 7. Residual structure. This figure is cited from [25].

multiply them with x to produce the final output, which is used for segmentation. This entire process is depicted in **Figure 5**.

ResU-Net

The design inspiration for ResU-Net is mainly derived from Residual Network and U-Net. Residual Network is a highly successful deep neural network structure that addresses the issues of gradient vanishing and gradient explosion during training by introducing residual connections. This innovation allows for the successful training of deeper network models.

As shown in **Figure 6**, the training and testing errors gradually increase as the network deepens. When a network is made deeper by adding more layers, issues of gradient vanishing or gradient explosion arise, hindering the model's convergence. Moreover, as the network continues to converge, it faces the problem of degradation, where the network's performance rapidly declines once a certain level of accuracy is achieved. To address these issues, a residual structure is introduced. The residual structure consists of two key components: the Primary Path and Shortcut Connection. As depicted in **Figure 7**, the left side represents the residual structure for networks with fewer layers, while the right side represents the residual structure employed in deeper networks [21].

In the residual structure, the primary pathway includes a convolutional layer followed by the ReLU activation function. The Shortcut Connection links the input directly to the output, bypassing the convolutional layer. This connection allows the input feature matrix to be added to the output feature matrix after it has passed through the primary pathway. Importantly, the feature matrix after passing through the Shortcut Connection must have the same dimensions as the output feature matrix of the primary pathway. The ResU-Net architecture integrates this residual connection structure into the standard U-Net framework. **Figure 8** delineates its configuration.

Application in ultrasound

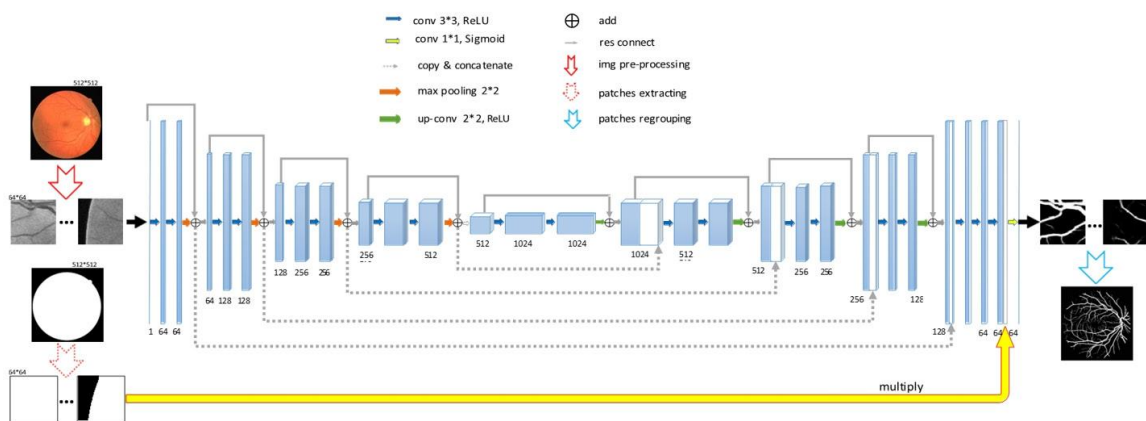
U-Net

Upon its proposal, U-Net has become a classical neural network in image segmentation, adept at extracting and processing distinct image features through the acquisition of feature representations. This capability has led to its widespread adoption in ultrasound imaging, with both the original U-Net architecture and its refinements being employed. Initially designed for the analysis of microscopic images, such as cell and tissue sample recognition, U-Net has since found significant applications in ul-

Table 4. The applications of U-Net in ultrasound image analysis

Presenters	Analytical objectives	Results
Zyuzin et al. [29]	Left ventricle endocardial border	Accuracy of LV segmentation is up to 92.3%
Wang et al. [30]	Brachial plexus	Dice coefficient got 0.6490
Tian et al. [31]	Brachial plexus	Accuracy with an IoU of 68.50%
Li et al. [32]	Levator hiatus	Gets a mean Dice of 96.4%±0.7%
Díaz-Vargas et al. [33]	Peripheral nerve	Average Dice coefficient of 0.70
Zou et al. [34]	Ovarian lesions	The AUC was 0.94, and the accuracy was 0.89.
Yang et al. [35]	Vascular lumen and medium-adventitia	0.87 JM, 0.82 mm HD for lumen and 0.86 JM, 1.07 mm HD for media (40 MHz dataset)
Ben et al. [36]	Breast cancer	Dice coefficient, Precision and Recall all reached over 80%.
Shodiq et al. [37]	Intravascular deep vein thrombosis segmentation	The mean IoU is 84.9% accurate, the measure of the mean Hausdorff distance is 4.17 of the score.
Suissa et al. [38]	Transversus abdominis plane	Global Dice score of 73.31%
Oghli et al. [39]	Automatic measurement of BPD and HC by segmentation of fetal head	Achieved a DSC of 0.95 and HD of 4.5
Lin et al. [40]	Laparoscopic ovarian	The ultrasound image showed no unnecessary interference noise, and the image clarity was significantly improved.

Note: LV, left ventricular; IoU, Intersection over Union; AUC, Area Under the Curve; JM, Jaccard Measure; HD, Hausdorff Distance; HC, Head circumference; BPD, Biparietal diameter; DSC, Dice similarity coefficient.

**Figure 8. ResU-Net network framework.** This figure is cited from [26].

trasound imaging. These applications include the segmentation of organs and tissues, the detection and segmentation of lesions, and cross-modal image registration. The following sections provide detailed exploration of these applications.

Organ and tissue segmentation

Ultrasound imaging, including liver and cardiac ultrasound, has gained popularity due to its real-time imaging capabilities, cost-effectiveness, and non-invasive nature, especially when compared to CT scans. However, ultrasound images tend to have high noise, which complicates the accurate identification of structures. Deep learning algorithms, notably U-Net, have proven effective in overcoming these challenges, enabling the segmentation and recognition of organs and tissues in real-time. This capa-

bility allows ultrasound machines to display the boundaries of these structures clearly, helping physicians detect them more efficiently and improving diagnostic and treatment efficiency.

Lesion detection and segmentation

Cancer, characterized by high morbidity and mortality, profoundly impacts individuals both psychologically and economically. Liver cancer, ranking as the second most common cancer after lung cancer, accounts for 11% of global cancer-related mortality [27]. Additionally, breast cancer remains a significant concern, making early screening essential to improve outcomes. Compared to other screening methods, ultrasound is less harmful to patients; however, manual segmentation and recognition of lesions can be time-consuming and inconsistent, with different clinicians potentially yielding

Table 5. The applications of U-Net++ in ultrasound image analysis

Presenters	Analytical objectives	Model	Results
Zhou et al. [41]	Carotid vessel wall	UNet++ ensemble approach	The Dice coefficients of the media-adventitia boundary (MAB) are 91.07%–91.56%, and those of the lumen-intima boundary (LIB) are 87.53%–89.44%.
Chen et al. [42]	Thyroid nodules	Based on the U-Net++ model, EfficientDet is used as the encoder, and the CSSE block is merged in the encoder and decoder.	Dice similarity coefficient of 0.9213
Orlando et al. [43]	Prostate	U-Net++	Dice similarity coefficient of 85%–89%
Zhu et al. [44]	lumen and median-adventitia	Feature pyramid network was added to the U-Net++ model.	Jaccard measure (JM) of 0.9080 ± 0.0321 and a Hausdorff distance (HD) of 0.1484 ± 0.1584 mm for the lumen border; JM of 0.9199 ± 0.0370 and HD of 0.1781 ± 0.1906 mm for the MA border.
Xu et al. [45]	Vessel	EAU-Net++: basis of U-Net++, EfficientNet is employed as an encoder.	The Accuracy of EAU-Net++ is 0.99.
Dai et al. [46]	Thyroid nodule	SK-Unet++: adds the adaptive receptive fields based on the selective kernel (SK) attention mechanisms into the Unet++ network.	Accuracy of 80.6%; Dice similarity coefficient of 84.7%
Zhou et al.[47]	Carotid plaques	UNet++ ensemble algorithm	The DSC in the SPARC dataset is 83.3%–85.7%; the DSC of the Zhongnan dataset is 88.6%.

Note: CSSE, Concurrent squeeze and spatial excitation; MA, Median Adventitia; DSC, Dice similarity coefficient.

varying results. Therefore, segmentation algorithms, like U-Net, have become crucial tools in standardizing and improving lesion detection [28]. A summary of U-Net's applications in ultrasound image analysis is presented in **Table 4**.

U-Net++

U-Net++, an advanced variant of the original U-Net, has demonstrated superior performance in various ultrasound applications. In obstetric ultrasound, the model excels in accurately outlining the morphology of fetal heart chambers and valves, as well as precisely delineating the placenta's boundaries. This significantly enhances the accuracy of fetal developmental assessments and the reliability of monitoring placental abnormalities. In abdominal ultrasound, U-Net++ effectively distinguishes tumors, cysts, and lesions in organs such as the liver and kidneys from normal tissues, providing a robust foundation for the development of subsequent treatment plans. Additionally, during ultrasound-guided puncture biopsy, U-Net++ can distinguish the target tissue from surrounding critical structures in real time, reducing procedure-related risks.

Furthermore, U-Net++ demonstrates high-precision segmentation in ultrasound images with blurred boundaries. For instance, in breast ultrasound tumor segmentation tasks, the model achieves a high Dice coefficient, enabling more

accurate tumor boundary delineation. Its multi-scale feature extraction capacity allows the model to capture both global information, such as the overall orientation of the carotid artery, and microscopic details, such as the fine structure of the intima, media, and plaques.

However, UNet++ also has limitations. The model requires substantial computational resources, and issues such as memory overflow may arise when processing large ultrasound image datasets on typical computer configurations, thereby limiting its applicability in resource-constrained environments. Additionally, the training process is also time-intensive, often taking days or even weeks to process large volumes of ultrasound images, significantly extending the research and development cycle and increasing costs. Furthermore, due to the nature of ultrasound images, U-Net++ may be susceptible to overfitting, leading to poor generalization performance when handling new or unseen images. A summary of U-Net++'s applications in ultrasound image analysis is presented in **Table 5**.

Attention U-Net

Compared to traditional U-Net models, Attention U-Net innovatively integrates an attention mechanism, which enhances region selectivity during the feature extraction. This innovation is particularly beneficial in ultrasound image analysis, such as thyroid nodule detection, where

Table 6. The applications of Attention U-Net in ultrasound image analysis

Presenters	Analytical objectives	Network	Results
Chen et al. [48]	Breast	Adaptive attention U-net	It has better generalization performance in the breast lesion segmentation.
Huang et al. [49]	Pleural effusion areas	Attention U-net	Average specificity of 91%, average Dice coefficient was 0.86.
Jain et al. [50]	Internal carotid artery and common carotid artery	Attention U-net	The correlation coefficient value f is 0.96; the AUC value is 0.97.
Katakis et al. [51]	Skeletal muscles	Attention U-net	Exhibits superb performance
Khaledyan et al. [52]	Breast	The novel approach by combining Sharp UNet and Attention UNet	The Dice coefficient, specificity, sensitivity, and F1 score values obtained were 0.93, 0.99, 0.94, and 0.94.
Shah et al. [53]	Brachial Plexus after clavicular surgeries	Attention U-net	The Dice coefficient improves about 1–2%.

Note: AUC, Area Under the Curve.

Table 7. The applications of ResU-Net in ultrasound image analysis

Presenters	Analytical objectives	Model	Results
Cai et al. [54]	Dipose layers	Attention Skip-Convolutions ResU-Net	It obtains a Dice of $99.06 \pm 0.95\%$ and an ASD of 0.19 ± 0.18 mm.
Dhar et al. [55]	Breast cancerous tissue	Connected ResU-Net (combine the algorithm of ResU-Net and Connected U-Net algorithms.)	IoU of 72.50% for the training and 63.10% for the testing dataset
Wang et al. [56]	Nerve of brachial plexus	ResU-Net	The dice coefficient 0.7093
Honghan et al. [57]	Breast tumor	Rms-se-U-Net	Dice similarity coefficient and intersection over union were 94.69% and 90.01%.

Note: ASD, Average Surface Distance.

the model can focus on key features like nodule boundaries, internal echogenic structures, and blood flow signals. For small papillary cancer nodules, Attention U-Net can accurately capture distinctive features such as gravel-like calcification and hypoechoic halos, significantly improving the accuracy of distinguishing between benign and malignant thyroid nodules and reducing the risk of misdiagnosis. Furthermore, Attention U-Net demonstrates strong resilience to noise and artifacts present in ultrasound images. In abdominal ultrasonography, noise introduced by factors like intestinal gas and patient movement can interfere with image clarity. The attention mechanism helps the model differentiate between true lesion features and noise artifacts, maintaining accurate lesion representation in organs such as the liver and pancreas, thus ensuring diagnostic accuracy.

Despite its advantages, Attention U-Net has certain limitations, particularly in its requirement for a large volume of accurately labeled ultrasound image data during training. Labeling ultrasound images is a challenging task that demands not only profound medical expertise but also a deep understanding of ultrasound imaging principles. Variations in labeling practices and standards among different personnel can lead to inconsistent data labeling quality, which may result in the model producing inac-

curate diagnostic results. A summary of Attention U-Net's application in ultrasound image analysis is presented in **Table 6**.

ResU-Net

ResU-Net is particularly effective in ultrasound applications that require high-precision segmentation, such as the automated segmentation of organs and lesions, including the liver, breast, heart, fetus, and kidneys. Ultrasound images often suffer from noise and artifacts, making traditional image segmentation methods struggle to accurately delineate the lesion regions, particularly in cases where tumors or organs exhibit fuzzy boundaries and minimal intensity differences. In such instances, ResU-Net leverages its residual connections to improve feature flow and gradient propagation, thereby enhancing the model's capacity to learn fine details. This results in more accurate segmentation, even in complex or noisy backgrounds with blurred boundaries. For instance, in the segmentation of liver ultrasound images, the boundaries between the liver and surrounding tissues are often unclear and prone to noise. ResU-Net effectively separates the liver contours from these blurred regions, supporting more accurate diagnosis of liver diseases, such as hepatocellular carcinoma and fatty liver. In breast ultrasound images, where tumor bound-

aries are similarly indistinct, ResU-Net, by learning residual features, not only achieves precise tumor segmentation but also demonstrates high sensitivity in detecting small lesions or early-stage tumors. In fetal ultrasound images, accurate segmentation of fetal organs is crucial. ResU-Net's ability to extract fine-grained features allows clinicians to better visualize critical organs such as the fetal heart and brain, providing valuable diagnostic information.

However, the application of ResU-Net in ultrasound images also encounters challenges. For example, when dealing with low-quality, low-resolution, or severely noise-contaminated ultrasound images, despite enhanced gradient flow, the model may still struggle with inaccurate segmentation, missed detections, or excessive artifacts when dealing with extreme blurriness or insufficient training samples. In such cases, combining ResU-Net with other image enhancement techniques or multimodal data fusion could improve its robustness and segmentation performance. A summary of ResU-Net's application in ultrasound image analysis is presented in **Table 7**.

III Conclusion

This review summarizes the application of U-Net-based models in ultrasound imaging, highlighting their significant contributions to medical diagnosis. Variants such as U-Net++, Attention U-Net, and ResU-Net have significantly enhanced both accuracy and efficiency in image segmentation tasks. Continuous advancements in segmentation-related research have been made, including the utilization of methods such as multi-scale feature pyramids and multi-level attention pyramids [58]. These approaches provide fine-grained feature extraction by focusing on local details and improving the representation of critical regions, thus offering valuable insights for enhancing segmentation performance.

Despite these advancements, challenges remain in ensuring the generalization and robustness of U-Net models across diverse datasets and varying imaging conditions. Future research should prioritize enhancing the adaptability of U-Net models to various ultrasound imaging environments while improving their interpretability for clinical use. Additionally, the integration of U-Net models with advanced techniques like transfer learning and multimodal imaging is anticipated to further expand their applicability and performance.

Author contributions: Yuxiang Wang: Respon-

sible for the general framework design of the article, the planning and organization of the research direction, led the writing and editing of the literature review section, and participated in the discussion and writing of the research methodology section. Miao Zhou: Provided medically relevant theoretical support and was responsible for the collection and organization of literature. Fangfang Chen: Participated in literature searches, provided experimental data support, and was responsible for producing some of the graphs. Jintao Duan: Partial data collection and organization and production of partial tables. Liangqing LIN: Provides theoretical support for relevant parts of the development of ultrasound medicine. Qinghua Wu: Language optimization and restructuring of the sections of the article. Wenhui Guo: The analysis of the research methodology and technical details is discussed in depth and the structure of the article is optimized. Haipo Cui: Led the revision of the article, participated in the final review and revision of the article, ensured that all authors' comments were integrated, and ultimately approved the article for submission.

References

- [1] Shin SJ, Jeong BJ. Principle and comprehension of ultrasound imaging. *J Korean Orthop Assoc* 2013;48(5):325-333.
- [2] Pehrson LM, Lauridsen C, Nielsen MB. Machine learning and deep learning applied in ultrasound. *Ultraschall Med* 2018;39(4):379-381.
- [3] Micucci M, Iulio A. Recent advances in machine learning applied to ultrasound imaging. *Electronics* 2022;11(11):1800.
- [4] Yap MH, Pons G, Marti J, et al. Automated breast ultrasound lesions detection using convolutional neural networks. *IEEE J Biomed Health Inform* 2017;22(4):1218-1226.
- [5] Almajalid R, Shan J, Du Y, et al. Development of a deep-learning-based method for breast ultrasound image segmentation. 2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA) 2018;1103-1108.
- [6] Shareef B, Vakanski A, Freer PE, et al. Estan: Enhanced small tumor-aware network for breast ultrasound image segmentation. *Healthcare (Basel)*. 2022;10(11):2262.
- [7] Hossain S, Azam S, Montaha S, et al. Automated breast tumor ultrasound image segmentation with hybrid UNet and classification using fine-tuned CNN model. *Heliyon* 2023;9(11):e21369.
- [8] Weng Y, Zhou T, Li Y, et al. Nas-unet: Neural architecture search for medical image segmentation. *IEEE Access* 2019;7:44247-

- 44257.
- [9] Zhong LF, Han M. A Unet-based method for Cell Segmentation Challenge. 36th Conference on Neural Information Processing Systems (NeurIPS 2022). 2022;1-5.
- [10] Zeng Z, Xie W, Zhang Y, et al. RIC-Unet: An improved neural network based on Unet for nuclei segmentation in histology images. IEEE Access 2019;7:21420-21428.
- [11] Tang P, Liang Q, Yan X, et al. Efficient skin lesion segmentation using separable-Unet with stochastic weight averaging. Comput Methods Programs Biomed 2019;178:289-301.
- [12] Ronneberger O, Fischer P, Brox T. U-net: Convolutional networks for biomedical image segmentation. Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18 2015;234-241.
- [13] Dubey SR, Singh SK, Chaudhuri BB. Activation functions in deep learning: A comprehensive survey and benchmark. Neurocomputing 2022;503:92-108.
- [14] Sharma S, Sharma S, Athaiya A. ACTIVATION FUNCTIONS IN NEURAL NETWORKS. Int J Eng Appl Sci Technol 2020;12(4):310-316.
- [15] Dumoulin V, Visin F. A guide to convolution arithmetic for deep learning. arXiv preprint arXiv:1603.07285 2016.
- [16] Goodfellow IJ, Pouget-Abadie J, Mirza M, et al. Generative adversarial nets. Proceedings of the 28th International Conference on Neural Information Processing Systems - Volume 2 2014;2672–2680.
- [17] Long J, Shelhamer E, Darrell T. Fully convolutional networks for semantic segmentation. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2015;3431-3440.
- [18] Noh H, Hong S, Han B. Learning Deconvolution Network for Semantic Segmentation. 2015 IEEE International Conference on Computer Vision (ICCV) 2015;1520-1528.
- [19] Zhou Z, Rahman Siddiquee MM, Tajbakhsh N, et al. UNet++: A Nested U-Net Architecture for Medical Image Segmentation. Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support 2018;3-11.
- [20] Lee CY, Xie S, Gallagher P, et al. Deeply-Supervised Nets. Proceedings of the Eighteenth International Conference on Artificial Intelligence and Statistics 2015;562-570.
- [21] Zhou Z, Siddiquee MMR, Tajbakhsh N, et al. Unet++: Redesigning skip connections to exploit multiscale features in image segmentation. IEEE Trans Med Imaging 2019;39(6):1856-1867.
- [22] Niu Z, Zhong G, Yu H. A review on the attention mechanism of deep learning. Neurocomputing 2021;452:48-62.
- [23] Guo MH, Xu TX, Liu JJ, et al. Attention mechanisms in computer vision: A survey. Comput Visual Media 2022;8(3):331-368.
- [24] Oktay O, Schlemper J, Folgoc LL, et al. Attention u-net: Learning where to look for the pancreas. arXiv preprint arXiv:1804.03999 2018.
- [25] He K, Zhang X, Ren S, et al. Deep residual learning for image recognition. Proceedings of the IEEE conference on computer vision and pattern recognition 2016;770-778.
- [26] Xiao X, Lian S, Luo Z, et al. Weighted res-unet for high-quality retina vessel segmentation. 2018 9th international conference on information technology in medicine and education (ITME) 2018;327-331.
- [27] International Agency for Research on Cancer. (2012). GLOBOCAN 2012: Estimated cancer incidence, mortality and prevalence worldwide in 2012. Available at: <https://gco.iarc.fr/en>
- [28] Li Y, Chouzenoux E, Charmettant B, et al. Lightweight u-net for lesion segmentation in ultrasound images. 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI) 2021;611-615.
- [29] Zyuzin V, Sergey P, Mukhtarov A, et al. Identification of the left ventricle endocardial border on two-dimensional ultrasound images using the convolutional neural network Unet. 2018 Ural Symposium on Biomedical Engineering, Radioelectronics and Information Technology (USBREIT) 2018;76-78.
- [30] Wang Y, Geng J, Zhou C, et al. Segmentation of ultrasound brachial plexus based on U-Net. 2021 international conference on communications, information system and computer engineering (CISCE) 2021;482-485.
- [31] Tian D, Zhu B, Wang J, et al. Brachial plexus nerve trunk recognition from ultrasound images: a comparative study of deep learning models. IEEE Access 2022;10:82003-82014.
- [32] Li X, Hong Y, Kong D, et al. Automatic segmentation of levator hiatus from ultrasound images using U-net with dense connections. Phys Med Biol 2019;64(7):075015.
- [33] Díaz-Vargas HM, Jimenez-Castaño CA, Cárdenas-Peña DA, et al. Peripheral Nerve Segmentation in Ultrasound Images Using Conditioned U-Net. Prog Artif Intell Pattern Recognit 2021;124-133.
- [34] Zou Y, Amidi E, Luo H, et al. Ultrasound-enhanced Unet model for quantitative photoacoustic tomography of ovarian lesions. Photoacoustics 2022;28:100420.
- [35] Yang J, Faraji M, Basu A. Robust segmentation

- of arterial walls in intravascular ultrasound images using Dual Path U-Net. *Ultrasonics* 2019;96:24-33.
- [36] Ben Ahmed I, Ouarda W, Ben Amar C. Hybrid UNET model segmentation for an early breast cancer detection using ultrasound images. *International Conference on Computational Collective Intelligence* 2022;464-476.
- [37] Shodiq MN, Yuniarno EM, Nugroho J, et al. Ultrasound image segmentation for deep vein thrombosis using UNet-CNN based on denoising filter. *2022 IEEE international conference on imaging systems and techniques (IST)* 2022;1-6.
- [38] Suissa N. Development and validation of automatic ultrasound nerve block guidance for regional anesthesia using deep learning. McGill University 2023.
- [39] Oghli MG, Moradi S, Sirjani N, et al. Automatic measurement of fetal head biometry from ultrasound images using deep neural networks. *2020 IEEE Nuclear Science Symposium and Medical Imaging Conference (NSS/MIC)* 2020;1-3.
- [40] Lin S, Zhao Y, Lei D, et al. Perioperative Nursing Management of Patients Undergoing Laparoscopic Ovarian Cystectomy Guided by Ultrasound Imaging under Intelligent Algorithm. *Comput Math Methods Med* 2022;2022:7193005.
- [41] Zhou R, Guo F, Azarpazhooh MR, et al. Carotid vessel-wall-volume ultrasound measurement via a UNet++ ensemble algorithm trained on small data sets. *Ultrasound Med Biol* 2023;49(4):1031-1036.
- [42] Chen C, Xu B, Wu Y, et al. Research on ultrasonic image segmentation of thyroid nodules based on improved U-net++. *2022 2nd International Conference on Bioinformatics and Intelligent Computing* 2022;532-536.
- [43] Orlando N, Gyacskov I, Gillies DJ, et al. Effect of dataset size, image quality, and image type on deep learning-based automatic prostate segmentation in 3D ultrasound. *Phys Med Biol* 2022;67(7):074002.
- [44] Zhu F, Gao Z, Zhao C, et al. A deep learning-based method to extract lumen and media-adventitia in intravascular ultrasound images. *Ultrason Imaging* 2022;44(5-6):191-203.
- [45] Xu X, Luan K, Liu J. EAU-Net++: An Efficient Segmentation Network for Ultrasound Vessel Images. *2023 IEEE International Conference on Mechatronics and Automation (ICMA)* 2023;1809-1814.
- [46] Dai H, Xie W, Xia E. SK-Unet++: An improved Unet++ network with adaptive receptive fields for automatic segmentation of ultrasound thyroid nodule images. *Med Phys* 2024;51(3):1798-1811.
- [47] Zhou R, Guo F, Azarpazhooh MR, et al. Deep learning-based measurement of total plaque area in B-mode ultrasound images. *IEEE J Biomed Health Inform* 2021;25(8):2967-2977.
- [48] Chen G, Li L, Dai Y, et al. AAU-net: an adaptive attention U-net for breast lesions segmentation in ultrasound images. *IEEE Trans Med Imaging* 2022;42(5):1289-1300.
- [49] Huang L, Lin Y, Cao P, et al. Automated detection and segmentation of pleural effusion on ultrasound images using an Attention U-net. *J Appl Clin Med Phys* 2024;25(1):e14231.
- [50] Jain PK, Dubey A, Saba L, et al. Attention-based UNet deep learning model for plaque segmentation in carotid ultrasound for stroke risk stratification: an artificial intelligence paradigm. *J Cardiovasc Dev Dis* 2022;9(10):326.
- [51] Katakis S, Barotsis N, Kakotaritis A, et al. Automatic extraction of muscle parameters with attention UNet in ultrasonography. *Sensors* 2022;22(14):5230.
- [52] Khaledyan D, Marini TJ, M. Baran T, et al. Enhancing breast ultrasound segmentation through fine-tuning and optimization techniques: sharp attention UNet. *Plos one* 2023;18(12):e0289195.
- [53] Shah N, Bhatia J, Vasavat N, et al. Ultrasound Nerve Image Segmentation Using Attention Mechanism. *International Conference on Advances and Applications of Artificial Intelligence and Machine Learning* 2022;789-802.
- [54] Cai R, Liu Y, Sun Z, et al. Deep-learning based segmentation of ultrasound adipose image for liposuction. *Int J Med Robot* 2023;19(6):e2548.
- [55] Dhar O, Yow K-C. Connected ResU-Net: A Deep Learning Model for Segmentation of Breast Cancer Ultrasound Images. *2023 IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)* 2023;88-93.
- [56] Wang R, Shen H, Zhou M. Ultrasound nerve segmentation of brachial plexus based on optimized ResU-Net. *2019 IEEE international conference on imaging systems and techniques (IST)* 2019;1-6.
- [57] Honghan Z, Liu DC, Jingyan L, et al. Rms-se-unet: A segmentation method for tumors in breast ultrasound images. *2021 IEEE 6th International Conference on Computer and Communication Systems (ICCCS)* 2021;328-334.
- [58] Tang H, Yuan C, Li Z, et al. Learning attention-guided pyramidal features for few-shot fine-grained recognition. *Pattern Recognit* 2022;130:108792.



Research progress of the biological clock gene in pancreatic cancer

Haoran Huang¹, Ge Yu², Rong Wan^{1,2*}

¹University of Shanghai for Science and Technology, Shanghai 200093, China. ²Department of Gastroenterology, Shanghai General Hospital, Shanghai Jiaotong University School of Medicine, Shanghai 200080, China.

Corresponding author: Rong Wan.

Declaration of conflict of interest: None.

Received May 20, 2025; Accepted August 4, 2025; Published December 31, 2025

Highlights

- Biological clock genes play a crucial roles in pancreatic cancer progression.
- Abnormal expression of biological clock genes is closely associated with malignant proliferation, invasion, metastasis, and resistance to chemotherapy.
- Multiple biological clock genes promote pancreatic cancer progression by modulating key signaling pathways.

Abstract

Biological clock genes, which regulate the body's circadian rhythm, play a crucial role in pancreatic cancer development. The abnormal expression of these genes is closely associated with malignant proliferation, invasion, metastasis, and resistance to chemotherapy. Several biological clock genes in pancreatic cancer tissues contribute to tumor progression by modulating key signaling pathways. Moreover, disruptions in circadian clock genes are significantly linked to poor prognosis and may serve as diagnostic markers and prognostic indicators. This review summarizes recent research on the regulatory mechanisms of biological clock genes in pancreatic cancer, emphasizing their potential clinical applications as diagnostic markers, therapeutic targets, and prognostic tools. These findings may lead to new approaches for personalized pancreatic cancer treatment.

Keywords: Biological clock genes, pancreatic cancer; biomarkers, therapeutic targets

Introduction

Pancreatic cancer is a highly malignant tumor of the digestive system and one of the most common malignant tumors worldwide, ranking 7th in cancer-related mortality [1]. Due to its subtle early symptoms, most patients are diagnosed at an advanced stage, making treatment challenging and the prognosis poor. Statistics show that pancreatic cancer was responsible for 466,003 deaths in 2020, making it the seventh leading cause of cancer-related deaths, with a 5-year survival rate of only 10% [2]. Projections suggest that by 2030, pancreatic cancer will become the second leading cause of cancer-related deaths [3, 4].

The biological clock is an internal rhythm developed over evolutionary time, allowing

organisms to adapt to periodic environmental changes such as temperature and light. It coordinates physiological, biochemical processes, and the metabolism of materials and energy across tissues, organs, cells, and molecules. In mammals, the biological clock consists of a central clock in the suprachiasmatic nucleus (SCN) of the hypothalamus and peripheral clocks distributed throughout various organs, including the brain, muscle, adipose tissue, liver, colon, and pancreas. The SCN detects the 24-hour light-dark cycle through the retina and transmits synchronized signals to peripheral clocks via neural, hormonal, and body temperature cues, maintaining bodily homeostasis [5, 6]. This biological clock is especially important for the endocrine system, as its function relies on the precise timing of communication among organs to ensure proper responses to environ-



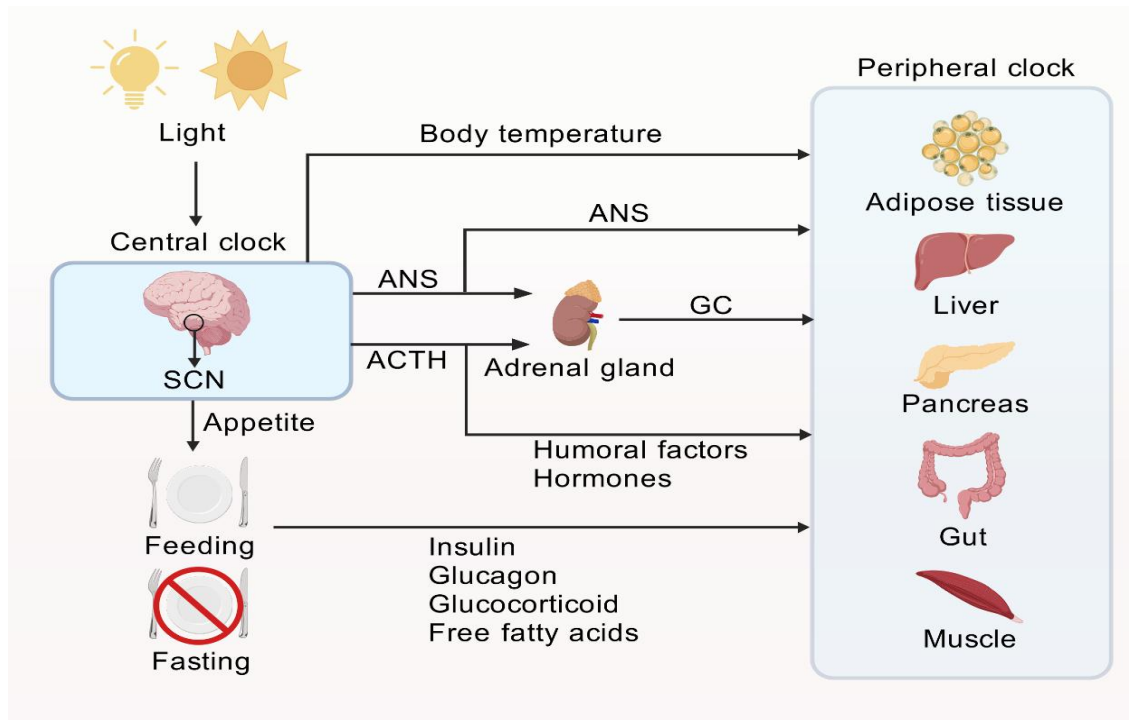


Figure 1. Biological clock regulatory network. SCN, Suprachiasmatic nucleus; GC, Glucocorticoid; ACTH, Adrenocorticotropic hormone; ANS, Autonomic nervous system.

mental changes or stressors [7]. Biological clock genes regulate circadian rhythms and processes such as cell metabolism and proliferation, and their dysregulation is closely linked to the occurrence, development, metastasis, and treatment resistance of many diseases, including cancers (**Figure 1**) [8-10].

As one of the organs most significantly affected by the biological clock system, the pancreas not only possesses its own intrinsic clock but also maintains normal physiological functions through regulation by the SCN and non-light signals. Given that pancreatic function is closely linked to the endocrine system, dysregulation of biological clock genes may greatly impact its physiological functions [11]. Based on this understanding, clinical trials have focused on optimizing treatment timing to enhance pancreatic function and improve therapeutic outcomes [12, 13].

The role of biological clock genes in pancreatic cancer is gradually being uncovered. This review focuses on the relationship between the biological clock system and the cell cycle, as well as the involvement of tumor suppressor genes and oncogenes in cancer development. Additionally, the latest research progress on biological clock disorders and their implications in pancreatic cancer are summarized, which may help discover new treatment strategies.

Molecular mechanisms of the biological clock gene

The mammalian biological clock primarily operates through a Transcription-Translation Feedback Loop, which regulates circadian rhythm and involves multiple key molecular genetic networks. The core transcription factors, CLOCK and BMAL1 form complexes that bind to the promoters of target genes, playing a positive regulatory role [14]. At the onset of the inactive phase, this complex induces the expression of inhibitory factors, including Period (PER1-3) and Cryptochrome (CRY1, CRY2). During this phase, other two core transcription factors, PER and CRY proteins form a complex that prevents the CLOCK: BMAL1 complex from entering the nucleus, inhibiting the transcriptional activity of their target genes and thus creating a feedback loop. At the end of the inactive phase, the expression levels of the target genes regulated by BMAL1/CLOCK reach their peak (**Figure 2A**).

Another feedback mechanism involves the retinoic acid receptor-related orphan receptor α (ROR α) and the heme-regulated nuclear receptor α/β (Reverse Erythroblastosis Virus α/β [Rev-Erb α/β]), which regulate BMAL1 expression by binding to ROR response elements [15]. The CLOCK: BMAL1 heterodimer promotes the transcription of Rev-Erb α , which then enters the nucleus and inhibits BMAL1 production by blocking its transcription. Meanwhile, the PER:

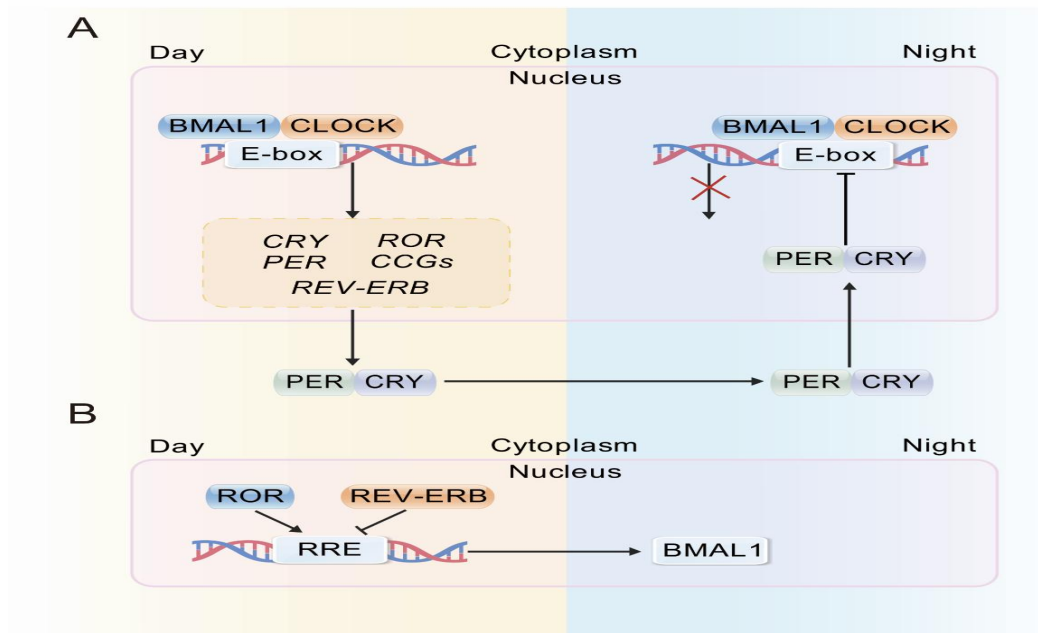


Figure 2. Molecular regulatory mechanisms of the biological clock. (A) Molecular mechanisms of Transcription/Translation Feedback Loop regulation; (B) Regulation of BMAL1 expression.

CRY protein complex also inhibits the transcription of Rev-Erb α . At the end of the night phase, ROR α binds to the BMAL1 promoter, promoting BMAL1 expression and creating favorable conditions for the cycle to restart. These molecular mechanisms oscillate over a roughly 24-hour period (**Figure 2B**) [16].

Under normal circumstances, biological clock genes support regular cell growth and metabolic functions. However, when the biological clock is disrupted, cell growth and metabolism may be adversely affected, potentially leading to tumor formation and progression [17]. Increasing evidence from studies shows that disruptions in the biological clock are closely linked to the development of pancreatic cancer [18, 19]. Abnormal expression of biological clock genes in pancreatic cancer tissues may disrupt tumor metabolism, alter the immune microenvironment, and promote tumor growth. For instance, the downregulation of the biological clock gene BMAL1 in pancreatic cancer tissues has been associated with the malignant progression of the disease [20]. Additionally, biological clock genes may influence pancreatic cancer development by regulating key signaling pathways, including the p53 tumor suppressor pathway and the transforming growth factor β (TGF- β)/MAD4 pathway [21, 22].

Research progress on the role of biological clock genes in pancreatic cancer

Abnormal expression of biological clock genes:

Alterations in core biological clock gene expression in pancreatic cancer tissues

Significant alterations in the expression of core biological clock genes have been observed in pancreatic cancer tissues. An analysis based on The Cancer Genome Atlas database revealed that the expression of multiple biological clock genes (e.g., CRY1, BTRC, CLOCK, CUL1, FBXL3, PER2, PER3, PRKAA1, PRKAA2, NR1D1, NR1D2, RRB, and NFIL3) was significantly higher in pancreatic cancer samples compared to normal samples [23]. Conversely, genes such as DBP, CSNK1D, ARNTL, CSNK1E, SKP1, and RORA were expressed at lower levels, while CRY2 showed no significant difference between normal samples and tumor samples. Additionally, research found that the expression of the BMAL1 gene was downregulated in pancreatic cancer tissues, acting as a tumor suppressor gene, with its expression correlating with clinicopathological characteristics of the disease [24]. Another study reported that the biological clock genes PER1 and PER2 were downregulated in human pancreatic cancer SW1990 cells, and the silencing of Epithelial Cell Transformation Sequence 2 via siRNA inhibited cancer cell proliferation and induced apoptosis by promoting the expression of PER1 and PER2 [25]. Moreover, the expression levels of biological clock genes such as PER1, PER2, PER3, CRY1, CRY2, TIPIN, TIM, CK1 β , BMAL-ARNTL, and CLOCK in pancreatic cancer tissues were significantly lower than those in matched adjacent tissues [26]. These changes in the expression

of biological clock genes may be closely related to the occurrence and development of pancreatic cancer. Abnormal expression of these biological clock genes may lead to metabolic disturbances within the tumor, alterations in the immune microenvironment, and other changes that accelerate cancer progression.

Biological clock genes and key signaling pathways in pancreatic cancer

Biological clock genes play a crucial role in the occurrence and development of pancreatic cancer, regulating the disease through multiple key signaling pathways. A prominent pathway is the p53 tumor suppressor pathway, which governs cell proliferation, survival, and metabolism [27]. Abnormal expression of biological clock genes, such as BMAL1, may impact the activity of this pathway. The study confirmed that overexpression of BMAL1 significantly inhibits the proliferation and invasion of pancreatic cancer cells while inducing G2/M phase cell cycle arrest [28]. The mechanism involves BMAL1 directly binding to the p53 gene promoter, promoting its transcription and activating downstream tumor suppressor pathways in a p53-dependent manner, thereby exerting anti-tumor effects. This interaction not only reveals a novel function of biological clock genes in pancreatic cancer but also provides a theoretical basis for developing new therapeutic strategies targeting this pathway (Figure 3).

The TGF- β /SMAD4 signaling pathway is also involved in regulating pancreatic cancer through biological clock genes. SMAD4 is a crucial component of the TGF- β signaling pathway, participating in the regulation of processes such as cell growth, differentiation, and apoptosis. SMAD4 mutation in pancreatic cancer disrupts normal TGF- β signaling, resulting in increased proliferation, migration, and invasion of pancreatic cancer cells. These mutations are significantly linked to local infiltration and distant metastasis of pancreatic cancer. In a study by Li et al., the interaction between circadian rhythm and the classical TGF- β signaling pathway was explored in SMAD4-positive and SMAD4-negative pancreatic ductal adenocarcinoma models [22]. They found that in SMAD4-positive pancreatic ductal adenocarcinoma cells, the expression of TGF β 1, SMAD3, SMAD4, and SMAD7 oscillates according to circadian rhythm pattern. Disruption of the biological clock altered the mRNA dynamics of these genes. Additionally, the expression of circadian clock genes such as DEC1, DEC2, and CRY1 also depends on the status of SMAD4. Activation of the TGF- β pathway leads to changes in the biological clock,

resulting in cell cycle arrest, accelerated apoptosis, increased invasiveness, and enhanced sensitivity to chemotherapy. The influence of TGF- β on the biological clock depends on SMAD4.

The miR-135b-BMAL1-YY1 signaling regulatory circuit also plays a key role in the development of pancreatic cancer. This pathway influences tumorigenesis and progression by regulating key processes like cell self-renewal, proliferation, and differentiation [29]. Studies have shown that miR-135b directly targets the 3' untranslated region of BMAL1, thereby disrupting the pancreatic biological clock [30]. This desynchronization between miR-135b and BMAL1 expression undermines the control of local circadian rhythm over tumor suppression, significantly promoting tumor growth in pancreatic cancer cells. YY1, a transcription factor, activates miR-135b and forms a feedback regulatory loop with BMAL1, creating a pathway with considerable predictive and prognostic value in patients with pancreatic cancer.

Therefore, biological clock genes play a complex role in regulating the occurrence and development of pancreatic cancer through multiple key signaling pathways, including the p53 tumor suppressor pathway, the miR-135b-BMAL1-YY1 signaling regulatory circuit, and the TGF- β /SMAD4 signaling pathway. These findings provide new ideas and targets for the precise diagnosis and treatment of pancreatic cancer.

Disruption of the biological clock and prognostic outlook in patients with pancreatic cancer

Abnormal expression of biological clock genes is closely linked to tumor invasiveness, chemotherapy resistance, and survival rates in pancreatic cancer patients. Recent studies have increasingly focused on the role of biological clock genes in predicting pancreatic cancer prognosis. BMAL1 promotes the proliferation of pancreatic cancer cells in vitro by binding to the p53 gene promoter. Research by Jiang et al. demonstrates that silencing BMAL1 in pancreatic cancer BXPC-3 cells induces anti-apoptotic and growth-promoting transcriptomic changes. Conversely, upregulation of BMAL1 inhibits the proliferation and invasion of pancreatic cancer cells, promotes apoptosis, and induces cell cycle arrest at the G2/M phase [24]. Subcutaneous tumor formation in nude mice confirmed that BMAL1 also suppresses the proliferation and invasion of pancreatic cancer cells in vivo [24].

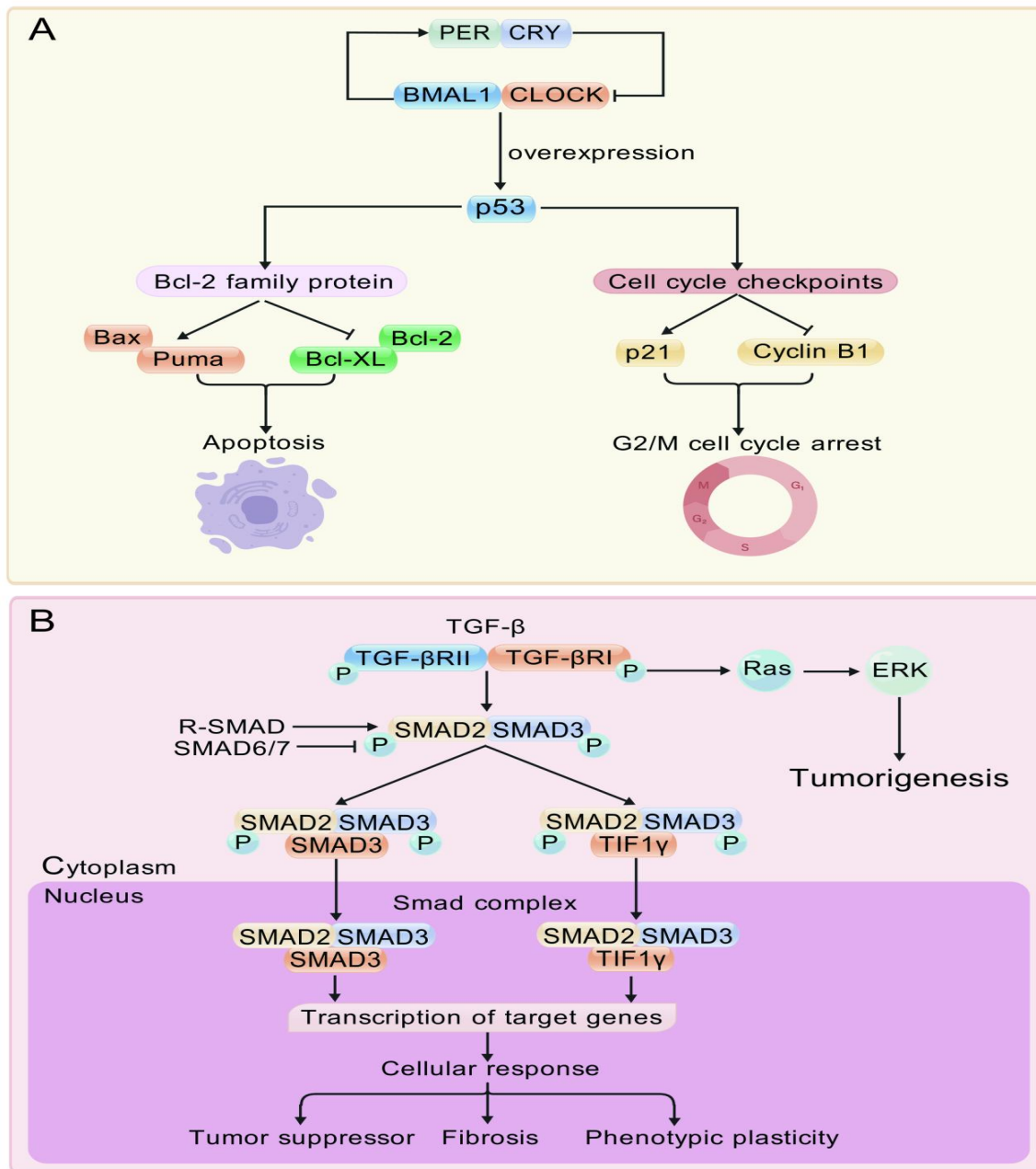


Figure 3. Biological clock genes and key signaling pathways in pancreatic cancer. (A) The p53 tumor suppressor pathway; (B) The TGF- β /Smad4 signaling pathway. TGF- β , transforming growth factor β .

The abnormal expression of biological clock genes also plays a significant role in chemotherapy resistance. A study showed that down-regulation of the BMAL1 gene significantly suppressed the expression of tumor suppressor genes such as p53, p16, RB1, BRCA1, and PTEN in pancreatic cancer cells. This suppression promoted proliferation, migration, and invasion of pancreatic cancer cells and enhanced chemotherapy resistance to gemcitabine [30]. This finding has important implications for pancreatic cancer treatment strategies, suggesting that monitoring the status of biological clock genes could help optimize chemotherapy regi-

mens and improve treatment outcomes.

The abnormal expression of biological clock genes is also closely associated with the survival rate of patients with pancreatic cancer (Table 1). A study confirmed that the expression levels of PER1, PER2, PER3, CRY2, Tipin, and CK1 β in tumor tissues from pancreatic cancer patients was significantly downregulated compared to benign lesion tissues [26]. Low expression levels of PER1, PER2, PER3, CRY2, TIPIN, CK1 β , CLOCK, and Bmal-ARNTL in tumors were significantly linked to decreased survival rates in these patients. Another study indicated that

Table 1. Summary of representative studies

Gene and/or model system	Finding	Reference
Kras-mutant mice	Chronic jet lag accelerated the development of pancreatic cancer precursor lesions and lesion grade. Non-cell-autonomous clock dysfunction was responsible for the expedited tumor development.	[31]
Bmal1 knockout mice	Knockout of the core clock gene Bmal1 in pancreatic cancer cells led to faster tumor growth, worse survival in mice, and enhanced chemotherapeutic resistance to standard chemotherapy agents used in the treatment of pancreatic cancer.	[26]
20 clock genes	A prediction model was constructed based on 20 biological clock genes, highlighting the important role of these biological clock genes in pancreatic cancer prognosis and the development of potential biomarkers.	[32]

variations in the CLOCK gene, including CLOCK-Gain and CLOCKLoss, were associated with tumor suppression and promotion, respectively. Among them, pancreatic cancer patients with high expression of the oncogene CLOCK had worse survival outcomes [33]. Zhang et al. investigated the relationship between abnormal expression of clock gene and the prognosis of common abdominal malignancies and found that high expression of the CRY2 gene and low expression of the DEC1 gene were linked to better prognosis, clinical stage, and tumor differentiation in patients with pancreatic cancer [32]. These findings underscore the crucial role of biological clock genes in the prognosis assessment of pancreatic cancer.

By constructing prognostic models based on the expression of biological clock genes, the survival rate of pancreatic cancer patients can be predicted more accurately, facilitating the development of personalized treatment plans. Zhang et al. analyzed clinical data and biological clock gene expression from 260 patients with pancreatic cancer in The Cancer Genome Atlas database and the International Cancer Genome Consortium database [33]. They identified 19 differentially expressed biological clock genes between pancreatic cancer tissues and normal tissues. Through LASSO regression analysis, 10 differentially expressed biological clock genes were identified as independent prognostic predictors for pancreatic cancer patients, leading to the development of a prognostic risk scoring model. A prognostic nomogram was then created, integrating clinicopathological features (e.g., pathological stage) and the risk score, demonstrating high accuracy in predicting the prognosis of pancreatic cancer patients.

Therefore, the abnormal expression of biological clock genes plays an important role in various aspects of pancreatic cancer, including tumor invasiveness, chemotherapy resistance, and patient survival rates. The development of prognostic models provides more accurate pre-

dictive tools, aiding in the optimization of treatment plans and improving both outcomes and prognosis for patients with pancreatic cancer.

Prospects for the clinical application of biological clock genes

As diagnostic and prognostic markers

Circadian rhythm genes have the potential to serve as both diagnostic and prognostic biomarkers for pancreatic cancer. Although the field remains in its early stages, next-generation sequencing technology offers a powerful tool for the rapid and accurate analysis of circadian rhythm biomarkers associated with the occurrence, prognosis, and therapeutic response of pancreatic cancer. Dagmura et al. investigated the association between PER2/PER3 gene tandem repeat polymorphisms (VNTR) and the susceptibility and clinical characteristics of pancreatic cancer. The results showed that the frequency of 4R/3R, 3R/3R genotypes, 3R alleles, and the combined genotype PER2/PER3 VNTR 4/5-4R/3R in peripheral blood mononuclear cells of pancreatic cancer patients was significantly higher than in healthy volunteers [34]. The PER2 VNTR 4/5 genotype was associated with peripheral nerve infiltration in pancreatic cancer patients. Additionally, the PER2 VNTR 3R allele may play a crucial role in the development of the disease and could serve as a key biomarker for predicting pancreatic cancer. In a separate study, Zhang et al. screened ten differentially expressed biological clock genes between pancreatic cancer tissues and normal tissues using The Cancer Genome Atlas and International Cancer Genome Consortium databases. A prognostic risk scoring model was developed and showed excellent predictive ability and discrimination for the prognosis of pancreatic cancer patients [32].

As a therapeutic target

Currently, the most effective treatment for pancreatic cancer combines surgical resection

with chemotherapy. Standard regimens include gemcitabine alone, gemcitabine with cisplatin, and combination therapies such as oxaliplatin, irinotecan, fluorouracil, and FOLFIRINOX. Other options include gemcitabine with docetaxel and capecitabine, or nanoparticle albumin-bound paclitaxel with gemcitabine. The primary mechanism by which gemcitabine works is by inhibiting DNA replication, thereby preventing tumor cell growth. However, clinical studies have shown that many patients develop resistance to gemcitabine, posing a significant challenge in treating pancreatic cancer [35]. Biological clock genes may influence the sensitivity of pancreatic cancer cells to chemotherapy drugs. For example, miR-135b promotes gemcitabine resistance by suppressing BMAL1. Schwartz et al. examined the effects of disrupting circadian rhythms using a BMAL1 knockout pancreatic cancer model and found that such disruption led to faster cancer growth, shorter survival, and increased resistance to chemotherapy [26]. Additionally, a study indicates that ASF1B may serve as a potential biomarker of pancreatic cancer and a novel therapeutic target. The study compared ASF1B expression levels in pancreatic cancer specimens with those in normal tissues, and found that ASF1B was commonly overexpressed in pancreatic cancer specimens [36]. By regulating the expression of the biological clock gene, it may be possible to improve the response of pancreatic cancer cells to chemotherapy, thereby enhancing treatment outcomes.

Study found that down-regulation of PER3 gene expression in the mouse pancreatic cancer cell line Pan02 inhibited cell proliferation and migration, reduced the stemness of pancreatic cancer stem cells, and suppressed the expression of stemness marker mRNA [37]. Therefore, biological clock genes may serve as potential therapeutic targets for pancreatic cancer. Additionally, a study showed that after human pancreatic cancer PANC-1 cells were treated with the biological clock protein REV-ERB agonist SR9009 for 72 hours, cell viability decreased significantly in a concentration-dependent manner, accompanied by an increased apoptosis rate, upregulated expression of caspase 3 protein and autophagy-related protein P62, and downregulated expression of autophagy-related protein 5 [38]. The observed cytotoxicity reduction further suggests that the biological clock protein REV-ERB agonist has a significant anti-tumor effect on human pancreatic cancer PANC-1 cells. Drug development targeting these biological clock genes may offer new therapeutic solutions for pancreatic cancer.

Circadian rhythm therapy

In clinical practice, all processes regulated by the circadian rhythm are potential targets for new tumor treatments. Circadian rhythm therapy involves targeting clock or clock-controlled gene products and timing treatments accordingly. The first approach aims to target circadian components or restore normal rhythms. For instance, a study showed that the circadian clock component of REV-ERBs is an effective target for alleviating cancer-induced tactile allodynia [39]. L-theanine has been shown to attenuate the viability, proliferation, and migration, while promoting apoptosis of melanoma cells. Additionally, L-theanine significantly enhances the expression of BMAL1, a clock gene in melanoma cells. Notably, the downregulation of BMAL1 suppresses the anti-cancer effects of L-theanine on melanoma cells [40]. A study developed a model to identify regulatory factors of the peripheral biological clock system, finding that modulation of Bmal1 or Per2 transcription, most likely by temperature or nutrient exposure cycles, is most effective in regulating the clock [41]. Another study has shown that cisplatin can stimulate multiple classified molecules, including DNA repair factors, DNA damage recognition factors, and transcription factors, which modulate the circadian clock through two mechanisms: the circadian clock control of DNA excision repair and the effect of circadian clock disruption on apoptosis [42]. Since most drug targets exhibit circadian rhythm behavior, identifying the best times of day and night for treatment application is important for enhancing efficacy and safety. It is worth noting that our understanding of circadian rhythms remains incomplete. The clinical application of circadian rhythm therapy remains complex, as tumors are heterogeneous. Therefore, drug treatment protocols need to be optimized to improve their effectiveness and efficiency.

Limitations of current research and future directions

Although some progress has been made in studying the relationship between biological clock genes and pancreatic cancer, several limitations remain. The primary issue is the relatively small sample size, which limits the generalizability and reliability of the findings. Additionally, while some studies have identified mechanisms through which biological clock genes influence pancreatic cancer, these mechanisms remain incompletely understood and require further research. Most importantly, there is a lack of large-scale, multicenter studies on biological clock genes in pancreatic cancer,

which impedes a comprehensive and in-depth understanding of their role.

Looking ahead, in-depth research on how circadian clock genes regulate metabolism and immune escape in pancreatic cancer will be a key focus. Metabolic reprogramming and immune escape are essential steps in the malignant progression of pancreatic cancer, and biological clock genes may play important roles in these processes. Exploring these mechanisms could identify new targets for pancreatic cancer treatment. Additionally, integrating multi-omics data (e.g., genomic, transcriptomic, proteomic, and other types of data) will create new opportunities for studying biological clock genes in pancreatic cancer. This approach will allow for a more comprehensive understanding of the regulatory networks of biological clock genes, helping to identify key genes and pathways involved in pancreatic cancer development. For preclinical research, developing treatment strategies that target biological clock genes will be a crucial task in the future. Constructing animal models with knocked-out or overexpressed biological clock genes will help evaluate the effectiveness and safety of drugs targeting these genes, supporting clinical translation. Ultimately, improving the translational use of biological clock genes in diagnosing and treating pancreatic cancer is our main goal. By investigating the functions and mechanisms of biological clock genes in pancreatic cancer, we aim to develop clinically applicable diagnostic markers and therapeutic targets, ultimately leading to better patient outcomes.

Conclusions

Disruption of biological clock genes expression plays a vital role in the development and progression of pancreatic cancer. An increasing body of evidence has demonstrated that abnormal expression of these genes is closely associated with malignant progression, chemotherapy resistance, and patient prognosis. A comprehensive investigation into the molecular mechanisms involving biological clock genes in pancreatic cancer will not only enhance our understanding of its pathogenesis but also open new avenues for targeted treatment. By examining how biological clock genes interact with key signaling pathways in pancreatic cancer, we can identify new therapeutic targets, refine treatment strategies, and improve patient outcomes. Furthermore, the potential of biological clock genes as diagnostic markers and prognostic indicators in clinical practice should not be underestimated. As research progresses and multi-omics data integration advances, bio-

logical clock genes are expected to play a significant role in supporting personalized treatment approaches for pancreatic cancer.

Author contributions: Haoran Huang was responsible for determining the review topic and framework, conducting the primary literature search and selection, writing the core sections, and integrating the initial draft. Yu Ge was responsible for supplementary literature retrieval, data/figure curation, writing specific sections, and reviewing and editing the entire manuscript. Rong Wan contributed to the formulation of the review topic, provided critical guidance on its overall direction and outline, and was responsible for overseeing the entire writing process. All three authors participated in the revision of the manuscript and approved the final version.

References

- [1] Sung H, Ferlay J, Siegel RL, et al. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA Cancer J Clin* 2021;71(3):209-249.
- [2] Cai J, Chen H, Lu M, et al. Advances in the epidemiology of pancreatic cancer: Trends, risk factors, screening, and prognosis. *Cancer Lett* 2021;520:1-11.
- [3] Shi Z, Lin J, Wu Y, et al. Burden of cancer and changing cancer spectrum among older adults in China: Trends and projections to 2030. *Cancer Epidemiol* 2022;76:102068.
- [4] Chung V, Sun V, Ruel N, et al. Improving Palliative Care and Quality of Life in Pancreatic Cancer Patients. *J Palliat Med* 2022;25(5):720-727.
- [5] Laothamatas I, Rasmussen ES, Green CB, et al. Metabolic and chemical architecture of the mammalian circadian clock. *Cell Chem Biol* 2023;30(9):1033-1052.
- [6] Kavakli IH, Ozturk N, Baris I. Protein interaction networks of the mammalian core clock proteins. *Adv Protein Chem Struct Biol* 2022;131:207-233.
- [7] Koop S, Oster H. Eat, sleep, repeat - endocrine regulation of behavioural circadian rhythms. *Febs j* 2022;289(21):6543-6558.
- [8] Ritonja JA, Aronson KJ, Leung M, et al. Investigating the relationship between melatonin patterns and methylation in circadian genes among day shift and night shift workers. *Occup Environ Med* 2022.
- [9] Papantoniou K, Konrad P, Haghayegh S, et al. Rotating Night Shift Work, Sleep, and Thyroid Cancer Risk in the Nurses' Health Study 2. *Cancers (Basel)* 2023;15(23):5673.
- [10] Lane JM, Qian J, Mignot E, et al. Genetics of

- circadian rhythms and sleep in human health and disease. *Nat Rev Genet* 2023;24(1):4-20.
- [11] Gomes PRL, Vilas-Boas EA, Leite EA, et al. Melatonin regulates maternal pancreatic remodeling and B-cell function during pregnancy and lactation. *J Pineal Res* 2021;71(1):e12717.
- [12] Hasenmajer V, De Alcubierre D, Ferrari D, et al. Exploring sexual function in adrenal insufficiency: findings from the Dual RElease hydrocortisone versus conventionAl glucocorticoid replaceMent therapy in hypocortisolism (DREAM) trial. *Andrology* 2025;13(2):302-313.
- [13] Guarnotta V, Amodei R, Giordano C. Metabolic comorbidities of adrenal insufficiency: Focus on steroid replacement therapy and chronopharmacology. *Curr Opin Pharmacol* 2021;60:123-132.
- [14] de Assis LVM, Demir M, Oster H. The role of the circadian clock in the development, progression, and treatment of non-alcoholic fatty liver disease. *Acta Physiol (Oxf)* 2023;237(3):e13915.
- [15] Schrader LA, Ronnekleiv-Kelly SM, Hogenesch JB, et al. Circadian disruption, clock genes, and metabolic health. *J Clin Invest* 2024;134(14):e170998.
- [16] Ono D, Honma KI, Schmal C, et al. CHRONO and DEC1/DEC2 compensate for lack of CRY1/CRY2 in expression of coherent circadian rhythm but not in generation of circadian oscillation in the neonatal mouse SCN. *Sci Rep* 2021;11(1):19240.
- [17] Yongzhuo Z, Dao W, Weiqing L, et al. The Expressional Circadian Rhythms of Clock Genes Bmal1 and Per2 in Raji, Hut -78, OCI-LY8 and HL-60 Cell Lines. *Pract J Cancer* 2024;39(04):530-533.
- [18] Pourali G, Ahmadzade AM, Arastonejad M, et al. The circadian clock as a potential biomarker and therapeutic target in pancreatic cancer. *Mol Cell Biochem* 2024;479(5):1243-1255.
- [19] Sharma D, Adnan D, Abdel-Reheem MK, et al. Circadian transcriptome of pancreatic adenocarcinoma unravels chronotherapeutic targets. *JCI Insight* 2024;9(9):e177697.
- [20] Farmakis D, Stravopodis DJ, Prombona A. TH301 Emerges as a Novel Anti-Oncogenic Agent for Human Pancreatic Cancer Cells: The Dispensable Roles of p53, CRY2 and BMAL1 in TH301-Induced CDKN1A/p21(CIP1/WAF1) Upregulation. *Int J Mol Sci* 2024;26(1):178.
- [21] Ma Z, Li Z, Wang S, et al. ZMAT1 acts as a tumor suppressor in pancreatic ductal adenocarcinoma by inducing SIRT3/p53 signaling pathway. *J Exp Clin Cancer Res* 2022;41(1):130.
- [22] Li Y, Basti A, Yalçın M, et al. Circadian Dysregulation of the TGFβ/SMAD4 Pathway Modulates Metastatic Properties and Cell Fate Decisions in Pancreatic Cancer Cells. *iScience* 2020;23(10):101551.
- [23] Lundø K, Dmytriyeva O, Spøhr L, et al. Lactate receptor GPR81 drives breast cancer growth and invasiveness through regulation of ECM properties and Notch ligand DLL4. *BMC Cancer* 2023;23(1):1136.
- [24] Jiang W, Jin L, Ju D, et al. The pancreatic clock is a key determinant of pancreatic fibrosis progression and exocrine dysfunction. *Sci Transl Med* 2022;14(664):eabn3586.
- [25] García-Montero C, Fraile-Martinez O, Cobo-Prieto D, et al. Abnormal Histopathological Expression of Klotho, Ferroptosis, and Circadian Clock Regulators in Pancreatic Ductal Adenocarcinoma: Prognostic Implications and Correlation Analyses. *Biomolecules* 2024;14(8):947.
- [26] Schwartz PB, Nukaya M, Berres ME, et al. The circadian clock is disrupted in pancreatic cancer. *PLoS Genet* 2023;19(6):e1010770.
- [27] Calheiros J, Corbo V, Saraiva L. Overcoming therapeutic resistance in pancreatic cancer: Emerging opportunities by targeting BRCA5 and p53. *Biochim Biophys Acta Rev Cancer* 2023;1878(4):188914.
- [28] Ou A, Zhao X, Lu Z. The potential roles of p53 signaling reactivation in pancreatic cancer therapy. *Biochim Biophys Acta Rev Cancer* 2022;1877(1):188662.
- [29] Gu W, Shen H, Xie L, et al. The Role of Feedback Loops in Targeted Therapy for Pancreatic Cancer. *Front Oncol* 2022;12:800140.
- [30] Jiang W, Zhao S, Shen J, et al. The MiR-135b-BMAL1-YY1 loop disturbs pancreatic clockwork to promote tumourigenesis and chemoresistance. *Cell Death Dis* 2018;9(2):149.
- [31] Schwartz PB, Walcheck MT, Nukaya M, et al. Chronic jetlag accelerates pancreatic neoplasia in conditional Kras-mutant mice. *Chronobiol Int* 2023;40(4):417-437.
- [32] Zhang LL, He QK, Lv YN, et al. Expression pattern and prognostic value of circadian clock genes in pancreatic adenocarcinoma. *Chronobiol Int* 2021;38(5):681-693.
- [33] Yao J, He C, Zhao W, et al. Circadian clock and cell cycle: Cancer and chronotherapy. *Acta Histochem* 2021;123(8):151816.
- [34] Dagmura H, Yiğit S, Nursal AF, et al. Possible Association of PER2/PER3 Variable Number Tandem Repeat Polymorphism Variants with Susceptibility and Clinical Characteristics in Pancreatic Cancer. *Genet Test Mol Biomarkers* 2021;25(2):124-130.
- [35] Chen JY, Yu PF, Liu P, et al. Expression of SH2 domain containing 2A in pancreatic ductal

- adenocarcinoma and its effect on gemcitabine resistance. *Chin J Exp Surg* 2024;41(4):728-731.
- [36] Kim JH, Youn Y, Lee JC, et al. Downregulation of ASF1B inhibits tumor progression and enhances efficacy of cisplatin in pancreatic cancer. *Cancer Biomark* 2022;34(4):647-659.
- [37] Wang JL, Zheng R, Wang MH, et al. Pancreatic cancer metastasis inhibition by circadian gene PER3 based on pancreatic cancer stem cells regulation. *J Dalian Univ Technol* 2024;64(3):236-242.
- [38] Fu X, Xu CY. Antitumor effect and mechanisms of circadian protein REV-ERB agonist on human pancreatic cancer PANC-1 cells in vitro. *Chin J Clin Pharmacol* 2022;38(22):2684-2687.
- [39] Yasukochi S, Yamakawa W, Taniguchi M, et al. The Circadian Clock Component REV-ERB Is an Analgesic Target for Cancer-Induced Tactile Pain Hypersensitivity. *J Neurosci* 2025;45(22):e1969242025.
- [40] Zhang R, Zheng S, Guo Z, et al. L-Theanine inhibits melanoma cell growth and migration via regulating expression of the clock gene BMAL1. *Eur J Nutr* 2022;61(2):763-777.
- [41] Martinelli J, Dulong S, Li XM, et al. Model learning to identify systemic regulators of the peripheral circadian clock. *Bioinformatics* 2021;37(Suppl_1):i401-i409.
- [42] Sadiq Z, Varghese E, Büsselberg D. Cisplatin's dual-effect on the circadian clock triggers proliferation and apoptosis. *Neurobiol Sleep Circadian Rhythms* 2020;9:100054.