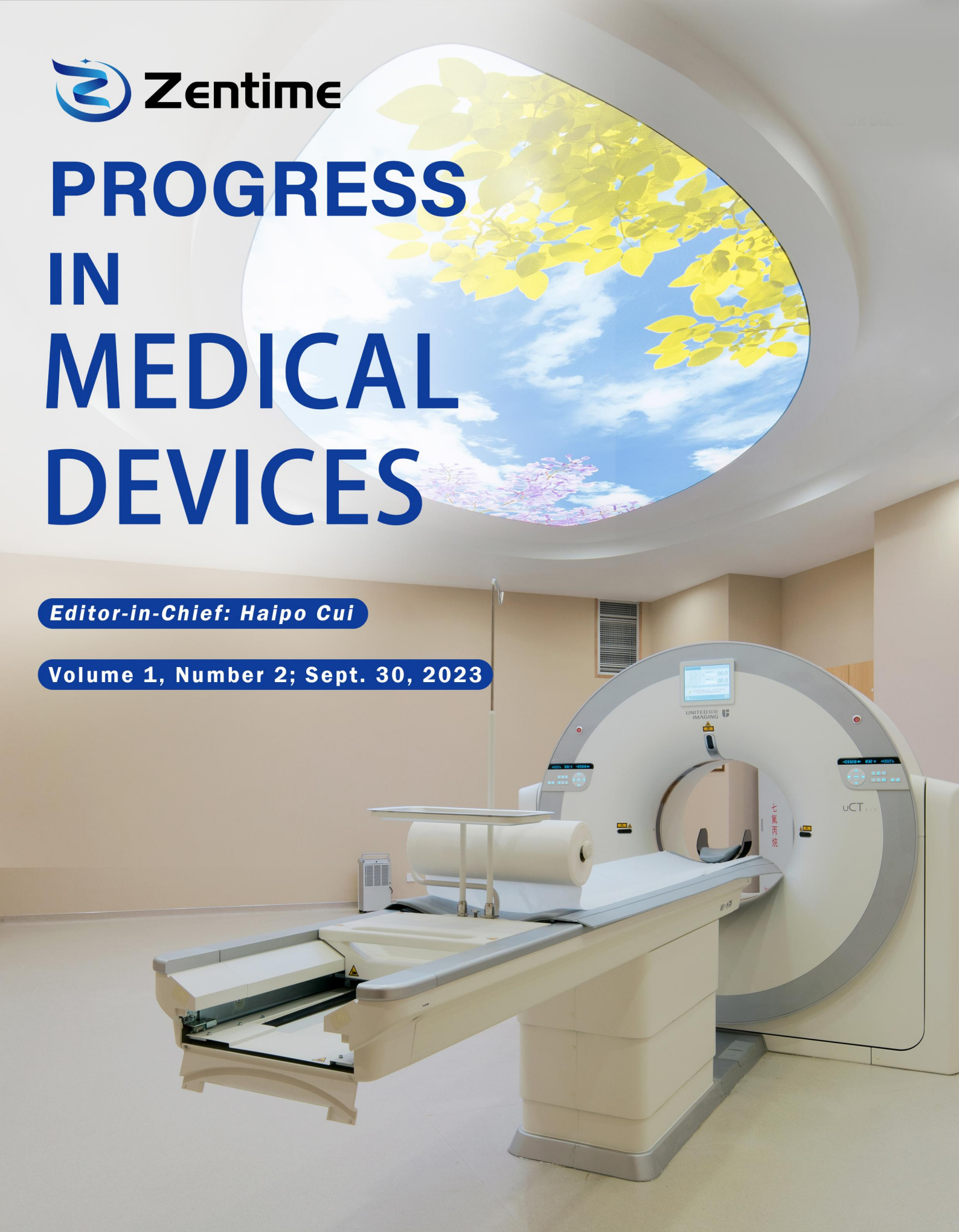
Zentime

PROGRESS IN MEDICAL DEVICES



Editor-in-Chief: Haipo Cui

Volume 1, Number 2; Sept. 30, 2023

ISSN: 2957-5478

Published by ZENTIME PUBLISHING CORPORATION

Progress in Medical Devices
eISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/473511cbaive

Research process on deep learning methods for heart sounds classification

Weifeng Wu, Yongqian Zhang, Qianfeng Xu, Jiuzhou Zhao, Rongguo Yan

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

Declaration of conflict of interest: None.

Received February 7, 2023; Accepted August 25, 2023; Published September 30, 2023

Highlights

- Denoising, segmentation, and feature extraction of heart sounds as well as its classification process are reviewed.
- A detailed exposition of diverse deep learning methods for heart sounds classification is presented.

Abstract

Cardiovascular diseases are still the primary threats to people's health around the world. Automatic heart sound classification technology, as a fast and efficient means for diagnosis and treatment, is of great clinical significance. With the rapid development of artificial intelligence technology, deep learning algorithms are widely used in automatic heart sound classification. This paper reviewed the key technologies related to the automatic classification of heart sounds in recent years, including heart sound denoising, segmentation, feature extraction, and classification recognition. The classification and recognition technologies related to deep learning are presented in detail, with a focus on the application and development of convolutional neural network and recurrent neural network, as well as various combination models for heart sound classification in the past five years.

Keywords: Cardiovascular disease, deep learning, heart sounds classification, convolutional neural network, recurrent neural network

Introduction

Cardiovascular diseases (CVDs) are associated with a high fatality rate and posing an escalating threat to public health [1]. In 2019, approximately 17.9 million global deaths were attributed to CVDs, with heart disease and stroke constituting 85% of it. In low- and middle-income countries, approximately 75% of the financial burden endured by individuals with cardiovascular conditions can be attributed to a lack of adequate medical resources [1].

CVDs, such as hypertension, coronary heart disease, and heart failure, are featured by rapid onset and substantial damage. Thus, early detection of heart diseases is of vital importance. A significant physiological signal directly associated with CVDs is the heart sound signal, a primary tool for diagnosing CVDs. Through analysis of cardiac sound waves, physicians can identify abnormal characteristics. However, heart sound auscultation currently presents two key limitations. Firstly, human auditory sensitivity constraints may complicate the discernment of faint physiological sounds from human internal organs [2]. Secondly, the accuracy of heart sound auscultation may be critically constrained when physicians lack comprehensive knowledge and clinical experience, as the diagnostic results are limited by doctors' subjective expertise and subjectivity [3]. As a consequence, the use of computer software-based heart sound analysis in the precise diagnosis of cardiovascular illnesses is in line with the trend.

Phonocardiogram (PCG) refers to audio data captured via the electronic stethoscope, and the heart sound signals can be classified by using computer-assisted techniques [4]. The parameters of these heart sound signals can vary based on the condition of the heart. Notably, there is a substantial divergence between normal and pathological cardiac sounds, as their corresponding PCG signals differ in character-

Address correspondence to: Rongguo Yan, School of Health Science and Engineering, University of Shanghai for Science and Technology, NO.516, Jungong Road, Shanghai 200093, China. Email: yanrongguo@usst.edu.cn.

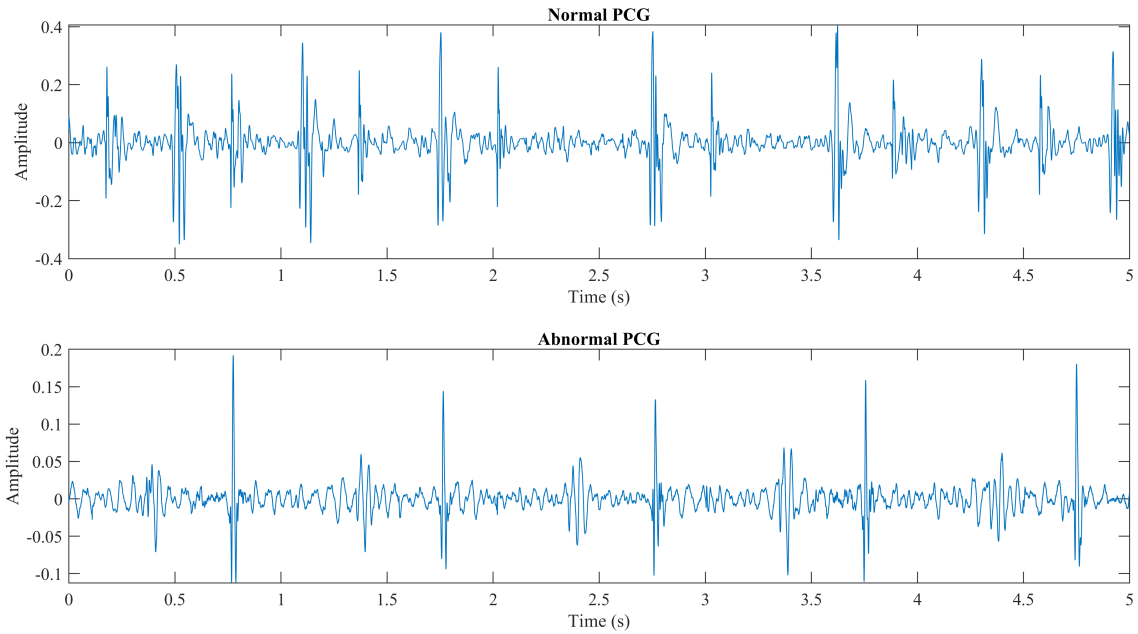


Figure 1. Samples of PCG records from PhysioNet /Computing in Cardiology Challenge 2016 dataset. PCG, Phonocardiogram.

istics such as magnitude, duration, intensity, spectrum, and uniformity [5].

With the precipitous advancement in digital signal processing and artificial intelligence technologies, deep learning has been extensively explored in the field of automatic heart sound analysis. It has the potential to reduce the labor costs associated with manual extracting features by employing data-driven feature learning. This approach is anticipated to bring about significant advances in the field of heart sound auscultation, and it is expected to be implemented in mobile devices.

Heart sounds and datasets

Heart sound signals are mechanical vibrations generated by the contraction and relaxation of the myocardium, as well as the interactions of blood flow with the valves, the atrioventricular walls, and vascular systems. These vibrations are transmitted through the surrounding tissues to the skin, generating sound signals on body surface. Heart sound signals are quasi-periodic. A typical cardiac cycle signal consists of four components: the first heart sound, the systolic phase, the second heart sound, and the diastolic phase [6]. In certain populations, the third and fourth heart sounds may also exist.

Heart murmur is an abnormal heart sound signal in the cardiac cycle, which contains pathological information of various CVDs. The most common heart murmurs are caused by valvular diseases, such as aortic stenosis, mitral regurgitation, mitral stenosis, mitral valve prolapse, and tricuspid regurgitation [7].

Among the existing studies, PhysioNet/Computing in Cardiology Challenge 2016 dataset is the most commonly used dataset to verify the performance of the heart sound classification algorithms, which contains only normal and abnormal heart sounds [8]. **Figure 1** shows examples of these two heart sounds. There is another dataset that is widely used in the multi-classification task of heart sounds [9]. This dataset contains five categories of sounds, namely normal, aortic stenosis, mitral regurgitation, mitral stenosis, and mitral valve prolapse. **Figure 2** shows the examples of these five types of heart sounds.

Process of heart sound classification

The automatic heart sound classification processes include denoising, segmentation, feature extraction, and classification.

Denoising

Due to the susceptibility to external factors, heart sound signals are commonly disrupted by disturbances, such as electromagnetic interference, random noise, and respiratory sounds. Therefore, the preliminary step in automatic heart sound classification involves the denoising of these signals. Two commonly employed techniques for heart sound signal denoising are discrete wavelet transform (DWT) and empirical mode decomposition (EMD).

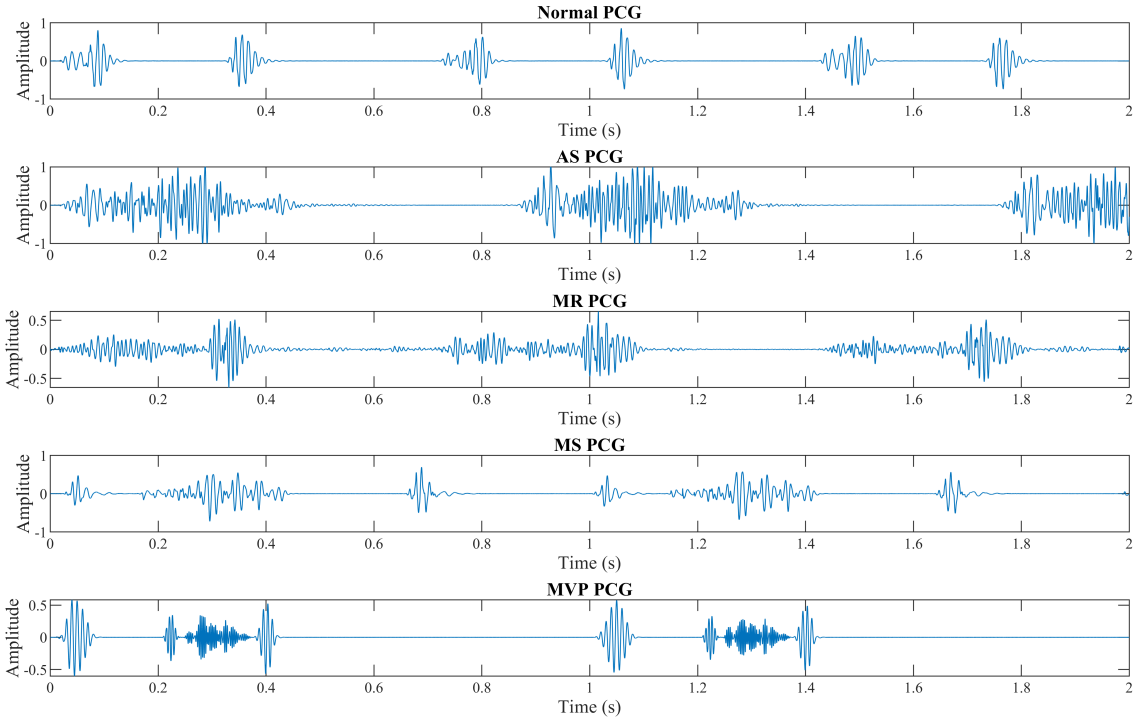


Figure 2. Samples of PCG records of five categories from the dataset. PCG, phonocardiogram; AS, aortic stenosis; MR, mitral regurgitation; MS, mitral stenosis; MVP, mitral valve prolapse.

DWT is a pivotal tool used for the analysis in the time-frequency domain. The core of this technique involves calculating the wavelet coefficients of both the noise and the heart sound signals. This is achieved by adjusting the scaling and shifting parameters of the so-called mother wavelet. Following this step, an appropriate threshold function can be identified and utilized to reconstruct the denoised heart sound signals [10]. Ali et al. conducted an examination on the influences of the wavelet basis function type, the threshold function type, and the signal decomposition level on the performance of the discrete wavelet denoising methodology [11]. In a distinct approach, Zhang et al. proposed an adjustable threshold function which can be fine-tuned using two parameters [12]. These adjustments alter the soft and hard threshold properties of the function, thus enhancing the integrity of the denoised heart sound signals. Despite the effectiveness of the DWT methodology, one significant drawback is that the denoised heart sounds may suffer from localized distortion, especially in the presence of anomalous heart sound signals.

Contrarily, EMD technique separates heart sound signals into several intrinsic mode functions, each comprising both heart sound components and noise components. By selectively eliminating the noise components, heart sound denoising can be achieved [13]. Furthermore, there have been experiments that combine the two methods. For instance, Dong et al. proposed a wavelet transform denoising technique grounded in complementary population EMD [14]. This method partitions the cardiac sound signals into intrinsic mode function components of varying frequencies. Then, the autocorrelation coefficient is deployed to distinguish between the pertinent signals and the noise within the range of the modal components. Following this, the heart sound signals are filtered and reconstructed employing the wavelet transform, resulting in the acquisition of the denoised heart sound signal.

The two methods have shown broad applicability in the field of heart sound denoising, although each of them faces some challenges. For DWT-based methods, their denoising effectiveness largely depends on thresholds, wavelet basis functions, and decomposition levels. The selection of these variables greatly affects the results, thereby leading to difficulties in parameter selection [15]. In contrast, EMD-based methods are highly sensitive to noise and suffer from issues such as the modal confusion [16]. Therefore, despite their wide application, these two strategies still have a large amount of research potential and room for optimization.

Segmentation

Heart sound segmentation is the division of the heart sound signals into multiple cardiac cy-

cles. In order to improve the accuracy of heart sound classification, automatic heart sound analysis focuses on segmentation of heart sounds. Currently, there are many successful heart sound segmentation methods, which are roughly classified into two categories. One is envelope-based segmentation method, and the other is based on statistical modelling.

The envelope-based segmentation is commonly employed in segmentation algorithms for heart sounds. The cardiac cycle is determined by detecting the peak value of the first or second heart sound following the extraction of the heart sound envelope. In the heart sound segmentation methods based on envelope, there are techniques, such as normalized average Shannon energy, Hilbert transform, homomorphic filtering, heart sound feature waveform extraction, Hilbert-Huang transform, and short-time corrected Hilbert transform [17-22]. In an in-depth research, Choi et al. examined and analyzed the envelope extraction techniques of the normalized Shannon envelope, the Hilbert transform envelope, and the heart sound characteristic waveform [23]. They found that the heart sound characteristic waveform envelope extraction method has uniform first heart sound and second heart sound information compared to the normalized Shannon envelope and Hilbert transform envelope curve.

In addition to using envelope analysis for heart sound segmentation, methods based on statistical models are also widely used, such as the hidden Markov model, the ensemble EMD method, the K-means clustering algorithm, and dynamic clustering [24-27]. These methods segment heart sound signals by utilizing different characteristics of the heart sound signals, such as the distribution of time-frequency energy, systole, and time-relatedness.

Both the two types of methods are widely used in the field of heart sound segmentation, but each has its own drawbacks. Envelope-based heart sound segmentation methods are susceptible to environmental noise interference, often missing the true peaks of the heart sounds but detecting noise peaks instead. Heart sound segmentation methods based on statistical models are constrained by the high degree of signal specificity in heart sounds across different individuals, making it challenging to model all heart sound signals with a unified model.

Feature extraction

The main purpose of feature extraction of heart sounds is to find some effective features to replace the high-dimensional original information. Presently, the characteristics employed for classifying heart sounds can be categorized as time domain, frequency domain, statistical domain, and time-frequency domain features. The methods including DWT, continuous wavelet transform, short-time Fourier transform, and Mel frequency cepstral coefficients (MFCC) are widely feature extraction [28-32]. Mel frequency spectral coefficient, as a MFCC without the discrete cosine transform, retains more of the original data [33-39]. MFCC and Mel frequency spectral coefficient can extract comparable characteristics of human perception of loudness and tone from audio data, hence they are commonly utilized in speech signal processing. In the field of heart sound feature extraction, several improved methods have been proposed. For instance, Deng et al. proposed an improved MFCC feature extraction method, which reduces the computational complexity and can dynamically represent the heart sound signals [40]. Likewise, Cheng et al. proposed an improved approach for classifying heart sounds by integrating MFCC with Gammatone Frequency Cepstral Coefficients [41].

In conclusion, due to the unique nature of heart sound signals, feature extraction is a critical step in the automatic classification of heart sounds, playing a significant role in the accuracy of subsequent heart sound classification results. Therefore, it is essential to continuously conduct research and improve methodology to optimize heart sound feature extraction.

Classification

Classification of heart sounds primarily provides qualitative data of heart sound detection and often separates heart sound signals into normal and pathological forms. Currently, machine learning techniques, such as support vector machine, hidden Markov model, artificial neural network, K-nearest neighbor method, and Euclidean distance, are utilized extensively in heart sound classification. By extracting the P value of features, Yadav et al. examined the performance of support vector machines, Naive Bayes model, random forest, and K-nearest neighbor in classifying heart sounds [42]. Deep learning techniques, such as convolutional neural networks (CNNs) and recurrent neural networks (RNNs), which are detailed in depth in Section 4, are the most popular approaches for classifying heart sounds.

Application of deep learning in heart sound classification

Table 1. Literature on heart sound classification using deep learning

No.	Year	Reference	Method	Input Features	Results		
1	2018	Meintjes et al. [29]	CNN	Scalograms	Sensitivity 87.4%	Specificity 86.7%	Accuracy 93.8%
2	2018	Bozkurt et al. [43]	CNN	Sub-band Envelopes	Sensitivity 84.5%	Specificity 78.5%	MAcc 81.5%
3	2019	Tan et al. [33]	CNN	MFSC		Accuracy 89.6%	
4	2019	Tan et al. [34]	CNN	MFSC	Sensitivity 91%	Specificity 88%	Accuracy 89.5%
5	2021	Xu et al. [44]	CNN	PSD	Sensitivity 77.6%	Specificity 94.6%	Accuracy 84.7%
6	2021	Boulares et al. [31]	CNN	MFCC	Sensitivity 94.6%	Specificity 94.6%	Accuracy 97%
7	2022	Li et al. [32]	CNN	MFCC	Sensitivity 89.5%	Specificity 89.7%	
8	2021	Shen et al. [45]	CNN	Spectrograms	Sensitivity 91.2%	Specificity 92.1%	Accuracy 91.1%
9	2021	Wang et al. [46]	CNN	Sub-band Envelopes		Accuracy 85.7%	
10	2022	Azam et al. [47]	CNN	MFCC&Fbank	Accuracy 86.2%	F1 84%	MAcc 85.1%
11	2022	Chen et al. [41]	CNN	MFCC&GFCC	Sensitivity 93.8%	Specificity 88.8%	Accuracy 91.3%
12	2021	Chen et al. [35]	CNN	MFSC		Accuracy 99.5%	
13	2021	Meng et al. [36]	CNN	MFSC	Sensitivity 98.5%	Specificity 97.2%	Accuracy 98.2%
14	2017	Maknickas et al. [37]	CNN	MFSC	Sensitivity 80.6%	Specificity 86.66%	Accuracy 93.7%
15	2022	Zhang et al. [38]	CNN	MFSC	Sensitivity 92.5%	Specificity 98.6%	Accuracy 96.6%
16	2022	Zhu et al. [39]	CNN-LSTM	MFSC	Sensitivity 84.3%	Specificity 85.6%	Accuracy 84.4%
17	2020	Li et al. [48]	CNN-LSTM	Spectrograms		Accuracy 85.7%	
18	2022	Liu et al. [49]	CNN-LSTM	Second order spectrum	Sensitivity 96.6%	Specificity 95.6%	Accuracy 95.3%
19	2022	Al-Issa et al. [50]	CNN-LSTM	FFT	Accuracy 93.8%	F1 85.6%	AUC 95.1%
20	2022	Chen et al. [51]	CNN-LSTM	Sequence	Sensitivity 87.0%	Specificity 89.0%	Accuracy 86.0%

Note: CNN, convolutional neural network; MFSC, Mel frequency spectral coefficient; PSD, power spectral density; MFCC, Mel frequency cepstral coefficients; GFCC, Gammatone Frequency Cepstral Coefficients; LSTM, long short-term memory networks; FFT, fast fourier transform; AUC, area under the curve; MAcc, mean accuracy.

This part reviews the applications of CNNs, RNNs, and Hybrid methods in heart sound classification. The research literature in recent years is shown in **Table 1**.

CNN methods for heart sound classification

CNN is a traditional model of deep learning. The LeNet is a typical CNN model proposed by Lecun et al. , as illustrated in **Figure 3** [52]. It consists of an input layer, two convolutional layers, two pooling layers, two fully connected layers, and an output layer. CNNs extract data features through the convolution layer based on local perception and decrease the parameter scale of the model via a weight-sharing technique, and the convolution process has local perception and weight sharing characteristics. The pooling layer, also known as the downsampling layer, efficiently reduces network parameters to minimize overfitting. The completely linked layer resides at the end of the CNNs. Each neuron of the fully connected layer is connected to the neurons of the previous layer, and the multi-dimensional feature vectors calculated by the previous layer are mapped into one-dimensional vectors, which reduces the influence of the feature position arrangement on the classification results and increases the network's robustness.

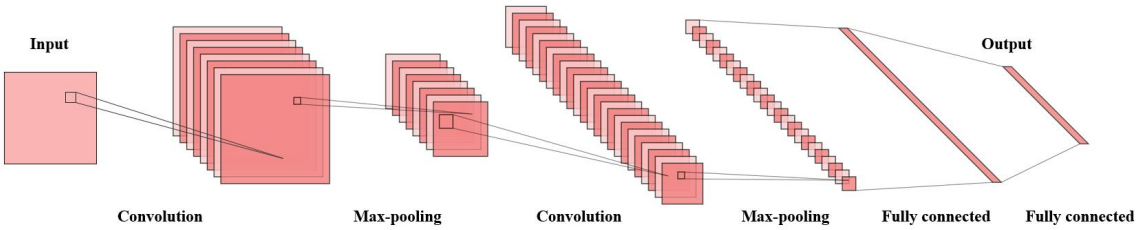


Figure 3. Network structure of LeNet.

Several optimization algorithms based on LeNet have been explored by researchers, with representative algorithms including AlexNet, GoogleNet, VGGNet, ResNet, and DenseNet. These algorithms have shown exemplary performance in heart sound classification tasks. For instance, Maknickas et al. proposed a heart sound classification model based on AlexNet, and obtained an accuracy of 93.7% [37]. Distinct from the most fundamental CNN model, AlexNet employs ReLU instead of Sigmoid as the activation function to avoid gradient vanishing when the network is deep and utilizes dropout to reduce the complexity of the model to prevent overfitting.

The performance of CNNs is affected by depth, width, and the size of the convolution kernel. The deepening of the number of network layers will have more linear transformations, which will help improve network performance. However, as the depth reaches a certain degree, model training becomes more difficult, culminating in the disappearance of gradients. In 2015, He et al. presented ResNet based on residual learning module, with network parameters of up to 152 layers, which can more effectively address the training challenges caused by network depth [53]. ResNet is currently more common in the classification of heart sounds.

Subsequent studies have built upon this foundation. For example, Zhang et al. presented a ResNet50 based on bilinear homologous transformation [38]. The model is comprised of multi-layer combinations of bottleneck structures with cross-layer connections and incorporates convolutional attention modules, better cross-entropy loss functions, and label smoothing techniques extensively. The greatest accuracy score reached 96.6%. Not resting on these laurels, further research pushed the boundaries of accuracy. Chen et al. proposed a heart sound classification model based on Inception-ResNet, using ResNet to increase the network's depth and convergence speed, and to improve the Inception network's accuracy [41]. The maximum possible score could reach 99.5% correctness.

However, the superiority of ResNet over other network models is not always clear-cut. As illustrated by Li et al., when comparing the performance of VGG and ResNet, the lack of a more complex network model can lead to an equivalent performance of the two [32]. In fact, the VGG model employing Focal loss and Log Mel features achieved an Unweighted Average Recall of 89.6% across all test groups.

Efforts to enhance CNN performance didn't stop at ResNet. Meng et al. compared the performance of CNN based on single connection (SingleCNN), skip connection (SkipCNN), and dense connection (DenseCNN) in the classification of heart sounds [36]. They investigated the effect of several convolutional layer connection methods on the classification of heart sounds. DenseCNN includes dense connections that enable the reuse of heart sound information, allowing the network to extract more heart sound features with fewer layers. DenseCNN had superior performance compared to the other two models, achieving 98.2% accuracy, 98.5% sensitivity, 97.2% specificity, and 0.0557 loss.

The application of CNNs in heart sound classification has witnessed considerable progression over the years, with successive studies contributing fresh perspectives and avenues for future exploration. These advancements underscore the extensive potential of CNNs within the realm of biomedical engineering. However, they also remind us of the crucial role that continuous inquiry and research play in further propelling the advancement within this discipline.

RNN methods and hybrid methods for heart sound classification

Compared with CNNs, RNNs have excellent time modeling ability and is widely used in natural language processing, speech recognition, and other fields. RNNs are feed-forward neural network variant with internal memory. With the

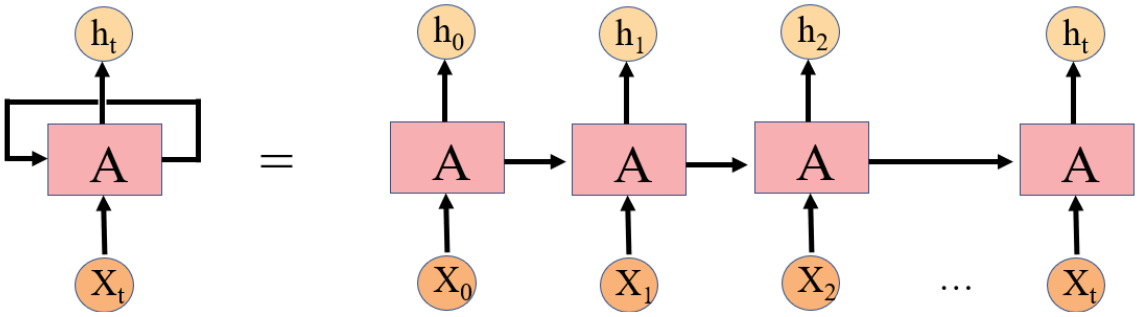


Figure 4. Network structure of RNN. RNN, recurrent neural network.

help of the internal memory structure, RNNs can memorize historical data and effectively predict future data. The typical structure of the RNN model is shown in **Figure 4**. The Xt is the input of the current module, and the ht is calculated as the output and participates in the input of the next module. Although RNNs have great advantages in time series modeling, it also has problems such as vanishing and exploding gradient. In order to solve these problems, researchers have proposed improved RNN models, including long short-term memory networks (LSTM) and gated recurrent units (GRU). LSTM is composed of forgetting gate, input gate and output gate, which make up for the deficiency of ordinary RNNs by introducing linear self-circulation unit. GRU is an important variant of LSTM. It uses update gate instead of forgetting gate and output gate, which makes GRU have fewer parameters and is easier to converge.

Further research by Bao et al. delved into the effects of different feature inputs on these network models [54]. In their study, the MFCC static features, and the extended MFCC with dynamic features, were used respectively as inputs to test their impacts on CNN and RNN models. Their findings suggest that the extended MFCC did not significantly improve the performance of RNN, although the accuracy of CNN using MFCC with dynamic features did show a slight improvement. Interestingly, when the static features of MFCC were used as input, RNN models outperformed CNN. However, among different RNN models, such as LSTM, bidirectional LSTM, GRU, and bidirectional GRU, no significant differences in performance were observed.

In audio processing, the CNN is more suitable for identifying specific portions of sounds, while the RNN is more excellent at handling sequences that vary over time. Each has its respective advantages, hence, the integration of CNN and RNN has become a prevalent direction in research.

A variety of studies have examined this approach. Chen et al. proposed the use of one-dimensional sequences as feature inputs, and the performance of a network combination of CNN and LSTM outperforms single networks, achieving an accuracy of 86.0% [36]. While using one-dimensional sequences as feature inputs does not have the best performance, it is beneficial for applications in real-time detection. On a similar note, Li et al. utilized CNN to extract the frequency domain features of data, complemented by the use of RNN to draw out the time domain features. This method yielded a classification accuracy rate of 85.7% [48]. In another study, Liu et al. explored the effectiveness of different CNN-LSTM structures in heart sound classification, demonstrating that the convolution kernel's size within the CNN can yield better outcomes within a certain range [49]. Additionally, they found that the single-layer unidirectional and single-layer bidirectional LSTM models were the most effective. However, they also noted that increasing the number of LSTM layers could impact the model's accuracy negatively. Despite this, their approach achieved a noteworthy classification accuracy of 95.2%.

Moreover, Al-Issa et al. designed a lightweight CNN-LSTM model to distinguish between five categories of heart valvular disorders [50]. They highlighted that using frequency domain features post-FFT transformation as input could enhance heart sound classification performance. They achieved an accuracy of 93.8%, a F1 score of 85.6%, and an AUC of 95.1% on the binary classification dataset. Additionally, on the enhanced five-class classification dataset, they attained an accuracy of 99.87%, a F1 score of 99.87%, and an AUC of 99.85%.

The integration of CNNs and RNNs, among other networks, is a well-adopted strategy, serving to unite the individual strengths of these architectures whilst offsetting their respective weaknesses. CNNs are particularly proficient in handling tasks related to image processing or spatial pattern recognition. On the other hand, RNNs are highly effective in dealing with sequence data and temporal pattern recognition. By combining these two, the stability and generalization capability of the model can be greatly enhanced. In real-world applications, this hybrid approach has been proven to be successful in numerous instances, demonstrating the flexibility offered by deep learning techniques and highlighting the potential of such combined models.

Summary and prospective

In recent years, as the frequency of CVDs has increased, automatic heart sound diagnosis technologies have enabled clinicians to diagnose patients more quickly and precisely, contributing to the healthcare system. This review summarizes the most recent research in the field of heart sound automatic classification during the past five years, the research status of denoising, segmentation, and feature extraction, and the use of a deep learning model in heart sound classification.

With the ongoing advancement of deep learning, numerous effective deep learning models, including CNN and RNN algorithms, have been presented and used in heart sound categorization. The algorithm model that combines the properties of the CNN and RNN is currently the study focus, and examination of datasets reveals that these algorithms have excellent performance and strong classification results. However, the complexity of the present heart sound categorization model needs to be reduced. The deployment of the lightweight model in mobile devices is anticipated to significantly improve the application of heart sound automatic categorization technologies.

Due to the lack of heart sound datasets, the majority of existing heart sound classification algorithms utilize both normal and abnormal heart sound datasets. However, various kinds of cardiac murmurs exist. Thus, to further assist physicians in the diagnosis of cardiovascular disorders, it is still important to increase and improve the heart sound datasets and investigate the multi-classification model in order to more precisely identify the types of heart murmurs.

References

[1] World Health Orgnization. Cardiovascular diseases (CVDs): Fact sheets. 2021.

[2] Xu WZ, Yu K, Ye JJ, et al. Automatic pediatric congenital heart disease classification based on heart sound signal. Artif Intell Med 2022;126:102257.

[3] Germanakis I, Petridou ET, Varlamis G, et al. Skills of primary healthcare physicians in paediatric cardiac auscultation. Acta Paediatr 2013;102(2):e74-e8.

[4] Fernando T, Ghaemmaghami H, Denman S, et al. Heart Sound Segmentation Using Bidirectional LSTMs With Attention. IEEE J Biomed Health Inform 2020;24(6):1601-1609.

[5] Wang M, Tang H, Feng TF, et al. Continuous and Noninvasive Estimation of Right Ventricle Systolic Blood Pressure Using Heart Sound Signal by Deep Bidirectional LSTM Network. Appl Sci (Basel) 2020;10(16):5466.

[6] Debbal SM, Bereksi-Reguig F. Time-frequency analysis of the second cardiac sound in phonocardiogram signals. Medical Phys 2005;32(9):2911-2917.

[7] Itagaki S, Adams DH, Anyanwu AC. Triggers for Surgical Referral in Degenerative Mitral Valve Regurgitation. Circu J 2013;77(1):28-34.

[8] Liu CY, Springer D, Li Q, et al. An open access database for the evaluation of heart sound algorithms. Physiol Meas 2016;37(12):2181-213.

[9] Yaseen, Son GY, Kwon S. Classification of Heart Sound Signal Using Multiple Features. Appl Sci (Basel) 2018;8(12):2344.

[10] Emmanouilidou D, McCollum ED, Park DE, et al. Adaptive Noise Suppression of Pediatric Lung Auscultations With Real Applications to Noisy Clinical Settings in Developing Countries. IEEE Trans Biomed Eng 2015;62(9):2279-2288.

[11] Ali MN, El-Dahshan ELSA, Yahia AH. Denoising of Heart Sound Signals Using Discrete Wavelet Transform. Circuits Sys Signal Process 2017;36(11):4482-4497.

[12] Zhang NN. The Application of an Improved Wavelet Threshold Function in De-noising of Heart Sound Signal. 2020 32nd Chinese Control And Decision Conference (CCDC) 2020;1768-1772.

[13] JiaRY, Jiann SS, Huang NE. Complementary ensemble empirical mode decomposition: a novel noise enhanced data analysis method. Adv Adapt Data Anal Theory Appl (Singapore) 2010;2(2):135-156.

[14] Dong L, Guo X, Zheng Y. Wavelet packet de-noising algorithm for heart sound signals based on CEEMD. J Vib Shock 2019;38(9):192-198,222.

[15] Li R, Bao C, Dou H. Speech enhancement algorithm based on wavelet transform. J Data Acquis Process (China) 2009;24(3):362-368.
[16] Gu J, Peng YX. An improved complementary ensemble empirical mode decomposition method and its application in rolling bearing fault diagnosis. Digital Signal Process2021;113:103050.
[17] Maglogiannis I, Loukis E, Zafiropoulos E, et al. Support Vectors Machine-based identification of heart valve diseases using heart sounds. Comput Methods Programs Biomed 2009;95(1):47-61.
[18] Huang NE, Shen Z, Long SR, et al. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. Proceedings of the Royal Society a-Mathematical Physical and Engineering Sciences. 1998;454(1971):903-995.
[19] Kumar D, Carvalho P, Antunes M, et al. Heart Murmur Recognition and Segmentation by Complexity Signatures.Annu Int Conf IEEE Eng Med Biol 2008; 2128-2132.
[20] Jiang ZW, Choi SJ. A cardiac sound characteristic waveform method for in-home heart disorder monitoring with electric stethoscope. Expert Syst Appl 2006;31(2):286-298.
[21] Tseng YL, Ko PY, Jaw FS. Detection of the third and fourth heart sounds using Hilbert-Huang transform. BioMed Eng OnLine 2012;11(1):8.
[22] Sun SP. An innovative intelligent system based on automatic diagnostic feature extraction for diagnosing heart diseases. Knowledge-Based Systems. 2015;75:224-238.
[23] Choi SJ, Jiang ZW. Comparison of envelope extraction algorithms for cardiac sound signal segmentation. Expert Syst Appl 2008;34(2):1056-1069.
[24] Gill D, Gavrieli N, Intrator N, et al. Detection and identification of heart sounds using homomorphic envelogram and self-organizing probabilistic model. Comput Cardiol 2005;957-960.
[25] Papadaniil CD, Hadjileontiadis LJ. Efficient Heart Sound Segmentation and Extraction Using Ensemble Empirical Mode Decomposition and Kurtosis Features. IEEE J Biomed Health Inform 2014;18(4):1138-1152.
[26] Gupta CN, Palaniappan R, Swaminathan S, et al. Neural network classification of homomorphic segmented heart sounds. Appl Soft Comput 2007;7(1):286-297.
[27] Tang H, Li T, Qiu TS, et al. Segmentation of heart sounds based on dynamic clustering. Biomed Signal Proce Control 2012;7(5):509-516.
[28] Coskun H, Deperlioglu O, Yigit T, et al. Implementation of Wavelet Transform Extrasystole Heart Sound With Convolution Method for Feature Extraction. 25th Signal Processing and Communications Applications Conference (SIU) 2017;1-4.
[29] Meintjes A, Lowe A, Legget M, et al. Fundamental Heart Sound Classification using the Continuous Wavelet Transform and Convolutional Neural Networks. 40th Annual International Conference of the IEEE-Engineering-in-Medicine-and-Biology-Society (EMBC) 2018;409-412.
[30] Chen QY, Zhang WB, Tian X, et al. Automatic Heart and Lung Sounds Classification using Convolutional Neural Networks. Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA) 2016;1-4.
[31] Boulares M, Alotaibi R, AlMansour A, et al. Cardiovascular Disease Recognition Based on Heartbeat Segmentation and Selection Process. Int J Environ Res Public Health. 2021;18(20).
[32] Li Z, iChang Y, Schuller BW. CNN-Based Heart Sound Classification with an Imbalance-Compensating Weighted Loss Function. Annual International Conference of the IEEE Engineering in Medicine and Biology Society IEEE Engineering in Medicine and Biology Society Annual International Conference 2022;4934-4937.
[33] Tan ZW, Wang WL, Zong R, et al. Research on Classification of Congenital Heart Disease Heart Sound Signal Using Convolutional Neural Network. Compu Eng Appl 2019;55(12):174-180.
[34] Tan ZW, Wang WL, Zong R, et al. Classification of heart sound signals in congenital heart disease based on convolutional neural network. J Biomedl Eng 2019;36(5):726-738.
[35] Chen W, Sun Q, Qi YY, et al. Deep Convolutional Neural Networks for Heart Sound Classification. Compu Eng Appl 2021;57(16):182-189.
[36] Meng LN, Xie HW, Ning C, et al. Heart Sound Classification Algorithm Based on Deep Convolutional Neural Network. Compu Meas Control. 2021;29(8):211-217,222.
[37] Maknickas V, Maknickas A. Recognition of normal-abnormal phonocardiographic signals using deep convolutional neural networks and mel-frequency spectral coefficients. Phys Meas 2017;38(8):1671-1684.
[38] Zhang JF, Zhang GY. Heart sound classification based on homologous bilinear residual network. Inf Technol 2022(5):36-42,47.
[39] Zhu LX, Qian K, Wang ZH, et al. Heart Sound Classification based on Residual Shrinkage Networks. 2022 44th Annual International Conference of the IEEE Engineering

in Medicine & Biology Society (EMBC) 2022;4469-4472.

[40] Deng MQ, Meng TT, Cao JW, et al. Heart sound classification based on improved MFCC features and convolutional recurrent neural networks. Neural Networks 2020;130:22-32.

[41] Chen C, Pan J, Sun J. Research on Heart Sound Classification of CHD Based on MFCC and GFCC Mixed Features. Compu Simul 2022;39(7):263-269.

[42] Yadav A, Singh A, Dutta MK, et al. Machine learning-based classification of cardiac diseases from PCG recorded heart sounds. Neural CompuAppl 2020;32(24):17843-17856.

[43] Bozkurt B, Germanakis I, Stylianou Y. A study of time-frequency features for CNN-based automatic heart sound classification for pathology detection. Compu Biol Med 2018;100:132-143.

[44] Xu CD, Xin PL, Zhou J, et al. Classification of Heart Sounds Using Power Spectral Density and Convolutional Neural Networks. Compu Eng App 2021;57(10):125-132.

[45] Shen YJ, Wang X, Tang M, et al. Recognition of S1 and S2 heart sounds with two-stream convolutional neural networks J Biomed Eng 2021;38(1):138-144.

[46] Wang XZ, Yang HB, Zong R, et al. Heart sound classification based on sub-band envelope and convolution neural network. J Biomed Eng 2021;38(5):969-978.

[47] Azam FB, Ansari MI, Nuhash SISK, et al. Cardiac anomaly detection considering an additive noise and convolutional distortion model of heart sound recordings. Artif Intell Med 2022;133:102417.

[48] Li W, Yang XD, Chen K. Heart sound classification based on CNN and RNN. Compu Eng Des 2020;40(1):46-51.

[49] Liu W, Sang S, Zhang H. Study on improved heart sound classification model based on CNN+LSTM. Electron Des Eng 2022;30(2):38-42.

[50] Al-Issa Y, Alqudah AM. A lightweight hybrid deep learning system for cardiac valvular disease classification. Sci Rep 2022;12(1):14297.

[51] Chen D, Xuan WP, Gu YX, et al. Automatic Classification of Normal-Abnormal Heart Sounds Using Convolution Neural Network and Long-Short Term Memory. Electronics 2022;11(8).

[52] Lecun Y, Bottou L, Bengio Y, et al. Gradient-based learning applied to document recognition. Proc IEEE 1998;86(11):2278-2324.

[53] He KM, Zhang XY, Ren SQ, et al. Deep Residual Learning for Image Recognition. IEEE Conference on Computer Vision and Pattern Recognition 2016;770-778.

[54] Bao XQ, Xu YJ, Kamavuako EN. The Effect of Signal Duration on the Classification of Heart Sounds: A Deep Learning Approach. Sensors 2022;22(6):2261.

Progress in Medical Devices
eISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/031576vfjfax

Research progress on biliary stents

Qi Zhang, Haipo Cui, Yan Zhang, Hexuan Jiang

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

Acknowledgment: This work was supported by the Shanghai Municipal Science and Technology Major Project (2021SHZDZX), Shanghai Industrial Collaborative Innovation Project (2021-cyxt1-kj07), Shanghai Science and Technology Innovation Action Plan (22S31902200), and Cooperation Fund of the Eighth Research Institute of China Aerospace Science and Technology Corporation (SAST2022-094), PR China.

Declaration of conflict of interest: None.

Received February 5, 2023; Accepted September 11, 2023; Published September 30, 2023

Highlights

- Plastic stents are more suitable for diseases such as benign bile duct stenosis.
- Metal-uncoated biliary stents are available for patients with malignant biliary obstruction.
- Degradability of biliary stents is a major research direction at present.

Abstract

Bile duct stenosis is a common condition in gastroenterology and hepatobiliary surgery and can be divided into benign stenosis and malignant stenosis according to different etiologies. The implantation of a gall stent into the site of the stenosis or obstruction is currently an important means of treating the bile duct stenosis. Biliary stents encompass two main types: plastic stents and metal stents. In recent years, biodegradable biliary stents and drug-eluting stents have also emerged. The material and structure of biliary stents have an important influence on their performance. In this paper, the research progress on biliary stent implantation technology in the treatment of biliary stenosis is reviewed. Besides, the advantages and disadvantages of biliary stents made from different materials and structures, along with their respective indications are summarized, and the development trend of degradable biliary stents is prospected.

Keywords: Bile duct stenosis, biliary stents, structural design, biodegradable biliary stents

Introduction

The bile system of the human body primarily consists of the gall bladder, the common liver canal, and the common bile duct. The common bile duct is a small tubular structure (average 7.5–11 cm long and 6–8 mm wide) that transports bile stored in the gallbladder to the duodenum [1]. Due to complex clinical pathologies, the lumen of the bile duct may be narrowed or structured, resulting in obstruction of normal bile flow [1]. Bile duct stenosis can be benign or malignant. Biliary stricture, without appropriate clinical intervention, can lead to impaired liver function, secondary biliary cirrhosis, and even death.

Li et al. reported that endoscopic retrograde cholangiopancreatography (ERCP) and percutaneous transhepatobiliary drainage are effective means for the treatment of biliary stricture and biliary obstruction [2]. ERCP is less aggressive with fewer complications than surgery, so endoscopic therapy is the preferred treatment for benign strictures [3]. In 1979, Soehendra et al. performed endoscopic biliary stenting for the first time using plastic stents (PSs), performing palliative drainage in patients with malignant biliary stenosis [4]. In patients with malignant biliary strictures such as distal or hilar stenosis, less than 20% of patients can undergo radical resection due to local spread and distant metastasis. Among these patients, 70% to 90% have biliary obstruction, resulting in jaundice, cholangitis, pruritus, malabsorption, coagulopathy, and hepatocellular dysfunction [3]. Due to the higher safety profile, ERCP is currently considered the optimal option for palliative care in patients with malignant distal biliary obstruction that cannot be resected or inoperable [5].

Address correspondence to: Haipo Cui, School of Health Science and Engineering, University of Shanghai for Science and Technology, NO.516, Jungong Road, Shanghai 200093, China. Tel: +86-21-55271290. E-mail: h_b_cui@163.com.

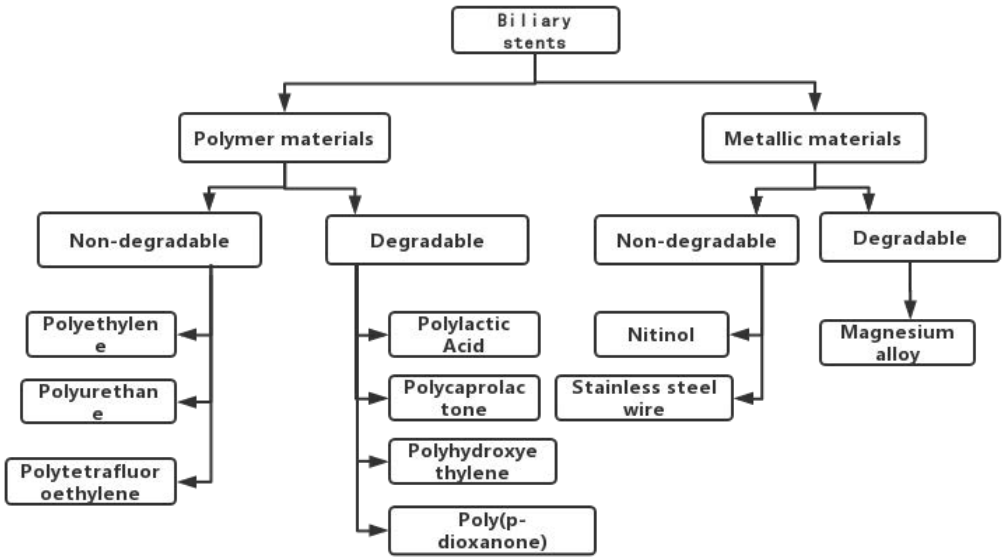


Figure 1. Classification of common biliary stents.

Biliary stents are tubular medical devices. According to the different stent materials, they are categorized into PSs and self-expanding metal stents (SEMSs). The plastic scaffold is mainly composed of polymer materials, including non-degradable polymer and degradable polymer. Metal scaffolds are also composed of non-degradable and degradable metals (**Figure 1**). Endoscopic stenting is the most commonly used method for the treatment of biliary stenosis, and there are two ways to insert biliary stents through the ERCP route: Endoscopic Retrograde Biliary Drainage and Endoscopic Metal Biliary Endoprosthesis. Of them, the PSs (**Figure 2A**) inserted through Endoscopic Retrograde Biliary Drainage is much cheaper than the metal stents, but it is prone to frequent occlusion in about three to six months after stenting and needs to be replaced in time. Metal stents (**Figure 2B**) can be divided into bare metal stents, partially coated stents, and fully coated stents [6, 7]. Bare metal stents should be considered when the expected survival period exceeds 6 months. Due to its large diameter and high resilience, it is often used in cases of malignant biliary stricture.

This article reviews the current progress on biliary stent implantation for the treatment of biliary strictures, compares the stent patency after the implantation of stents made from different materials and structures, and outlines the necessity for re-intervention and the incidence of stent dysfunction. In addition, this review summarizes the advantages and disadvantages of biliary stents made from different materials and structures, along with their respective indications, and finally prospects the development trend of degradable biliary stents.

PSs

Non-degradable PSs

The first literature on PSs was reported in 1979. Soehendra et al. successfully unobstructed the biliary system through percutaneous transhepatic bile duct implantation of a PS and achieved palliative bile drainage in patients with malignant biliary stenosis [4]. At that time, surgery such as biliary anastomosis was still the primary palliative means of reducing jaundice, while PSs, on the other hand, have a limited patency period and seem to be more suitable for preoperative or temporary internal drainage of benign diseases. Since 1979, researchers have continued to improve the function of plastic biliary stents. The advantages of PSs are low cost and easy to remove [8]. Commonly used non-degradable polymer materials are polyethylene (PE), polyurethane (PU), and polytetrafluoroethylene. Due to the limitation of endoscopic attachment access, the diameter of a stent is usually 3 - 5 mm, and therefore the stents tent to clog 3 to 6 months after placement [1, 6]. According to the characteristics of PSs, they are generally used in patients with malignant obstructive jaundice with an estimated survival of less than 3 months, or for patients with benign bile duct obstruction, who have an expected

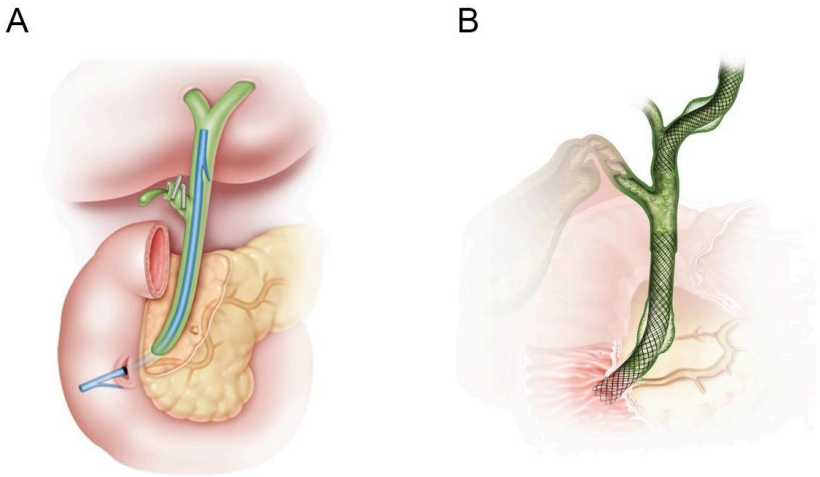


Figure 2. Schematic diagram of insertion of stents in biliary duct. (A) plastic stents; (B) metal stents. This figure is cited from [7].

life of over 3 months, for temporarily drainage along with surgical treatment [9, 10].

In terms of PS structure, depending on the degree of stricture, the length of non-degradable plastic biliary stent is about 5-20 cm and the diameter is about 3-5 mm. Although the PS can be bent, its diameter cannot be changed, so the type of stent needs to be determined according to intraoperative angiography. Notably, due to the limitation of the duodenoscopic channel, the maximum stent can only pass through 14F catheter at present. In 1982, Huibregtse et al. first placed a large-bore (10F) PS endoscopically into the bile duct, and the stent used single or double flaps instead of braids to prevent displacement [11]. Although this bracket has a unique flap design to eliminate the side holes, its performance was still not significantly different from that of traditional brackets. In 1988, Speer et al. demonstrated that the 10F stent had a much greater flow capacity than physiological bile flow due to its large diameter, so it was recommended to use a stent with a diameter of at least 10F to alleviate biliary obstruction caused by malignant tumors [12].

In terms of materials used in PSs, the first material widely used for stents in clinical treatment of biliary stenosis was PE in 1979. PE stents have been shown to have better plasticity and are easier to adapt to the shape of the biliary tract. However, due to the limitation of the diameter, PSs are prone to blockage. Vaishnavi et al. systematically characterized the complex chemical biofilm formed on the surface of the stent and clarified that smaller stent diameter, longer residence time (more than 6 months), and the presence of cholangitis at the time of implantation, could all lead to an increased risk of biofilm formation [13]. Easy displacement is also a major problem of PSs that needs to be improved. Johanson et al. retrospectively analyzed 322 cases of biliary stenting and found a distal displacement rate of 5.9% and a proximal displacement rate of 4.9% [14]. Cheon et al. compared a new PU rack (Cotton-Leung SOF-Flex bracket; Cook Endoscopy, Winston-Salem, NC, USA) and a standard 10F PE stent (Cook Endoscopy) [15]. The two types of scaffolds were found to have similar structures but different fabrication materials. Their results showed that there was no difference in patency between the two, but the total mobility and distal mobility of the PU rack were significantly lower than those of the PE stent (4.5% versus 29%, P=0.032 and 4.5% vs. 26.1%, P=0.049, respectively). Overall, PU stents are made of medical-grade PU, which has excellent biocompatibility, is not easy to scab, prevents against both distal and proximal stent migration, and has a smooth surface for easy placement and extraction.

In summary, the affordability of non-degradable PSs makes it widely used in clinical practice, although there are still shortcomings such as short patency time and easy displacement. There are also clinical reports of new non-degradable PSs. Researchers have modified the material to reduce the friction coefficient of the inner surface, improve the anti-bacterial adhesion performance, and increase the diameter of the stent, thereby prolonging the patency time of the PSs and reducing the occurrence of stent displacement.

Biodegradable PSs

Biodegradable biliary stent (BDBS), as a novel type of stent, is safe and effective and does not need to be replaced, with great applica-

Table 1. Comparison of stents made of different polymer materials

Polymer	Degradation time	Advantages	Deficiencies
PLA	About two years	Less bile drains and drains can be removed earlier; Easy deployment; Good immediate self-deployment	The radial support force is much lower than that of the metal stents [23].
PDX	In 3 months	8 mm self-expanding PDX stents with radial force up to 90% of normal metal stents [24]; Better flexibility, slower hydrolysis rate, and longer retention of mechanical properties in bile [25].	Cholangitis is prone to occur after surgery.
PDO	In 3-6 months	Superior flexibility and elasticity; Longer retention of mechanical properties.	Mild cholangitis

Note: PLA, polylactide; PDX/PDO, polydioxanone.

Figure 3. ARCHIMEDES biliary stent. This figure is cited from [21].

tion prospects. At present, the materials used in degradable plastic biliary stents include polylactide (PLA), polycaprolactone (PCL), polydioxanone (PDX or PDO), polyglycolic and their copolymers, etc. Since 2003, Ginsberg et al. reported the first in vivo experiment on BDBS. They successfully implanted PLA stents in the normal bile ducts of pigs without obvious stent complications, and the stents were still unblocked after six months [16]. In 2007, Laukkarinen et al. explored the therapeutic role of BDBS after cholecystectomy in minipigs, showing that compared with PSs, the PLA stent resulted in less bile drainage and earlier drainage tube extubation [17]. In 2010, Petry et al. pioneered the use of biodegradable stents in human bile ducts and successfully placed the PDX stent using percutaneous hepatic puncture cholangiogram in two patients with biliary anastomotic stenosis [18]. Cholangitis occurred after surgery, but no recurrence of stenosis during 2-year follow-up. Mauri et al. reported a study of PDO biodegradable biliary stents in 107 patients in 2013 and 2016, respectively [19]. They concluded that percutaneous placement of biodegradable biliary stents is a viable and safe strategy for the treatment of benign biliary strictures. PDO stents are superior in flexibility and can degrade within 3-6 months, and they can maintain their mechanical properties longer than most polymers such as PLA [20]. A degradable plastic biliary stent, ARCHIMEDES™ (**Figure 3**), has obtained Conformité Européenne certification in February 2019, which is a biliary and pancreatic drainage device made of degradable PDX. ARCHIMEDES is available in three different degradation rates to address different indications, offering products in sizes ranging from 2-3.4 mm in diameter and 40-225 mm in length [21]. The stent degrades rapidly within a few weeks of implantation and loses its mechanical strength, but its safety and efficacy still need to be further demonstrated in large-scale clinical trials. PLA typically takes an average of 2 years to fully degrade in the body, but the advantage of absorbable polymers is that the degradation rate can be adjusted to some extent by copolymerizing the appropriate polymer [1]. For example, PLA, polycaprolactone (PCA) and PGA, i.e., synthetic copolymers (lactide-co-glycolide-co-caprolactone)/PLGCL are completely absorbed within six months [22]. **Table 1** compares the degradation rates and advantages and disadvantages of PLA, PDX and PDO. However, at present, there is relatively little research on polycaprolactone stents, which means that there is still a lot of room for exploration in the future.

Metal biliary stents

Non-biodegradable metal stents

At the end of 1980s, SEMSs were introduced, and Nam et al. proved its superior patency than PSs [26]. With the development of minimally invasive technology, interventional therapies tend to be primary palliative treatment for

Table 2. Bismuth classification of hilar cholangiocarcinoma

Type	Feature
1	The tumor is located in the common hepatic duct and does not invade the confluence.
2	The tumor invades the common hepatic duct and the confluence of the left and right hepatic ducts.
3a	The tumor invades the common hepatic duct, the confluence of the left and right hepatic ducts, and the right hepatic duct.
3b	The tumor invades the common hepatic duct, the confluence of the left and right hepatic ducts, and the left hepatic duct.
4	The tumor invades the common hepatic duct, the confluence of the left and right hepatic ducts, and simultaneously invades the left and right hepatic ducts.

malignant bile duct obstruction in the hepatic portal [26]. There are two ways that metal stents unfold in the body: self-expanding and balloon-expanding. SEMSs can be divided into uncoated stents, partially coated stents, and fully coated stents, with metal uncoated stents being the first to be applied in practice, which are not easy to shift, but have the disadvantage of being difficult to take out due to tissue ingrowth. The fully coated metal stents have a non-porous membrane coating designed to reduce occlusion rates and are easy to remove, but with a higher risk of stent displacement [27]. Metal stents are often metal mesh cylinders, designed primarily by laser cutting, and some by knitting or braiding [1]. Metal stents have an inner diameter of up to 8-12 mm and a common length 4-12 cm [28]. When being implanted into the body, the metal stent is completely expanded to its full diameter. Therefore, the patency time of metal brackets is longer than that of PSs, usually 10 months.

With good expansion, better drainage effect, and larger lumen than PSs, metal biliary stents are commonly used in malignant biliary stenosis. Malignant bile duct stenosis is mainly caused by cholangiocarcinoma, periampullary cancer, pancreatic head cancer, tumor metastasis, or lymph node compression of the bile duct. Among the many types of cholangiocarcinoma, there is a unique type - high-lying cholangiocarcinoma (hilar cholangiocarcinoma of the liver). Liang et al. classified hilar cholangiocarcinoma into type 1, 2, 3, and 4 by Bismuth classification method (as shown in **Table 2**), of which type 3 can be divided into type 3a and 3b [29]. Patients with type 1 cholangiocarcinoma should be treated with ERCP for drainage or internal stent therapy, and patients with type 2 should be treated with double stenting.

Wagner et al. noted that in hilar tumors of the liver, metal stents improved patency and reduced the incidence of cholangitis compared to PSs [30]. The study of Chang et al. pointed out that double-branch drainage significantly reduced the occurrence of sepsis, decreased the mortality, and prolonged survival compared with single-branch drainage [31]. However, due to the difficulty of bilateral stent implantation and high operational technical requirements in clinical practice, it has not been popularized. The Y-shaped metal biliary stent, on the one hand, combines the advantages of metal stent and double branch drainage, and on the other hand is less difficult to implant than previous parallel double stents. Y-shaped metal biliary stent, characterized by a more open wire-woven mesh in the central section to facilitate the placement of the contralateral stent through the central section. Kim et al. retrospectively investigated the technical approach and clinical efficacy of implanting a new self-expanding biliary stent with nitinol material in the treatment of malignant hilar occlusion, which is a transverse stent with Y-shaped structure that combined of spiral stent and Z-shaped stent, called Y-shaped stent (NITI-S biliary Y stent, MI Tech, Heol, Korea) [32].

Bilateral stenting in patients with hilar cholangiocarcinoma requires multiple intubations with many instruments such as catheters, guidewires, dilators, and stents. Firstly, the first stent should be placed on the more curved side of the bile duct for successful access to the contralateral side. Secondly, the newly designed Y-shaped stent has a longer central mesh segment, and the braided part of the main stent is about 10-25 mm, allowing for better placement of the central part of the stent in the hilar bifurcation. The desired catheter can be deflected from the lateral wall of the common hepatic duct or common bile duct to the desired catheter through the middle and distal openings of the three-lumen catheter or the distal guidewire of the rotatable ERCP cannula. Finally, a slimmer open-hole stent can be used as the second scaffold to help the second scaffold pass more easily through the tight central mesh of the first stent. Tex et al. performed Y-shaped metal biliary stenting on 8 patients with hilar cholangiocarcinoma, 4 cases by external cardiopulmonary resuscitation, 3 cases by percutaneous transhepatic cholangio drainage, and 1 case by combining the two, of which 7 cases achieved good treatment results [33]. Kim et al. per-

formed Y-shaped metal biliary stenting on 12 patients with hepatic hilar cholangiocarcinoma, and the success rate of the operation reached 83.3% (10/12) [34]. In China, Y-shaped metal biliary stenting for the treatment of hilar malignant bile duct obstruction started late, and the patient data in the study of Li et al. showed that after Y-shaped metal biliary stent drainage, jaundice subsided rapidly and obviously, and there were no early complications such as cholangitis and bleeding [35].

Mukai et al. showed that uncoated metal stents, with a wire mesh structure, did not block the para-bile duct branches near the liver portal lesion, and the patency was better than that of the PSs [36, 37]. Srinivasan et al. pointed out that after the stent implantation, epithelial hyperplasia embedded in the tissue wall and blocked the stent, making it easier for the tumor to grow inward between the stent reticular filaments and overgrow proximal or distal to the stent as the main mechanism of SEMS obstruction [38, 39]. The inability to change or remove the stent after implantation is a major disadvantage of uncoated metal stents. Studies are underway to overcome these limitations to extend stent patency time. The Bonastent M-Hilar scaffold (Standard Sci Tech, Seoul, Korea) has a mesh structure with a smaller mesh size (1.6 mm * 1.6 mm) in the 25 mm central section, which Kwon et al. noted that could reduce tumor ingrowth [40]. At present, the commonly used materials for metal biliary stents are platinum (platinum core and nitinol sleeve), stainless steel, and nitinol. The study of Park et al. showed that compared with uncoated metal stents, coating silver nanoparticles in nitinol stents and then implanting them into extrahepatic bile ducts in rabbits significantly reduced submucosal fibrosis and inflammation, which can improve silt accumulation and epithelial hyperplasia [41]. But at present, related research is only limited to in vitro experiments and animal experiments. Emerging partial coverage (PCSEMS) and full-coverage self-expanding metal scaffolds (FCSEMS) are covered with a thin polymer membrane that prevents tumor growth inward and is easy to remove [42]. At present, commonly used polymer coating materials include degradable or non-degradable materials such as polytetrafluoroethylene, PU, and silicone. Metal-covered stents are commonly used for malignant hilar biliary obstruction. Naitoh et al. reported a significant advantage of CSEMS, which was their removability during reintervention, they also found that FCSEMS was easier to remove than PCSEMS during reintervention [8]. FCSEMS was easily removed in all re-intervention cases, while PCSEMS could not be removed in all cases. Kullman et al. proposed and explained that the biggest problem with coated stents compared to non-coated stents is that the covering material can block the opening of the cystic and pancreatic ducts, resulting in cholecystitis and pancreatitis, which in turn can lead to cell epithelial hyperplasia around FCSEMS [43, 44].

In the case of choosing a metal stent, there is still some controversy over whether to choose a coated stent or an uncoated stent. The development of metal-coated stents overcame the problem of luminal stenosis caused by intratumor growth after bare stent implantation. Isayama et al. have confirmed that although covered stents reduce stent occlusion caused by intratumor growth and are relatively easier to replace and remove, they also increase the possibility of stent displacement. In addition, there are cases of re-blockage after implanting covered stents, mainly due to external growth of tumors, bile sludge formation, and stent displacement [45]. The main evaluation indicators of biliary stents include stent patency and patient survival. The results of Almadi et al. showed that there were no differences in stent patency and survival at 6 and 12 months after implantation between patients with the two stents [46]. It was suggested that coated stents had lower occlusion rates than bare stents, and there were no significant differences in survival and adverse events [47]. Therefore, without considering economic factors, Isayama et al. believe that choosing a coated stent is more likely to be beneficial [45].

Drug-eluting stents

At present, the drug-eluting stent is still in its infancy. Chun et al. showed that for malignant biliary obstruction, drug-eluting stents coated with antitumor agents could inhibit tumor ingrowth, thereby improving stent patency [48]. Paclitaxel has shown apoptotic, antiproliferative, and angiogenic activities in a variety of tumor-related chemotherapy. In basic studies, paclitaxel demonstrated dose-dependent inhibition against the proliferation of gallbladder epithelial cells, fibroblasts, and pancreatic cancer cells. In a canine ureteral model, Shin et al. concluded that paclitaxel - eluting stents exhibited inhibitory effects on proliferative tissue [49]. Shi et al. of Xi'an Jiaotong University developed paclitaxel-coated PLA stents and applied it in pig biliary anastomosis model. Their results showed that compared with stent-free and simple BDBS, paclitaxel-coated BDBS significantly reduced biliary anastomotic tissue hyperplasia, showing that paclitaxel-coated BDBS

Figure 4. Unity-B magnesium alloy degradable biliary stent. This figure is cited from [55].

can reduce anastomosis granulation formation and excessive proliferation of extracellular matrix, thereby preventing biliary-intestinal anastomosis stenosis [50]. Lee et al. also showed that after implanting drug-eluting stents in the bile ducts of pigs, only local inflammatory reactions were manifested, and there were no complications such as perforation and necrosis [51]. However, the study using paclitaxel-eluting CSEMS by Song et al. did not show clear advantages in stent patency and patient survival [52]. In the study by Jang et al., they mainly analyzed the effects of drug concentration, stent membrane type, and shape on the effect of stent use [53, 54]. They concluded that the novel paclitaxel-eluting scaffold containing 10% Pluronic F-127 was safe and provided enhanced topical drug delivery.

In drug loading methods using polymer materials as drug carriers, the adverse consequences of limited drug loading, uneven distribution, and increased wall thickness of the stent due to the use of polymer materials have been demonstrated. The maintenance and efficacy of the drug-eluting stents in long term are still not clear, and there are few corresponding studies in this field. By summarizing the relevant studies of drug-eluting vascular stents, it is shown that directly coating the drug on the stent cannot produce obvious anti-restenosis effect due to the "storm effect" released in the early stage. The first problem is that the polymer drug-loaded coating is unfirmly combined with the stent, and the poor adhesion can easily lead to coating falling off, and the drug coating may be damaged when the stent is stretched. The second is whether the biocompatibility of the scaffold can be maintained for a long time. Third, for biodegradable polymer drug-loaded scaffolds, the degradation rate of scaffolds in the generation of new cells is also a problem that needs to be explored. So, a great deal of work needs to be done in the research and application of drug-eluting scaffolds, but there is no doubt that combining drug elution characteristics with absorbability to have coordination advantages will be a major development direction of biliary stents in the future.

Biodegradable metal stents

Mg, Fe, Zn and their alloys are the most favorable metals in the research of degradable medical implants, including stents. The Unity-B magnesium alloy degradable biliary stent has obtained Conformité Européenne certification in June 2021 (**Figure 4**) [55]. According to different degradation rates, it is divided into three types: (1) rapid degradation stent (1-3 months), (2) moderate degradation stent (3-6 months), and (3) slow degradation stent (more than 6 months). The stent diameter ranges from 5-10 mm (1 mm apart) and the length ranges from 17-77 mm. The radial retraction of the stent is 2.8-4.1%. This type of stent can be implanted via three routes, namely ERCP, percutaneous intervention, and hepatic, biliary and pancreatic surgical placement. The degradable metal stent avoids secondary surgical removal, greatly reducing the incidence of complications and alleviating the burden on patients.

Conclusion

Since its first clinical application in 1979, biliary stenting has made great progress. Biliary stents have also evolved from the early days of using only PSs without any customization, to patient-specific self-expanding metal and absorbable stents. The cumulative patency of metal stents generally exceeds that of PSs. Although new PSs continue to be introduced, there have been no reports found that the cumulative patency of PSs exceeds that of metal stents, and there is no significant difference in cumulative survival between the two. In terms of the incidence of complications and adverse events, there is no significant difference in the early adverse events between the plastic and metal stents, while PSs are associated with higher late adverse event and total adverse events, and the main differences were reflected in sepsis, cholangitis, obstruction caused by bile sludge and reintervention caused by various causes. The advantages of PSs are low cost, simple operation, and easy removal and replacement in the event of blockage. However, its disadvantages are also obvious, including high incidence of re-obstruction in about 3

months and easy to cause cholangitis. For the above reasons, PSs are more suitable for benign bile duct stenosis, sclerosing cholangitis, and biliary fistula, which can be used after endoscopic dilation, and their cost is much lower than that of metal stents. The advantages of metal stents are long stent patency time and low incidence of re-obstruction, but they are expensive and difficult to remove after implantation. The leading causes of restenosis after bare metal stent implantation include intratumoral growth, cholestasis, and granulation tissue hyperplasia. Management of restenosis after stent implantation include (1) re-implantation of the stent, (2) changing the structure of the scaffold, such as inserting metal-coated stents, drug-eluting stents, double-layer stents, and biodegradable coated stents (still under study), (3) combination therapy, such as combined radiotherapy, combined radiofrequency ablation, combined photodynamic therapy, and combined percutaneous perfusion chemotherapy. Although USEMS metaloid stent implantation has obvious advantages over PSs in terms of surgical safety, long-term patency rate, and total medical cost during survival, its disadvantage is that it is difficult to remove after implantation, and permanent metal stent implantation should be contraindicated for those who cannot identify benign and malignant lesions and who intend to undergo surgery. Where economic permits, fully coated metal stents are also indicated for patients with benign biliary stenosis or whose nature is pending. Moreover, non-degradable PSs and metal stents have a common disadvantage, that is, secondary occlusion when the stent is placed for a long time, so the degradability of biliary stents is a major research direction at present.

Prospect

With the continuous update of biliary stent design and materials, various types of SEMS, such as anti-migration, anti-reflux, drug elution, radioactivity, 3D printed scaffolds, and bioresorbable scaffolds, have been introduced. Although many emerging stents are still in the research stage, it is certain that the performance of stents will be further optimized, and more high-performance and targeted stents will be developed to benefit patients. It is foreseeable that more research on drug-eluting stents and drug-eluting degradable stents will be conducted in the future to play their synergistic advantages, reduce the risk of secondary stenosis, and reduce the cost of coated stents. Besides, we also look forward to the development of 3D and 4D printing micro-size stent technology.

References

[1] Choudhury S, Asthana S, Homer-Vanniasinkam S, et al. Emerging trends in biliary stents: a materials and manufacturing perspective. Biomater Sci 2022;10(14):3716-3729.

[2] ERCP Group, Chinese Society of Digestive Endoscopology; Biliopancreatic Group, Chinese Association of Gastroenterologist and Hepatologist; National Clinical Research Center for Digestive Diseases. Guidelines for ERCP in China (2018 edition). Chin J Dige Endoscopy. 2018;35(11):777-813.

[3] Davids PH, Tanka AK, Rauws EA, et al. Benign biliary strictures. Surgery or endoscopy? Ann Surg 1993;217(3):237-243.

[4] Soehendra N, Reynders-Frederix V. Palliative biliary duct drainage. A new method for endoscopic introduction of a new drain. Dtsch Med Wochenschr 1979;104(6):206-207.

[5] Cabral JGG, Taglieri E, Pelosof A, et al. Comparative Study between Plastic and Metallic Stents for Biliary Decompression in Patients with Distal Biliary Obstruction. Gastroenterol Res Pract 2017;2017:7621821.

[6] Isayama H, Nakai Y, Kogure H, et al. Biliary self-expandable metallic stent for unresectable malignant distal biliary obstruction: which is better: covered or uncovered? Dig Endosc 2013;25 Suppl 2:71-74.

[7] Boston Scientific. Advanix™ Biliary Stent with NaviFlex™ RX Delivery System. Available at: https://www.bostonscientific.com/en-US/products/stents--gastrointestinal/advanix-biliary-stent-with-naviflex-rx-delivery-system.html

[8] Naitoh I, Inoue T. Optimal endoscopic drainage strategy for unresectable malignant hilar biliary obstruction. Clin Endosc 2023;56(2):135-142.

[9] Sawas T, Al Halabi S, Parsi MA, et al. Self-expandable metal stents versus PSs for malignant biliary obstruction: a meta-analysis. Gastrointest Endosc 2015;82(2):256-67.e7.

[10] Almadi MA, Barkun A, Martel M. Plastic vs. Self-Expandable Metal Stents for Palliation in Malignant Biliary Obstruction: A Series of Meta-Analyses. Am J Gastroenterol 2017;112(2):260-273.

[11] Huibregtse K, Tytgat G. Palliative treatment of obstructive jaundice by transpapillary introduction of large bore bile duct endoprosthesis. Gut 1982;23(5):371-375.

[12] Speer AG, Cotton PB, MacRae KD. Endoscopic management of malignant biliary obstruction: stents of 10 French gauge are preferable to stents of 8 French gauge. Gastrointest Endosc 1988;34(5):412-417.

[13] Vaishnavi C, Samanta J, Kochhar R. Characterization of biofilms in biliary stents and potential factors involved in occlusion. World J Gastroenterol 2018;24(1):112-123.
[14] Johanson JF, Schmalz MJ, Geenen JE. Incidence and risk factors for biliary and pancreatic stent migration. Gastrointest Endosc 1992;38(3):341-346.
[15] Cheon YK, Oh HC, Cho YD, et al. New 10F soft and pliable polyurethane stents decrease the migration rate compared with conventional 10F polyethylene stents in hilar biliary obstruction: results of a pilot study. Gastrointest Endosc 2012;75(4):790-797.
[16] Ginsberg G, Cope C, Shah J, et al. In vivo evaluation of a new bioabsorbable self-expanding biliary stent. Gastrointest Endosc 2003;58(5):777-784.
[17] Laukkarinen J, Sand J, Leppiniemi J, et al. A novel technique for hepaticojejunostomy for nondilated bile ducts: a purse-string anastomosis with an intra-anastomotic biodegradable biliary stent. Am J Surg 2010;200(1):124-130.
[18] Petrtýl J, Brůha R, Horák L, et al. Management of benign intrahepatic bile duct strictures: initial experience with polydioxanone biodegradable stents. Endoscopy 2010;42 Suppl 2:E89-90.
[19] Mauri G, Michelozzi C, Melchiorre F, et al. Benign biliary strictures refractory to standard bilioplasty treated using polydoxanone biodegradable biliary stents: retrospective multicentric data analysis on 107 patients. Eur Radiol 2016;26(11):4057-4063.
[20] Siiki A, Rinta-Kiikka I, Sand J, et al. A pilot study of endoscopically inserted biodegradable biliary stents in the treatment of benign biliary strictures and cystic duct leaks. Gastrointest Endosc 2018;87(4):1132-1137.
[21] Medtronic C. ARCHIMEDES™* Biodegradable Biliary and Pancreatic Stent. Available at: https://www.medtronic.com/covidien/en-gb/products/therapeutic-endoscopy/archimedes-biliary-and-pancreatic-stent.html#
[22] Boland EL, Shine R, Kelly N, et al. A Review of Material Degradation Modelling for the Analysis and Design of Bioabsorbable Stents. Ann Biomed Eng 2016;44:341-356.
[23] Ginsberg G, Cope C, Shah J, et al. In vivo evaluation of a new bioabsorbable self-expanding biliary stent. Gastrointest Endosc 2003;58(5):777-784.
[24] Grolich T, Crha M, Novotný L, et al. Self-expandable biodegradable biliary stents in porcine model. J Surg Res 2015;193(2):606-612.
[25] Sabino MA, González S, Marquez L, et al. Study of the hydrolytic degradation of polydioxanone PPDX. Polym Degrad Stab 2000;69(2):209-216.
[26] De Palma GD, Pezzullo A, Rega M, et al. Unilateral placement of metallic stents for malignant hilar obstruction: a prospective study. Gastrointest Endosc 2003;58(1):50-53.
[27] Isayama H, Komatsu Y, Tsujino T, et al. A prospective randomised study of "covered" versus "uncovered" diamond stents for the management of distal malignant biliary obstruction. Gut 2004;53(5):729-734.
[28] Pfau PR, Pleskow DK, Banerjee S, et al. Pancreatic and biliary stents. Gastrointest Endosc 2013;77(3):319-327.
[29] Liang DB, Dai J, Hu BD. Research progress of endoscopic treatment of malignant bile duct stenosis. J Minimally Invasive Med 2021;16(02):157-161.
[30] Wagner HJ, Knyrim K, Vakil N, et al. Plastic endoprostheses versus metal stents in the palliative treatment of malignant hilar biliary obstruction. A prospective and randomized trial. Endoscopy 1993;25(3):213-218.
[31] Chang WH, Kortan P, Haber GB. Outcome in patients with bifurcation tumors who undergo unilateral versus bilateral hepatic duct drainage. Gastrointest Endosc 1998;47(5):354-362.
[32] Kim DU, Kang DH, Kim GH, et al. Bilateral biliary drainage for malignant hilar obstruction using the 'stent-in-stent' method with a Y-stent: efficacy and complications. Eur J Gastroenterol Hepatol 2013;25(1):99-106.
[33] Tex S, Hahn E, Hochberger J. Endoscopic retrograde implantation of Y-configurated metal stents in the hilus-an effective drainage for malignant hilar stenosis. Gastrointest Endosc 2000;51(4):181.
[34] Kim JH, Yoo BM, Choi JH, et al. Y-Shaped Endoscopic Bilateral Stenting with New Wider-Mesh Biliary Stent in Malignant Hilar Obstruction: Long-Term Results. Gastrointest Endosc 2005;61(5):193.
[35] Li ZS, Cao J. Y-shaped metal stent drainage in the treatment of malignant bile duct obstruction in the hilar part of the liver. Chin J Dig Endoscopy 2006;23(1):2.
[36] Mukai T, Yasuda I, Nakashima M, et al. Metallic stents are more efficacious than PSs in unresectable malignant hilar biliary strictures: a randomized controlled trial. J Hepatobiliary Pancreat Sci 2013;20(2):214-222.
[37] Sangchan A, Kongkasame W, Pugkhem A, et al. Efficacy of metal and PSs in unresectable complex hilar cholangiocarcinoma: a randomized controlled trial. Gastrointest Endosc 2012;76(1):93-99.
[38] Srinivasan I, Kahaleh M. Biliary stents in the millennium. Adv Ther 2011;28(11):960-972.

[39] Nam HS, Kang DH. Current Status of Biliary Metal Stents. Clin Endosc 2016;49(2):124-130.

[40] Kwon C, Jeong S. Choice of stents for distal biliary obstruction. Clin Endosc 2013;46(46):295-302.

[41] Park W, Kim KY, Kang JM, et al. Metallic stent mesh coated with silver nanoparticles suppresses stent-induced tissue hyperplasia and biliary sludge in the rabbit extrahepatic bile duct. Pharmaceutics 2020;12(6):563.

[42] Hair CD, Sejpal DV. Future developments in biliary stenting. Clin Exp Gastroenterol 2013;6:91-99.

[43] Kullman E, Frozanpor F, Söderlund C, et al. Covered versus uncovered self-expandable nitinol stents in the palliative treatment of malignant distal biliary obstruction: results from a randomized, multicenter study. Gastrointest Endosc 2010;72(5):915-923.

[44] Telford JJ, Carr-Locke DL, Baron TH, et al. A randomized trial comparing uncovered and partially covered self-expandable metal stents in the palliation of distal malignant biliary obstruction. Gastrointest Endosc 2010;72(5):907-914.

[45] Isayama H, Nakai Y, Toyokawa Y, et al. Measurement of radial and axial forces of biliary self-expandable metallic stents. Gastrointest Endosc 2009;70(1):37-44.

[46] Almadi MA, Barkun AN, Martel M. No benefit of covered vs uncovered self-expandable metal stents in patients with malignant distal biliary obstruction: a meta-analysis. Clin Gastroenterol Hepatol 2013;11(1):27-37.e1.

[47] Moole H, Bechtold ML, Cashman M, et al. Covered versus uncovered self-expandable metal stents for malignant biliary strictures: A meta-analysis and systematic review. Indian J Gastroenterol 2016;35(5):323-330.

[48] Chun HJ, Kim ES, Hyun JJ, et al. Gastrointestinal and biliary stents. J Gastroenterol Hepatol 2010;25(2):234-243.

[49] Shin JH, Song HY, Choi CG, et al. Tissue hyperplasia: influence of a paclitaxel-eluting covered stent--preliminary study in a canine urethral model. Radiology 2005;234(2):438-444.

[50] Shi J, Lv Y, Yu L, et al. Interest of a new biodegradable stent coated with paclitaxel on anastomotic wound healing after biliary reconstruction. Eur J Gastroenterol Hepatol 2013;25(12):1415-1423.

[51] Lee DK, Kim HS, Kim KS, et al. The effect on porcine bile duct of a metallic stent covered with a paclitaxel-incorporated membrane. Gastrointest Endosc 2005;61(2):296-301.

[52] Song TJ, Lee SS, Yun SC, et al. Paclitaxel-eluting covered metal stents versus covered metal stents for distal malignant biliary obstruction: a prospective comparative pilot study. Gastrointest Endosc 2011;73(4):727-733.

[53] Jang SI, Kim JH, Kim M, et al. Porcine feasibility and safety study of a new paclitaxel-eluting biliary stent with a Pluronic-containing membrane. Endoscopy 2012;44(9):825-831.

[54] Chung MJ, Kim H, Kim KS, et al. Safety evaluation of self-expanding metallic biliary stents eluting gemcitabine in a porcine model. Journal of Gastroenterology Hepatolog. 2012;27(2):261-267.

[55] Medical_Device_News M. UNITY-B Percutaneous Balloon Expandable Biodegradable Biliary Stent Receives European CE Mark. Available at: https://infomeddnews.com/unity-b-percutaneous-balloon-expandable-biodegradable-biliary-stent/

Progress in Medical Devices
eISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/264742zadfwq

Research progress on vascular clips for minimally invasive surgery

Jiaqi Jiang[1], Haipo Cui[1], Fang Geng[2], Xudong Guo[1]

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. [2]Medtronic China Research and Development Center, Shanghai 200100, China

Acknowledgment: This work was supported by the Research Project of Medtronic China Research and Development Center (20S31906800), Shanghai Municipal Science and Technology Major Project (2021SHZDZX), Shanghai Industrial Collaborative Innovation Project (2021-cyxt1-kj07), and Cooperation Fund of the Eighth Research Institute of China Aerospace Science and Technology Corporation (SAST2022-094), PR China.

Declaration of conflict of interest: None.

Received March 13, 2023; Accepted August 29, 2023; Published September 30, 2023

Hightlights

- Material and mechanical structures of absorbable vascular clips affect its hemostatic efficiency.
- The tooth shape, tail end, and closure of vascular clips are the main direction of mechanical structure improvement.
- The coating of the vascular clips can regulate the degradation rate and improve hemostatic efficiency.
- Reasonable mechanical properties and degradation rate are the main development direction for absorbable vas cular clips.

Abstract

Minimally invasive surgeries are widely applied due to the advantages of small surgical wound, short postoperative recovery period, and low surgical infection rate. In minimally invasive surgery, vascular clips play a significant role in hemostasis and managing the direction of blood flow, ensuring the success of surgical procedures. The mechanical structure and manufacturing materials of vascular clips have an important influence on its clinical application effect. In this paper, we classify and summarize the mechanical structure and manufacturing materials of the currently available vascular clips worldwide, and then analyze the hemostatic coating materials of vascular clips. Besides, we also summarize the shortcomings of the existing vascular clips and propose a coat-absorbable vascular clip with future research potential.

Keywords: Vascular clips, mechanical structure, absorbable material, mechanical properties, coating

Introduction

Hemostasis is one of the core techniques in surgical operations. Hemostatic technology has developed from simple instrumental hemostasis measures in the past to a complete technical system, including drugs, materials, equipment, and anesthesia. Finding the bleeding site and treating the bleeding vessels are the key steps of hemostasis. Traditional methods, such as ligation, suture, packing, and heat hemostasis, are still in effect today. Besides, vascular clips, hemostatic gauze, staplers, high-frequency electrosurgical units, ultrasonic scalpels, and ligation bundles for vascular closure systems have been developed and widely used in surgery. Given the increasing preference for minimally invasive surgical procedures, laparoscopes have gained widespread utilization in surgeries, and vascular ligation during various laparoscopic surgeries has always been a challenge in clinical operations. Due to the limited space of laparoscopic surgery and the heat sensitivity of some deep small and medium blood vessels, vascular clips are more convenient and precise than traditional sutures of blood vessels under laparoscope. At present, the materials of tissue vascular

Address correspondence to: Haipo Cui, School of Health Science and Engineering, University of Shanghai for Science and Technology, NO.516, Jungong Road, Shanghai 200093, China. Tel: +86-21-55271290. E-mail: h_b_cui@163.com.

clips are mostly metal or polymer [1]. Common metal clips include silver clips, titanium clips, and tantalum clips. Due to the high price of silver and tantalum, titanium clips are commonly used in clinical practice. Especially, U-shaped titanium clips are widely used clinically, which clamp and close blood vessel by metal plastic deformation. Polymer clips mainly include polyoxymethylene non-absorbable clips, etc. The main products are Hem-o-lok vascular clips and Teleflex products of the same type. Due to its non-absorbability, this type of vascular clips is eventually wrapped by tissue and left as a foreign body in patients, which may cause foreign body reactions or complications, such as inflammation, pain, and stones. Besides, metal vascular clips also affect the conduction of postoperative X-ray, CT, MRI and other examinations.

Absorbable vascular clips can be dissolved through a simple hydrolysis reaction, leading to no foreign body residue in the body, which can effectively avoid the above problems, and at the same time, the distal closure technique can avoid cutting the tissue [2]. The current absorbable metal clips mainly include pure magnesium vascular clips invented by Jorgensen et al. and magnesium alloy vascular clips developed by Ikeo et al. [3, 4]. The marketed absorbable polymer vascular clips mainly include Medtronic's Laproclip and Johnson & Johnson's Absolok vascular clips. At present, the standards for tissue clips are only for metal clips and clip appliers, but there is no standard for absorbable tissue clips. With the gradually increased use of absorbable tissue clips, relevant standards in this field need to be established and strengthened.

In this review, we classify and summarize the mechanical structure and materials of vascular clips in clinical use and analyze the research progress of hydrogel materials used as vascular clip coatings. Besides, we also summarize the shortcomings of existing vascular clips and point out the main research directions of vascular clips in the future.

Clinical application of vascular clips

Samir et al. used Endoloop sutures, staplers, and Hem-o-lok clips, respectively, to ligate the base of the appendix in patients after laparoscopic appendectomy [5]. Through statistical comparison of the application time, postoperative and intraoperative complications, and related costs of the three medical devices, it was found that the hospital stay of patients using the three devices was not significantly different, but the operation time in those using stapler was significantly shorter than that in those receiving suture, while the operation time was not different statistically between those using ligation clip and suture. The time for ligature in those receiving Endoloop suture was significantly longer than that in those using the stapler and the ligation clip, while the ligature time in those using the stapler was significantly shorter than that in those using the ligation clip. However, in terms of cost, the Endoloop suture costed €57.7, the stapler costed €230.77, while the ligation clip only costed €4.75. These results indicate that Endoloop suture requires more manipulation and is relatively complex for appendix base ligation after laparoscopic appendectomy, hence more time-consuming. The application of staplers or ligature clips is relatively simple compared to sutures, and have lower requirements for manipulation. However, the ligature clip requires accurate application size and ligature position, which is an important reason why it is more time-consuming than the stapler. Chang et al. believed that compared with endoloop sutures, the staplers were simple to operate but expensive, and Hem-o-lok ligature clips were the optimal choice for closing the appendiceal stump after laparoscopic appendectomy [6].

Thereafter, the effectiveness, safety, and limitations of ligature clips were studied. By comparing the clinical data of Endoloop sutures and Hem-o-lok clips, it was found that there was no statistical significance in the incidence of postoperative complications between the two. However, ligation clips were superior to sutures in terms of operation time and cost. Nevertheless, sutures have to be used when the diameter of the appendix stump is greater than 10mm and without inflammation. Therefore, the ligation clips should be designed with a larger clamping range and a more reasonable hinge to meet the stump closure requirements in laparoscopic appendectomy. Colak et al. and Andrea et al. verified the safety and effectiveness of Hem-o-lok clips by comparing the complications, operation time, and hospital stay between Endoloop and Hem-o-lok closure modes for appendectomy remnant after laparoscopic appendectomy [7, 8]. They found that the operation time and cost of the ligature clip group were shorter and lower than those of the suture group. There were no complications or deaths in both groups, and the hospital stay was not significantly different, which proved that the safety and effectiveness of Hem-o-lok were consistent with sutures. Hem-o-lok was also found to be safe for appendiceal stump closure in stump and gangrenous appendicitis. Soll et al. studied the incidence of

Table 1. Classification of laparoscopic nephrectomy types using Hem-o-lok clips

Procedure	Right side	Left side	Total
Simple nephrectomy	2	8	10
Radicel nephrectomy	23	24	47
Nephroureterectomy	3	4	7
Live donor nephrectomy	11	55	66
Total procedures	39	91	130

postoperative abdominal abscess after Hem-o-lok ligation of appendiceal stump after laparoscopic appendectomy, and they reported that the abdominal abscess was mainly affected by three factors: surgical resection technique, histology, and postoperative antibiotic administration [9]. In addition to the short ligature time and low cost, the reverberation sound generated by the ligature clamping can also enable doctors to grasp the ligation situation in time, which greatly reduces the incidence of postoperative abdominal abscess. Therefore, in the design of a ligature clip, the clasp should be retained as much as possible to ensure that the doctor gets feedback from the closing sound. Sadat-Safavi et al. compared the disadvantages of sutures, staplers, and ligation clips in closing the stump of the appendix when performing laparoscopic appendectomy [10]. They concluded that sutures had the potential to break during tension, with the possibility of knot loosening; the ligature clip couldn't be reopened once closed; and the improper design of the ligature clip teeth could result in slipping and detachment. However, it was noted that the procedure was simple and cost-effective. As for the stapler, it couldn't be applied for the appendix stump with large diameter due to its small anastomosis range. Therefore, only suture and clips were compared. Through comparison of clinical data, it was found that there was no significant difference in postoperative complications and recovery time between the two methods, but the ligation time and cost of clips were much lower than that of sutures, and stump suture was highly dependent on the doctor's technique and proficiency in laparoscopic operations. However, enhancements in the design of ligature clips, specifically the tooth shape, are required to prevent slipping and detachment while preserving the integrity of the appendix stump.

Yoshida et al. studied the application of Mg-Zn-Ca metal clips and traditional titanium clips for ligature after cholecystectomy in dogs [11]. By CT scan and analysis of the stained tissue sections 6 months after surgery, it was found that compared with non-degradable titanium clips, magnesium clips produced smaller metal artifacts and less fibrosis of the surrounding tissues. Tissue adhesion was only observed in the titanium clips group, but there was no difference in inflammatory components and interstitium between the two groups. For magnesium alloy clips, the effect of degradation products on the surrounding tissues should be further studied, and the degradation rate should be properly managed to prevent excessive hydrogen resulting in gas cavities, so as to replace the traditional titanium clips in clinical surgery.

A large number of clinical trials have shown that, the use of vascular clips as a means of hemostasis in gastric and colorectal polypectomy can effectively prevent postoperative bleeding and digestive tract perforation [12-17].

In laparoscopic nephrectomy, Herve et al. and Simforoosh et al. studied the effect of traditional staplers and ligation clips on the closure of renal arteries and veins [18, 19]. In living nephrectomy, vascular clip ligation can preserve longer renal veins and renal arteries for transplanted kidneys, thus ensuring faster and better recovery of kidney function after transplantation. Another benefit of using a vascular clip is that the vascular clip can be passed through a smaller port when comparing with a stapler. **Table 1** presents the classification of laparoscopic nephrectomy types using Hem-o-lok clips [18].

Ai et al. compared the effect of ultrasonic scalpel and traditional metal clips on closing the gallbladder duct during laparoscopic cholecystectomy [20]. They found that although ultrasonic scalpel shortened the surgical time of vascular and gallbladder duct closure and the length of hospital stay, there was no difference in the incidence of postoperative complications between patients undergoing the two procedures, and additional ligation was required if the gallbladder had benign disease or the diameter of the cystic duct was greater than 6mm. For small blood vessels with narrow cavities and high heat sensitivity, ultrasonic scalpel is obviously superior to traditional metal clips.

Material and mechanical structures of vascular

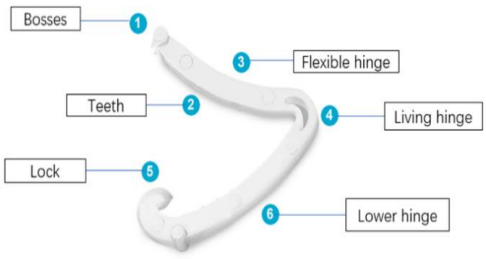


Figure 1. Schematic diagram of Hem-o-lok clip structure. This figure is cited from [23].

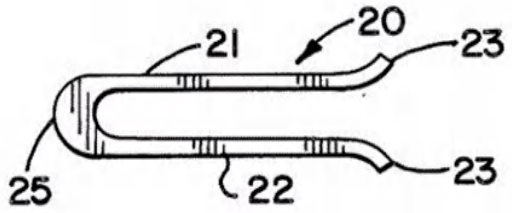


Figure 2. Schematic diagram of the degradable pure magnesium vascular clip.

clips

In terms of structural design, vascular clips require to have clamping stability, clamping durability, and user-friendliness. That is, vascular clips should be stable and firm in clamping the tissue, not easy to slip and loosen, and work well with clip appliers. Under a certain water pressure, the clamped tissue should be free of leakage.

Non-absorbable vascular clips

Titanium clips are the earliest vascular clips that were mass-produced and used on a large scale in modern laparoscopic surgery. They are widely used in the treatment of gastroduodenal hemorrhage and the clamping of lymphatic stumps after lung cancer surgeries. However, due to the inherent properties of the metal material, titanium clips can damage the tissue through an electric current when it touches an electrode during the operation. As a result, it loosens and falls off prematurely due to tissue inflammation after the operation, leading to bile leakage or bleeding. It can also interfere with diagnostic accuracy.

Hem-o-lok vascular clips are disposable and non-absorbable polymer vascular clips made of acetal diethyl alcohol. The overall arc design of the clip makes the ligation range larger and increases the external tension. The clamped blood vessel is also evenly stressed due to the curved design that reduces the risk of rupture. The buckle design at the front end of the clip not only provides a secure closure mechanism when clamping blood vessels, but also gives tactile feedback to doctors. The crescent-shaped air cavity design at the end of the clip can disperse the stress on the tail when the clip is closed, preventing fatigue fracture at the end of the clip. The anti-slip teeth on the upper and lower arms of the clip can effectively prevent blood vessels from slipping [21]. **Figure 1** shows the structure of the Hem-o-lok clip. Since the material of the vascular clip is inert, the clip is non-conductive and will not affect CT, MRI, and X-ray examinations. However, before the clip closes the blood vessel, the vessels can easily slip from the gap between the upper and lower arms of the clip. Since it is not absorbable, it can remain permanently in the body, causing potential harm and bringing psychological burden to the patient [22]. If the material of the Hem-o-lok vascular clips is simply replaced by absorbable polymer without modifying the structure, the tail end with crescent-shaped design will easily degrade before the vascular wound heals due to the thinner inner wall. This in turn sabotages the effect of clamp and leads to bleeding.

Absorbable vascular clips

Richard et al. designed the first absorbable metal hemostatic clip using pure magnesium with high toughness but high brittleness [3]. As shown in **Figure 2**, the hemostatic clamp is designed to enable bending of the clamp by combining the outward bending of part 23 with the clamp structure. This design allows the parts 21 and 22 to remain straight and provides better clamping paths for the blood vessels. This, in turn, facilitates better alignment with the gaps in the unclosed blood vessel clamp [3]. By increasing the thickness of the tail end (25), it can prevent the tail end from breaking and failing when the vascular clamp is closed.

Naoko et al. adjusted the composition of different elements in Mg-Zn-Ca alloy and processed it into bars using hot extrusion technology [4]. The extruded bars were annealed at different temperatures and times, and finally manually bent into U-shaped clamps. The U-shaped clamps underwent dynamic recovery during deformation, resulting in a high compressive fracture strain of 0.40, which is greater than the maximum fracture strain of the vascular clamp. This proves that annealing can reduce strain hardening rate and improve material ductility by weakening the substrate texture. After 12 weeks of degradation, the vascular clamps

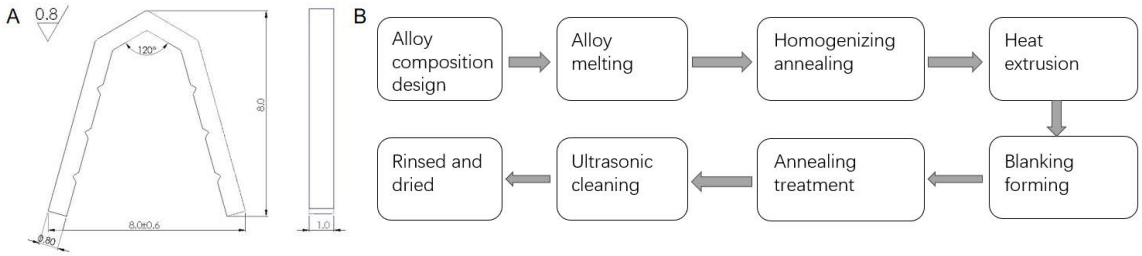


Figure 3. Degradable magnesium-based composite vascular clip. (A) Schematic diagram of the structure; (B) Processing flow chart. This figure was cited from [24].

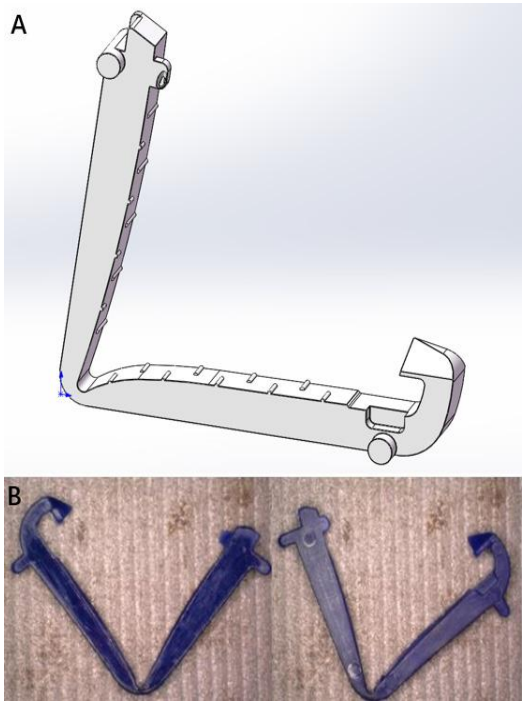


Figure 4. Vascular clip clamping. (A) American Johnson & Johnson ligature clip; (B) Sichuan Guona absorbable vascular clip. This figure was cited from [25].

still maintained a U shape, indicating a uniform degradation. In animal experiments, the air cavity gradually decreased with the passage of implantation time, and there was no obvious air cavity around the surgery. Therefore, adjusting the composition of Zn and Ca in the alloy can help to achieve uniform degradation of the alloy and gradual gas generation.

Bai et al. developed a new type of V-shaped vascular clamp using Mg-Zn-Ca alloy [23]. As shown in **Figure 3A**, the vascular clamp is designed with a 120° arc at the tail end and an angle at the connection between the tail end and the vascular clamp arm, which expands the diameter of the vessel that can be closed and reduces the possibility of tail end fracture caused by the deformation of the vascular clamp. The grooves machined on the inner surface of the clamp prevent the blood vessel from slipping behind the clamp teeth. In terms of materials, the yield strength and tensile strength of the material have been significantly improved by squeezing the plate, and the punching process along the extrusion direction has better adapted the microstructure characteristics of the material to the deformation of vascular clamps. When heated at 380 ℃, the strength of the material decreases, and the elongation reaches maximum then decreases with time. Therefore, the heat treatment time should be controlled to better meet the deformation needs of the vascular clamps. **Figure 3B** is a schematic structural view and a processing flow chart of a biodegradable magnesium-based composite material vascular clip [24].

The absorbable vascular clip of Sichuan Guona Technology Co., Ltd. and the ABSOLOK ligature clip of American Johnson & Johnson are both integrally formed with polydioxanone, and snap locks are set at the front ends of the upper and lower arms of the clip to form a joint. The gripping surfaces thereby close the vessel or tissue. The lock and linear design of the upper and lower arms of the vascular clip can effectively prevent blood vessels from slipping along the upper and lower arms, but the sharp hook design at the front end may cause damage to blood vessels. **Figure 4** shows vascular clip clamping [25].

Covidien absorbable ligation clip adopts the mechanical structure of a double-layer crocodile mouth, which is assembled with four parts: an inner clip, an outer clip, a piston, and a jacket. The outer layer material is polyglycolic acid, and the inner layer material is polyglycol carbonate. After assembly, it forms a double-layer crocodile mouth structure. The jacket has a fixed U-shaped 1cm insertion opening. The mechanism is that under the action of a special clamping tool, the inner piston of the jacket is used to push out the outer layer clips one by one according to the set track, and the duct tissue of the pipe is flattened and blocked. Due to its structural characteristics, the double-layer crocodile mouth absorbable vascular clip

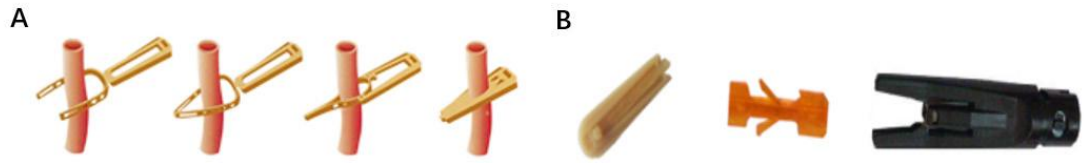


Figure 5. Covidien vascular clip. (A) Process of clamping blood vessels; (B) Picture of the clip. This figure was cited from [25].

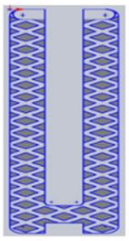
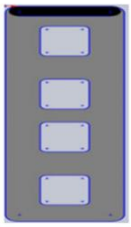
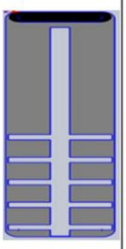
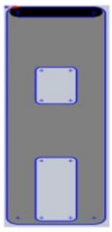
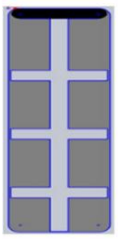
Product	SLS™ clip	Hem-o-lok®	Double-shank	Laparo-clip™	TiGold®	Endo clip™	Ligaclip®
Material	Titanium	Polymer Non-absorbable	Titanium	Polymer absorbable (2-component)	Titanium-gold	Titanium	Titanium
Profile							
Hinge end Jaw End							
Contact surface	Rhomboid	Transverse Raised jaw	Rhomboid w/o latch	Notches	Transverse Linear	Lock-Super Interlock™	Transverse Linear

Figure 6. Tooth profiles of different vascular clips.

can be closed without completely freeing the duct tissue. To use, the operator only needs to buckle the coat directly to the special clamping tool, which enters the body through the laparoscopic trocar, then insert the tissue intended for clamping into the U-shaped opening of the coat, and quickly close it. Therefore, only two steps are required to complete a clamping. The advantage of this structural design is twofold. Firstly, during the clamping process, the inner clip effectively prevents the blood vessel from slipping along the upper and lower arms of the vascular clip. Secondly, the inner clip is constructed from a relatively soft material where it makes direct contact with the blood vessel, minimizing the risk of blood vessel damage. The material of the outer clip is relatively hard, and it is connected with the inner clip to enhance the clamping performance and provide a closure buckle for the overall clip structure. However, because the clip has a double-layer structure, the overall size is relatively large, so it is difficult to clamp the blood vessel in the narrow and closed human body cavity. **Figure 5** shows the design and mechanism of a Covidien vascular clip [25].

Teeth design of vascular clips

Slipping and dislocation of vascular clips in the process of clamping blood vessels is likely to cause serious complications [26-28]. Reasonable tooth shape design can effectively avoid this scenario. Both the French SLS™-Clip and the German Aesculap vascular clip adopt a rhombus design, which not only increases the contact area with the blood vessel, but also enhances the clamping safety to ensure interlocking and non-slip. TiGold® and Ligaclip® have a transverse and a single longitudinal tooth design that prevents vessels from slipping in both directions. Lapro-clip™ ensures non-slippage through the gap in the inner clip for the jaw of the outer clip to pass through [28]. **Figure 6** shows the tooth profile of different vascular clips.

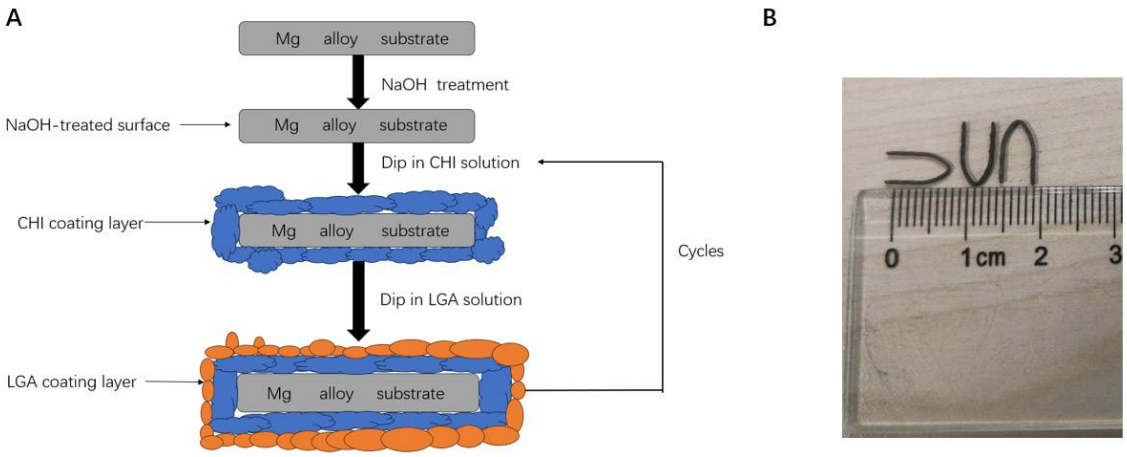


Figure 7. Hemostatic Magnesium Clip. (A) Processing flow of hemostatic coating; (B) Physical image. This figure was cited from [29]. NaOH, sodium hydroxide; CHI, chitosan; LGA, L-glutamic acid.

Coating for vascular clips

Since the hemostasis related to vascular clips is mainly aimed at the internal bleeding of blood vessels in various organs of the human body, many researchers have focused their research on coatings of vascular clips. Chang et al. applied chitosan /L-glutamic acid multilayer coating (CHI/LGA) on the surface of metal clip by adjusting the pH of solution using the principle of layer-by-layer self-assembly technology [29]. CHI/LGA can improve the corrosion resistance of the metal clamp to the expected level through the assembly of 7 layers of coating. This ensures no major mass loss and mechanical property change within three weeks after clinical surgeries, but the magnesium clamp can lose its original shape and function after five weeks of stress corrosion. In future experiments, the corrosion resistance of magnesium clips can also be improved by adjusting the thickness and number of coating layers by manipulating the concentration of CHI and LGA solution and dipping times. **Figure 7** shows the processing flow and samples of the hemostatic coating of the hemostatic magnesium clip [29].

Zheng et al. applied plasma electrolytic oxidation/phytic acid coating on the surface of hemostatic clips made of Mg-Zn-Y-Nd alloy [30]. This coating improves the early corrosion resistance of the clamp by using plasma electrolytic oxidation with a porous surface as the first layer of the composite coating. Phytic acid, used as a chelating agent, forms a conversion film with metal ions such as magnesium and zinc. This has a certain insulation effect on air and water in the middle and later stages of degradation. All these improves corrosion resistance and prevents emphysema caused by gas accumulation and tissue cell damage attributed to the rapid degradation of high concentrations of magnesium ions. Although the biocompatibility and uniform degradation performance of plasma electrolytic oxidation/phytic acid coated vascular clamps are good, in the long-term clinical application process, it is also necessary to observe whether their mechanical properties meet the mechanical requirements after three weeks.

Summary

Judging from the research status in recent years, absorbable metal and polymer vascular clips are developing rapidly. Current research primarily focuses on the improvement of vascular clips from the selection of mechanical structure and absorbable materials, so that vascular clips can be degraded after clamping the bleeding blood vessel, and not remain in the human body. Although the research on vascular clips has made great progress, there is still room for further research:

Mechanical structure: Although currently aviliable vascular clips have improved the general structure and their own tails, as well as increased the overall size of the vascular clips to varying degrees, some of the improved designs did not take the degradation of the vascular clips into account.

Degradable materials: After the degradation in vivo, the degradation products are likely to cause inflammation and adverse reactions in human organs. In addition, the degradation rate of the material needs to be regulated to meet the time and mechanical performance requirements for vascular clamping of the blood vessel.

Hemostasis efficiency: The hemostasis efficiency of vascular clips still needs to be improved. The existing vascular clips may still cause bleeding at the hemostatic site during clamping, which affects the surgical field of view.

In order to make vascular clips better adapt to the needs of laparoscopic clinical surgery, future research and development should focus on reasonable mechanical structure and degradation rate, good biocompatibility of degradation products, and high hemostasis efficiency of vascular clips. At the same time, various hemostatic coatings have been applied on vascular clips in animal experiments [29, 30]. With the continuous improvement of hemostatic efficiency, mechanical structure, and degradation rate of vascular clips, future vascular clips will also have higher clinical application value.

References

[1] Bencini L, Boffi B, Farsi M, et al. Laparoscopic cholecystectomy: retrospective comparative evaluation of titanium versus absorbable clips. J Laparoendosc Adv Surg Tech A 2003;13(2):93-98.

[2] Yano H, Okada K, Kinuta M, et al. Efficacy of absorbable clips compared with metal clips for cystic duct ligation in laparoscopic cholecystectomy. Surg Today 2003;33(1):18-23.

[3] Jorgensen R. Bio absorbable metal hemostatic clip. US patent US19830560909. 1983/12/14.

[4] Ikeo N, Nakamura R, Naka K, et al. Fabrication of a magnesium alloy with excellent ductility for biodegradable clips. Acta Biomater 2016;29:468-476.

[5] Delibegović S. The use of a single Hem-o-lok clip in securing the base of the appendix during laparoscopic appendectomy. J Laparoendosc Adv Surg Tech A 2012;22(1):85-87.

[6] Hue CS, Kim JS, Kim KH, et al. The usefulness and safety of Hem-o-lok clips for the closure of appendicular stump during laparoscopic appendectomy. J Korean Surg Soc 2013;84(1):27-32.

[7] Colak E, Kement M, Ozlem N, et al. A comparison of nonabsorbable polymeric clips and endoloop ligatures for the closure of the appendicular stump in laparoscopic appendectomy: a prospective, randomized study. Surg Laparosc Endosc Percutan Tech 2013;23(3):255-258.

[8] Lucchi A, Berti P, Grassia M, et al. Laparoscopic appendectomy: Hem-o-lok versus Endoloop in stump closure. Updates Surg 2017;69(1):61-65.

[9] Soll C, Wyss P, Gelpke H, et al. Appendiceal stump closure using polymeric clips reduces intra-abdominal abscesses. Langenbecks Arch Surg 2016;401(5):661-666.

[10] Sadat-Safavi SA, Nasiri S, Shojaiefard A, et al. Comparison the effect of stump closure by endoclips versus endoloop on the duration of surgery and complications in patients under laparoscopic appendectomy: A randomized clinical trial. J Res Med Sci 2016;21:87.

[11] Yoshida T, Fukumoto T, Urade T, et al. Development of a new biodegradable operative clip made of a magnesium alloy: Evaluation of its safety and tolerability for canine cholecystectomy. Surgery 2017;161(6):1553-1560.

[12] Zhan X, Wang B, Di D, et al. Endoscopic closure of gastric tube perforations with titanium clips: a four-case report. World J Surg Oncol 2015;13:25.

[13] Mu C, Linghu E. Pretreating Colonic Polyps Under Endoscopy to Prevent Bleeding After Polyp Resection: 2749. Off J A Coll Gastroenterol | ACG 2017;112.

[14] Jaruvongvanich V, Prasitlumkum N, Assavapongpaiboon B, et al. Risk factors for delayed colonic post-polypectomy bleeding: a systematic review and meta-analysis. Int J Colorectal Dis 2017;32(10):1399-1406.

[15] Ayoub F, Westerveld DR, Forde JJ, et al. Effect of prophylactic clip placement following endoscopic mucosal resection of large colorectal lesions on delayed polypectomy bleeding: A meta-analysis. World J Gastroenterol 2019;25(18):2251-2263.

[16] Elliott TR, Tsiamoulos ZP, Thomas-Gibson S, et al. Factors associated with delayed bleeding after resection of large nonpedunculated colorectal polyps. Endoscopy 2018;50(8):790-799.

[17] Okugawa T, Oshima T, Nakai K, et al. Effect of Instruction on Preventing Delayed Bleeding after Colorectal Polypectomy and Endoscopic Mucosal Resection. J Clin Med 2021;10(5).

[18] Baumert H, Ballaro A, Arroyo C, et al. The use of polymer (Hem-o-lok) clips for management of the renal hilum during laparoscopic nephrectomy. Eur Urol 2006;49(5):816-819.

[19] Simforoosh N, Basiri A, Tabibi A, et al. Pedicular vascular control in laparoscopic living donor nephrectomy: The use of clips instead of stapler in 341 donors. Transplant Proc 2006;38(2):390-391.

[20] Ai XM, Ho LC, Yang NY, et al. A comparative study of ultrasonic scalpel (US) versus conventional metal clips for closure of the cystic duct in laparoscopic cholecystectomy (LC): A meta-analysis. Medicine (Baltimore) 2018;97(51):e13735.

[21] Kennedy DL, Donald F. Wilson J, et al. Non-snag polymer ligating clip. US patent US20090088783. 2009/04/02.

[22] Baek M, Chun H, Oh SJ, et al. Open conversion from laparoscopic nephrectomy: slippage of surgical clips ligating the renal artery affected by atherosclerosis. J Urol 2004;171(1):333-334.
[23] Bai H, He X, Ding P, et al. Fabrication, microstructure, and properties of a biodegradable Mg-Zn-Ca clip. J Biomed Mater Res B Appl Biomater 2019;107(5):1741-1749.
[24] Weck® Hem-o-lok® Polymer Locking Ligation System. Available at: https://www.teleflex.com/usa/en/product-areas/surgical/ligation-solutions/weck-polymer-ligation/index.html
[25] Sunstone. Available at: www.hzsskj.com.
[26] Saito Y, Shibata S, Akayama K, et al. Thoracic empyema due to migrated endoclips after laparoscopic cholecystectomy. Asian J Endosc Surg 2012;5(2):89-92.
[27] Matsumoto H, Ikeda E, Mitsunaga S, et al. Choledochal stenosis and lithiasis caused by penetration and migration of surgical metal clips. J Hepatobiliary Pancreat Surg 2000;7(6):603-605.
[28] Mehmood T, Saxena AK. Clip options in endoscopic surgery: contact surface profile and material variations. J Pediatr Endosc Surg 2021;3(3):175-177.
[29] Chang YH, Tseng CC, Chao CY, et al. Mg-Zn-Ca Alloys for Hemostasis Clips for Vessel Ligation: In Vitro and In Vivo Studies of Their Degradation and Response. Materials (Basel) 2020;13(13).
[30] Zheng Q, Sun Z, Wang Z, et al. Corrosion and biocompatibility behaviours of microarc oxidation/phytic acid coated magnesium alloy clips for use in cholecystectomy in a rabbit model. RSC Adv 2021;11(34):20730-20736.

Progress in Medical Devices
eISSN: 2957-5478

Zentime

Research Article https://doi.org/10.61189/594641kmfbkw

A method for identifying pleural lines in B-mode ultrasound images

Tingting Zhou[1], Haozhe Zhuang[1], Shiju Yan[1], Erze Xie[1], Yibo Ma[2], Tao Zhang[1], Tianxiang Yu[1], Shuang Deng[1]

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. [2]Department of Ultrasound, the Third Affiliated Hospital of Soochow University, Changzhou 213000, Jiangsu, China.

Declaration of conflict of interest: None.

Received March 17, 2023; Accepted August 28, 2023; Published September 30, 2023

Highlights

- Automated pleural line identification: A method was introduced to automatically identify pleural lines in lung ultrasound images, ensuring diagnosis speed and accuracy.
- High reliability: An average of 90.45% identification rate of pleural lines was achieved in a comprehensive experiment on 890 ultrasound videos, highlighting its broad applicability and reliability.
- Efficient integration: The algorithm's rapid processing (1.36 seconds for a 5-second video) makes it suitable for seamless integration into ultrasound instrument software, aiding clinicians in diagnosing conditions like pneumothorax more efficiently.

Abstract

In ultrasound imaging, the pleura is visualized as echo reflection formed by the echoes of the interface between the pleura and the lung surface. Three major signs of pleura determines whether the patient has pneumothorax. In this paper, we propose a method to identify pleural line for the diagnosis of pneumothorax. Firstly, the gray threshold of ultrasonic image is properly classified by pre-experiment. Secondly, possible pleural line regions are identified based on threshold classification. Thirdly, the region of pleural line is identified based on the known characteristics of pleural line. The last step is to consider whether it is necessary to modify the threshold to accurately identify the pleural line region. Moreover, we tested 890 ultrasound samples, which included three categories: lung sliding, lung point, and lung sliding disappearance. Each category of samples was divided into two subsets, typical and atypical. The average identification rate reached 90.45%. According to the test results, the advantages and disadvantages of the proposed method as well as the further improvement direction were analyzed. This method for identifying pleural line can serve as the groundwork for developing automatic algorithm for diagnosing pneumothorax.

Keywords: Pleural line identification, lung ultrasound, pneumothorax, automatic identification algorithm, image processing

Introduction

The pleura is a thin, smooth serous membrane adhering to the inner chest wall and the lung surface. In ultrasound images, the pleura, also known as the pleura line, is presented as echo reflection formed by the echoes of the interface between the pleura and the lung surface. Serving as a key anatomical marker, the pleural line plays an important role in chest ultrasound, and it is the basis for the evaluation of several basic signs in chest ultrasound, such as pleural sliding sign, seashore sign, stratospheric sign, A-lines, and B-lines. Disappearance of pleural line and static, rough, thickened or irregular pleural line are all abnormal signs [1, 2]. The pleural line is associated with the diagnosis and interventional treatment of many pulmonary diseases, such as pneumothorax, pulmonary atelectasis, pleural effusion, pleural hemorrhage, pleural tumor, and ultrasound-guided thoracic drainage [3, 4]. Especially in the diagnosis of pneumothorax, the identification of pleural line plays a key role [1-4].

In ultrasound diagnosis of pneumothorax, there

Address correspondence to: Shiju Yan, School of Health Science and Engineering, University of Shanghai for Sci-ence and Technology, No.516 Jungong Road, Shanghai 200093, China. Email: yanshiju@usst.edu.cn; **Yibo Ma**, Department of Ultrasound, the Third Affiliated Hospital of Soochow University, No.185 Juqian Street, Changzhou 213000, Jiangsu, China. Email: mayibo@czfph.com.

are mainly three signs of pleural line: (1) complete presence of lung sliding, namely normal pleural sliding sign, which shows that the pleura moves with breathing; (2) partial presence of lung sliding, which is characterized by partial movement of the pleura, and the junction of pneumothorax with and without lung sliding is lung point; (3) disappearance of lung sliding, namely abnormal pleural sliding sign, which shows that the pleura does not move. The above signs are the main basis for the diagnosis of pneumothorax. The evaluation of pleural sliding sign and lung point sign is based on the accurate identification of pleural line [1, 2, 5-9]. Therefore, identifying the pleural line is of great significance in the ultrasonic diagnosis of pneumothorax.

At present, the identification of pleural lines mainly relies on the clinicians' observation and experience, which leads to certain limitations [10-12]. (1) Inexperienced clinicians may easily mistake the pleural line for myofascial or the interface between muscle and adipose layer. (2) Some pleural lines may exhibit irregular shapes in ultrasonic examinations, leading to a time-consuming identification process that demands experienced clinicians for accurate recognition. (3) Using manual observational methods to identify pleural line lacks objectivity. Automatic pleural line identification can detect the position of pleural line quickly and quantitatively without the use of human labor. The results can be applied not only for pneumothorax diagnosis but also for intelligent diagnosis of other pulmonary diseases.

A number of intelligent algorithms have been developed for automatic pleural line identification. Moshavegh et al. used the random walk-based segmentation method proposed by Grady to compute confidence map. The pleural line is delineated on the confidence map by thresholding the confidence map to be higher than the global threshold of the entire map [13, 14]. Anantrasirichai et al. proposed a linear restoration method of speckle image [15]. Based on Radon transform and sparse regularization, the linear pleural line can be accurately extracted. However, this method still has limitations in identifying irregular pleural lines. Furthermore, Chen et al. proposed an automatic pleural line identification method based on Radon transform [16]; however, this method is limited to the use of ultrasound video for verification, and the effect of identifying pleural lines for videos not within this range is unknown. Carrer et al. proposed an automatic pleural line identification algorithm based on hidden Markov model and Viterbi algorithm [17]. The algorithm can be used to accurately extract the pleural lines in most videos, but it still misestimates the gray threshold of some ultrasound images with low contrast. The above-mentioned traditional algorithms still have a long way to further improve the accuracy. Although the accuracy of neural network and other machine learning algorithms is significantly better than that of traditional algorithms, such as the identification of abnormalities in ultrasound lung images proposed by Veeramani, machine learning algorithms still need a lot of data training to ensure their accuracy [18-20]. However, medical related resources are limited and private, so it is difficult to obtain data. Therefore, it is still necessary to study the traditional algorithms to further improve the pleural line identification rate. In this study, we analyzed and compared the pleural lines in B-mode ultrasound videos of pneumothorax patients and non-pneumothorax patients. Based on the characteristics of linear strong echo in ultrasonic identification of pleural line, an automatic identification algorithm for pleural line in pneumothorax diagnosis is proposed in order to improve the accuracy of pleural line identification, so as to improve the work efficiency of clinicians and the accuracy and efficiency of diagnosis.

Methods

The gray threshold was determined to be 100 by pre-experiments. Based on the area S with pixel gray values greater than 100 in the gray histogram of an ultrasound video, the gray threshold Y of the pleural line area was determined, which was taken as a variable that could be adjusted. Because the pleural line is smooth, uniform and linear with high echo, the correct pleural line region can be extracted based on this characteristic. The adequacy of the threshold value can be evaluated by calculating the extracted pleural line area, thereby determining whether any adjustments to the threshold value are necessary. This process establishes a closed loop of identification system, thereby extracting the correct pleural line region. A flowchart of the proposed pleural line identification method is shown in **Figure 1**.

Threshold classification

In ultrasound videos, pleural lines usually have high brightness (gray value) due to differences in acoustic impedance between the pleura and the lung surface interface. Theoretically, as long as the appropriate gray threshold is set, the pleural line regions in ultrasound video can be separated and extracted. Nevertheless, in practice due to different ultrasonic equipment,

Table 1. Gray threshold classification

Area S	S≤90k	90k<S≤150k	150k<S≤400k	400k<S
Gray threshold *y*	120	150	160	210

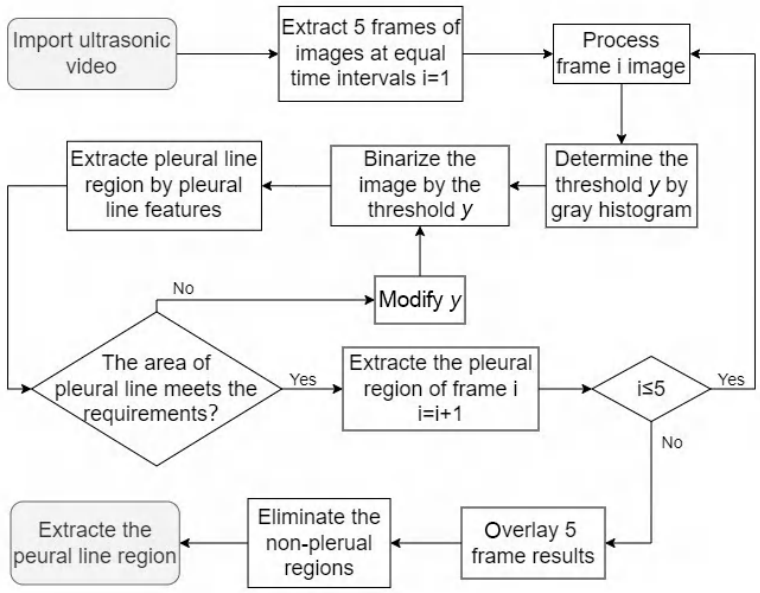


Figure 1. Flowchart of pleural line identification.

diagnostic frequency and other factors, it is impossible to set one gray threshold applicable to all videos, as shown in **Figure 2**. Therefore, before the extraction of pleural line area, it is necessary to carry out a regular threshold classification, so that specific gray thresholds can be set for target videos to extract pleural line. Particularly, the threshold classification is not based on biophysical basis, but on the global gray value of the images.

Results of pre-experiments showed that the area with pixel values larger than the threshold was positively correlated with the selected threshold. The larger the area in which pixel values exceed the threshold, the larger the bright line's area in the image, necessitating the setting of a higher gray threshold. By observing gray histograms of different videos, it was found that the pixel value threshold of 100 was proper for extracting regions incorporating pleural lines, so the threshold of 100 was set as the classification standard. Based on this threshold, the regions with pixel values greater than 100 were obtained, and the pixel value thresholds were classified. The classification results are shown in **Table 1**. Additionally, the threshold was slightly larger so that the area would be slightly smaller, simplifying subsequent steps, where the focus is on determining whether to decrease the threshold instead of increasing it.

Pleural line extraction

In the algorithm of automatic pleural line identification, 5 frames were extracted from each 5-second ultrasonic video at equal intervals, and the 5 frames were superimposed into one image. The purpose of this operation was to avoid the disappearance or blurring of pleural lines in a few frames of 5-second ultrasound videos. Then, regions with pixel values greater than 100 were extracted, and the gray threshold was classified based on areas of these regions, as shown in **Table 1**.

Then each frame of a 5-second ultrasonic video was processed as follows.

Step 1: Process the initial image (**Figure 3A**) as follows. Obtain the binary image. The gray value greater than the initial threshold *y* was set to 1, and the gray value less than the threshold *y* was set to 0. All connected regions were extracted with pixel values greater than the gray threshold *y* (**Figure 3B**). Then, the threshold *y* was adjusted according to all the connected areas and the relationship between the area and the threshold obtained in **Table 1**. As shown in **Figure 3C**, the binary graph was extracted based on the modified threshold *y*.

Step 2: Remove the non-pleural line regions. The noise regions were deleted according to their small areas, and the attribute set of the remaining regions were obtained. The attribute set contained the measured values of several attributes of each 8-connected object in the binary image, including the center of mass and the range of the horizontal and vertical coordinates of the region. Only the potential pleural line area was left (**Figure 3D**).

Step 3: Extract the pleural line based on the characteristics. The pleural line is usually long and almost runs through the whole frame width. Moreover, the pleural line is always located at a moderate height of the image. Ac-

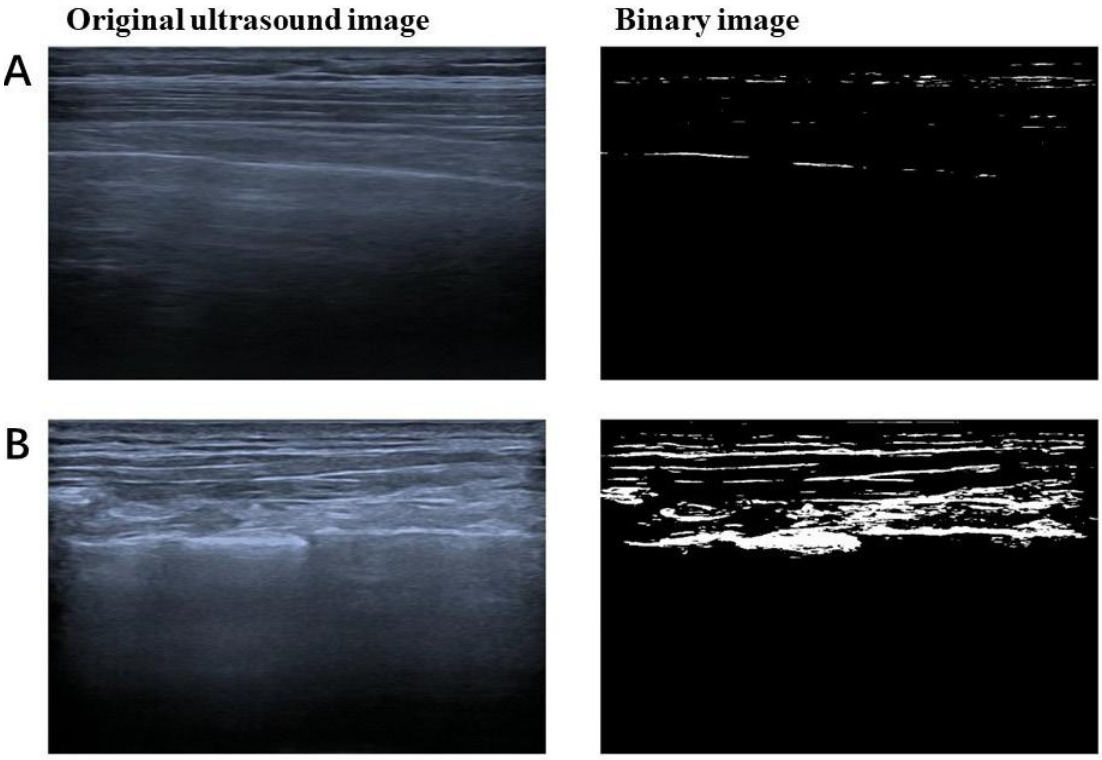


Figure 2. Threshold classification explanatory figure. The gray threshold for both figures were set to be the same, making it impossible to accurately separate the pleural line later. (A) The pleural line is incompletely extracted due to the too high threshold (under segmentation); (B) Other highlighted areas adhering to pleural line area are extracted as pleural line due to the too low threshold (over segmentation).

cording to the characteristics of pleura line, the values of the upper and lower fifths of the binary image were set to 0, which could exclude the regions with very low possibility of pleura line. Referring to the attribute set obtained from step 2, the regions with transverse lengths less than two-fifths of the image width were excluded. Because the muscular fascia above the pleural line usually has a high gray value, the lowest connected region is preliminarily determined as the pleural line region. The area of this region was calculated to determine whether it belonged to pleural line. If the area is too small, the threshold value must be too high. Then, the threshold value was reduced by 20 to repeat the above two steps (**Figure 3E**).

Step 4: Superimpose the results of the 5 frames of binary images on one image. As shown in **Figure 4A**, the darker the color, the more likely this region is to be a pleural line region. The region with lighter color is the movement of pleural line in the direction of non-pleural line caused by factors such as probe instability, which should be eliminated. In the superimposed image, only a superimposed value greater than 2 was left. Finally, the connected component with the largest area was determined as the pleural line region (**Figure 4B**).

Experiments

Materials

Imaging samples used in this study were all from the Department of Ultrasound of the Third People's Hospital Affiliated to Soochow University. SONIMAGE HS1color ultrasound diagnostic instrument of Konica Minolta was used for ultrasonic examinations, with a probe frequency of 4-18 mhz. During ultrasonic examinations, the subjects were in supine or sitting position. The second intercostal longitudinal section of the central clavicular line and the transverse section of the intercostal space were taken as the routine sections, and the ultrasonic probe was placed perpendicular to the patient's skin during the examinations. The machine was adjusted so that the pleura could be clearly displayed. Video acquisition, pleural line identification, annotation, typical or atypical judgment, and pneumothorax diagnosis were all performed by clinicians with more than 5 years of experience in ultrasound diagnosis. These samples cover three categories: lung sliding, lung point, and lung sliding disappearance. Each category of samples was divided into two

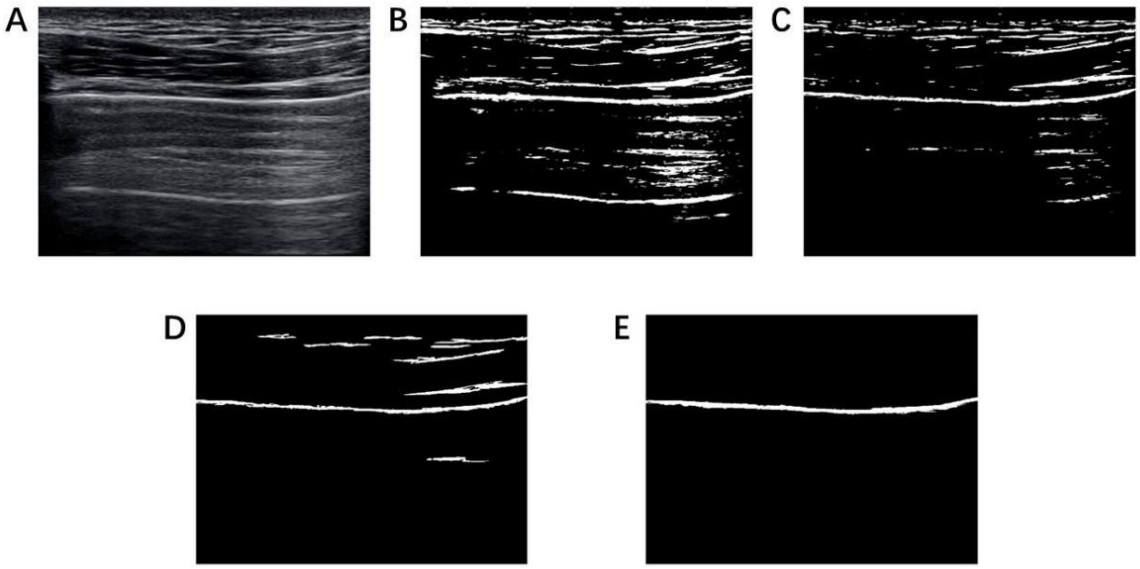


Figure 3. Illustration of pleural line extraction process. (A) Original image; (B) Binary image; (C) Binary image based on *y*; (D) Image of possible pleural line; (E) Image of recognized pleural line.

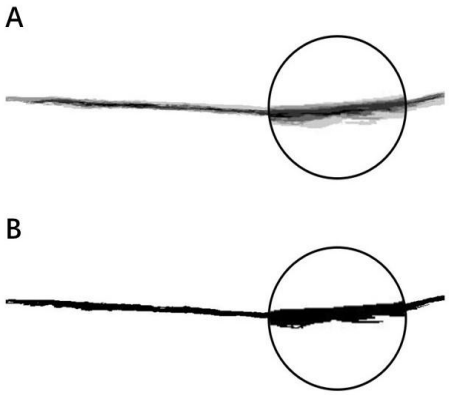


Figure 4. Processed diagram of stacking results (The inside circle is the enlarged image). (A) Superimposing the images; (B) Result image after removing regions generated by motion.

subsets, typical and atypical, according to the significance of signs determined by ultrasound clinicians. A total of 890 5-second B-mode lung ultrasound videos were collected, including 261 lung sliding videos, 331 lung point videos, 298 lung sliding disappearing videos. The 890 videos included 715 typical ones and 175 atypical ones. The study was approved by the Ethics Committee of Changzhou First People's Hospital. All patients provided written informed consents. The study is retrospective.

Experiments

In order to verify the performance of the proposed pleural line identification method, clinicians with over 5 years of experience in ultrasound diagnosis tried to identify the pleural line using this method. The clinicians manually delineated the pleural line area on the superimposed image, as shown in **Figure 5A**. The image delineated by the clinicians was saved as a binary image, and the delineated region was set as the gold standard of the pleural line region, with ±10 pixels as the allowable error range. In **Figure 5B**, the black region is the delineated region, and the gray region is the allowable error range. The image was used as the gold standard for determining the position of the pleural line in the corresponding 5-second ultrasound video. The pleural line region image obtained by the proposed pleural line identification method was compared with the gold standard. Within the allowable range of error, a 95% agreement was found between the proposed method and the gold standard.

Results

A total of 890 ultrasound videos were analyzed by the proposed pleural line identification method in 1,212.62 seconds, and the pleural lines of 805 ultrasound videos were identified correctly. The average identification rate reached 90.45%. The detailed identification rates of pleural lines in various ultrasound videos are shown in **Figure 6**. The identification success rate of each category was calculated by the formula: success recognition rate = number of successful recognition times / total number of samples under this classification * 100%.

Discussion

The results of this study indicate that the proposed identification method exhibits good accu-

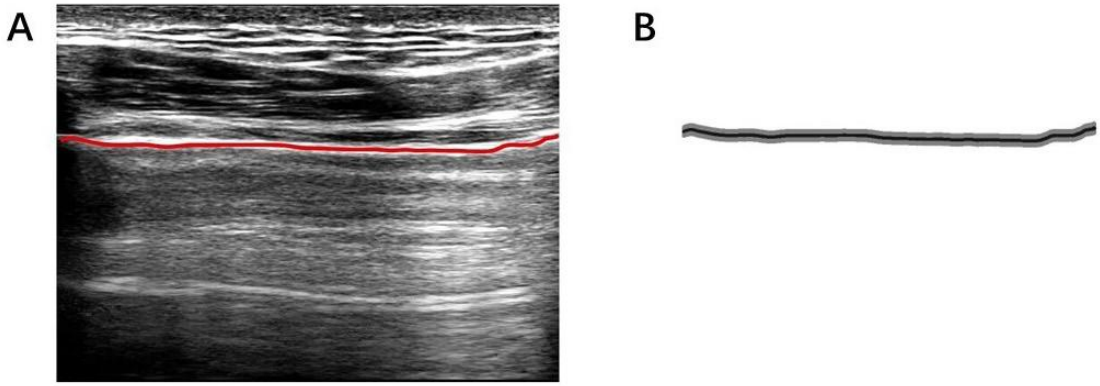


Figure 5. Pleural line obtained by gold standard. (A) The clinician manually outlined the pleural line. (B) The final pleural line as gold standard.

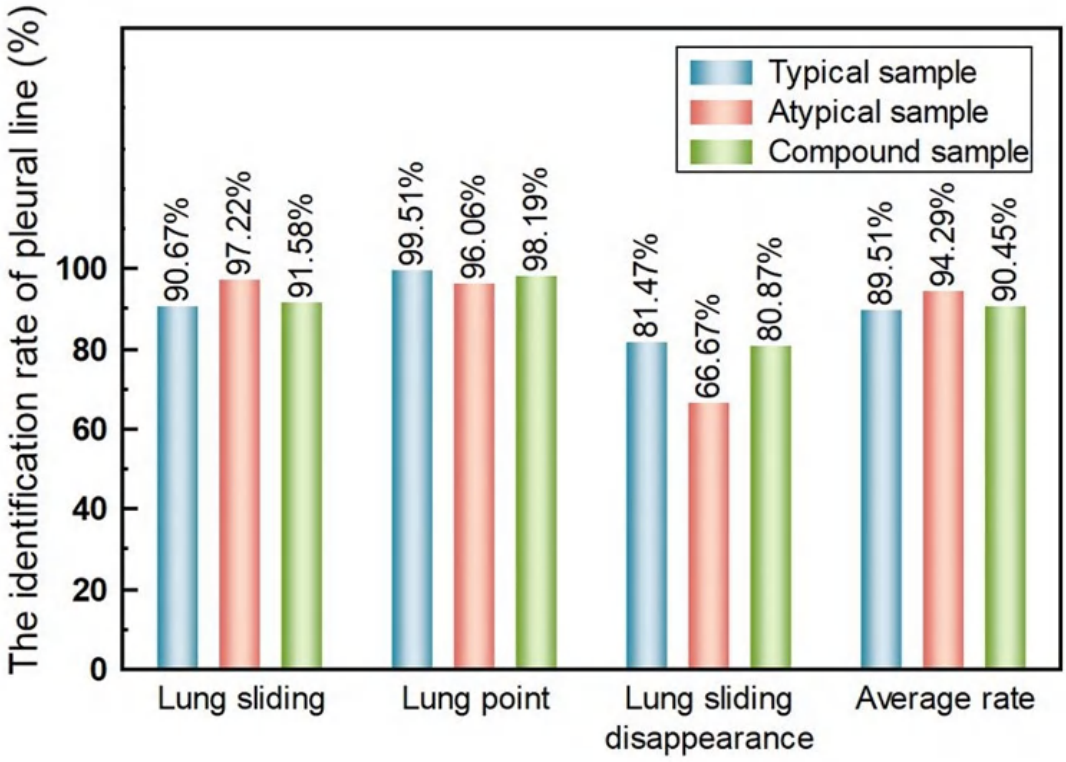


Figure 6. The identification rate of pleural line.

racy, universality and reliability. (1) The pleural line region identified by the proposed method and the pleural line delineated by the clinicians had an agreement of over 95%, indicating the high accuracy of the proposed method. (2) In general, the method has a good identification rate for the videos with lung sliding and lung point, but the identification rate of lung sliding disappearance video was 80.87%, which is lower than the average identification rate of 90.45%. That is, the method can better identify the pleura line in cases of lung sliding and lung point, with a certain universality. (3) When testing the typical and atypical ultrasound videos, there was a difference in the identification rate between typical and atypical ultrasound videos in the category of lung sliding disappearance, which may be caused by the small number of atypical test sets. However, with a larger number in the atypical subsets, there were no significant differences in the identification rate of pleura line between typical and atypical ultrasound videos in the categories of lung sliding and lung point. Generally, the algorithm has a certain reliability.

The identification rate of the pleural lines in the ultrasound videos with the disappearance of lung sliding is slightly lower. This may be attributed to the absence of lung sliding, because long and bright linear strong echo lines are often associated with the pleural line. This echo line looks very similar to the pleura line, so the method is prone to misjudgment (**Figure 7A**).

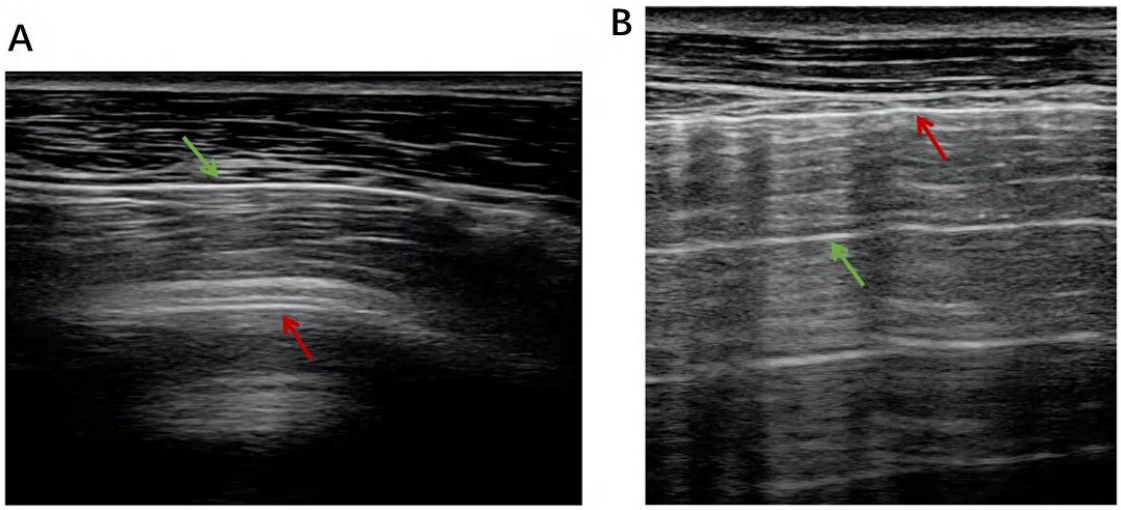


Figure 7. Wrong identification of pleural line. The green arrows point to the correct pleural line area, and the red arrows point to the wrongly identified pleural line area. (A) Error condition 1; (B) Error condition 2.

However, there were only 12 ultrasound videos of atypical lung sliding disappearance, so the results with low recognition rate may not be representative. Besides, the identification rate of pleural lines in ultrasound videos of lung sliding was slightly low than that of lung point. This may be due to the fact that some flow point on the pleural line diluted the highlighted state of the pleural line during lung sliding. It may also be related to equipment, frequency and other factors, resulting in unclear pleural lines in some videos, making the identification more difficult. Pleura line is then mistakenly identified as muscle fiber, myofascial membrane and other tissues with high gray values in ultrasound, as shown in **Figure 7B**. In general, the identification rate of typical videos was even slightly lower than that of atypical videos, which means whether the sign is typical has little influence on the identification performance of the proposed pleural line identification method.

In a clinical context, clinicians require extensive experience and advanced skills to accurately identify pleural lines. The proposed method can be used to identify pleural lines accurately and rapidly, thus improving the diagnostic efficiency and reducing eye fatigue. The method proposed in this paper can assist clinicians lack of clinical experience to identify pleural lines, and at the same time accumulate experience and improve confidence. However, the identification rate of the method is not very high in cases of lung sliding disappearance, so a secondary manual identification may be needed for this type of video in clinical use.

The method described in this paper provides a basis for the research on automatic diagnosis of pneumothorax and other pleural diseases. Accurate identification of pleural lines can enhance the diagnosis efficiency of pleural diseases such as pneumothorax, pleural effusion, pleural bleeding, and pleural tumor.

However, there are still some limitations in this method. There are only a few atypical samples in case of lung sliding and disappearance of lung sliding, so more clinical data are needed to verify the performance of the proposed method. The identification rate of pleura line in case of lung sliding disappearance is low, so the algorithm needs to be further improved. For example, by learning more about the characteristics of the pleura lines, more precise rules can be set for the extraction of pleura line.

In our future research, we will combine more clinical applications and data to comprehensively verify the universality of the proposed algorithm and improve its identification rate. Meanwhile, the algorithm will be integrated into the software system of the ultrasound mainframe to form an automatic recognition module for pleural lines. This will enable the ultrasound device to automatically or manually activate this module, to facilitating automatic recognition of pleural line.

Conclusion

In pulmonary ultrasound, the identification of pleural line is the most basic and important. This paper presents an automatic pleural line identification method for the diagnosis of pneumothorax. The algorithm realizes fast and accurate automatic identification of pleural line regions in ultrasound videos by proper gray threshold classification based on the high brightness of pleural lines in ultrasound

videos. In the verification experiment using 890 ultrasound videos of three categories, the results showed that the average identification rate of pleural lines was 90.45%, which verifies the universality and reliability of the proposed method. The algorithm took an average of 1.36 seconds to process a 5-second ultrasound video, proving its fast speed. As a means to assist clinicians to detect pleural line, it can be integrated into the software system of ultrasound instruments to reduce the pressure of clinicians and improve the diagnosis efficiency of diseases such as pneumothorax.

References

[1] Gargani L. Ultrasound of the Lungs: More than a Room with a View. Heart Fail Clin 2019;15(2):297-303.

[2] Wangüemert Pérez AL. Clinical applications of pulmonary ultrasound. Med Clin (Barc) 2020;154(7):260-68.

[3] Wernecke K. Ultrasound study of the pleura. Eur Radiol 2000;10(10):1515-23.

[4] McLoud TC, Flower CD. Imaging the pleura: sonography, CT, and MR imaging. Am J Roentgenol 1991;156(6):1145-53.

[5] Salah AI, El-Serwi HBe-D, Ansary AMA, et al. Pleural ultrasonography versus chest radiography for the diagnosis of pneumothorax. QJM-INT J MED 2021;114(Supplement_1).

[6] Ramos HC, Núñez DM, Botana RM, et al. Validity of lung ultrasound to rule out iatrogenic pneumothorax performed by pulmonologists without experience in this procedure. Rev Clin Esp 2021;221(5):258-63.

[7] Soldati G, Demi M, Smargiassi A, et al. The role of ultrasound lung artifacts in the diagnosis of respiratory diseases. Expert Rev Resp Med 2019;13(2):163-72.

[8] Luna G, Marica D, Luigia D, et al. Ultrasound lung comets in systemic sclerosis: a chest sonography hallmark of pulmonary interstitial fibrosis. Rheumatology 2009;48(11):1382-87.

[9] Mearelli F, Casarsa C, Trapani A, et al. Lung ultrasound may support internal medicine physicians in predicting the diagnosis, bacterial etiology and favorable outcome of community-acquired pneumonia. Sci Rep 2021;11(1):17016.

[10] Marchetti G, Arondi S, Baglivo F, et al. New insights in the use of pleural ultrasonography for diagnosis and treatment of pleural disease. Clin Respir J 2018;12(6):1993-2005.

[11] Anderson KL, Fields JM, Panebianco NL, et al. Inter-Rater Reliability of Quantifying Pleural B-Lines Using Multiple Counting Methods. J Ultras Med 2013;32(1):115-20.

[12] Koenig SJ, Narasimhan M, Mayo PH. Thoracic Ultrasonography for the Pulmonary Specialist. CHEST. 2011;140(5):1332-41.

[13] Moshavegh R, Hansen KL, Sorensen HM, et al. Novel automatic detection of pleura and B-lines (comet-tail artifacts) on in vivo lung ultrasound scans. SPIE 2016;9790:1-7.

[14] Grady L. Random walks for image segmentation. IEEE T Pattern Anal 2006;28(11):1768-83.

[15] Anantrasirichai N, Hayes W, Allinovi M, et al. Line detection as an inverse problem: application to lung ultrasound imaging. IEEE T Med Imaging 2017;36(10):2045-56.

[16] Chen J, Li J, He C, et al. Automated pleural line detection based on radon transform using ultrasound. Ultrasonic Imaging 2021;43(1):19-28.

[17] Carrer L, Donini E, Marinelli D, et al. Automatic pleural line extraction and COVID-19 scoring from lung ultrasound data. IEEE T Ultrason Ferr 2020;67(11):2207-17.

[18] Roy S, Menapace W, Oei S, et al. Deep Learning for Classification and Localization of COVID-19 Markers in Point-of-Care Lung Ultrasound. IEEE T Med Imaging 2020;39(8):2676-87.

[19] Wu Y, Yi Z. Automated detection of kidney abnormalities using multi-feature fusion convolutional neural networks. Knowl-Based Sys 2020;200:105873.

[20] Veeramani SK, Muthusamy E. Detection of abnormalities in ultrasound lung image using multi-level RVM classification. J Matern-Fetal Neo M 2016;29(11):1844-52.

Progress in Medical Devices
eISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/494562hpikfi

Dipstick color recognition in dry chemical urinalysis: A mini review

Qianfeng Xu, Rongguo Yan, Weifeng Wu, Yongqian Zhang

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

Declaration of conflict of interest: None.

Received April 3, 2023; Accepted August 25, 2023; Published September 30, 2023

Highlights

This review presents four urinalysis methods, the naked-eye method, the integrating sphere method, the color sensor method and the image sensor method, based on color recognition technology, and discusses the techniques of each method.

Abstract

Urinalysis is an essential diagnostic tool for urinary tract infections, kidney disease, diabetes, and other clinical conditions. Dipsticks, which allow for quick screening of urine specimens, are used in the clinic settings to identify the presence and concentration of labeled substances such as urine pH, urine protein, urine glucose, urine ketone, and urine nitrite. This paper reviews four urine dry chemical analysis methods, which are based on human eyes, integrating sphere, color sensors, and image sensors, respectively. The techniques of each method are also discussed.

Keywords: Urinalysis, dry chemical analysis, color recognition, RGB color sensor, image sensor

Introduction

As a laboratory test, urinalysis is an essential diagnostic tool for kidney disease, urinary tract infections, and diabetes, which are highly prevalent diseases in clinic [1-3]. Urinalysis is the most common medical test performed on urine specimens, and the methods for urine composition analysis include physical analysis, formed element analysis, and dry chemical analysis [4].

Physical analysis of urine is the direct visual observation of urine with eyes and/or nose, such as its quantity, turbidity, color, and odor. Taking urine turbidity as an example, it may be clear, cloudy, turbid, or layered, all of which are observable with human eyes. Microscopy can be employed for the physical analysis of urine, enabling the differentiation of components such as red blood cells, white blood cells, epithelial cells, casts, crystals, sperms, and mucous threads.

Formed element analysis is performed on urine sediments that contain cells (red blood cells, white blood cells, epithelial cells, fat oval bodies), casts (cellular casts with cells inside, and noncellular casts with hyaline, granular, waxy, and fatty shapes), crystals (normal and abnormal), organisms (bacteria, yeast cells), and others such as sperms. Formed elements analysis is usually conducted using a flow cytometer and an impedance sensor.

Urine dry chemical analysis uses dipsticks, also called reagent strips, to identify the presence and concentration of labeled substances based on color recognition [5]. Urine dry chemical analysis is crucial for clinical diagnosis since it keeps track of the patient's state, which assists the physician in formulating treatment plans. It also aids disease screening and enables hospitals to conduct high-volume-sample testing. As the currently most used test in urine composition analysis, dry chemical analysis acquires the relevant test findings by recognizing the color of the test items. Chemical analysis of urine includes urine specific gravity, urine pH, urine leukocytes, urine nitrite, urine protein, urine glucose, urine ketone, urobilinogen, urine biliru-

Address correspondence to: Rongguo Yan, School of Health Science and Engineering, University of Shanghai for Science and Technology, NO.516, Jungong Road, Shanghai 200093, China. Email: yanrongguo@usst.edu.cn.

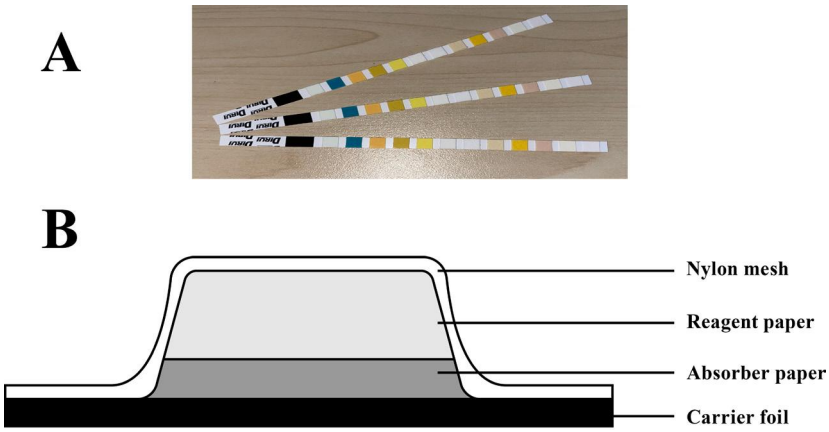


Figure 1. Urine dipsticks and structure of a reagent pad. (A) Urine dipsticks; (B) Structure of a reagent pad.

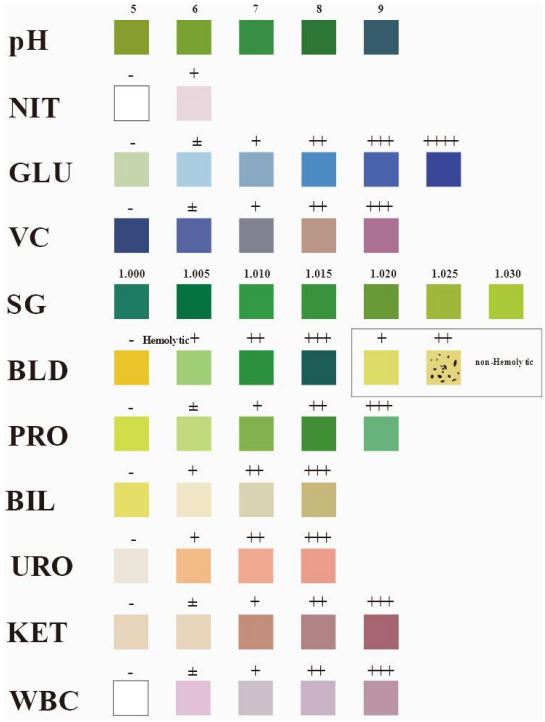


Figure 2. Color chart of an 11-pad reagent dipstick. pH, Urine pH; NIT, urine nitrite; GLU, urine glucose; VC, ascorbic acid; SG, specific gravity; BLD, urine blood; PRO, urine protein; BIL, urine bilirubin; URO, urobilinogen; KET, urine ketone; WBC, white blood cell.

bin, urine blood and so on [4, 6].

This a mini review summarizes multiple newly developed techniques in urine dry chemical analysis, aiming at providing ideas and references to inspire research and technological advancements in this field.

Principles of dry chemical analysis

Dry chemical analysis uses dipsticks, also called reagent strips, as shown in **Figure 1A**, to determine the presence and concentration of the labeled substances. A urine dipstick is composed of multiple reagent pads attached to the plastic carrier foil. Each reagent pad typically contains a four-layer structure, a nylon mesh, a piece of reagent paper, a piece of absorbent paper, and a plastic carrier foil, as shown in **Figure 1B**. After dipping the dipstick into urine, the nylon mesh on the surface prevents biomolecules from entering the reagent paper and protects it from contamination, while the absorbent paper beneath the reagent paper absorbs excess urine, and blocks incident light transmission. The chemical reaction occurs in the reagent paper, resulting in a color change on the pad's surface corresponding to the substance that measured.

Figure 2 shows the color chart of an 11-pad dipstick, as well as qualitative and semi-quantitative descriptions of each item, with the lightest color on the left side and the darkest color on the right side. Generally, to each item, the deeper the color of an item, the higher the concentration, whereas the lighter the color, the lower the concentration.

Thus, the qualitative and semi-quantitative descriptions of each item can be determined based on color recognition. The following are methods that are used to recognize the color of each item.

Color recognition methods

Urinalysis based on viewing with naked eyes

Some color differences can be distinguished by viewing with naked eyes. When it comes to the recognition of color, human eyes can distinguish around 100 shades of grey values and approximately 7 million different colors. However, eyes - the human vision system - are often not as accurate as digital machine vision. Moreover, as for urinalysis based on viewing with eyes, the results may vary from person to person, causing recognition errors. Besides, checking over ten parameters in each test makes it extra challenging. **Figure 3A** shows that the reagent pads are being compared to the color chart sticked on the canister.

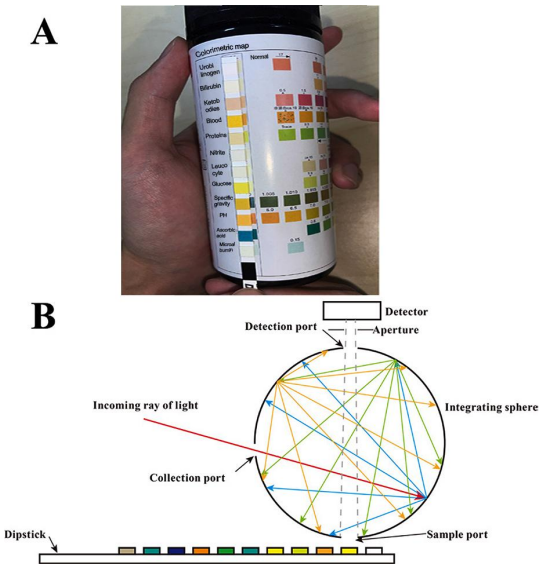


Figure 3. Conventional urinalysis with color detection methods. (A) Urinalysis based on naked eyes; (B) Urinalysis based on the integrating sphere.

Urinalysis based on an integrating sphere

An integrating sphere is a highly reflective device that is placed above the pad of the dipstick so that the incoming ray of light enters the sphere, bounces around the highly reflective diffuse surface of the sphere wall, and finally impinges upon the detector, which is also a part of the integrating sphere assembly. The name, integrating sphere, derives from one of the primary functions of the device, namely that it spatially integrates the light flux, in our case, the light reflected from the pad of the dipstick.

As is shown in **Figure 3B**, when there is no pad placed under the sample port, the amount of light obtained by the detector is relatively large. However, when there is a reagent tape placed under the sample port, the amount of light obtained by the detector drops, which is proportional to the color depth of the measured pad, that is, inversely proportional to the concentration of the substance contained in the pad.

Urinalysis based on color sensors

A color sensor is a device that can measure the RGB color (red, green and blue) values of the object being assessed [7-9]. It has two measurement modes, the first of which involves determining the proportion of red, green, and blue light. Changes in light intensity are the only variables that have an impact on this measurement method, and changes in detecting distance have no impact. The second mode utilizes the intensity of reflected red, green, and blue light to detect target substances, although the sensor's position may have an impact on the outcome [10]. This is generally tested by using a white light source to illuminate the urine reagent pad and using a color sensor to receive the reflected light to obtain the color value of the corresponding substance in the test pad. The exact working principle is shown in **Figure 4**. The color sensor is particularly suitable for tasks that require color recognition. Numerous researchers have developed color sensor-based urine test strips for applications involving color recognition [11-15].

Urinalysis based on image sensors

Image sensors have become a crucial component of cell phones, as much as they are in professional cameras, observatories, and space telescopes. According to the design of the sensor's pixel layout, image sensors can be classified into surface array sensors or line array sensors. The surface array sensor is made up of a two-dimensional sensor element that can capture an entire two-dimensional image at once, but the image quality can be easily influenced by the amount of light present, the angle from which it is being taken, and other variables. It is also susceptible to geometric distortion [16]. Using a power unit to drive the line array image sensor or target object to generate a two-dimensional image by scanning line by line, this method produces images with high resolution and no geometric distortion. Line array sensors are made up of an array of one-dimensional photoreceptors [17].

Figure 5 shows the principle of urinalysis based on image sensors. Linear light sources, generally composed of three LED lights, irradiate the reagent pad, and the reflected light is captured by the charge-coupled device image sensors, which produce an image depicting the color distribution of the tested reagent pad.

There are numerous studies that use image sensors for urine composition analysis. They primarily use smartphone cameras as a medium for color recognition of urine dry chemical reagent bands [18-23]. Again, relying on the benefits of smartphones, some researchers have developed supporting artificial intelligence techniques to replace human judgment, enabling prompt generation of test results [24-26]. However, due to the limitations of smartphones and development platforms, most of these studies still rely on image sensors for urine composition analysis [27, 28].

Discussion

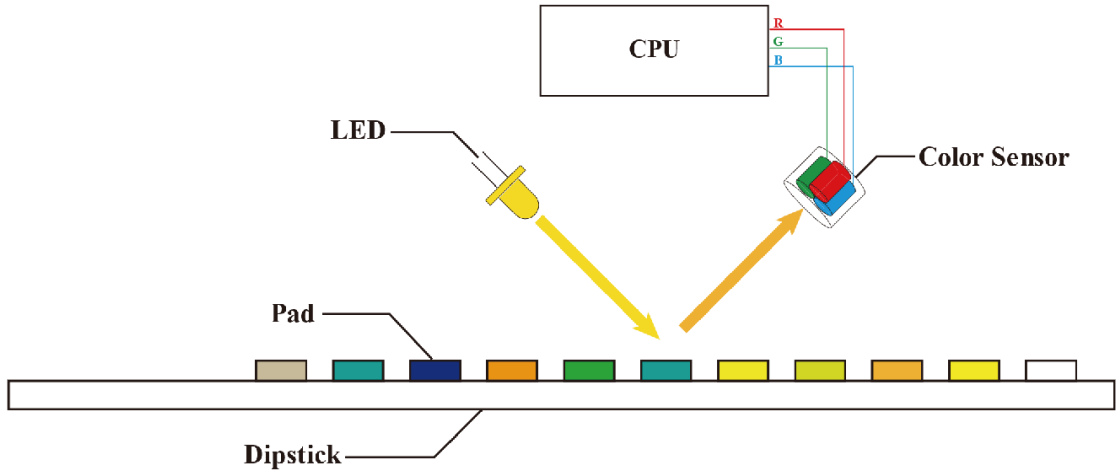


Figure 4. Working principle of color sensor-based dry chemical analysis. The white light emitted from the LED hits the reaction pad and the resulting reflected light is picked up by the color sensor. The color filter in the color sensor identifies the light intensity of the three colors, which is then processed by the CPU to obtain the color of the reaction pad.

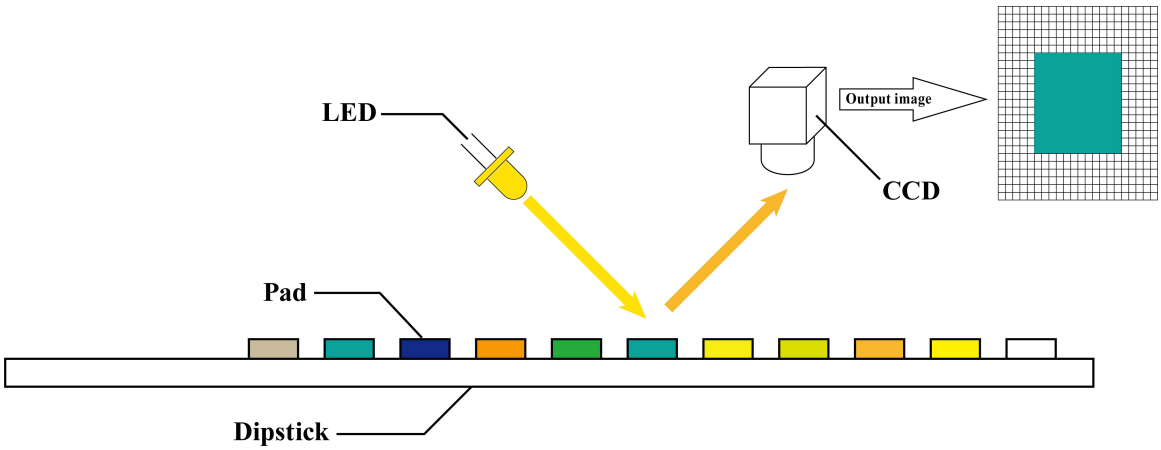


Figure 5. Urinalysis based on image sensors. CCD, charge-coupled device.

Urinalysis is the examination of a urine sample for the diagnosis of urinary tract disorders as well as a variety of metabolic, systemic, and endocrine disorders. This test is usually performed by means of a dipstick loaded with several square pads containing regents that correspond to different objective compositions to be tested.

In section 3, four methods for color recognition of the reagent dipstick after immersion in urine are presented.

Regarding urinalysis based on viewing with eyes, the sensitivity of the human eyes to variations in color brightness and chromaticity can vary. Different individuals may have distinct responses to the same reagent pad due to both subjective and objective factors. Therefore, this method is only used in situations when the use of instruments is inconvenient or only at home. But some people with color weakness or color blindness may produce inaccurate results [29].

Currently, urinalysis based on an integrating sphere is widely used in clinical settings. However, this method requires an ideal integrating sphere, wherein: (1) the inner surface of the integrating sphere is a complete geometric sphere with equal radii everywhere; (2) the inner wall of the sphere is a neutral and uniformly diffusing surface that maintains consistent diffuse reflectance for incident light across various wavelengths; (3) there is no object in the sphere, and the light source should be regarded as an abstract light source that only emits light and has no entity. Therefore, this method has a high requirement for optical equipment.

Urinalysis based on a color sensor can recognize subtle color differences with a high degree of sensitivity, offering a user-friendly and straightforward operation that doesn't necessitate specialized expertise. However, it requires

white balance processing and calibration. In order to further improve its performance, color sensor recognition technology based on optical fiber transmission has been studied recently [14].

In the context of urinalysis utilizing image sensors, it should be noted that both charge-coupled device (CCD) sensors (discussed in this paper) and complementary metal-oxide-semiconductor (CMOS) sensors (not covered in this paper) primarily measure the intensity of light at individual pixels. When adopted, illumination correction is needed, considering different light source environments in image acquisition module of charge-coupled device, as this can lead to deviations in the color image of the tested reagent pad. Since both urine and reagent pads possess colors, color calibration is also essential.

Conclusion

In this paper, we summarize four urine dry chemical analysis methods, including urinalysis based on human eyes, integrating sphere, color sensors, and image sensors. Urinalysis based on an integrating sphere is widely employed in clinical settings. However, with the progress of science and technology, urinalysis based on color sensors and image sensors will become increasingly popular due to their lower costs, higher accuracy, and easier maintenance, especially for home settings.

References

[1] Mahalingasivam V, Booth J, Sheaff M, et al. Nephrotic syndrome in adults. Acute Med 2018;17(1):36-43.

[2] Coresh J, Heerspink HJL, Sang Y, et al. Change in albuminuria and subsequent risk of end-stage kidney disease: an individual participant-level consortium meta-analysis of observational studies. Lancet Diabetes Endocrinol 2019;7(2):115-127.

[3] Pasternak M, Liu P, Quinn R, et al. Association of Albuminuria and Regression of Chronic Kidney Disease in Adults With Newly Diagnosed Moderate to Severe Chronic Kidney Disease. JAMA Netw Open 2022;5(8):e2225821.

[4] Haibo S, Jian'er Y. In Vitro Diagnostic Industry in China. Springer Singapore 2021.

[5] Cyriac J, Holden K, Tullus K. How to use urine dipsticks. Arch Dis Child Educ Pract Ed 2017;102(3):148-154.

[6] Lin F. Design of Comprehensive Urine Detection Equipment. Chongqing Univ 2021.

[7] Wu WY, Gong YC, Mao YM, et al. Study on wavelength-measured by a color sensor. J Optoelectron·Laser 2003;14(3):285-287.

[8] Liu GH, Yu JX, Ren LJ. Progress of RGB Detection Methods. J Anal Sci 2020;36(04):591-596.

[9] Woodstock TK, Karlicek RF. RGB Color Sensors for Occupant Detection: An Alternative to PIR Sensors. IEEE Sensors Journal 2020;20(20):12364-12373.

[10] Wang ZX, Zhong RY, Li YE. Design of Dynamic Color Monitoring and Control System Based on Light Reflection and Transmission and Internet of Things Technology. Chin J Sens Actuators 2021;34(11):1549-1554.

[11] Yong H, Kai D, Yongheng H, et al. Colorimetric recognition for urinalysis dipsticks based on quadratic discriminant analysis. Annu Int Conf IEEE Eng Med Biol Soc 2017;2017:3902-3905.

[12] Xie DD. Design and Realization of A Portable Urine Analyzer Based on Micropython. Chongqing Univ 2018.

[13] Wei RF, Yan RG, Shen JK, et al. Urine component analysis based on RGB color recognition. Electron Meas Technol 2020;43(13):161-166.

[14] Zhang HW, Yang RG, Xu RX, et al. urine composition analysis based on an fiber color sensor. Intell Comput Appl 2021;11(10):82-88.

[15] de Carvalho Oliveira G, Machado CCS, Inácio DK, et al. RGB color sensor for colorimetric determinations: Evaluation and quantitative analysis of colored liquid samples. Talanta 2022;241:123244.

[16] Zhang C. Design,Implementation and Application Research of High Speed Liear CCD Camera. Wuhan Univ 2017.

[17] Liu L. RMB-banknote Discriminating Device Based on Image and Magnetism Technology. South China Univ Technol 2013.

[18] Karlsen H, Dong T. Smartphone-Based Rapid Screening of Urinary Biomarkers. IEEE Trans Biomed Circuits Syst 2017;11(2):455-463.

[19] Huang X, Xu D, Chen J, et al. Smartphone-based analytical biosensors. Analyst 2018;143(22):5339-5351.

[20] Xu Z, Liu Z, Xiao M, et al. A smartphone-based quantitative point-of-care testing (POCT) system for simultaneous detection of multiple heavy metal ions. Chem Eng J 2020;394:124966.

[21] Baek SH, Choi HR, Kim KS. Urinalysis Screening Application based on Smartphone. J Inst Internet, Broadcast Commun 2021;21(5):95-102.

[22] Flaucher M, Nissen M, Jaeger KM, et al. Smartphone-Based Colorimetric Analysis of Urine Test Strips for At-Home Prenatal Care. IEEE J Transl Eng Health Med 2022;10:2800109.

[23] Geng Z, Miao Y, Zhang G, et al. Colorimetric

biosensor based on smartphone: State-of-art. Sensors and Actuators A: Physical 2023;349:114056.

[24] Guo D, Li G, Miao JQ et al. A smartphone-based calibration-free portable urinalysis device. J Cent South Univ 2021;28(12):3829-3837.

[25] Zhou J, Dong T. Design of a wearable device for real-time screening of urinary tract infection and kidney disease based on smartphone. Analyst 2018;143(12):2812-2818.

[26] Thakur R, Maheshwari P, Kumar Datta S, et al. Smartphone-based, automated detection of urine albumin using deep learning approach. Measurement 2022;194:110948.

[27] Lu X. Research on Color Image Reconstruction for Contact Image Sensor. Wuhan Univ 2017.

[28] Thakur R, Bhatt S, Dubey SK. Effect of smartphone camera settings in colorimetric measurements under controlled illumination. 2021 14th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI) 2021 23-25.

[29] Kibria IE, Ali H, Khan SA. Smartphone-Based Point-of-Care Urinalysis Assessment. Annu Int Conf IEEE Eng Med Biol Soc. 2022 Jul;2022:3374-3377.

Progress in Medical Devices
eISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/145362zgyopx

Research progress and applications of image defogging algorithms

Yi Chen[1,*], Shiju Yan[1,*], Yunhua Xu[1], Linping Gu[2]

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China. [2]Shanghai Lung Cancer Center, Shanghai Chest Hospital, School of Medicine, Shanghai Jiao Tong University, Shanghai 200025, China.
[*]The authors contribute equally.

Declaration of conflict of interest: None.

Received April 25, 2023; Accepted August 28, 2023; Published September 30, 2023

Highlights

- The mainstream defogging algorithms can be classified into three categories based on their principles: image enhancement-based, physical model-based, and neural network-based.
- This paper aims to introduce and explore these categories, as well as to provide an outlook on the application and possible future development directions of defogging algorithms.

Abstract

Images taken under hazy weather conditions suffer from problems such as blurring, low contrast, and low saturation due to the scattering of atmospheric light by aerosol particles in the air, which affects the performance and judgment of image analysis equipment. With the rapid development of image processing technology and computer vision technology, researchers have proposed a large number of targeted haze removal algorithms to improve the quality of images taken under hazy weather conditions. According to the haze removal principle, mainstream haze removal algorithms can be classified into three categories: image enhancement-based, physics model-based, and neural network-based. This paper introduces and explores classic haze removal algorithms from the perspectives of principles, development, advantages, and disadvantages, and outlines the prospects for the future development and application direction of haze removal algorithms.

Keywords: Haze removal algorithm, image enhancement, physics model, neural network.

Introduction

With the rapid development of computer vision technology, image processing techniques such as object detection, classification, and tracking have become research hotspots [1, 2]. Image processing systems require high image quality, as image quality directly affects system operating efficiency and output [3]. Under special weather conditions such as haze, a large number of aerosol particles suspended in the air refract and scatter atmospheric light, making it difficult for imaging devices to obtain clear reflected light signals. The resulting images may exhibit loss of details, low contrast and saturation, color shifts, and other defects, which are not conducive to image information extraction and assessment [4]. Therefore, using image processing techniques to obtain high-quality images is of great research value.

Over the years, researchers have proposed various defogging strategies. Early defogging algorithms aimed to enhance the global contrast of images and highlight their details to achieve the defogging effect, such as histogram equalization, Retinex algorithm, and wavelet transform [5-10]. With the proposal of atmospheric scattering models, researchers have established the mapping relationship between hazy images and haze-free images by analyzing atmospheric light scattering and attenuation, and reversed this relationship to obtain dehazed images [11-14]. There are also defogging algorithms based on neural networks. With the rapid development of deep learning technology, two strategies of neural network-based defogging algorithms have emerged: one is the end-to-end model, which directly finds the mapping relationship between hazy and haze-free images

Address correspondence to: Shiju Yan, School of Health Science and Engineering, University of Shanghai for Science and Technology, No.516 Jungong Road, Shanghai 200093, China. Tel: 18956153985, E-mail: yanshiju@usst.edu.cn.

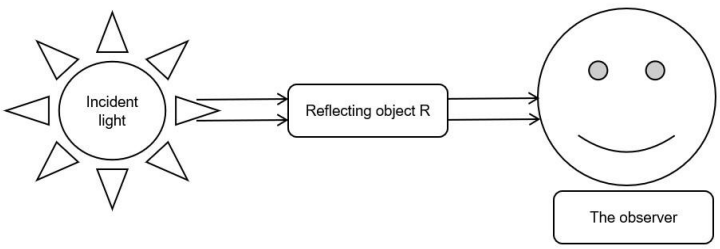


Figure 1. Theoretical model of Retinex.

through the powerful analytical capabilities of neural networks and outputs haze-free images; the other is the non-end-to-end model, which restores haze-free images by inferring key parameters of the atmospheric physical model [15-23].

Defogging algorithms have been widely used in target detection, intelligent driving, industrial image dust removal and other fields, but the application of defogging algorithm in the field of medical image is relatively simple. According to the differences in implementation principles, we divide the existing mainstream defogging methods into three categories: image enhancement-based, physical model-based and neural network-based, summarize the principles of the three defogging methods, and outline the advantages and disadvantages of the three defogging methods. In this paper, we expound the application status of defogging algorithm in industrial field and look forward to the research and development direction of defogging algorithm in medical image field.

Review of three kinds of defogging algorithms

Defogging algorithm based on image enhancement

Early haze removal algorithms aimed to enhance the global contrast of the image and highlight its details, such as histogram equalization, Retinex algorithm, and wavelet transform. The Retinex algorithm decomposes the image into an illumination component and a reflectance component, and eliminates the influence of the reflectance component according to the atmospheric scattering model to achieve haze removal. Histogram equalization stretches the image to improve its contrast and detail information, while wavelet transform enhances the image's temporal and frequency resolution, enlarging useful information to achieve haze removal [5-10].

Histogram equalization algorithm

The grayscale histogram is a graphical representation of the distribution of pixel grayscale values in an image, which can display information about the brightness, pixel values, and other characteristics of the image. The main idea of histogram equalization is to transform the distribution of the image histogram into an approximately uniform distribution through a cumulative distribution function, thereby enhancing the image contrast. To expand the brightness range of the original image, a mapping function is required to map the pixel values of the original image to the histogram in the range of (0, 255). The linear transformation process can be expressed as follows:

$$h(v) = round(\frac{s_x - s_{\min}}{N - s_{\min}} \times (L-1)) \quad (1)$$

s_x is the gray level distribution function, and L is the maximum gray level of the image. The classical global histogram equalization algorithm has good contrast enhancement and fog removal effects on images with single depth of field variations. However, it often results in the merging of low-frequency gray levels, excessive enhancement of high-frequency gray levels, and halo artifacts, leading to loss of image details.

Retinex algorithm

The Retinex theory holds that the incident light determines the dynamic range of all pixels in an image, while the inherent invariant reflection coefficient of the object itself determines the inherent properties of the image. That is, the image we perceive is formed by the light reflected off an object according to its reflection coefficient, as shown in **Figure 1** [24].

The core mathematical expression of the theory is:

$$S(x,y) = R(x,y) \times L(x,y) \quad (2)$$

S(x,y) represents the final image result, R(x,y) represents the reflectance function of the object to light, and L(x,y) represents the illumination function. The single-scale Retinex algorithm assumes that the final reflectance image is a smoothed image, and convolves

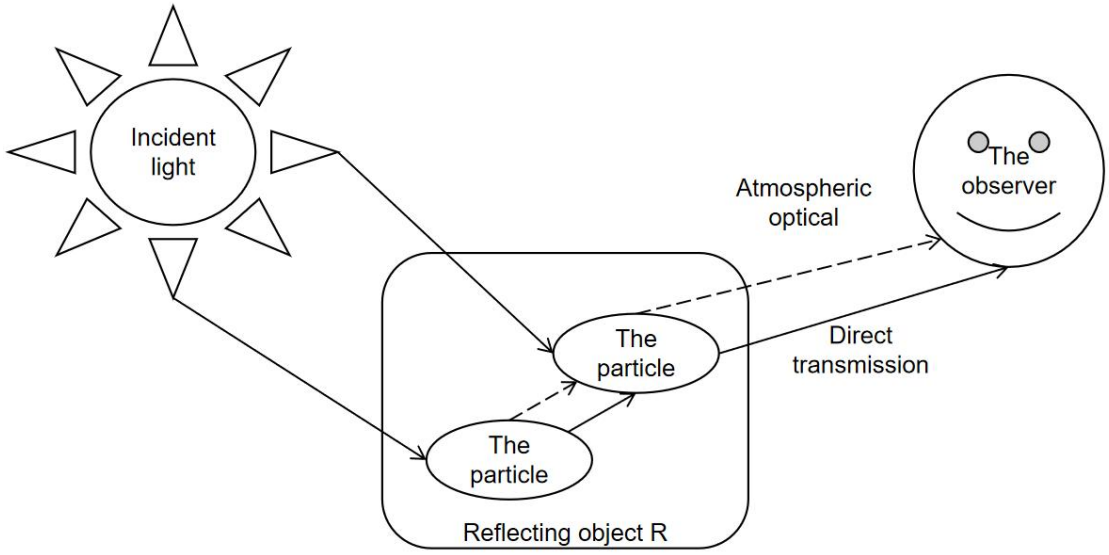


Figure 2. Atmospheric scattering model.

the original image with a Gaussian filter. It estimates the illumination changes in the image by calculating the weighted average of the pixels in the surrounding area, and finally retains the reflection properties of the object in the image, which is equivalent to performing low-pass filtering on the original image to obtain the low-pass filtered image. However, this algorithm has drawbacks such as high complexity, easy occurrence of halos, and easy distortion.

Wavelet transform algorithm

The wavelet transform algorithm decomposes the original image signals into high-frequency and low-frequency images at multiple scales. Since haze mainly affects the low-frequency part of the image, this algorithm enhances the contrast of the image by enhancing the high-frequency coefficients, thus achieving the effect of haze removal. The algorithm processes multi-channel images and merges them. However, there are issues such as possible distortion and high computational complexity.

Defogging algorithm based on physical model

Most image enhancement-based defogging algorithms propose defogging strategies from the perspectives of image brightness, contrast, saturation, etc. The defogging effect of these algorithms is often limited. In comparison, physical model-based defogging algorithms have more stable and efficient characteristics.

Atmospheric scattering model

In 1999, Narasimhan et al. proposed an atmospheric scattering model based on the theory of atmospheric scattering, as shown in **Figure 2** [25, 26]. The model assumes that the light received by the imaging system mainly comes from two parts: one is the light reflected by the target which attenuates and reaches the imaging system, and the other is the atmospheric light formed by the particle scattering of the light source.

The mathematical model for foggy image formation obtained through this physical model is as follows:

$$I(x) = J(x)t(x) + A(1-t(x)) \quad (3)$$

where I(x) is the captured foggy image, J(x) is the corresponding clear image, A is the atmospheric light intensity, and t(x) is the haze transmission rate, which can be expressed as:

$$t(x) = e^{-\beta \bullet d(x)} \quad (4)$$

where β is the atmospheric scattering coefficient, and d(x) represents the depth of field information of the pixel. Formula (3) can be divided into two parts. The first part, J(x)·t(x), represents the direct attenuation term, indicating that the reflected light of the target scene exponentially decays with the increase of depth of field. The second part, A(1-t(x)), represents the atmospheric light that enters the imaging device after being scattered by the haze. By inversely solving the atmospheric light intensity A and the transmission rate t(x) in the model, the resulting image can be defogged. Based on this, researchers have proposed three defogging strategies, which are based on additional information, image differences, and single-image, respectively.

Extra information-based defogging

In some specific situations, the key parameters in the atmospheric scattering model can be obtained by measuring the real-world haze condition, which can be used to recover a haze-free image. For example, the interactive defogging strategy proposed by Narasimhan et al. establishes a data transmission channel between the user and the system, where the user provides information such as atmospheric light intensity and sky position, and the key parameters in the model can be solved to obtain the recovered image [27]. Kopf et al. obtained the key parameters by acquiring a 3D model of the target location, which includes important information such as depth of field, from the Internet, and then used the model to generate a haze-free image [28]. Although extra information-based defogging algorithms produce good results, the acquisition of geographic and lighting information in real-world scenarios is difficult, which limits their practical application.

Image-difference-based defogging

The image-difference-based defogging algorithms estimate the model parameters by studying multiple sets of hazy or haze-free images of the same target scene, thus obtaining the haze-free image. For example, Shwartz et al. proposed a self-calibrated parameter defogging algorithm based on polarization characteristics. By installing a polarizing filter on the imaging device and using different polarization angles to capture three images with different optical characteristics, the atmospheric light intensity can be derived to obtain haze-free images [29]. Although the image-difference-based defogging algorithms perform well, they are cumbersome and expensive, which is not conducive to practical applications. Therefore, researchers have focused on single-image-based defogging algorithms.

Single-image-based defogging

The most widely used defogging algorithm based on atmospheric physics models is the dark channel prior algorithm proposed by He et al. in 2009 [30]. The authors observed the dark channel of more than 5,000 haze-free images and found that about 75% of the pixel values were 0, and 90% of the pixels had very low values [31]. They proposed the dark channel prior theory, which states that for a haze-free image, its dark channel can be represented as:

$$J_{dark} = \min_{y\in\Omega(x)}(\min_{c\in\{r,g,b\}}(J^c(y))) \quad (5)$$

where J represents the haze-free image, J_{dark} represents the dark channel image, C represents any channel of the R, G, and B channels of the image, and Ω(x) represents the window at pixel point x. According to the dark channel prior theory, the transmission rate t(x) can be derived as:

$$J_{dark} \longrightarrow 0 \quad (6)$$

Assuming that the transmission rate t(x) is constant, the atmospheric light intensity A can be obtained by selecting the top 0.1% brightest pixels in the dark channel image of the hazy image and locating the corresponding brightness point with the highest value in the original hazy image, as shown in Eq. (7):

$$\min_{y\in\Omega(x)}(\min_{c}\frac{Ic(y)}{Ac}) = t(x)\min_{y\in\Omega(x)}(\min_{c}\frac{Ic(y)}{Ac}) + 1 - t(x) \quad (7)$$

According to the dark channel prior theory, the transmission rate t(x) can be derived as:

$$t(x) = 1 - \omega\min_{y\in\Omega(x)}(\min_{c}\frac{I^c(y)}{A^c}) \quad (8)$$

He et al. argued that a clear sky can be regarded as a state of light fog, and depth of field information can be reflected through the density of the fog [30]. Therefore, a density factor, ω, is introduced to retain a certain degree of fog, and typically set to 0.95.

The dark channel prior defogging algorithm can restore clear, haze-free images with stability and high efficiency. However, there is a problem of distortion in defogging the sky region, leading to the emergence of a series of improved algorithms.

Defogging algorithm based on neural network

With the rapid development of deep learning technology, researchers have proposed a series of neural network-based defogging algorithms, which can be divided into two categories based on different defogging principles: non-end-to-end algorithms based on atmospheric physics models, such as the classic DehazeNet algorithm, and end-to-end defogging algorithms, such as the classic GCANet algorithm.

DehazeNet, proposed by Cai et al. in 2016, is a non-end-to-end defogging algorithm, and its network structure is shown in **Figure 3** [32]. The network consists of four parts: feature ex-

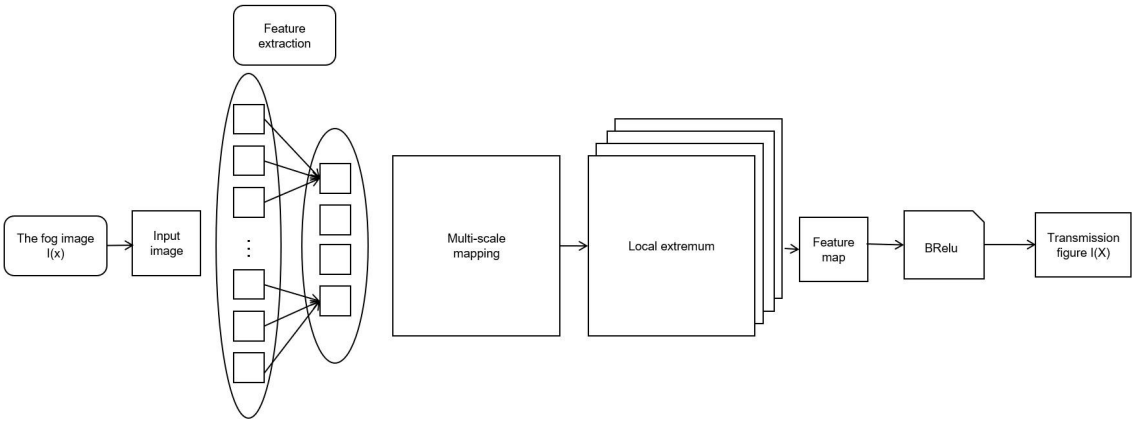


Figure 3. DehazeNet.

traction, multi-scale mapping, local extremum, and non-linear regression. Its input is a hazy image, and the output is the image's transmission rate map, which is then used to recover haze-free images based on the atmospheric physical model. Specifically, the structural design of the feature extraction layer of this algorithm combines four traditional algorithms: dark channel prior defogging algorithm, maximum contrast defogging algorithm, color attenuation prior algorithm, and defogging algorithm based on chromatic inconsistency. This fusion of neural networks and traditional algorithms leverages the powerful performance of the neural network to improve the traditional algorithms, providing new ideas for defogging algorithm research.

GCANet is an end-to-end image defogging algorithm based on the Generative Adversarial Network proposed by Chen et al. [33]. The network takes hazy images as input and directly outputs clear haze-free images. The algorithm consists of three convolutional blocks as the encoder part, one deconvolutional block, and two convolutional blocks as the decoder part. The algorithm focuses on using smooth convolution instead of dilated convolution to solve the problem of grid artifacts, and improves the defogging effect of the image by combining more contextual information and fusing different levels of features.

Current status of defogging algorithms

Based on image enhancement

To address the problem of excessive enhancement in the histogram equalization-based defogging algorithm, Kim et al. added contrast limiting modules to limit the enhancement, but leading to high algorithm complexity [34]. Soni et al. proposed an adaptive histogram equalization algorithm, which effectively improved the algorithm's efficiency and output results [35]. However, the histogram equalization-based algorithm only relies on the image's grayscale features, which are less robust and prone to additional noise.

Jobson et al. proposed a multi-scale Retinex algorithm based on the single-scale Retinex algorithm that effectively solved the halo problem but caused color distortion in the image [36]. To address this issue, researchers have proposed many improved algorithms. For example, Jobson et al. added a color restoration module based on multi-scale Retinex to propose a multi-scale Retinex with color restore algorithm, which effectively solved the color distortion problem in the defogging process [37]. The Retinex algorithm has made remarkable progress in addressing halo and color distortion problems, but with high complexity [6, 38].

Liu et al. proposed to use wavelet transform on the brightness information in the HSV space of the image, enhancing the brightness of high-frequency part of the image and suppressing that of the low-frequency part, and then restoring the color to achieve defogging [31]. This approach can reduce the algorithm's complexity. Wavelet transform has the advantage of processing images at multiple scales, and some researchers have fused wavelet transform algorithms with other algorithms for defogging. For example, Wang et al. processed foggy images with wavelet transform, sharpened the image with high-frequency signals, and then enhanced the image with server side render, effectively improving the defogging efficiency of the algorithm [39].

Histogram equalization, Retinex algorithms, and wavelet transforms, which are based on image enhancement, mainly adjust pixel values

based on the image's statistical information or highlight local details in the image to achieve defogging effects. Essentially, they are all image enhancement techniques. The histogram equalization algorithm has low complexity, can enhance the image's brightness and contrast, but its robustness is poor and can easily generate additional noise. This algorithm is suitable for thin fog images that require brightness enhancement or have high resolution, but its limitations are significant when enhancing haze removal mages. Future research directions aim to enhance the algorithm's flexibility to apply to images with significant differences, but this could potentially further increase the algorithm's complexity. The Retinex algorithm and its modification algorithms are relatively comprehensive, especially the multi-scale Retinex with color restore algorithm, which performs well in haze removal, thick fog, and low light conditions and has good visual outcomes. However, the Retinex series algorithms have high complexity. Wavelet transforms depend on parameter selection, and the parameters are different under different light intensities and background conditions. This approach generally cannot be applied solely to defogging.

Based on physical models

Xiao et al. proposed a scene-aware defogging algorithm by using gamma correction based on sky segmentation, which solved the color distortion problem but could cause excessive enhancement in high-light areas [40]. Li et al. proposed a threshold-based sky region segmentation algorithm in 2021, which has good defogging effects on the sky region but has a high computational complexity [41].

Zhu et al. proposed a defogging strategy based on the color attenuation prior in 2015 [42]. The color attenuation prior theory suggests that the difference between the brightness and saturation of outdoor images is positively correlated with the haze concentration. This algorithm creates a linear model to simulate the scene depth of a hazy image and uses supervised learning methods to learn the model parameters, effectively restoring the image depth information. Then, the atmospheric scattering model is used to derive parameters such as the transmission rate to obtain a dehazed image. Yang et al. proposed using machine learning methods to derive depth-of-field information and achieve defogging effects based on the color prior theory, but the algorithm's stability is poor [43].

Defogging algorithms based on physical models rely on the imaging principles of images and atmospheric light scattering models and have some theoretical support from physics. Although defogging strategies based on additional information and image differences can achieve good defogging effects, they rely heavily on geographic information and databases, which limits their practical applications. Defogging algorithms based on the dark channel prior and color attenuation prior have less dependence on additional information. The dark channel-based defogging algorithm performs well on thin haze images, but its performance is poor in strong light, thick fog, and mist conditions. The color attenuation prior-based algorithm performs poorly on images with severe color distortion and unclear image gradient structure. Overall, defogging algorithms based on physical models have better visual outcomes, and the defogging process does not introduce additional noise, but they are severely limited by the scene.

Based on Neural network

Wang et al. proposed the STCSDN (semi training color stripping dehaze-net) defogging algorithm based on DehazeNet [44]. The algorithm assumes that convolutional neural networks have two properties in the defogging process: 1) they learn contour and shadow information faster than color information, and 2) the network is not sensitive to the concentration of haze. STCSDN uses cycle generative adversarial network and a semi-trained generator as a feature extraction module to extract haze-free grayscale images from blurry color images. The algorithm can handle haze images with different concentrations, and can effectively restore image details and enhance visual outcomes. Yang et al. proposed a hybrid iterative model that combines the dark channel prior theory and DehazeNet algorithms to better restore haze-free images that are closer to real scenes, but the algorithm is computationally complex [45]. Non-end-to-end defogging algorithms have the support of physical theory, which makes the dehazed images closer to real scenes, but they have higher parameter count and lower algorithmic simplicity.

Similarly, Engin et al. enhanced the cycle generative adversarial network formula by combining cycle consistency and perceptual loss to improve the model's ability to recover texture information, so as to obtain haze-free images with better visual outcomes [46]. Li et al. proposed the all-in-one network defogging algorithm based on convolutional neural networks, which can directly generate clear haze-free images according to the atmospheric scattering

model [47]. End-to-end defogging strategies have better defogging effects, but they also suffer from issues such as more model parameters and difficulty in obtaining datasets.

Two deep learning-based defogging strategies are currently the research focus: (1) inferring certain parameters in the atmospheric scattering model using neural networks to obtain dehazed images, and (2) directly inputting hazy images and outputting haze-free images. The former can combine image enhancement algorithms with neural networks to enhance the visual outcomes of dehazing, but the parameter complexity also increases exponentially, and the stability of the algorithm is relatively poor. The latter eliminates the model's dependence on prior conditions and reduces a significant number of parameters, but the complexity increases further due to the intricate nature of real scenes and an excessive reliance on samples.

The application field of fog removal algorithm

With the thriving development of defogging algorithms based on three different strategies, defogging algorithms have been widely applied in various fields such as target detection in foggy weather, intelligent driving in foggy weather, and industrial image de-dusting. In the field of intelligent driving, autonomous driving relies on the detection of objects in the target scene. Haze can change the color of objects and weaken the brightness of the scene, affecting the performance of autonomous driving equipment. To address this issue, Ma et al. employed the Retinex algorithm to estimate the atmospheric light intensity and transmission rate based on physical properties such as color consistency in hazy images, thereby improving the road segmentation accuracy of the target scene [48]. Xu et al. replaced the soft matting process in the dark channel prior algorithm with two different filters to improve the real-time performance of dehazing, significantly improving the speed of video dehazing [49]. El-Hashash et al. also designed a real-time video defogging system based on the dark channel prior algorithm, greatly improving the performance of automatic driving in foggy weather [50].

In the field of remote sensing images, haze can cause imaging defects such as low contrast and low clarity, which are not conducive to the analysis of detailed information of images captured by devices such as drones and optical satellites. To address this issue, Huang et al. proposed an adaptive transmission rate estimation method based on the dark channel prior algorithm, according to the histogram of the dark channel image and the maximum transmission rate constraint, to obtain high-fidelity optical satellite images [51]. Li et al. proposed an end-to-end convolutional neural network defogging algorithm, which first improved the atmospheric scattering model into an end-to-end defogging model, unified multiple unknown parameters, and used a multi-scale convolutional neural network to estimate the unknown parameters, and then input the estimated values into the model to obtain haze-free images [52]. This algorithm has different degrees of improvement in visual outcomes and objective indicators for hazy remote sensing images.

In the industrial field, underground coal mine images are characterized by low illumination, high dust, and high noise intensity, which makes it difficult to extract effective image information. To address this issue, Zhang et al. enhanced the underground images using histogram equalization, obtained the dark channel image from the thumbnail image, and restored the pixel values of the image using bilinear interpolation to improve the algorithm's running speed [53]. This algorithm improved the defogging effect of underground coal mine images and increased the algorithm processing speed. Wu et al. also proposed a defogging method for degraded underground coal mine images based on the dark channel prior algorithm [54]. The algorithm weakened the influence of dust and water mist in the environment, achieving image enhancement. Similarly, researchers from Anna University implemented underwater image defogging and color correction based on the single-scale Retinex algorithm with color correction [55].

In the field of medicine, the visibility and analyzability of medical images have a great impact on the diagnosis of diseases. Defogging algorithms are often used for denoising and enhancement in the medical image field, but the application scenarios are relatively single. For example, Chen et al. combined the dark channel prior algorithm and the Retinex algorithm in a continuous sequence of medical images to achieve denoising of CT images and MRI images, effectively improving the image contrast [56]. Justin et al. proposed a local contrast mapping scheme based on the histogram equalization algorithm, achieving controllable denoising of MRI images [57].

In special scenarios, medical images may exhibit fog-like appearances. Applying physical model-based defogging algorithms to such medical images can provide critical diagnos-

tic indicators, enabling the quantification and grading of corresponding medical features through artificial intelligence. For example, the imaging-based features of interstitial pneumonia may include ground-glass opacities, consolidations, fibrotic stripes, thickened interlobular septa, traction bronchiectasis, nodules, and reticular opacities, which appear as fog-like features [58]. By applying defogging algorithms to pneumonia CT images, the X-ray transmission rate information of the local pneumonia medium can be obtained to quantify the severity of pneumonia in this area, ultimately achieving a scientific and automatic quantification and grading definition.

Summary

Currently, there are three main approaches for foggy image restoration: image enhancement-based, physics-based, and deep learning-based methods. Image enhancement-based strategies mainly enhance image contrast and highlight details using mathematical methods to achieve the goal of defogging. Physics-based image defogging strategies are mainly based on atmospheric scattering models, in which key physical parameters are derived to obtain haze-free images by inverse calculation. Deep learning-based methods utilize powerful computational and fitting abilities of neural networks to estimate the parameters in atmospheric scattering models or directly find the mapping relationship between hazy and clear images for defogging.

The defogging strategy based on image enhancement does not require prior information and is easy to apply, yet it often introduces noise and yields constrained defogging outcomes. In contrast, the physical model-based defogging strategy is supported by physical theory. The defogging algorithm based on this model has better visual outcomes, and introduces no additional noise during the defogging process. It is more stable and efficient, but requires a lot of prior information, making it difficult to apply, and severely constraining the applicable scenes. The defogging strategy based on deep learning can enhance the visual outcome of image defogging, but how to reduce the parameters of the model and reduce the excessive dependence on the sample is an urgent problem to be solved.

All three types of defogging algorithms have significant defogging effects, but there are still some research difficulties and challenges in the development of defogging algorithms, including:

(1) High spatio-temporal complexity of algorithms. Existing algorithms have good defogging outcomes, but their high complexity poses challenges in adapting them to complex scenes such as video defogging or real-time defogging, which is not conducive to the practical application of the algorithms.

(2) Difficulties in obtaining reliable datasets. Due to the uncontrollable factors such as lighting and fog density in real-life scenes, it is difficult to obtain high-quality pairs of hazy and clear images. The robustness and stability of deep learning algorithms require high-quality datasets, which limits the development of neural network-based defogging algorithms. To address the issue of data quantity and quality, considerable human and material resources must be invested. Alternatively, transfer learning or other methods can be used to alleviate dataset limitations.

(3) Difficulty in ensuring algorithm robustness. Existing defogging algorithms have good visual outcomes when applied to one or several specific application scenarios. How to further improve the robustness of the models is the focus of future research. Combining multiple defogging algorithms for different scenarios can be a solution to increase algorithm robustness.

(4) Further expansion of application scenarios. Defogging algorithms have been applied in intelligent driving, unmanned inspection, and other fields. However, in the field of medical image processing, defogging is still at the image enhancement stage, and there has been no research on fog-like symptoms such as pulmonary pneumonia. It is possible to model based on medical imaging principles and explore methods to solve the severity grading of pneumonia.

References

[1] Karkera T, Singh C. Autonomous Bot Using Machine Learning and Computer Vision. SN Comput Sci 2021;2(4):251.

[2] Zhang X, Xu S. Research on Image Processing Technology of Computer Vision Algorithm. 2020 International Conference on Computer Vision, Image and Deep Learning (CVIDL) 2020;122-124.

[3] Ju M, Ding C, Guo YJ, et al. Remote Sensing Image Haze Removal Using Gamma-Correction-Based Dehazing Model. IEEE Access 2019;7:5250-5261.

[4] Parihar AS, Gupta YK, Singodia Y, et al. A Comparative Study of Image Dehazing

Algorithms. 2020 5th International Conference on Communication and Electronics Systems (ICCES) 2020;766-771.
[5] Sarkar M, Sarkar PR, Mondal U, et al. Empirical wavelet transform-based fog removal via dark channel prior. IET Image Process 2020;14(6):1170-1179.
[6] Zhou J, Zhang D, Zou P, et al. Retinex-Based Laplacian Pyramid Method for Image Defogging. IEEE Access 2019;7:122459-122472.
[7] Zhang W, Dong L, Pan X, et al. Single Image Defogging Based on Multi-Channel Convolutional MSRCR. IEEE Access 2019;7:72492-72504.
[8] Kansal I, Kasana SS. Fusion-based image de-fogging using dual tree complex wavelet transform. Int J Wavelets, Multiresolution Inf Process 2018;16(06):1850054.
[9] Wu H, Lan J. A Novel Fog-Degraded Image Restoration Model of Golden Scale Extraction in Color Space. Arabian J Sci Eng 2018;43(12):6801-6821.
[10] Fu F, Liu F. Wavelet-Based Retinex Algorithm for Unmanned Aerial Vehicle Image Defogging. 2015 8th International Symposium on Computational Intelligence and Design (ISCID) 2015;426-430.
[11] Guo F, Lan G, Xiao X, et al. Parameter Selection of Image Fog Removal Using Artificial Fish Swarm Algorithm. Intell Computing Theor Appl 2018;25-37.
[12] Zhang W, Hou X. Light source point cluster selection-based atmospheric light estimation. Multimedia Tools Appl 2018;77(3):2947-2958.
[13] Qu C, Bi D. Novel Defogging Algorithm Based on the Joint Use of Saturation and Color Attenuation Prior. IEICE Trans Inf Syst 2018;E101.D(5):1421-1429.
[14] Shiau YH, Kuo YT, Chen PY, et al. VLSI Design of an Efficient Flicker-Free Video Defogging Method for Real-Time Applications. IEEE Trans Circuits Syst Video Technol 2019;29(1):238-251.
[15] Li B, Zhao J, Fu H. DLT-Net: deep learning transmittance network for single image haze removal. Signal, Image Video Process 2020;14:1245-1253.
[16] Ren W, Pan J, Zhang H, et al. Single Image Dehazing via Multi-scale Convolutional Neural Networks with Holistic Edges. Int J Comput Vision 2020;128(1):240-259.
[17] Ren W, Zhang J, Xu X, et al. Deep Video Dehazing With Semantic Segmentation. IEICE Trans Image Process 2019;28(4):1895-1908.
[18] Liu Y, Pan J, Ren J, et al. Learning Deep Priors for Image Dehazing. 2019 IEEE/CVF International Conference on Computer Vision (ICCV) 2019;2492-2500.
[19] Jiang X, Sun J, Li C, et al. Video Image Defogging Recognition Based on Recurrent Neural Network. IEEE Trans Ind Inf 2018;14(7):3281-3288.
[20] Shi Y, Jiang X. Deep quality assessment toward defogged aerial images. Signal Process: Image Commun 2020;83:115808.
[21] Qian W, Zhou C, Zhang DY. FAOD-Net: A Fast AOD-Net for Dehazing Single Image. Math Probl Eng 2020;2020:1-11.
[22] Liu W, Yao R, Qiu G. A physics based generative adversarial network for single image defogging. Image Vision Comput 2019;92:103815.
[23] Santra S, Mondal R, Chanda B. Learning a Patch Quality Comparator for Single Image Dehazing. IEICE Trans Image Process 2018;27(9):4598-4607.
[24] Land EH, McCann JJ. Lightness and Retinex Theory. J Opt Soc Am 1971;61(1):1-11.
[25] McCartney EJ. Optics of the atmosphere: Scattering by molecules and particles 1976.
[26] Narasimhan SG, Nayar SK. Chromatic framework for vision in bad weather. Proceedings IEEE Conference on Computer Vision and Pattern Recognition CVPR 2000 (Cat NoPR00662) 2000;598-605 vol.591.
[27] Narasimhan SG, Nayar SK. Contrast restoration of weather degraded images. IEEE Trans Pattern Anal Mach Intell 2003;25(6):713-724.
[28] Kopf J, Neubert B, Chen B, et al. Deep photo: model-based photograph enhancement and viewing. ACM Trans Graph 2008;27(5):116.
[29] Shwartz S, Namer E, Schechner YY. Blind Haze Separation. 2006 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'06) 2006;1984-1991.
[30] He K, Sun J, Tang X. Single Image Haze Removal Using Dark Channel Prior. IEEE Trans Pattern Anal Mach Intell. 2011 Dec;33(12):2341-53.
[31] Liu F, Cao L, Shao X, et al. Polarimetric dehazing utilizing spatial frequency segregation of images. Appl Opt 2015;54(27):8116-8122.
[32] Cai B, Xu X, Jia K, et al. DehazeNet: An End-to-End System for Single Image Haze Removal. IEEE Trans Image Process 2016;25(11):5187-5198.
[33] Chen D, He M, Fan Q, et al. Gated Context Aggregation Network for Image Dehazing and Deraining. 2019 IEEE Winter Conference on Applications of Computer Vision (WACV) 2019;1375-1383.
[34] Joung-Youn K, Lee-Sup K, Seung-Ho H. An advanced contrast enhancement using partially overlapped sub-block histogram equalization. IEEE Trans Circuits Syst Video

Technol 2001;11(4):475-484.
[35] Soni B, Mathur P. An Improved Image Dehazing Technique using CLAHE and Guided Filter. 2020 7th International Conference on Signal Processing and Integrated Networks (SPIN) 2020;902-907.
[36] Jobson DJ, Rahman Z, Woodell GA. Properties and performance of a center/surround retinex. IEEE Transact Image Process 1997;6(3):451-462.
[37] Jobson DJ, Rahman Z, Woodell GA. A multiscale retinex for bridging the gap between color images and the human observation of scenes. IEEE Transacti Image Process 1997;6(7):965-976.
[38] Rahman ZU, Jobson DJ, Woodell GA. Retinex processing for automatic image enhancement. IS&T/SPIE Electron Imaging 2002.
[39] Wang M, Zhou S-d, Huang F, et al. The study of color image defogging based on wavelet transform and single scale retinex 2011;81940F.
[40] Xiao J, Zhu L, Zhang Y, et al. Scene-aware image dehazing based on sky-segmented dark channel prior. IET Image Processing [Internet]. 2017; 11(12):[1163-1171 pp.]. Available at: https://digital-library.theiet.org/content/journals/10.1049/iet-ipr.2017.0058.
[41] Li WX, Jie W, Zadeh SM. Single Image Dehazing Algorithm Based on Sky Region Segmentation. International Conference on Advanced Data Mining and Applications 2020.
[42] Zhu Q, Mai J, Shao L. A Fast Single Image Haze Removal Algorithm Using Color Attenuation Prior. IEEE Trans Image Process 2015;24(11):3522-3533.
[43] Yang A, Wang Q, Ji Z, et al. Single underwater image restoration based on adaptive color correction and adaptive transmission fusion. J Electron Imaging 2020;29:043006-043006.
[44] Wang Y, Ji D. Single Image Dehazing Method Based on Semi-Training Color Stripping. J Phys: Conf Ser 2021;1982(1):012057.
[45] Yang S, Sun Z, Jiang Q, et al. A Mixed Transmission Estimation Iterative Method for Single Image Dehazing. IEEE Access 2021;9:63685-63699.
[46] Engin D, Genç A, Ekenel HK. Cycle-Dehaze: Enhanced CycleGAN for Single Image Dehazing. 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW) 2018:938-9388.
[47] Li B, Peng X, Wang Z, et al. AOD-Net: All-in-One Dehazing Network. 2017 IEEE International Conference on Computer Vision (ICCV) 2017;4780-4788.
[48] Ma N, Xu J, Li H. A Fast Video Haze Removal Algorithm Via Dark Channel Prior. Procedia Comput Sci 2018;131:213-219.
[49] Xu J, Ma N, Ke J, et al. A fast video haze removal algorithm via mixed transmissivity optimisation. Int J Embedded Syst 2019;11(1):84-93.
[50] El-Hashash MM, Aly HA. High-speed video haze removal algorithm for embedded systems. J Real-Time Image Process 2019;16(4):1117-1128.
[51] Huang W, Wang Y, Wang R. A high fidelity haze removal algorithm for optical satellite images using progressive transmission estimation based on the dark channel prior. Int J Remote Sens 2019;40(9):3486-3503.
[52] Li Y, Ren J, Huang Y. An End-to-End System for Unmanned Aerial Vehicle High-Resolution Remote Sensing Image Haze Removal Algorithm Using Convolution Neural Network. IEEE Access 2020;8:158787-158797.
[53] Zhang X, Guo H. Research on an improved algorithm for image dehazing in underground coal mine. J Phys: Conf Ser 2020;1693(1):012153.
[54] Wu T, Yang X, Li J, et al. Image haze removal and computer-generated holographic display method in coal mine. International Conference on Automation Control, Algorithm, and Intelligent Bionics (ACAIB 2022) 2022.
[55] Jia JL, Amuzie AU, Lester JC. Encouraging the next generation of skin of color researchers: Funding skin of color research in medical school and residency. J Am Acad Dermatol 2021;85(3):e161-e163.
[56] Chen L, Tang C, Xu M, et al. Enhancement and denoising method for low-quality MRI, CT images via the sequence decomposition Retinex model, and haze removal algorithm. Med Biol Eng Comput 2021;59(11):2433-2448.
[57] Joseph J, Periyasamy R. A fully customized enhancement scheme for controlling brightness error and contrast in magnetic resonance images. Biomed Signal Process Control 2018;39:271-283.
[58] Wang DY, Salem JE, Cohen JV, et al. Fatal Toxic Effects Associated With Immune Checkpoint Inhibitors: A Systematic Review and Meta-analysis. JAMA Oncol 2018;4(12):1721-1728.

Progress in Medical Devices
eISSN: 2957-5478

Zentime

Review Article https://doi.org/10.61189/848897sppwnn

Preparation and application of electro-conductive hydrogels in biomedical engineering

Wenwen Zhang, Lin Mao, Zhongxin Hu, WanwenYang, Linying Zhang, Chengli Song

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

Declaration of conflict of interest: None.

Received May 7, 2023; Accepted August 25, 2023; Published September 30, 2023

Highlights

- The conductivity of hydrogel can be realized by introducing conductive medium.
- Various properties of hydrogel can be adjusted by changing the proportion of ingredients.
- Electro-conductive hydrogels are widely used in biomedical engineering as biomaterials.

Abstract

Electro-conductive hydrogel is a new composite hydrogel with high electrical conductivity, extraordinary mechanical properties, and controllability. It can be equipped with specific materials to obtain specific properties, which has attracted considerable research attention due to its wide range of application as a biomaterial in the biomedical engineering field. It can be used as a patch to promote healing and as a dressing for burns and pressure wounds. In pharmaceutical science, it is widely studied as a carrier of targeted and quantitative drug release. In surgery, it can serve as a cell scaffold for implantable therapy. Currently, increasing research is focusing on the development of sensors as simulated skin for robots, as well as in health monitors for human activity and well-being. In this review, we summarize the classification and selection of conductive mediums, natural polymer hydrogel matrices, and the application of electro-conductive hydrogels in the fields of bioengineering and biomedicine. We delve into the properties of different conductive mediums and the principle of combining different hydrogel matrices and discuss the advantages and disadvantages of emerging composite hydrogels. Our focus extends to the forefront of conductive hydrogel research in various fields, while also elucidating the current problems and challenges in terms of performance integration, preparation process and practical applications.

Keywords: Conductive hydrogel, biomedical engineering, conductive mediums, natural polymer, flexible sensors

Introduction

Hydrogels, polymer with three-dimensional (3D) networks formed through physical or chemical cross-linking, can contain a large amount of water while remaining insoluble in water [1]. The network pore size of the polymer is about 10 nm, much larger than the size of water molecules, which allows the water molecules in the hydrogel to maintain the same physical and chemical properties as liquid water [2]. At the same time, the microscopic network structure like sponge ensures that the hydrogel has a certain toughness. Hydrogel is an excellent biomaterial with excellent water absorption, controllable mechanical properties, biocompatibility, and biodegradability.

Nowadays, electricity is indispensable in human activities. If hydrogel is given a conductive property, the application range of hydrogels can be greatly improved. However, hydrogels made from natural polymers are not conductive or only weakly conductive, so combining hydrogel matrices and conductive materials provides a simple and efficient route to fabricate strong conductive hydrogels [3]. Conductive materials can be classified as nanomaterials, ionic materials and conducting polymers. **Figure 1** shows the composite hydrogel synthesized by incorporating conductive polymers, namely, electrically conductive hydrogels (ECHs). By harnessing the individual advantages of conductive materials and natural polymer hydrogel matrices, composite hydrogels not only maintain superior biocompatibility and biodegradability, but also

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

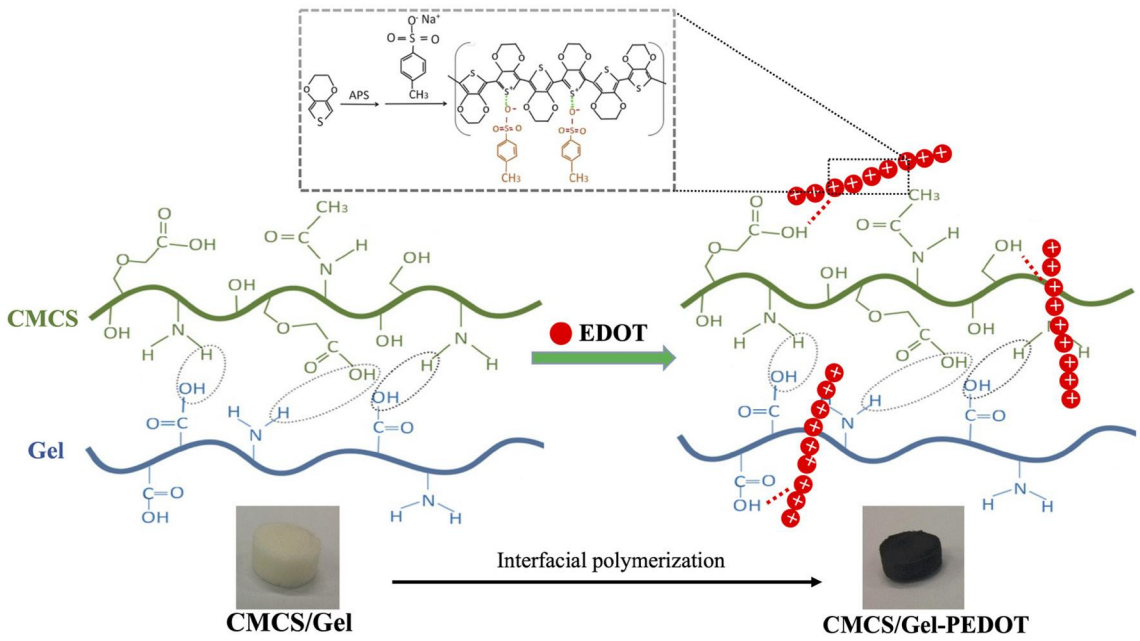


Figure 1. Conductive polymer ECHs. This figure is cited from [16]. ECHs, electrically conductive hydrogels; PEDOT, poly (3,4-ethylenedioxythiophene).

have high electrical conductivity. Many studies have shown that introducing different conductive materials into hydrogels can also improve their strength, toughness, and other mechanical properties. Powered by these biological and tunable features, ECHs have recently attracted growing attention in biomedical engineering.

Classification of conductive materials

Nanomaterials

The nanomaterials used to prepare ECHs mainly contain carbon-based nanomaterials, such as graphene and carbon nanotubes (CNTs), and metallic nanomaterials. Introducing nanomaterials to hydrogels not only enhances the electrical conductivity of ECHs, but also improves their mechanical, optical and thermal properties. Nanomaterials have attracted growing interest as conductive mediums for ECHs and have become one of the preferred components for the preparation of ECHs.

The physical interaction between hydrogel polymer chains and carbon-based nanomaterials contributes to the mechanical properties of hydrogels and makes the ECHs exhibit high electrical conductivity and excellent environmental stability [4]. However, the preparation process of combining nanomaterials and hydrogels is difficult to control because nanomaterials are hydrophobic and tend to aggregate in water, which may lead to uneven distribution of nanomaterials in the manufactured hydrogels, resulting in discontinuous electrical conductivity and inhomogeneous mechanical properties.

Metallic nanomaterials, which are inherently conductive, can conduct electricity under the action of electric current when they are free in the hydrogel network. However, metallic materials are unstable in humid environments and are susceptible to oxidation. For example, MXene, which has been studied by many researchers as an electrode material, has excellent conductivity and rich varieties, but its brief shelf life, susceptibility to oxidation, and use of environmentally harmful reagents in synthesis impose many limitations on its application [5].

Ionic materials

The high-water contents and 3D networks of hydrogels permit ions to migrate between the electrodes in the presence of electric current. Current research generally uses salts as a source of ions. Salts ionize into metal ions and anions in water, and both of which can act as conductors. The ECHs is prepared using an immersion method that utilizes the water absorption and swelling properties of hydrogels to incorporate the ions into the polymeric network. Metal ions can form complex coordination structures with functional groups on the soft chains of hydrogels, including monodentate coordination, bidentate chelate bidentate bridging and interchain cross-linked structures. The hydrogen bonds between the soft and rigid chains of hydrogels, as well as the stronger physical cross-linking points formed between the ions and chains make the ECHs have high

Table 1. Comparison of different conductive mediums

Type	Materials	Advantage	Disadvantage
Nanomaterials	Carbon-based (graphene, CNTS)	Excellent mechanical, optical, and thermal properties, chemical stability	Hydrophobic, complex synthesis process
	Metal-based (MXene)	High conductivity, self-healing ability, ductility	Unstable, harmful synthesis process
Ionic materials	Fe^{3+}, Ag^{+}, Ca^{2+}	High conductivity, low cost, simple synthesis process, functional specificity	Toxic at high concentrations
Conducting polymers	PPy, PANI, PEDOT, PEDOT: PSS	Excellent mechanical, optical, and thermal properties, high conductivity	Non-degradability, low solubility

Note: CNTS, carbon nanotubes; PPY, polypyrrole; PANI, olyaniline; PEDOT, poly (3,4-ethylenedioxythiophene); PEDOT: PSS, polystyrene sulfonate.

tensile strength while ensuring the toughness. Meanwhile, the reversible association and dissociation abilities of the ligand bonds provide ECHs excellent recoverability and durability [6]. Therefore, ions combine with water molecules to form hydrated ions. This hydration effect usually changes the freezing point of the hydrogels and increases their low temperature resistance [7, 8]. Different ions can also bring different specific properties to hydrogels. For instance, introducing Ag^{+} can endow hydrogels with antibacterial properties, and introducing Ca^{2+} can contribute to hemostasis of wound healing and promote cell proliferation and migration [9, 10].

Compared with ordinary ionic gels, the hydrogels based on salt electrolytes have the advantages of high conductivity, controllable cost, and simple preparation, with an irreplaceable practical application value.

Conducting polymers

Conducting polymers, including polypyrrole (PPy), polyaniline (PANI), poly (3,4-ethylenedioxythiophene) (PEDOT), and polystyrene sulfonate (PEDOT: PSS), are new generation of organic materials characterized by electrical conductivity. They have the electrical property of metallic and optical property of inorganic semiconductors, known for their redox stability, hydrophobicity, ability for reversible oxidation, and ease of synthesis and processing [11]. These properties, which give them advantages and potentialities in tissue engineering and regenerative medicine, have attracted great interest from researchers.

Therefore, the introduction of conducting polymers into hydrogels is an effective strategy to obtain ECHs. This approach highlights the idiosyncratic functions of polymers while leveraging the biocompatible and degradable qualities of natural hydrogels, creating a suitable milieu for applications in biomedical engineering.

The physical and chemical properties of conducting polymers can be designed by surface functionalization technology and electronic means, and the intensity and length of the electrical stimulus can be controlled when conducting polymers are used [11]. However, conducting polymers have two major disadvantages, non-degradability and low solubility in organic solvents and water, which greatly limits their processability, versatility, and applicability [12]. As shown in **Table 1**, we summarized and compared the advantages and disadvantages of introducing different conductive media into hydrogels.

Classification of natural polymer hydrogel matrices

Sodium alginate (SA)

SA is a natural polysaccharide separated from seaweed, which is composed of polymannuronic acid and polyguronic acid segments. SA forms a stable network by replacing sodium ions on the carboxyl group of the skeleton with other divalent metal ions. It has been widely used in medicine, biology and other fields due to it is non-irritating and biocompatible nature [13]. However, the mechanical properties of pure alginate networks are poor. Recently, researchers introduced other networks into alginate to form interpenetrating double network hydrogels with stable mechanical properties through hydrogen bonding.

Mo et al. crosslinked Zn^{2+} with SA networks, and then introduced it into the covalently crosslinked polyacrylamide (PAAm) networks to form an interpenetrating double-network structure [14]. The ionic crosslinking points formed by Zn^{2+} and SA can effectively dissipate energy when they are broken under strain, and then recover bond after stress unloading. Additionally, surplus free Zn^{2+} in the hydrogel networks

provides conductivity to hydrogels. These double-network ECHs present porous micro morphology, which is conducive to the transport of ions. The ionic conductivity of the hydrogel is as high as 32.4 MS/cm. A wearable application test showed that the hydrogel exhibited a sensitive and repeatable response to the movement of human limb joints, showing great potential in this field.

Liu et al. developed an alginate-PAAm ECHs with Fe^{3+} secondary crosslinking to adjust local stiffness [15]. Because the trivalent Fe^{3+} combines three different alginate polymer chains by interacting with the carboxyl group on each chain, a 3D compact structure is established, which improves the mechanical stiffness of the hydrogel. The local stiffness is related to the concentration of Fe^{3+}. So, changing the local stiffness of the hydrogel by controlling the concentration of Fe^{3+} and the immersion position is helpful to promote the coupling of flexible rigid electronic components and soft matrices materials. This design strategy provides an effective solution to the problems in the field of soft and stretchable electronics.

Based on double-network hydrogel, SA based ECHs has been developed and applied to flexible sensors, drug delivery and release, cell scaffold and other scenarios.

Gelatin

Gelatin is a hydrolysate biopolymer of collagen. A large number of carboxyl groups in gelatin is very suitable for molecular grafting, while the high level of amino groups easily interacts with aldehyde groups, making them very suitable for producing characteristic hydrogels. At the same time, gelatin is cheap, and has the properties of biocompatibility and biodegradability, making them good matrices for cell adhesion and proliferation. Therefore, gelatin-based ECHs have been widely used in cell scaffolds, 3D biological printing, and have great potential in complex tissue engineering such as bone, muscle, and nerve regeneration.

Guan et al. developed a carboxymethyl chitosan/gelatin composite hydrogel [16]. PEDOT nanoparticles were assembled on the hydrogel surface by in-situ chemical polymerization to obtain a new type of electroactive hydrogel scaffold. The conductivity of this ECH can reach 1.5210^{-3} S/cm, and its mechanical properties are similar to brain tissue [16]. The incorporation of PEDOT nanoparticles significantly improved the adhesion of neural stem cells and supported long-term cell growth and proliferation in a 3D microenvironment. It is an attractive conductive matrix for nerve tissue repair and regeneration.

In addition, gelatin-based conductive hydrogels can also be used as flexible sensors to monitor human movement. Gao et al. designed a humidity- and temperature- sensitive chitosan/gelatin/glycerin/NaCl organic ion conductive hydrogel film with only 0.1 mm in thickness [17]. The sensor prepared with this ECH film has a wide detection range of relative humidity (20-90%) and temperature (-30-50 °C), and also shows good antifreeze and water retention capabilities, which can accurately record and real-time feedback the changes in humidity and temperature on skin surface, holding great application potential in detecting human respiratory activity.

The hydrogels formed only by gelatin have disadvantages of brittleness and poor flexibility, which greatly limits their practical application. They can only play their role through cross-linking modification or nano composition. Therefore, it is particularly important to improve the mechanical strength of gelatin-based ECHs.

Cellulose

Cellulose is the largest polysaccharide material among natural polymers in the world. It not only exists in plants, but also can be synthesized by some specific microorganisms, such as Acetobacter xylinum. As a renewable organic substance in nature, cellulose is used to prepare natural hydrogels. The design of conductive cellulose hydrogel functional materials based on cellulose's good mechanical, dielectric, and piezoelectric properties has become a recent research hotspot [18].

The main raw materials for preparing cellulose based ECHs include cellulose molecules, nano cellulose, and cellulose derivatives. Wu et al. used Tannic acid (TA) to modify bacterial cells (BCs) to achieve an even distribution within the hydrogel, forming a stable double-network structure with the P(AA-AMPS) chain. The hydrogen bond between TA and PAAM chains also improved the mechanical properties of the hydrogels [19]. Then a high concentration of Ca^{2+} was introduced to obtain P(AA-AMPS)-TA@BA-Ca^{2+} ECHs, which exhibited excellent conductivity (1.92 S/m at room temperature), mechanical properties, and freezing resistance (0.36 S/m at -35 °C).

Bian et al. used carboxyl cellulose nanofibers to help conductive PEDOT: PSS to be evenly dis-

tributed in the PAAm network [20]. In addition to its conductivity, the addition of carboxyl cellulose nanofibers makes ECHs have excellent humidity sensitivity and strain sensing performance. It has quick response, reliable stability and reproducibility in a wide range (0-837%), showing great potential as a human breath detection sensor.

All these have proven that cellulose-based ECHs have great potential as a new generation material for flexible wearable electronic equipment. However, the crystalline region of cellulose is difficult to dissolve, which is a limitation in the preparation of cellulose based ECHs. Currently, cellulose is mainly dissolved in water or common organic solvents by polar composite solvents or modification methods, which provides an available strategy for preparing cellulose based ECHs.

Chitosan

Chitosan is the deacetylated derivative of chitin, the second largest natural polymer after cellulose, and the only natural alkaline polysaccharide found at present. Due to its excellent properties such as low cytotoxicity, biocompatibility, as well as antibacterial, degradation and regeneration ability, it has been developed in the fields of drug release, wound healing and tissue engineering [6].

Au-Yong et al. crosslinked chitosan and polyethylene glycol into a double network [21]. They introduced PPy derivatives to synthesize ECHs that capable of carrying the cancer drug pemetrexed through photopolymerization and oxidative chemical polymerization. Experiments showed that under electrical stimulation, the release of pemetrexed was increased by about 10-15%. If this ECH is integrated into coatings and electrodes for electrochemical therapy, the triggered release may be helpful for surgery.

Cong et al. developed a dual network ECH with PANI doping, which is dynamically crosslinked by chitosan and PAAm [22]. The synergy of hydrogen bonding involving hydroxyl, amide, and aniline groups, along with the interaction of high concentrations of ions in both networks, bestows the resulting ECH with excellent mechanical properties, amazing conductivity, and impressive frost resistance. The highest conductivity of PCPD hydrogel can reach 4.83 S/m, but the PANI concentration can impact the degree of crosslinking of the hydrogel, and then affect the movement of free ions in the network, that is, there is a certain competition between the two conductive mechanisms. The tensile strength of this ECH is 2.62 MPa, and the excellent fatigue resistance and long-term durability of PCPD hydrogel are verified through cyclic tensile test, which is very important for chitosan-based ECHs as flexible sensors, soft robots, and wearable devices materials.

But like cellulose, chitosan is insoluble in water and most organic solvents, which severely limits its application scope. Therefore, modification of chitosan is an important aspect of its research, which can be cross-linked in the form of esterification and etherification to fabricate chitosan-based ECHs [23].

Others

In addition to the polymers mentioned above, there are also hyaluronic acid, starch, agarose, collagen, and other nature polymers that can be used as ECH matrices.

Sharifisistani et al. designed a hyaluronic acid-based ECH [24]. They introduced phenol-substituted gelatin for cell adhesion and phenol-substituted CNT as conductive medium through laccase-mediated cross-linking, creating a microcapsule and microcarrier for myocardial cell encapsulation. Their experiment showed that cell viability was affected during the encapsulation process, and myocardial cells proliferated well and expressed proteins and genes. Besides, the cardiogenic characteristics of the developing spherical tissue structure could be regulated by electrical stimulation.

Zhou et al. developed an AA-AM hydrogel by cross-linking starch as a conductive hydrogel matrix with acrylamide [25]. They found that starch improved the adhesion of hydrogel to various materials, and the adhesion strength did not decrease with time. This starch-based ECH has good flexibility, and the maximum strain reach 1290%. At the same time, as a biosensor, it shows excellent conductivity and strain response sensitivity and can accurately detect various human movements in a wide strain range of 0-1200%. It provides a new idea for the development of ECH materials in the field of new-generation electronic skin.

In general, natural polymer based ECHs have advantages of non-toxic, biodegradable, excellent flexibility, and conductivity, so they are very promising in practical applications. However, most natural hydrogels have disadvantages of low mechanical strength and lack of toughness. Therefore, in order to promote the practical application and improve their poor mechanical

Table 2. Comparison of different natural polymer hydrogel matrices

Materials	Advantage	Disadvantage
Sodium alginate	Biocompatibility, biodegradability, low cost	Poor mechanical properties
Gelatin	Tunable degradation, low cost, specific processing	Brittleness, poor mechanical properties
Cellulose	Good mechanical properties, dielectric properties	Difficult to dissolve
Chitosan	Antibacterial ability, regeneration ability, adsorption capacity, thermosensitive properties	Difficult to dissolve
Others	Adhesion, high amenability, multifunctional properties	Poor mechanical properties

properties, complementary matrices are constructed to study the conductivity and mechanical properties of ECHs through dual networks and interlocking cross-linking networks strategies. As shown in **Table 2**, we summarized and compared the properties advantages and disadvantages of different hydrogel matrices.

Applications

Bioelectronic flexible sensors

The popularity of smart terminals has greatly contributed to the development of wearable electronic devices. Traditional electronic devices made of rigid materials cannot achieve close contact with the skin, resulting in a lack of wear resistance and functionality [26]. Therefore, flexible and comfortable electronic devices have attracted extensive research attention and have been applied in health monitoring, medical diagnosis, and human health care. Among these devices, flexible sensors, as the most indispensable part, limit the functional development and future design of wearable devices.

ECHs have mechanical properties similar to human skin, such as notable stretchability, elasticity, recoverability, and fatigue resistance, and also have high electrical conductivity, sensitive responsiveness, and biocompatibility, which are preferred properties for the fabrication of flexible sensors. Currently developed ECHs for flexible sensors can effectively detect human activities such as finger bending, elbow bending, walking, and running. They are also sensitive to small movements, such as talking, swallowing, and small pulses [27].

You et al. prepared a porous organic Gel/PPy/rGO hydrogel with stretchability, high electrical conductivity, ultra-permeability, low-temperature tolerance, and long-term stability [28]. The sensor made from this porous organic hydrogel exhibits excellent sensing sensitivity, fast response capability, and good endurance. It can effectively monitor various human activities, such as finger flexion, elbow bending, walking, running, as well as small pulse beats.

Sun et al. prepared a PAAm-oxCNTs hydrogel by in situ free radical polymerization, using gelatin to modify oxidized CNTs by hydrogen bonding, and homogeneously introducing oxidized CNTs into the PAAm matrices [27]. The mechanical properties and electrical conductivity of this hydrogel are directly related to the content of oxidized CNTs. Therefore, the mechanical properties and electrical conductivity of Gel/PPy/rGO hydrogel can be adjusted within a certain range according to the demand. In addition to that, it has excellent durability, recoverability, and fast mechanical response when detecting human motions without significant delay. The PAAm-oxCNTs hydrogel has a sensitive response to minor strain variations, allowing for accurate articulation tracking, which has great potential for applications in voice recovery and articulatory readaptation.

Skin adhesion is a basic requirement for hydrogels as sensors, and the currently developed hydrogels have achieved this. However, since most of the flexible sensors are designed for in vitro adhesion, where one side adheres to the skin and the other side is susceptible to external interferences like dust, grass, or paper, the sensor's signal detection and analysis may be compromised. Gong et al. developed a PAA/TA@HC/Fe^{3+} hydrogel with great mechanical and tissue adhesive properties and electrical conductivity, which can be used as a flexible sensor [29]. In addition, they treated the hydrogels with asymmetric adhesion. A 1 M/L solution of Fe^{3+} was applied to the top of one side of the hydrogel, and the adhesion force of the treated hydrogel side was 0 kPa. Fe^{3+} served as an effective sealant for the treated hydrogel, effectively chelating available carboxylates and other functional groups, which prevented signal disruption and analysis interference from external factors on the other side of the hydrogel.

The sensors can be applied not only for health

monitoring but also for medical diagnosis and treatment. Gao et al. simultaneously improved the mechanical properties and electrical conductivity of a wearable hydrogel sensor by introducing sodium carboxymethyl starch, which has ultra-sensitivity and fast response [30]. Assembling multiple sensors into a 3D sensor array allows the detection of stress and strain distributions in three dimensions. This capability facilitates the recording of the ECG signal peaks and waveforms, showing significant potential in human medical diagnostics.

The electrical conductivity and mechanical properties of ECHs can be affected by the environment, such as temperature, pH, and light, and these issues need to be addressed for sensors as long-term use medical products. Chen et al. developed hydrogels with excellent freeze protection (-33 °C) by introducing zinc chloride into cellulose-based hydrogels, which are of great value in extreme environmental applications [31]. Lu et al. developed SPAE hydrogels with reduced swelling properties using electrostatic repulsion, which can be applied to underwater wearable flexible sensors [32]. Chen et al. developed a chitosan-P(AM-MPC-$AA_{0.20}$)-Fe^{3+} hydrogel with anti-fouling and anti-bacterial properties [33]. Shen et al. developed a PEDOT:PSS-PVA hydrogel that can be combined with 3D printing [34].

However, the integration of hydrogel-based flexible sensing systems still presents challenges in terms of cost, multifunctionality, and device fabrication, hindering application potential in specific scenarios.

Wound healing patch

Wound healing is a common health burden affecting all humans, and the lengthy healing process of severe wounds can be both physical and economical burdens. Accelerating wound healing and restoring normal activities is beneficial to both patients and the society. To date, various methods have been developed to treat wounds, including micro needling, wound dressings, negative pressure, and ultrasound, but the average time to complete wound closure is still long.

Currently, researchers have made progress in the application of hydrogels in the field of wound healing by combining hydrogels with specific materials, such as silver and its derivatives with antibacterial properties and calcium ions with hemostatic properties, or therapeutic drugs, such as collagen and dopamine [35-38]. Those composite patches can effectively facilitate wound healing [39]. Electrical field stimulation is a novel method to accelerate wound healing. This method can limit the side effects associated with conventional therapies. Studies have demonstrated that electrical field stimulation at the wound site can guide the migration and proliferation of skin cells through electrical signals [40]. In addition, electrical stimulation increases angiogenesis and immunomodulation, both of which play important roles in tissue remodeling [41].

Wang et al. synthesized silver nanowires and methacrylate alginate to make a flexible electronic patch [42]. The silver nanowires have electrical conductivity and antibacterial properties, while the methacrylate alginate has clinical applicability in wound healing. The hydrogels were shown to promote epithelial re-formation, enhance angiogenesis, and mediate immune response. The composition of this ECHs was optimized to be printed on medical-grade patches for personalized wound treatment. Animal experiments confirmed that the wounds that originally required 20 days to heal, were completely healed in only 7 days after patch treatment, and the skin thickness and other characteristics after healing were similar to those of healthy skin.

Pressure wounds are a worldwide concern because they are difficult to heal and easily infected due to repeated pressure over time [43]. ECHs can be used not only for the prevention of pressure injuries but also for long-term treatment [44].

Liu et al. designed a pressure-resistant amphoteric CH-OSNP ionic hydrogel to monitor wound conditions and prevent injury progression [45]. This hydrogel can detect and differentiate multiple parameters for analysis due to its sensitivity to temperature, pressure, and exudate. Additionally, its robust swelling properties keep the wound environment moist, regulate oxygen tension and foster capillary formation while aiding necrotic tissue dissolution. Moreover, the hydrogel preserves active exudate substances, giving a favorable environment for wound healing.

Drug release

Hydrogels have been extensively studied for drug release, and the rate of drug release can be increased by applying external stimulations such as pH, temperature, electric or magnetic fields. Among these stimulations, electrical stimulation is not only easy to generate and control, but also can avoid frequent invasive

surgery. The application of small current stimulation on human skin allows for remote application without the need for large and specialized devices. In addition, programmable drug release can allow repeated drug delivery and even wireless implants [46].

The solubility of ECHs is affected by electric current, a property that has attracted researchers to delve more deeply into its application in drug release. The drug release behavior of ECHs can be controlled by three forces: polymer affinity, ionic pressure, and rubber elasticity, which are collectively referred to as osmotic pressure. The osmotic pressure difference between the hydrogel and the solution is the driving force for drug release from electro-responsive hydrogels [47]. Under electrical stimulation, the balance of these forces is disrupted, causing the dissolution or deswelling of the hydrogel. When an electric field is applied to an electro-responsive hydrogel in an aqueous medium, the ions on the polymer chains shift to positions with opposite charges, resulting in an uneven ion distribution and increased osmotic pressure within the polymer. This leads the hydrogel to an environment where osmotic pressure is not in equilibrium, whereupon the volume of the hydrogel is transformed [48].

Injectable ECHs

Injectable ECHs, a promising candidate for topical drug and cellular delivery, are now attracting increasing interest. It allows for the release of drugs through electrical stimulation and can be adapted to the needs of individual patient. This type of drug delivery offers new avenues for chronic diseases that require frequent injections or precise doses of drugs.

In fact, there are several degrees of freedom to optimize the effective release and delivery of drugs. For example, Qu et al. developed a dual-response hydrogel for electrical stimulation and pH [49]. This hydrogel was made by mixing oxidized dextran (OD) with chitosan grafting polyaniline (CP) onto the amino group of chains. Among them, chitosan and polyaniline are inherently conductive and biocompatible, and exhibit good antimicrobial properties. The Schiff base bond formed between chitosan and dextran can be decomposed in an acidic environment, which make them good candidates for preparing injectable antimicrobial conductive hydrogels with on-demand drug release. Experiments have shown that the release rate of model drugs carried in CP/OD hydrogels is positively correlated with the voltage. Upon reduction of CP under electrical stimulation, the positive charge in the CP/OD hydrogel is reduced, and the absorbent molecules are expelled from the hydrogel, then the hydrogel shrinks. After the drug molecules are released from the CP/OD hydrogel through the electrochemical reduction process, the charged object with the drug moves to the electrode, realizing drug escape. Meanwhile, the positively charged groups in CP/OD hydrogels (including amino and acid-doped polyaniline) can break through the bacterial wall and release intracellular fluid through electrostatic adhesion between polymer chains and bacteria.

Ge et al. developed a temperature and electric field stimulated dual-responsive hydrogel consisting of a conducting polymer PPy and a thermosensitive polymer poly(lactic-co-glycolic acid) (PLGA)- polyethylene glycol -PLGA [50]. The erythromycin was polymerized with PPy nanoparticles and then uniformly distributed in the PLGA- polyethylene glycol -PLGA. The hydrogel is in a liquid state at room temperature and can be accurately injected subcutaneously. After entering the body, the hydrogels form rapidly into a gel state at body temperature and remain immobile at the target location. Under the stimulation of electric field, erythromycin is released from PPy nanoparticles by electrochemical oxidation. Simultaneously, PPy nanoparticles move towards the electrode driven by electric field, and drug release is achieved in this process.

Transdermal ECHs

Transdermal drug delivery is a technique that introduces drugs into the blood system through the skin. This method not only avoids gastric degradation and hepatic first-pass metabolism, but is also very helpful in popular chronic treatments, eliminating the phobias and pain associated with injections [51]. Transdermal drug delivery is expected to be an alternative to subcutaneous injections due to it is non-invasiveness. It allows prolonged release of the drug, can be applied by the patients themselves, and can improve patient compliance.

Sangsuriyonk et al. investigated the effect of cross-linker concentration and electric field strength on drug release [52]. They found that the higher the concentration of cross-linker, the higher the cross-linked density of the hydrogel, and the lower the mass ratio and release rate of carried drug. The higher the electric field strength, the higher the rate of drug release. The amount of drug release may be different at the cathode and the anode because different drugs and drug-carrying particles have different

electrical properties. Under the action of electric current, the negatively charged drug is released by reduction of the conducting polymer, while the positively charged drug is released by oxidation [46].

In other words, in the absence of electrical stimulation, the slow release of the drug is due to the concentration difference through free diffusion. On the other hand, the electrically driven drug release from the ECHs is directly related to the electric field that drives the migration of charged objects and the change in the total charge within the polymer upon reduction or oxidation [53]. Moreover, the rate of drug release is directly related to the strength of current and voltage, the stronger the current and voltage, the faster the rate of drug release. However, at very high voltages, the drug may be susceptible to degradation. Although it is possible to reduce the release time at high voltages, this approach is also undesirable because high voltages may cause adverse side effects in complex mediums in vivo.

Implantable biomaterials

ECHs has the potential to be implantable biomaterials in the field of medical engineering due to its biocompatible and degradable properties. It may be applied to neuronal tissue with intrinsically weak electrical fields to induce the release of specific neurotransmitters or combined with electrical activity of brain regions.

Ding et al. designed a hydrogel matrix consisting of silk fibroin compounded with aldehyde-hyaluronic acid, forming a dynamic network that delays or interrupts the hardening induced by silk fibroin chain β-sheet formation [54]. This results in a hydrogel matrix that mechanically aligns with biological tissues. Functionalized CNTs were incorporated to impart electrical conductivity to the hydrogel, which was experimentally demonstrated through in vivo implantation, showing effective stimulation and recording of neural signals. Han et al. developed an SG hydrogel by blending poly with graphene, creating a physically cross-linked structure that mimics human tissue mechanics [55]. This hydrogel exhibits strong adhesion to living tissue in humid environments and can be cleanly peeled off without leaving residue. In particular, the hydrogels do not change its volume in aqueous solution. In the other words, it does not swell, which helps the hydrogels to maintain stable when implanted in the body and do not cause stress to the surrounding tissues. It has been experimentally demonstrated that the hydrogel is electrically conductive and can be implanted in the sciatic nerve of rats to achieve neuromodulation by low-current electrical stimulation.

In recent years, cell scaffolds have also been increasingly investigated. ECHs have been shown to be used as biological scaffolds for culturing cardiomyocytes, providing a suitable microenvironment for cell attachment, spreading and proliferation, especially for cardiac tissue engineering, because the expression of linker proteins in cardiomyocytes can be enhanced in response to electrical stimulation [24, 56].

Conclusion

When talking about the prospective of conductive hydrogels, we cannot help but notice the great potential they show in several fields. Conductive hydrogels have unique physical, chemical and electrical properties that make them attractive in many applications like never before.

Their biocompatible and conductive properties make them ideal candidates for biosensors, smart drug delivery systems, and biological tissue engineering. Conductive hydrogels are capable of integrating sensor technologies to enable real-time monitoring of physiological parameters in living organisms, such as blood glucose levels, heart rate and muscle activity, to support personalized medicine. With the development of wearable technology, conductive hydrogels can be used to create flexible and comfortable biosensors and health monitoring devices. These devices can fit closely to the skin, enabling seamless monitoring of body status and transmission of data to mobile devices for analysis and diagnosis.

Meanwhile, conductive hydrogels, as a new type of material with multifunctionality and tunability, have also been widely researched and applied in the fields of environmental detection, energy reserve and conversion, showing a wide range of application prospects. However, integrating the properties of ECHs, their fabrication, and practical applications in humans still faces significant challenges.

ECHs have the physicochemical properties of hydrogels and also have excellent electrical conductivity. The structure and properties of ECHs can be imparted and adjusted by controlling the type of synthetic materials, composition ratio, and cross-linking method. Therefore, it is important to select appropriate polymer matrices and conductive mediums for ECHs in different application fields, so as to leverage

their synergistic benefits.

For example, introducing conductive mediums usually increases the crosslink density and strength of the hydrogels, but decreases the toughness and tensile properties. However, the conductivity of ECHs is positively correlated with the content of conductive mediums, and only trade-offs can be made between the two. In addition, although many self-healing, recoverable ECHs have been reported, it takes 30 minutes or more to recover 90% properties of the original hydrogel.

The synthetic production of ECHs is also a major challenge because the properties of hydrogels are easily affected by the composition ratio, production process, and external environment, so even a small change in the composition ratio may largely affect the cross-linking of hydrogels. The combination between the hydrophobic materials such as metal nanoparticles and the polymer matrices also requires special introduction methods to ensure uniform structure, which poses challenges for stable mass production of ECHs.

Research on ECHs in the field of biomedical engineering is developing rapidly, extending from external applications to internal use. However, numerous studies are confined to animal experiments and have progressed to clinical trials. The functionalities and potential side effects are yet to be thoroughly explored. Therefore, there is still a long way to go before ECHs can be used as bioelectronic therapeutic tools.

References

[1] Yang C, Suo Z. Hydrogel ionotronics. Nature Reviews Materials 2018;3(6):125-142.

[2] Yang J, Bai R, Chen B, et al. Hydrogel Adhesion: A Supramolecular Synergy of Chemistry, Topology, and Mechanics. Advanced Functional Materials 2020;30(2):1901693.

[3] Rong Q, Lei W, Liu M. Conductive Hydrogels as Smart Materials for Flexible Electronic Devices. Chemistry 2018;24(64):16930-16943.

[4] Shao H, Wu YC, Lin Z, et al. Nanoporous carbon for electrochemical capacitive energy storage. Chem Soc Rev 2020;49(10):3005-3039.

[5] He Y, Deng Z, Wang YJ, et al. Polysaccharide/Ti(3)C(2)T(x) MXene adhesive hydrogels with self-healing ability for multifunctional and sensitive sensors. Carbohydr Polym 2022;291:119572.

[6] Li X, Liu Z, Liang Y, et al. Chitosan-based double cross-linked ionic hydrogels as a strain and pressure sensor with broad strain-range and high sensitivity. J Mater Chem B 2022;10(18):3434-3443.

[7] Li S, Pan H, Wang Y, et al. Polyelectrolyte complex-based self-healing, fatigue-resistant and anti-freezing hydrogels as highly sensitive ionic skins. Journal of Materials Chemistry A 2020;8(7):3667-3675.

[8] Hu O, Chen G, Gu J, et al. A facile preparation method for anti-freezing, tough, transprent, conductive and thermoplastic poly(vinyl alcohol)/sodium alginate/glycerol organohydrogel electrolyte. Int J Biol Macromol 2020;164:2512-2523.

[9] Jiang P, Huang L, Wang J, et al. Carboxymethyl chitosan-based multifunctional hydrogels incorporated with photothermal therapy against drug-resistant bacterial wound infection. Int J Biol Macromol 2022;209(Pt A):452-463.

[10] Li S, Lu D, Tang J, et al. Electrical Stimulation Activates Fibroblasts through the Elevation of Intracellular Free Ca(+2+): Potential Mechanism of Pelvic Electrical Stimulation Therapy. Biomed Res Int 2019;2019:7387803.

[11] Ferrigno B, Bordett R, Duraisamy N, et al. Bioactive polymeric materials and electrical stimulation strategies for musculoskeletal tissue repair and regeneration. Bioact Mater 2020;5(3):468-485.

[12] Kenry, Liu B. Recent Advances in Biodegradable Conducting Polymers and Their Biomedical Applications. Biomacromolecules 2018;19(6):1783-1803.

[13] Dodero A, Donati I, Scarfì S, et al. Effect of sodium alginate molecular structure on electrospun membrane cell adhesion. Mater Sci Eng C Mater Biol Appl 2021;124:112067.

[14] Mo F, Huang Y, Li Q, et al. A Highly Stable and Durable Capacitive Strain Sensor Based on Dynamically Super-Tough Hydro/Organo-Gels. Advanced Functional Materials 2021;31(28):2010830.

[15] Liu H, Li M, Liu S, et al. Spatially modulated stiffness on hydrogels for soft and stretchable integrated electronics. Materials Horizons 2020;7(1):203-213.

[16] Guan S, Wang Y, Xie F, et al. Carboxymethyl Chitosan and Gelatin Hydrogel Scaffolds Incorporated with Conductive PEDOT Nanoparticles for Improved Neural Stem Cell Proliferation and Neuronal Differentiation. Molecules 2022;27(23).

[17] Gao Y, Jia F, Gao G. Ultra-thin, transparent, anti-freezing organohydrogel film responded to a wide range of humidity and temperature. Chemical Engineering Journal 2022;430:132919.

[18] Zhao D, Zhu Y, Cheng W, et al. Cellulose-Based Flexible Functional Materials for

Emerging Intelligent Electronics. Adv Mater 2021;33(28):e2000619.
[19] Wu Y, Liu J, Chen Z, et al. High Multi-Environmental Mechanical Stability and Adhesive Transparent Ionic Conductive Hydrogels Used as Smart Wearable Devices. Polymers (Basel) 2022;14(23).
[20] Bian Z, Li Y, Sun H, et al. Transparent, intrinsically stretchable cellulose nanofiber-mediated conductive hydrogel for strain and humidity sensing. Carbohydr Polym 2023;301(Pt A):120300.
[21] Au-Yong S, Firlak M, Draper ER, et al. Electrochemically Enhanced Delivery of Pemetrexed from Electroactive Hydrogels. Polymers (Basel) 2022;14(22).
[22] Cong J, Fan Z, Pan S, et al. Polyacrylamide/Chitosan-Based Conductive Double Network Hydrogels with Outstanding Electrical and Mechanical Performance at Low Temperatures. ACS Appl Mater Interfaces 2021;13(29):34942-34953.
[23] Wang W, Meng Q, Li Q, et al. Chitosan Derivatives and Their Application in Biomedicine. Int J Mol Sci 2020;21(2).
[24] Sharifisistani M, Khanmohammadi M, Badali E, et al. Hyaluronic acid/gelatin microcapsule functionalized with carbon nanotube through laccase-catalyzed crosslinking for fabrication of cardiac microtissue. J Biomed Mater Res A 2022;110(12):1866-1880.
[25] Zhou Y, Fei X, Tian J, et al. Biomass-based hydrogels with high ductility, self-adhesion and conductivity inspired by starch paste for strain sensing. Int J Biol Macromol 2022;222(Pt A):1211-1220.
[26] Lee Y, Song WJ, Sun JY. Hydrogel soft robotics. Materials Today Physics 2020;15:100258.
[27] Sun X, Qin Z, Ye L, et al. Carbon nanotubes reinforced hydrogel as flexible strain sensor with high stretchability and mechanically toughness. Chemical Engineering Journal 2020;382:122832.
[28] You L, Shi X, Cheng J, et al. Flexible porous Gelatin/Polypyrrole/Reduction graphene oxide organohydrogel for wearable electronics. J Colloid Interface Sci 2022;625:197-209.
[29] Gong X, Fu C, Alam N, et al. Preparation of Hemicellulose Nanoparticle-Containing Ionic Hydrogels with High Strength, Self-Healing, and UV Resistance and Their Applications as Strain Sensors and Asymmetric Pressure Sensors. Biomacromolecules 2022;23(6):2272-2279.
[30] Gao Y, Zhang Z, Ren X, et al. A hydrogel sensor driven by sodium carboxymethyl starch with synergistic enhancement of toughness and conductivity. J Mater Chem B 2022;10(30):5743-5752.
[31] Chen D, Zhao X, Wei X, et al. Ultrastretchable, Tough, Antifreezing, and Conductive Cellulose Hydrogel for Wearable Strain Sensor. ACS Appl Mater Interfaces 2020;12(47):53247-53256.
[32] Lu L, Huang Z, Li X, et al. A high-conductive, anti-freezing, antibacterial and anti-swelling starch-based physical hydrogel for multifunctional flexible wearable sensors. Int J Biol Macromol 2022;213:791-803.
[33] Chen D, Zhao X, Gao H, et al. High-Strength, Conductive, Antifouling, and Antibacterial Hydrogels for Wearable Strain Sensors. ACS Biomater Sci Eng 2022;8(6):2624-2635.
[34] Shen Z, Zhang Z, Zhang N, et al. High-Stretchability, Ultralow-Hysteresis ConductingPolymer Hydrogel Strain Sensors for Soft Machines. Adv Mater 2022;34(32):e2203650.
[35] Huang W-C, Ying R, Wang W, et al. A Macroporous Hydrogel Dressing with Enhanced Antibacterial and Anti-Inflammatory Capabilities for Accelerated Wound Healing. Advanced Functional Materials 2020;30(21):2000644.
[36] Yu J, Su H, Wei S, et al. Calcium content mediated hemostasis of calcium-modified oxidized microporous starch. J Biomater Sci Polym Ed 2018;29(14):1716-1728.
[37] Xu L, Liu Y, Tang L, et al. Preparation of Recombinant Human Collagen III Protein Hydrogels with Sustained Release of Extracellular Vesicles for Skin Wound Healing. Int J Mol Sci 2022;23(11).
[38] Rao KM, Narayanan KB, Uthappa UT, et al. Tissue Adhesive, Self-Healing, Biocompatible, Hemostasis, and Antibacterial Properties of Fungal-Derived Carboxymethyl Chitosan-Polydopamine Hydrogels. Pharmaceutics 2022;14(5).
[39] Liu J, Qu M, Wang C, et al. A Dual-Cross-Linked Hydrogel Patch for Promoting Diabetic Wound Healing. Small 2022;18(17):e2106172.
[40] Talikowska M, Fu X, Lisak G. Application of conducting polymers to wound care and skin tissue engineering: A review. Biosens Bioelectron 2019;135:50-63.
[41] Luo R, Dai J, Zhang J, et al. Accelerated Skin Wound Healing by Electrical Stimulation. Adv Healthc Mater 2021;10(16):e2100557.
[42] Wang C, Jiang X, Kim HJ, et al. Flexible patch with printable and antibacterial conductive hydrogel electrodes for accelerated wound healing. Biomaterials 2022;285:121479.
[43] Hajhosseini B, Longaker MT, Gurtner GC. Pressure Injury. Ann Surg 2020;271(4):671-679.
[44] Li Y, Fu R, Guan Y, et al. Piezoelectric Hydrogel for Prophylaxis and Early Treatment of Pressure Injuries/Pressure Ulcers. ACS Biomater Sci Eng 2022;8(7):3078-3086.
[45] Liu X, Tian S, Xu S, et al. A pressure-resistant zwitterionic skin sensor for domestic real-time

monitoring and pro-healing of pressure injury. Biosens Bioelectron 2022;214:114528.

[46] Samanta D, Hosseini-Nassab N, Zare RN. Electroresponsive nanoparticles for drug delivery on demand. Nanoscale 2016;8(17):9310-9317.

[47] Merino S, Martín C, Kostarelos K, et al. Nanocomposite Hydrogels: 3D Polymer-Nanoparticle Synergies for On-Demand Drug Delivery. ACS Nano 2015;9(5):4686-4697.

[48] Deng Z, Hu T, Lei Q, et al. Stimuli-Responsive Conductive Nanocomposite Hydrogels with High Stretchability, Self-Healing, Adhesiveness, and 3D Printability for Human Motion Sensing. ACS Appl Mater Interfaces 2019;11(7):6796-6808.

[49] Qu J, Zhao X, Ma PX, et al. Injectable antibacterial conductive hydrogels with dual response to an electric field and pH for localized "smart" drug release. Acta Biomater 2018;72:55-69.

[50] Ge J, Neofytou E, Cahill TJ 3rd, et al. Drug release from electric-field-responsive nanoparticles. ACS Nano 2012;6(1):227-233.

[51] Sabbagh F, Kim BS. Recent advances in polymeric transdermal drug delivery systems. J Controll Release 2022;341:132-146.

[52] Sangsuriyonk K, Paradee N, Sirivat A. Electrically controlled release of anticancer drug 5-fluorouracil from carboxymethyl cellulose hydrogels. Int J Biol Macromol 2020;165(Pt A):865-873.

[53] Pérez-Martínez CJ, Morales Chávez SD, del Castillo-Castro T, et al. Electroconductive nanocomposite hydrogel for pulsatile drug release. Reactive and Functional Polymers 2016;100:12-17.

[54] Ding J, Chen Z, Liu X, et al. A mechanically adaptive hydrogel neural interface based on silk fibroin for high-efficiency neural activity recording. Mater Horiz 2022;9(8):2215-2225.

[55] Han IK, Song KI, Jung SM, et al. Electroconductive, Adhesive, Non-Swelling, and Viscoelastic Hydrogels for Bioelectronics. Adv Mater 2023;35(4):e2203431.

[56] Alvarez-Primo F, Anil Kumar S, Manciu FS, et al. Fabrication of Surfactant-Dispersed HiPco Single-Walled Carbon Nanotube-Based Alginate Hydrogel Composites as Cellular Products. Int J Mol Sci 2019;20(19).

Progress in Medical Devices
eISSN: 2957-5478

Zentime

Research Article https://doi.org/10.61189/780603itrglq

Precision cooling radiofrequency ablation under tumor boundary temperature control

Ruiyan Qian, Haixiao Lin, Dandan Gu, Difang Liu, Haitao Yao, Danni Rui, Yu Zhou

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

Declaration of conflict of interest: None.

Received June 8, 2023; Accepted August 28, 2023; Published September 30, 2023

Highlights

- The outcomes derived from simulation and ex vivo experiments exhibit precise control over the range of radiofrequency ablation when employing boundary temperature control.
- It was observed in this study that a single ablation procedure can achieve an ablation diameter of up to 30 mm.
- The regular index and elliptic index of the radiofrequency ablation area tend to be a spherical shape, which holds potential benefits for the radiofrequency ablation surgeries.

Abstract

Radiofrequency ablation (RFA) represents a convenient, minimally invasive, and cost-effective approach for the treatment of small liver cancers measuring less than 3 cm in diameter. Nonetheless, the existing RFA techniques encounter challenges in precisely controlling the extent of ablation and the risk of overheating, which can lead to damage on the surrounding tissues. The ability to control the ablation area also plays a crucial role in the success of RFA procedures. To address these issues, we introduce a novel method that utilizes tumor boundary temperature monitoring to achieve precise control over the RFA process. Through the utilization of COMSOL Multiphysics simulation software, the proposed method was verified by comprehensive simulation and modeling and its efficacy in achieving regional control was testified. Subsequently, ex vivo experiments were conducted employing a custom-designed cooling RFA instrument. The experimental results demonstrated that the RFA controlled by boundary temperature yielded an ablation area with a diameter approaching 30 mm. In comparison to the standard spherical solidification zone (characterized by both the ellipticity and regularity indices of 1), the cooling RFA discoloration zone, under boundary temperature control, exhibited a maximum deviation of 7% in the ellipticity index and 17% in the regularity index. The average ellipticity index was determined to be 0.96, while the average regularity index was 0.86. Collectively, these findings underscore the capability of the proposed temperature-controlled cooling RFA method to attain precise control over the ablation area, contributing to comprehensive ablation of tumor tissue within the target region.

Keywords: Cooling radiofrequency ablation, simulation, temperature controlled, regional control

Introduction

Cancer is one of the leading causes of death globally [1]. To effectively improve patient well-being, cancer treatment technologies have constantly been advancing. Radiofrequency ablation (RFA) has emerged as a safe and effective treatment method that plays a vital role in combating cancer [2]. RFA utilizes radiofrequency energy to generate thermal effects in tumor tissues, resulting in coagulation necrosis and the successful treatment of diseased tissues [3]. It is suggested that when the tumor cells are exposed to temperatures exceeding 42 °C, their biochemical reactions and metabolic processes become disrupted. Cell apoptosis is observed within 8 minutes at 46 °C, and in just 2 minutes at 51 °C. Temperature surpassing 60 °C can prompt rapid protein denaturation and necrosis within the cells, while temperature exceeding 100 °C can result in cell water boiling, leading to tissue gasification or carbonization [4, 5]. To achieve complete coagulation and necrosis of tumor tissues, the traditional multi-polar RFA technique employed in the treatment of small- and medium-sized liver cancers often leads to tissue carbonization, reducing tissue conductivity and hindering

Address correspondence to: Yu Zhou, School of Health Sciences and Engineering, University of Shanghai for Science and Technology, No.516 Jungong Road, Shanghai 200093, China. Tel: 18021042556, E-mail: zhouyu_working@163.com.

the transfer of radiofrequency energy to distant lesions. The use of unipolar RFA requires multiple ablations, which faces technical challenges, such as the determination of the optimal number of ablations and the precise placement of the electrodes [6]. While expanding the ablation range is important, precise control of the ablation range remains a significant challenge in current radiofrequency therapy.

In response to these challenges, cooling RFA (CRFA) has emerged as a promising technology. The safety of CRFA is comparable to that of traditional RFA [7]. Nevertheless, in comparison to conventional approaches, CRFA confers a multitude of advantages [8]. By effectively reducing the contact temperature between electrodes and tissues, CRFA mitigates the risk of tissue carbonization or gasification, expands the ablation range, and improves the therapeutic effect on liver cancer [9]. CRFA introduces cooling functionality to the existing RFA technology by incorporating a cooling device into the electrode needle. Some attempts have been made to expand the ablation range by regulating CRFA output through impedance control. However, these methods often result in irregular ablation areas and pose risks to healthy tissues [10].

In order to expand the ablation range, reduce treatment frequency, and optimize patients' treatment experience, this study introduces an independently developed CRFA system. Meanwhile, in order to enhance the precision of ablation range in clinical practice, mitigate the risk of injury to healthy tissues, and maximize the safety of patients, a novel tumor boundary temperature-controlled ablation algorithm is proposed. The power output of RFA is controlled based on the tumor boundary temperature, which is monitored in real-time to achieve more regular and precise ablation range. This study employed the multi-physics simulation software COMSOL to model the physical phenomena of RFA under boundary temperature control. Furthermore, ex vivo experiments were conducted using a self-developed CRFA device to analyze and evaluate the RFA outcomes under boundary temperature control, so as to verify the effectiveness of the algorithm.

Material and Methods

CRFA mathematical model

The main idea of supplying electric current in the tumor area is to convert the electrical energy into thermal energy according to the law of conservation of energy. The primary mechanism of energy propagation takes place through electrical conduction, which can be mathematically characterized as a coupled phenomenon involving quasi-static conduction and thermal conduction. The mathematical representation of the governing equation for the current interface can be expressed as Equation 4.1.

$$-\nabla(\sigma\nabla V - J^e) = Q_j \tag{4.1}$$

where σ is the electrical conductivity (S/m), J^e is an externally generated current density (A/m^2), and Q_j is the current source (A/m^3). In the CRFA model, there are no other applied current sources other than the electrodes. J^e and Q_j are zero, substituting these values to simplify the electric field distribution in the CRFA model, as shown in Equation 4.2 [11].

$$-\nabla(\sigma\nabla V) = 0 \tag{4.2}$$

The bio-heat transfer equation employed is the classical Pennes bio-heat transfer equation [12]. This heat transfer equation is often used in biomedical engineering to simulate hyperthermia for cancer treatment, which is described as follow:

$$\rho c\frac{\partial T}{\partial t} = \nabla\cdot(k\nabla T) - \rho_b c_b \omega_b (T - T_b) + Q_m + Q_{hs} \tag{4.3}$$

where ρ and ρ_b are the density of tissue and blood, respectively (kg/ m^3), c and c_b are the specific heat capacity of tissue and blood, respectively (J/kg · °C), T and T_b are the temperatures of tissue and blood (°C), ω_b is the volume blood perfusion rate (s^{-1}), Q_m is the metabolic heat generation rate of tissue metabolic processes (W/m^3), and Q_{hs} is the heat generated by the electrodes of the RFA device (W/m^3).

CRFA incorporates a cooling water circulation system to cool the electrode needle, thereby engaging both solid and fluid heat transfer phenomena. The fluid heat transfer component encompasses the convection and heat conduction processes within the cooling water, while the solid heat transfer pertains to the heat transfer mechanism occurring within the electrode needle tube. The heat exchange between the fluid cooling water and the solid electrode needle is mathematically modeled using a fluid-solid coupling equation:

$$\rho c\frac{\partial T}{\partial t} + \rho c_p u\nabla T + \nabla q = Q \tag{4.4}$$

$$q = -k\nabla T \tag{4.5}$$

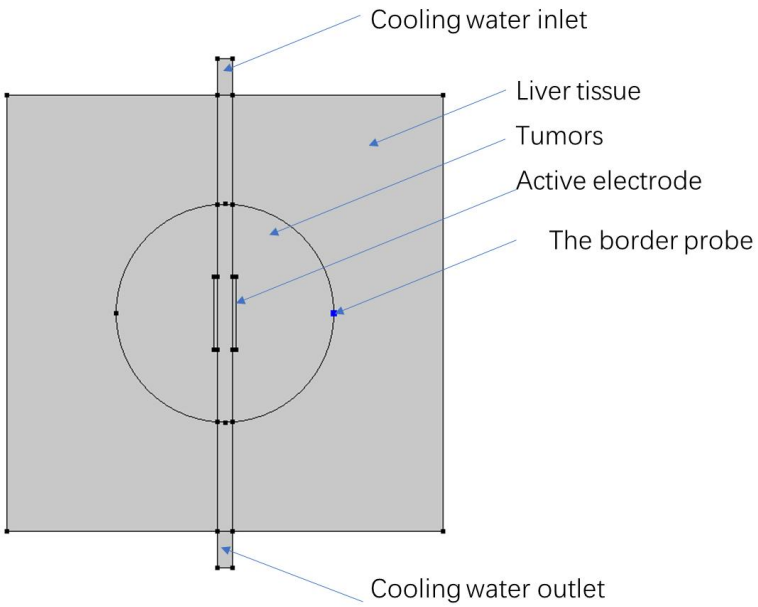


Figure 1. 2D geometry of the CRFA. CRFA, cooling radiofrequency ablation.

where $\rho c\frac{\partial T}{\partial t}$ is the transient term, $\rho c_p u\nabla T$ is the convection term, ρ is the electrode needle density (g/cm^3), c_p is the specific heat capacity of the substance (J/kg · °C), T is the temperature field (°C), u is theflow rate field of cooling water (m/s), ∇q is the diffusion term, Q is the heat source term (W), k is the thermal conductivity, and ∇T is the temperature difference (°C).

At each time step, the cumulative damage integral is computed using the well-established Arrhenius equation:

$$\Omega(t)=\ln\frac{c(0)}{c(t)}=A\int_0^T e^{-\Delta E/RT(t)}dt \qquad (4.6)$$

where c(t) is the concentration of living cells, R is the universal gas constant, A is a "frequency" factor for the kinetic expression (s^{-1}), and ΔE is the activation energy for the irreversible damage reaction (J/mol). During the finite element simulation of tissue damage, the critical threshold of the damage integral, denoted as Ω=1, corresponds to a 63% probability of cellular demise at a specific spatial location. Consequently, when the level of damage surpasses the 63% threshold, cell damage within the tissue reaches a critical threshold, resulting in the coagulation of the tissue and cessation of tissue perfusion [13].

The mathematical representation of the temperature dependency of electrical and thermal conductivity frequently employed in computational modeling of RFA can be approximated by a growth rate of +2% below the temperature threshold of 100 °C. However, once the temperature exceeds 100 °C, the effect of dehydration on the conductivity can be expressed by employing a sudden decrease of two or four orders of magnitude, while maintaining a constant value for thermal conductivity [14]. Consequently, in the context of modeling CRFA, a linear rate of change of -2%/°C for impedance was adopted, as outlined in Conway's investigation, to establish the mathematical function of impedance, denoted as R (T) [15].

$$R(T)=\begin{cases}R_0[1-(T-T_0)\times 2\%] & T<100℃\\ 1700 & T>100℃\end{cases} \qquad (4.6)$$

Establishment of the CRFA geometric model

In this study, the application of CRFA in the treatment of liver tumors involved the insertion of an active electrode into the tumor tissue for current delivery. Numerical simulations based on the theoretical model of the RFA process were developed to depict the corresponding geometry. As depicted in **Figure 1**, the model consisted of a circular region representing the liver tumor tissue, where the tumor size could be adjusted by controlling its diameter. The simulated RFA targeted tumors with diameters of 30 mm or less. The square area symbolized healthy liver tissue and possessed dimensions of 60 mm×60 mm. The electrode needle tube had a diameter of 2 mm and a length of 70 mm. The central part of the electrode, connected to the tumor center, functioned as the active electrode, while the remaining portion was electrically insulated. The active electrode itself had a diameter of 2 mm and a length of 10 mm. To monitor the temperature of the tumor boundary and the extent of necrotic tissue, a boundary probe was incorporated.

Considering the typical round or oval shape of tumor tissues, the ground pad in clinical practice is usually positioned at a considerable distance from the active electrode to ensure a uniform current distribution from the active electrode outward. As the ex vivo experiments exhibited axial symmetry in terms of geometry,

Table 1. Values of physical constants

Tissue Density	ρ=1060（kg/m^3）
Thermal conductivity	0.5020（W/K·m）
Tissue heat capacity	c=3600（J/kg·K）
Tissue electrical conductivity	σ=0.148(S/m)
Blood density	ρ_b= 1000(kg/m^3)
Blood heat capacity	c_b=4180（J/kg·K）
Perfusion coefficient	ω_b=6.4e^{-3}(s^{-1})

the CRFA simulation was simplified as a 2D model to significantly reduce the computational burden. The overall model closely resembled the insertion of a hollow electrode needle into the central region of liver tissue, with only the active electrode being electrically active as the current conductor while cooling water flowed through the interior of the electrode needle conduit. The top of the electrode needle served as the cooling water inlet, while the bottom functioned as the outlet.

Regarding the initial conditions of the model, the electrically insulating boundary conditions were applied to the nonconducting section of the electrode needle, and the four sides of the liver tissue were set to the ground state. The initial voltage of the entire simulation model was set to 0 V, while the cooling water temperature was maintained constant at 20 °C. Since the room temperature during the ex vivo experiments was approximately 23 °C, the initial tissue temperature was set to 23 °C.

The values of the individual physical constants are listed in **Table 1** [16].

The simulation was performed using COMSOL Multiphysics software. The radiofrequency output was controlled by monitoring the temperature at the tumor boundary using a probe. The observed ablation results were analyzed and evaluated. A bovine liver was used as the experimental subject, and a negative plate were affixed to the lateral side of the bovine liver, while a positioning plate was positioned directly above the bovine liver. The electrode needle was carefully inserted into the bovine liver via the central aperture of the positioning plate. Four temperature-measuring needles were introduced through the small aperture on the positioning plate, maintaining a distance of 15 mm from the central aperture. The thermocouple-based temperature measuring points were situated at the same depth as the exposed segment of the electrode needle. The radiofrequency energy generator operated in a constant power mode, and the experimental setup employed a cooling needle electrode measuring 2 mm in diameter and 250 mm in length, with a 15 mm bare leakage and uninsulated segment. The output duration was set to 900 s, and a total of five experimental iterations were conducted to document the size of the ablation area and the temperature within the ablation zone. The experimental protocol entailed activating the cooling apparatus. Once the cooling fluid reached the target temperature, the ex vivo CRFA procedure was initiated. Throughout the experiment, adaptive power output was realized by the control algorithm.

Development of the CRFA system

The design of the CRFA device necessitates adherence to the control algorithm's parameters, which involve temperature detection of biological tissue and adjustment of the output power based on impedance information. Therefore, the device must incorporate the corresponding circuitry to accurately detect temperature and impedance information, so as to ensure minimal errors and compliance with industry standards. Additionally, stable closed-loop control is essential, with impedance and temperature information serving as feedback variables to regulate the magnitude of radiofrequency energy and ensure controllable output power.

Following the design requirements, the hardware circuitry of the CRFA device primarily consists of the human-computer interaction interface, the main control module, the power supply module, the radiofrequency power amplifier module, and the interface module. External components include the electrode needle, temperature measuring needle, and negative electrode plate, as illustrated in **Figure 2A**. The human-machine interface comprises a programmable touch capacitive screen capable of displaying the operational status, by which users can select the output mode and control the device's switch. The power supply module furnishes power to the entire hardware system by initially converting 220 V alternating current to 48 V direct current (DC), which is then adjustable from 0-70 V DC to provide energy for the radiofrequency power amplifier. The main control module employs the STM32F103RCT6 MCU to oversee intelligent algorithm operations, power module control, output parameter control, temperature and impedance information detection, and other functions. Simultaneously, the 48 V DC output of power module can be further converted to 12 V, 3.3 V, and 8 V DC to accommodate the equipment control and detection sections. The radiofrequency power amplifier module incorporates a symmetric class E power amplifier that transforms the adjust-

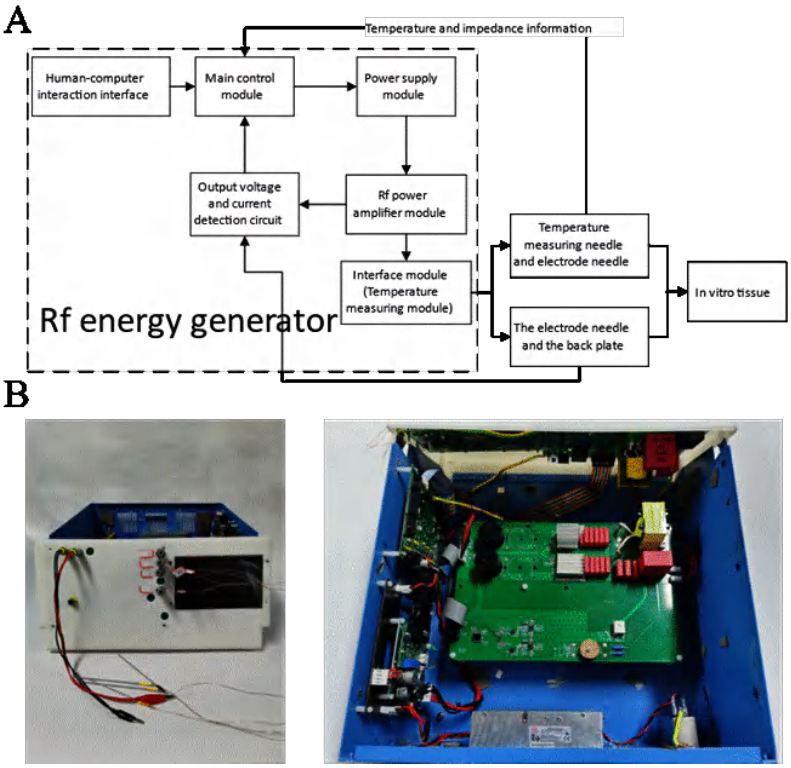


Figure 2. CRFA device. (A) Structure block diagram; (B) Appearance of the CRFA device. RF, radiofrequency; CRFA, cooling radiofrequency ablation.

able DC output from the power supply module into alternating current radiofrequency energy through a resonant circuit. The voltage and current detection circuit captures the voltage and current outputs of the radiofrequency power amplifier and converts them into low-voltage signals, proportionally feeding them back to the main control module. The interface module includes a temperature measuring circuit, a negative plate detection circuit, and an isolation transformer. The temperature measuring circuit amplifies the analog signals acquired by the temperature measuring needle and feeds them back to the main control module, while the negative plate detection circuit verifies the proper alignment of the negative plate with the human body. The isolation transformer ensures electrical isolation between the application and intermediate parts and is connected to the radiofrequency power amplifier, with its secondary side linked to the electrode needle and negative plate. During operation, radiofrequency current flows between the negative plate and the electrode needle, and real-time tissue temperature is collected using a K-type thermocouple. The overall appearance diagram and internal structure of the device are presented in **Figure 2B**.

Figure 3A presents the structural diagram of the cooling equipment, delineating the constituent elements such as the water pump, the refrigeration device, the radiator water discharge, and the control module. These components collaboratively facilitate the operation of the cooling system, enabling the circulation and cooling of the coolant. The primary objective of this cooling equipment is to effectively mitigate the elevated temperature of the electrode needle by dissipating heat through the coolant circulation process. **Figure 3B** shows the actual appearance of the water-cooling device.

Algorithm design and mechanism

The cooling apparatus facilitates the infusion of coolant into the electrode needle, thereby establishing a circulating mechanism for cooling the radiofrequency electrode. This mechanism effectively delays the onset of overheating and tissue carbonization. To further augment the cooling efficacy while ensuring uniform ablation, an enhanced output mode of radiofrequency energy was implemented, incorporating a duty cycle lasting 15 minutes. Within each cycle, the delivery of radiofrequency energy was sustained for 50% of the time, with the remaining 50% of the time allocated for output cessation. Additionally, the utilization of backflowing tissue fluid allows for the propagation of radiofrequency energy to more distal regions from the electrode [17].

The regional control algorithm plays a crucial role in governing the activation and power output of the system, responding to variations in tissue impedance and temperature. This dynamic regulation helps to prevent eschar formation near the electrode needle while ensuring compliance with the ablation area standard. The dimensions of the ablation zone are primarily dictated by the spatial separation between the temperature-measuring needle and the electrode needle.

Figure 4 provides a visual representation in the form of a flowchart, illustrating the process of power adjustment during radiofrequency en-

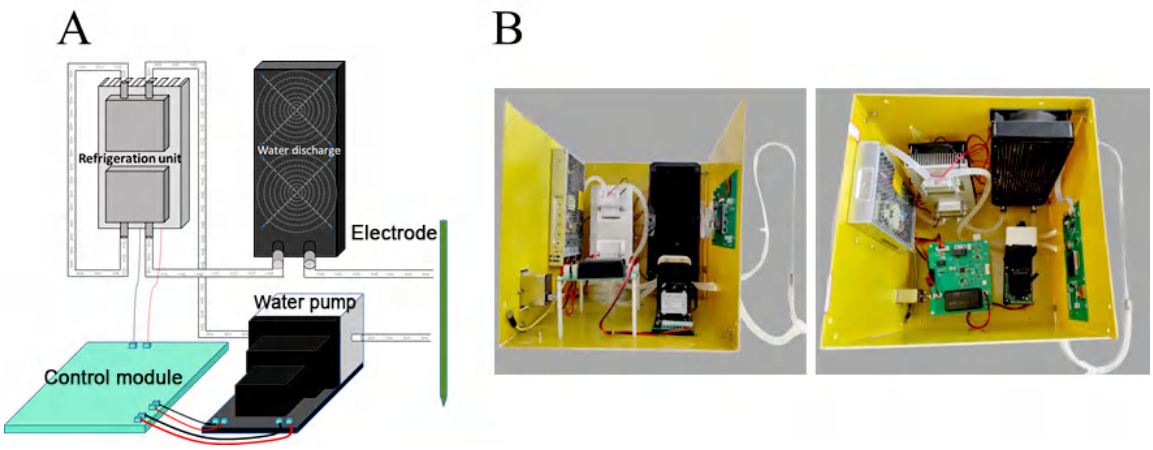


Figure 3. Schematic diagram of cooling equipment. (A) Structure; (B) Appearance of the cooling equipment.

ergy output based on impedance information. The initial cellular composition comprises the cell membrane and intracellular fluid, while extracellular fluid and intercellular stroma encompass the cellular structure. Upon introduction of an excitation current into the biological tissue, the current follows various pathways, bypassing the cells. Predominantly, the current flows within the extracellular fluid, with a fraction penetrating the cell membrane and permeating the intracellular fluid [18]. Ablation is conducted at higher power levels, resulting in cell destruction as the temperature rises, accompanied by fluid outflow from the tissue. Consequently, the tissue impedance progressively diminishes from a relatively high value to a minimum value, stabilizing for a certain duration. During this stable period, the administration of high-power energy does not induce tissue overheating. Excessive ablation boundary temperatures can inflict unwarranted thermal damage to neighboring tissues or blood vessels. Hence, the algorithm continuously monitors the real-time ablation boundary temperature recorded by the temperature-measuring needle. If the temperature exceeds 60 °C, the device halts the output and awaits tissue to cooldown. Simultaneously, when the device output is stopped, the reflux of tissue fluid reduces impedance near the electrode needle, facilitating the formation of a larger ablation area during subsequent heating and ensuring control over the ablation area size. Radiofrequency energy output continues when the boundary temperature is below 51 °C.

Methods for evaluating the effect of CRFA ex vivo

RFA involves the utilization of high-frequency alternating current, typically within the frequency range of 375-500 kHz. The radiofrequency applicators, referred to as electrodes, can be categorized as monopolar electrodes with a single active electrode applicator, where current dissipation occurs at one or more negative plates [19]. By delivering this energy through the electrodes into the patient's tissues, heating and subsequent destruction of the targeted tissue are achieved through variations in tissue conductivity. Radiofrequency electrodes generate an alternating electric field, leading to the oscillation of ions within the targeted tissue and the generation of heat. Consequently, cellular protein denaturation occurs, resulting in the loss of cellular viability [20, 21].

When monopolar RFA is utilized for tumor ablation in organs such as the liver, kidney, heart, and prostate, the evaluation of the CRFA takes place after the diseased tissue being coagulated. Complete and uniform destruction of cellular structures is observed within the pale region of the ablated lesion tissue, while the surrounding normal tissue exhibits a dark red appearance [22]. Several studies have summarized evaluation indices for assessing the efficacy of RFA ex vivo, including the ellipticity index (EI) and the regularity index (RI).

$$EI = \frac{2AD}{TD_{\max} + TD_{\min}} \tag{4.8}$$

$$RI = \frac{R_{\min}}{R_{\max}} \tag{4.9}$$

These parameters are measured within the ablated tissue, where an EI and RI value of 1 represents the standard sphere. The axial plane is defined as the plane aligned with the electrode axis (**Figure 5B**), while the transverse plane is perpendicular to the electrode and possesses the largest transverse diameter (**Figure 5C**). Maximum transverse diameter corresponds to

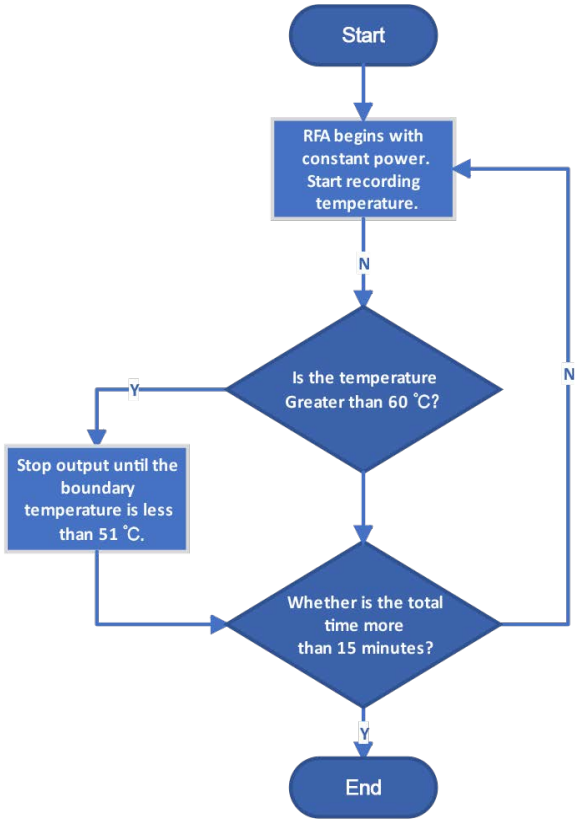


Figure 4. Flow chart of the control algorithm. RFA, radiofrequency ablation.

the maximum distance between two opposing edges, whereas minimum transverse diameter signifies the minimum distance between the edges along a relative unit in millimeters. Half of the maximum transverse diameter line is used as an intersection for measurement. AD denotes the distance between the proximal and distal edges of the coagulation zone within the axial plane along the electrode axis. Maximum radius represents the maximum distance between the edge of the solidified zone and the electrode axis within the horizontal plane. Conversely, minimum radius corresponds to the minimum distance between the edge of the solidified zone and the electrode axis within the horizontal plane (**Figure 5D**) [23].

Statistical analysis

For statistical analysis, we utilized deviation analysis simulation to compare the tissue damage diameter differences across various models. Deviation, mean value, and variance were applied to analyze the EI and RI values of the ablation range under temperature control. Deviation assessed the discrepancies between the EI and RI values and the standard values, while mean value gauged the central tendency of the RI and EI variance quantified the dispersion of EI and RI index around the mean.

Results

The simulation results of CRFA and temperature-controlled CRFA were compared

To preliminarily verify the feasibility of the temperature-controlled mode, a series of simulations were conducted to evaluate the experimental effects of CRFA in different ablation modes. Two modes of CRFA, namely cooling constant power RFA and boundary temperature-controlled CRFA, were simulated to assess their ablation effects. In the constant power mode, an active electrode length of 10 mm was employed with a continuous power output of 900 s, while the tumor diameter was set at 30 mm. To explore the feasibility of temperature-controlled CRFA, temperature control was incorporated into the CRFA simulation. The geometry of the temperature-controlled CRFA simulation model mirrored that of the CRFA, and the heat source electrodes played a consistent role.

During the simulation process, the algorithm monitored the temperature at the boundary of the tumor. When the boundary temperature exceeded 60 °C, the energy output from the electrode was paused until the temperature dropped below 51 °C, after which the energy output was resumed.

$$ONOFF = \begin{cases} 1 & T < 51℃ \\ 0 & T > 60℃ \end{cases} \quad (1.1)$$

$$V = V(T) \times ONOFF \quad (1.2)$$

During the simulation procedure, the discrete state ON/OFF of the voltage was determined by monitoring the temperature at the boundary of the tumor. The voltage output of the active electrode, denoted as V(T), represented the heat source for the actual ablation process. In the CRFA simulation, the formula $V(T) = \sqrt{(P \times R(T))}$ was employed when the power was set to a specific value. The active electrode was configured as a temperature-dependent voltage source. The ablation outcomes are presented in **Figure 6**.

Experimental results of temperature-controlled CRFA

Figure 7 depicts the ex vivo bovine liver experiment incorporating the temperature control algorithm. Following ablation, the bovine liver specimens were dissected along the electrode

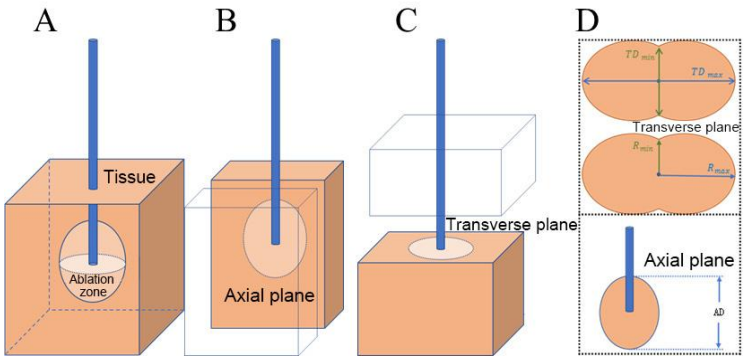


Figure 5. Schematic diagram of the ablation range. (A) Overall tissue; (B) Axial plane; (C) Transverse plane; (D) Parameters of the ex vivo experiment. TD_{max}, maximum transverse diameter; TD_{min}, minimum transverse diameter; AD, axial diameter.

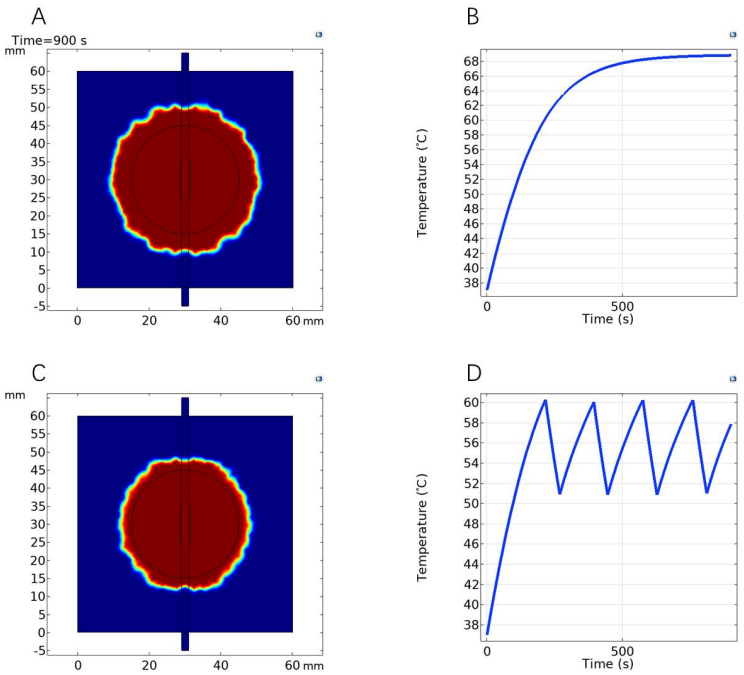


Figure 6. Simulation results of CRFA and boundary temperature-controlled CRFA. (A) Proportion of necrotic tissue in CRFA; (B) Tumor boundary temperature in CRFA; (C) Proportion of necrotic tissue in boundary temperature-controlled CRFA; (D) Tumor boundary temperature in boundary temperature-controlled CRFA. CRFA, cooling radiofrequency ablation.

needle axis, and the discolored region resulting from the ablation was quantified using a vernier caliper.

A vernier caliper was used to measure the results temperature-controlled CRFA, including the maximum transverse diameter, the minimum transverse diameter, the axial diameter, the maximum radius, and the minimum radius. **Table 2** shows the measured parameters in the discolored area in temperature-controlled CRFA.

Discussion

This study comparatively analyzed the simulation results between CRFA and temperature-controlled CRFA. During the simulation, a consistent output power of 35W was employed across all cases. **Figure 6** presents the simulation results, revealing that both CRFA and temperature-controlled CRFA achieved similar uniformity and regularity in the ablation range, surpassing the tumor boundary. The simulation results demonstrated that CRFA with the cumulative damage integral of 1 yielded ablation diameters of 43 mm, while under temperature control, the diameter reduced to 33 mm. Therefore, the extent of tissue damage was significantly reduced in temperature-controlled CRFA compared to CRFA. Under the temperature control algorithm, the temperature at the tumor tissue boundary fluctuated within the range of 50 °C to 60 °C, thereby preventing over ablation caused by excessive heating. This indicates that the theoretical control of the CRFA range under boundary temperature control is able to achieve favorable outcomes, which provides valuable guidance for ex vivo experiments.

Table 2 presents the data of various parameters related to the discolored area from temperature-controlled CRFA. Through verification via ex vivo experiments, it was observed that the diameter size of the ablation region was approximately 30 mm. In comparison to the standard spherical coagulation zone (with EI and RI both equal to 1), the maximum deviation of the EI for the traditional constant power CRFA discolored area was 7%, while the maximum deviation of the RI was 17%. The average EI was measured as 0.96, and the average RI was 0.86. The RI of the ablation results in this study is close to that of other CRFA ex vivo experiments, and the EI of the ablation range in this study is closer to the standard value than that of other CRFA, which indicates the superior performance of temperature controlled CRFA in EI [24-26]. This disparity can be attributed to the fact that the electrode surface temperature was consistently maintained below 20 °C, effectively reducing

Figure 7. Experimental plots of five groups of boundary temperature-controlled ex vivo bovine livers. (A) The axial plane of the first ablation region; (B) The transverse plane of the first ablation area; (C) The axial plane of the second ablation area; (D) The transverse plane of the second ablation area; (E) The axial plane of the third ablation region; (F) The transverse plane of the third ablation area; (G) The axial plane of the fourth ablation area; (H) The transverse plane of the fourth ablation area; (I) The axial plane of the fifth ablation region; (J) The transverse plane of the fifth ablation area.

the temperature of the surrounding tissue and mitigating the risk of excessive heating. The active cooling of the ablation electrode facilitates the delivery of higher energy to the target tissue, resulting in more effective ablation [27, 28]. The boundary temperature control regulates the output of radiofrequency energy, ensuring that the temperature at the boundary

Table 2. Parameters of the discoloration area during boundary temperature-controlled CRFA

Group	TD_{max}/mm	TD_{min}/mm	AD/mm	R_{max}/mm	R_{min}/mm	EI	RI
1	34.98	30.14	31.10	16.00	13.98	0.96	0.87
2	32.02	30.12	30.62	17.10	14.90	0.99	0.87
3	31.45	30.28	30.12	16.85	14.52	0.98	0.86
4	30.71	31.97	29.20	16.50	14.00	0.93	0.85
5	31.08	32.82	30.96	18.00	15.01	0.97	0.83

Notes: CRFA, cooling radiofrequency ablation; Td_{max}, maximum transverse diameter; Td_{min}, minimum transverse diameter; AD, axial diameter; R_{max}, maximum radius; R_{min}, minimum radius; EI, ellipticity index; RI, regularity index.

does not harm the healthy tissues surrounding it.

To conclude, in this boundary temperature-controlled CRFA, the maximum deviation between the EI and RI was minimal, with the mean value approaching a unity. Additionally, the variance values of the EI and RI were also small. These findings indicate that the self-designed CRFA instrument, coupled with the developed algorithm, enables stable control of CRFA, yielding spherical ablation patterns with a high degree of regularity. This allows doctors to better predict and plan the ablation range in clinical applications, and at the same time, the boundary temperature of the ablation area meets the condition of cell necrosis, which can ensure cell necrosis in the ablation area. The limitations of this study mainly lie in the inability to entirely replicate real-life conditions and dynamic changes, as well as the inability to obtain real-time biofeedback information, which may limit its accuracy and reliability in clinical application. Therefore, the results of ex vivo trials and clinical practice need to be considered comprehensively when evaluating RFA techniques. Despite the absence of clinical data validation analysis in this study, the ex vivo experiments exhibited promising outcomes. These findings offer a prospective avenue for further exploration and advancement of RFA in future research endeavors.

Conclusion

In this study, we initially established a simulation model using an ex vivo bovine liver model to preliminarily validate the efficacy of the boundary temperature control algorithm, which was designed to ensure the therapeutic effect while minimize unnecessary thermal damage. Subsequently, we successfully constructed the hardware platform for cooling the CRFA device following the design requirements. Finally, we conducted an ex vivo experiment to validate the results under the designed temperature control algorithm. Following the verification of the ex vivo experiment, the ablation region exhibited a diameter of approximately 30 mm, and the resulting spherical damage was stable and well-controlled. The high consistency observed between the ex vivo experiment and the simulation model results indicates the foreseeable capability of the simulation model. Furthermore, boundary temperature-controlled CRFA demonstrates improved potential for precise ablation, which establishes a foundation for further research and clinical applications, thus offering significant prospects in the field of tumor treatment.

References

[1] Cao W, Chen HD, Yu YW, et al. Changing profiles of cancer burden worldwide and in China: a secondary analysis of the global cancer statistics 2020. Chin Med J (Engl) 2021;134(7):783-791.

[2] Tatli S, Tapan U, Morrison PR, et al. Radiofrequency ablation: technique and clinical applications. Diagn Interv Radiol 2012;18(5):508-516.

[3] Habibi M, Berger RD, Calkins H. Radiofrequency ablation: technological trends, challenges, and opportunities. Europace 2021;23(4):511-519.

[4] Ni Y, Mulier S, Miao Y, et al. A review of the general aspects of radiofrequency ablation. Abdom Imaging 2005;30(4):381-400.

[5] Nath S, Lynch C, Whayne JG, et al. Cellular electrophysiological effects of hyperthermia on isolated guinea pig papillary muscle. Implications for catheter ablation. Circulation 1993;88(4 Pt 1):1826-1831.

[6] Chen MH, Yang W, Yan K, et al. Large liver tumors: protocol for radiofrequency ablation and its clinical application in 110 patients--mathematic model, overlapping mode, and electrode placement process. Radiology 2004;232(1):260-271.

[7] Taneja T, Huang S. Irrigated and Cooled-Tip Radiofrequency Catheter Ablation. 2019. p. 27-38.e23.

[8] Taneja T, Shoei K, Stephen H. Irrigated and Cooled-Tip Radiofrequency Catheter Ablation. Catheter Ablation Card Arrhythmias 2019.

[9] Cedeno DL, Vallejo A, Kelley CA, et al. Comparisons of Lesion Volumes and Shapes

Produced by a Radiofrequency System with a Cooled, a Protruding, or a Monopolar Probe. Pain Physician 2017;20(6):e915-e922.

[10] Zheng DP, Zhu MR, Liu WB, et al. Cool-tip Radiofrequency Ablation Therapy Instrument Based on Impedance Control Algorithm. J Biomed Eng 2015;32(04):905-909.

[11] Hamza NH. On the cooling effect of flowing blood on hepatic tumor ablation process. J Med Eng Technol 2018;42(6):475-481.

[12] Wang X, Gao H, Wu S, et al. RF ablation thermal simulation model: Parameter sensitivity analysis. Technol Health Care 2018;26(S1):179-192.

[13] Chang IA, Nguyen UD. Thermal modeling of lesion growth with radiofrequency ablation devices. Biomed Eng Online 2004;3(1):27.

[14] Trujillo M, Berjano E. Review of the mathematical functions used to model the temperature dependence of electrical and thermal conductivities of biological tissue in radiofrequency ablation. Int J Hyperthermia 2013;29(6):590-597.

[15] Nguyen DM, Andersen T, Qian P, et al. Electrical Impedance Tomography for monitoring cardiac radiofrequency ablation: a scoping review of an emerging technology. Med Eng Phys 2020;84:36-50.

[16] Chang I. Finite element analysis of hepatic radiofrequency ablation probes using temperature-dependent electrical conductivity. Biomed Eng Online 2003;2:12.

[17] Zhang B, Moser MA, Zhang EM, et al. A review of radiofrequency ablation: Large target tissue necrosis and mathematical modelling. Phys Med 2016;32(8):961-971.

[18] Chen XZ, Ma QY, Yang L, et al. Temperature Dependent Electrical Impedance Characteristics for Porcine Liver in Vitro. J Nanjing Norm Univ (Nat Sci Ed) 2012;35(02):32-38.

[19] Crocetti L, de Baére T, Pereira PL, et al. CIRSE Standards of Practice on Thermal Ablation of Liver Tumours. Cardiovasc Intervent Radiol 2020;43(7):951-962.

[20] Lencioni R, Crocetti L, Cioni D, et al. Percutaneous radiofrequency ablation of hepatic colorectal metastases: technique, indications, results, and new promises. Invest Radiol 2004;39(11):689-697.

[21] Venkatesan AM, Gervais DA, Mueller PR. Percutaneous radiofrequency thermal ablation of primary and metastatic hepatic tumors: current concepts and review of the literature. Semin Intervent Radiol 2006;23(1):73-84.

[22] Ng KK, Lam CM, Poon RT, et al. Porcine liver: morphologic characteristics and cell viability at experimental radiofrequency ablation with internally cooled electrodes. Radiology 2005;235(2):478-486.

[23] Mulier S, Ni Y, Frich L, et al. Experimental and clinical radiofrequency ablation: proposal for standardized description of coagulation size and geometry. Ann Surg Oncol 2007;14(4):1381-1396.

[24] Zhou W, Liang M, Pan H, et al. Comparison of ablation zones among different tissues using 2450-MHz cooled-shaft microwave antenna: results in ex vivo porcine models. PLoS One 2013;8(8):e71873.

[25] Rempp H, Mezger D, Voigtlaender M, et al. A comparison of internally water-perfused and cryogenically cooled monopolar and bipolar radiofrequency applicators in ex vivo liver samples. Acad Radiol 2014;21(5):661-666.

[26] Sommer CM, Bryant M, Kortes N, et al. Microwave ablation in porcine livers applying 5-minute protocols: influence of deployed energy on extent and shape of coagulation. J Vasc Interv Radiol 2012;23(12):1692-1699.

[27] d'Avila A, Houghtaling C, Gutierrez P, et al. Catheter ablation of ventricular epicardial tissue: a comparison of standard and cooled-tip radiofrequency energy. Circulation 2004;109(19):2363-2369.

[28] Kapural L, Deering JP. A technological overview of cooled radiofrequency ablation and its effectiveness in the management of chronic knee pain. Pain Manag 2020;10(3):133-140.