Perioperative Precision Medicine
eISSN: 2957-5443

Review Article https://doi.org/10.61189/663074tcakcn

Research progress of frontier image processing in medical endoscopes

Jinjing Wu[1,*], Yang Yuan[2,*], Long Liu[1], Haipo Cui[1], Tianying Xu[3], Miao Zhou[4], Zhanheng Chen[3], Bing Xu[3]

[1]School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China.[2]School of Computer Science and Artificial Intelligence, Changzhou University, Jiangsu 213164, China.[3]School of Anesthesiology, Second Military Medical University/Naval Medical University, Shanghai 200433, China. [4]Department of Anesthesiology, the Affiliated Cancer Hospital of Nanjing Medical University, Jiangsu Cancer Hospital, Jiangsu Institute of Cancer Research, Nanjing Medical University, Nanjing 210009, Jiangsu, China.
[*]The authors contribute equally.

Acknowledgements: This work was supported by the National Natural Science Foundation of China (62002297, 62073225, 61836005), the Science and Technology Commission of Shanghai Municipality (20XD1434400), the Talent Development Fund of Shanghai (2020075), Medical-Engineering Cross Fund of Shanghai Jiao Tong University (YG2022QN043), and the Guangxi Science and Technology Base and Talent Special Project (2021AC19394). The authors would like to thank all the guest editors and anonymous reviewers for their constructive comments.

Declaration of conflict of interest: None.

Received July 19, 2023; Accepted September 6, 2023; Published September 30, 2023

Highlights

- Medical endoscopic images can provide doctors with more accurate, visualized, and three-dimensional views of various internal tissues.
- Image processing techniques such as image denoising, image deblurring, image enhancement, and image segmentation can improve the imaging quality of endoscopes.

Abstract

In the modern medical diagnosis, digital medical images can provide physicians with a more accurate, visualized, and three-dimensional view of various tissues. These images assist in predicting, diagnosing, and treating diseases. However, medical images are highly susceptible to noise contamination from the influence of imaging equipment and the capture process, which poses a significant challenge in the analysis of medical images. This review summarizes the image processing technologies applied in endoscopy, such as image denoising, image deblurring, image enhancement, and image segmentation, involving traditional computational models and deep learning algorithms used in these technologies. Additionally, the clinical applications of these techniques are also discussed.

Keywords: Endoscopy, image denoising, image deblurring, image enhancement, image segmentation, artificial intelligence

Introduction

The development history of endoscopy is long and rich. In 1806, Philipp Bozzini developed the first endoscope, known as the "Lichtleiter," which was successfully used for examinations of bladder, esophagus, and urethra. Subsequently, in 1954, Harold Hopkins from the United Kingdom invented fiber-optic endoscope, marking the inception of modern gastrointestinal endoscopy. In 1983, Welch-Allyn, a company based in the United States, introduced the world's first electronic endoscope, which quickly found applications in clinical practice. Since then, electronic endoscopy has become a focal point of research, particularly with the dominance of the Japanese company Olympus in the electronic endoscope market. However, traditional endoscopes have limitations in image quality due to factors such as hardware

Address correspondence to: Haipo Cui, School of Health Science and Engineering, University of Shanghai for Science and Technology, NO.334, Jungong Road, Shanghai 200093, China. E-mail: h_b_cui@163.com, Tel: +86-21-55271290; **Bing Xu**, School of Anesthesiology, Second Military Medical University/Naval Medical University, No.800 Xiangyin Road, Shanghai 200433, China. E-mail: mzxubing1992@163.com, Tel: +86-21-81872030; **Zhanheng Chen**, School of Anesthesiology, Second Military Medical University/Naval Medical University, No.800 Xiangyin Road, Shanghai 200433, China. E-mail: chenzhanheng17@mails.ucas.ac.cn, Tel: +86 21 81872034.

constraints, which restrict their accuracy in disease diagnosis. To improve the quality of endoscopic images, modern endoscope researchers are dedicated to developing better techniques to enhance the applicability of endoscopy in detecting hard-to-detect lesions like adenomas. Image processing methods play a crucial role in endoscopy, encompassing research in areas such as image denoising, image deblurring, image enhancement, and image segmentation. These studies aim to make the images clearer and more visible, thus enhancing the applicability of endoscopy in the diagnosis of hard-to-detect lesions like adenomas.

With the advancement of artificial intelligence (AI) technology, image processing algorithms have also been continuously updated and iterated. The integration of endoscopy with AI has gained increasing attention. In recent years, deep learning has shown promising applications in the field of image processing, including image detection, image segmentation, and image restoration. Tola et al. proposed a convolutional neural network (CNN) for the detection of brain diseases such as inflammatory diseases, cerebrovascular diseases, and tumor diseases in magnetic resonance imaging (MRI) images [1]. Gupta et al. developed an effective computer-aided diagnosis system that includes four main processing steps: image preprocessing, lung image feature extraction, feature selection using bio-inspired evolutionary algorithms, and potential disease classification [2]. Xue et al. proposed two generative adversarial networks (GANs), namely Segmentor and Critic, for learning the transformation between brain MRI images and binary segmentation maps of brain tumors [3]. Schlegl et al. trained GANs using healthy patches in the retinal region to learn the data distribution of healthy tissues for anomaly detection in retinal images [4].

This paper provides a comprehensive overview of the research progress on the application of advanced image processing methods in endoscopy. In Section 2, a concise description of the process of literature search is provided. In Section 3, a brief introduction to deep learning methods is presented. In Section 4, the existing image processing techniques are outlined, including image denoising, image deblurring, image enhancement, and image segmentation. In Section 5, we summarize the content of this paper and discuss the current research challenges. In Section 6, future research directions are proposed.

Literature research

A literature search was conducted on PubMed, Web of Science, and Science Direct using the following search terms: “endoscope and deep learning,” “medical and image restoration,” “endoscope and image denoising,” “endoscope and image deblurring,” and “endoscope and image enhancement.” The search covered publications from 2009 to 2022. Initially, a total of 150 studies were identified. After screening for relevance, 31 studies were included. **Table 1** summarizes the key findings and main conclusions of the included literature.

Classical deep learning methods for image processing

Deep learning, a branch of machine learning, is an algorithm that utilizes artificial neural networks to automatically extract features from data. Unlike traditional machine learning methods that require manual feature extraction, deep learning algorithms perform both feature extraction and classification in an end-to-end manner, which makes deep learning particularly effective in various domains. The application scenarios of deep learning can be categorized into three major domains: (1) computer vision, such as image recognition and object detection; (2) natural language processing, including machine translation and text processing; (3) speech technology, such as speech recognition [5-13].

Convolutional neural networks (CNNs)

CNNs are widely based on deep learning models specifically designed for processing image and audio data. The core principle of CNNs is to extract local features of an image through convolutional operations and reduce the dimensionality of features through pooling operations. Classification or regression tasks are then performed through fully connected layers. The key components of CNN models include convolutional layers, pooling layers, and fully connected layers. The structure of a CNN is illustrated in **Figure 1**. The convolutional layer is the core of CNN, which extracts features from the input image through sliding convolutional kernels. The parameters of each convolutional kernel can be learned, so that the network can automatically extract features with different levels and abstractions. Pooling is a technique used to downsample the input image, reducing the number of pixels while retaining essential information. Its primary purpose is to decrease computational complexity. The two main types of pooling are max pooling and average pool-

Table 1. Summary of literature

Authors	Focus of study	Principal conclusions
Gopi et al. [14]	Denoising capsule endoscope images using double density dual tree complex wavelet transform.	The proposed method effectively reduces noise in capsule endoscopic images, improving their quality and enhancing diagnostic accuracy.
Patil et al. [15]	Denoising medical images using wavelet transform and singular value decomposition.	Wavelet transform and singular value decomposition effectively denoise medical images, preserving important features and improving image quality for enhanced diagnostic accuracy.
Qu et al. [16]	Real-time removal of stripe noise in endoscope images.	The article presents a real-time method for effectively removing stripe noise in endoscope images, improving image quality and facilitating accurate diagnosis.
Ning et al. [17]	Developing effective methods for removing noise in medical images, aiming to enhance image quality and improve the accuracy of medical diagnosis.	The research successfully investigates and proposes effective techniques for removing noise in medical images, leading to improved image quality and facilitating accurate medical diagnosis.
Magadalena et al. [18]	RGB-based compressed medical imaging using sparsity averaging reweighted analysis for wireless capsule endoscopy images.	The article proposes an RGB-based compressed medical imaging method for wireless capsule endoscopy images, utilizing sparsity averaging reweighted analysis. The approach effectively reduces data size while preserving image quality, enabling efficient transmission and storage of medical images.
Zhang et al. [19]	Deep CNN-based image denoising beyond Gaussian denoising.	The article introduces a CNN approach for image denoising, surpassing traditional Gaussian denoising methods. The proposed method achieves superior denoising performance by learning residual information, enhancing image quality, and preserving fine details.
Izadi et al. [20]	Blind image denoising using noise whiteness prior in the WhiteNNer method.	The article proposes the WhiteNNer method, which utilizes noise whiteness priors for blind image denoising. The approach effectively reduces noise and improves image quality without requiring prior knowledge of noise characteristics, making it suitable for various real-world denoising applications.
Zou et al. [21]	CNN-based blind denoising method for endoscopic images.	In this article, a novel CNN-based blind denoising method is proposed for the denoising of endoscopic images. The method demonstrates effective noise reduction and enhances image quality, leading to improved visibility of important structures and enabling more accurate diagnosis in endoscopy applications.
Ahmed et al. [23]	The denoising CNN can remove the noise on the image and improve the clarity and contrast by learning the noise distribution in the endoscope.	The study introduces an enhanced classification method using transfer learning and well-trained CNN models. It incorporates a noise removal stage as a pre-processing step, leading to improved performance on multiclass medical problems.
Peng et al. [25]	They present a frame work that can restore blurry frames by synthesizing image details from the nearby sharp frames.	The proposed synthesis-based approach successfully deblurs endoscopic videos, improving visual quality and enhancing image clarity, contributing to diagnostic and surgical applications.

Authors	Focus of study	Principal conclusions
Tchaka et al. [26]	The article focuses on chromaticity-based smoke removal in endoscopic images, aiming to effectively eliminate smoke artifacts and enhance image visibility for accurate diagnosis and better guiding surgical procedures.	Smoke artifacts in endoscopic images are effectively reduced using achromaticity-based approach, resulting in improved visibility and enhanced image quality for better diagnostic accuracy.
Wang et al. [27]	Reduction of bubble-like frames in wireless capsule endoscopy videos using an RSS filter.	The article demonstrates that employing an RSS filter effectively reduces bubble-like frames in wireless capsule endoscopy videos. The proposed method enhances the quality of the videos, improving visual clarity and facilitating accurate diagnosis.
Lin et al. [30]	Desmoking algorithm for endoscopic images using an improved U-Net model.	The article proposes an improved U-Net model for desmoking endoscopic images, effectively reducing smoke artifacts and improving image visibility. The method enhances the quality of images, facilitating accurate diagnosis in endoscopy.
Yang et al. [31]	Deep learning-based methods for deblurring and super-resolution reconstruction of endoscopic images.	The article presents deep learning-based techniques for deblurring and super-resolution reconstruction of endoscopic images, achieving improved image quality and enhanced visual details, and enabling better visualization and accurate diagnosis in endoscopy.
Hai et al. [32]	Image enhancement using Contrast Limited Adaptive Histogram Equalization for 3D images in stereoscopic endoscopy.	The article proposes using Contrast Limited Adaptive Histogram Equalization for enhancing 3D images in stereoscopic endoscopy. The method effectively improves image quality, enhances visual details, and enables better visualization of anatomical structures for accurate diagnosis.
Sato et al. [33]	Texture and color enhancement imaging for improving endoscopic image quality and enhancing visual details.	The article presents the Texture and color enhancement imaging method for endoscopic image enhancement, effectively improving image quality, enhancing texture and color, and facilitating better visualization of important details, leading to improved diagnostic capabilities in endoscopy.
Zhang et al. [34]	Improved weighted guided filtering for enhancing medical endoscope images.	The article proposes an improved weighted guided filtering method for enhancing medical endoscope images, effectively reducing noise, enhancing image details, and improving overall visual quality, leading to improved diagnostic capabilities in medical applications.
Wu et al. [35]	Quality enhancement of endoscopic images using Retinex and Pseudo-HDR synthesis algorithms.	The article presents Retinex and pseudo-HDR synthesis algorithms for enhancing the quality of endoscopic images, resulting in improved visual perception, enhanced details, and better image quality for accurate diagnosis and analysis in endoscopy.
An et al. [36]	Development of the EIEN, based on the Retinex theory, for endoscopic image enhancement.	The article introduces the EIEN, based on Retinex theory, for endoscopic image enhancement, effectively improving image quality, enhancing details, and providing better visual perception for improved diagnostic accuracy in endoscopy.
Shen et al. [37]	Content-aware specular reflection suppression in endoscopic images using adaptive image inpainting and neural network techniques.	The article presents a method for content-aware specular reflection suppression in endoscopic images using adaptive image inpainting and neural network, effectively removing reflections and enhancing image quality, and enabling better visualization and accurate diagnosis in endoscopy.

Authors	Focus of study	Principal conclusions
Gómez et al. [38]	Low-light image enhancement of high-speed endoscopic videos using a CNN.	The article presents a CNN-based method for low-light image enhancement of high-speed endoscopic videos, effectively improving visibility, enhancing details, and providing better image quality for accurate diagnosis in endoscopy.
Huang et al. [39]	Deep unsupervised endoscopic image enhancement through multi-image fusion.	The article demonstrates that deep unsupervised endoscopic image enhancement based on multi-image fusion can effectively improve the visual quality and enhance the details of endoscopic images without the need for manual annotations.
Li et al. [40]	Endoscopy image enhancement using adversarial training with generalized imaging defect models.	The article proposes an endoscopy image enhancement method based on adversarial training and generalized imaging defect models, effectively reducing image artifacts and improving the visual quality of endoscopic images.
Jha et al. [43]	Real-time polyp segmentation in video endoscopy using NanoNet, achieving accurate and efficient detection of polyps in capsule endoscopy and colonoscopy.	The article presents NanoNet, a real-time polyp segmentation method for video endoscopy, enabling accurate and efficient detection of polyps in capsule endoscopy and colonoscopy procedures.
Laves et al. [44]	Laryngeal endoscopic image dataset and CNN-based segmentation for improved analysis.	A comprehensive dataset of laryngeal endoscopic images was created for the study, and CNN-based semantic segmentation techniques were applied. The results demonstrated improved analysis and segmentation accuracy in laryngeal endoscopy.
Iwasa et al. [45]	Deep learning-based automatic segmentation of pancreatic tumors using contrast-enhanced endoscopic ultrasound video images.	Deep learning enables accurate segmentation of pancreatic tumors in contrast-enhanced endoscopic ultrasound videos, providing potential for improved diagnosis and treatment planning.

Note: CNN, convolutional neural network; RSS, recursive shadow suppression; EIEN, environmental information exchange network.

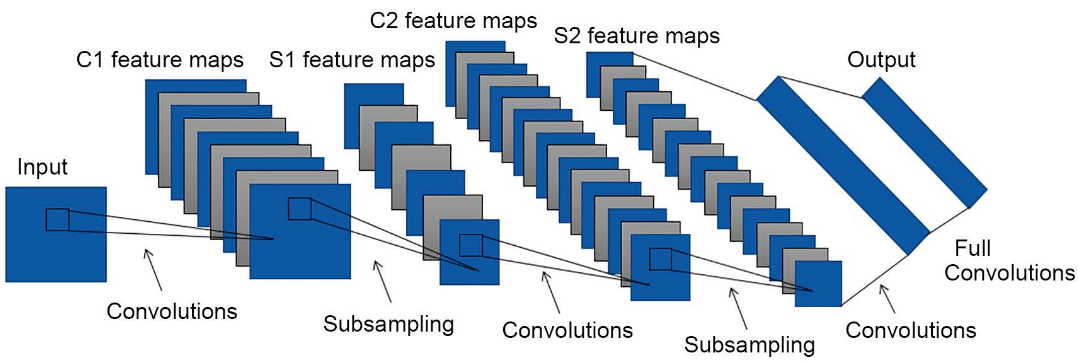

Figure 1. Structure of a CNN. CNN, convolutional neural network.

ing. The fully connected layer plays a role of classifier in the entire CNN. Before passing to fully connected layer, the outputs from previous layers need to be flattened.

Additionally, CNN models incorporate nonlinear activation functions, such as the rectified linear unit function, to introduce nonlinearity without changing the input shape, thereby enhancing the expressive power of the model. The training process of CNN models typically involves the use of a backpropagation algorithm to optimize the loss function and adjust the model parameters, enabling accurate classification or regression. Due to the local connectivity and weight-sharing properties of CNN models, they have shown excellent performance in image processing tasks and are commonly applied in image classification, object detection, and image segmentation.

Transfer learning

Transfer learning is a machine learning method

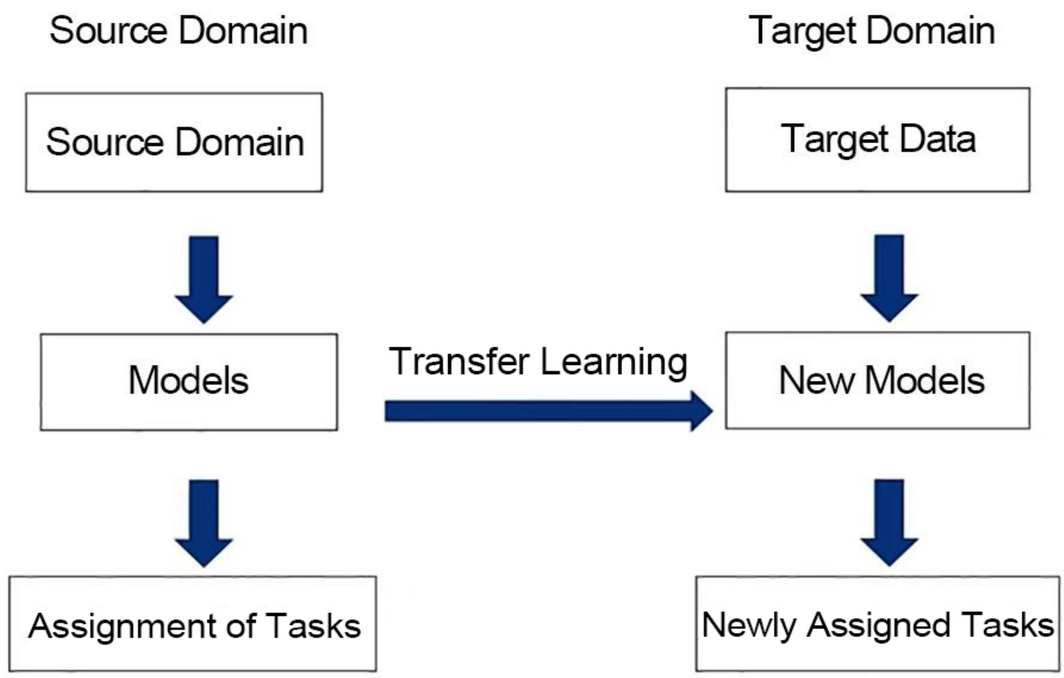

Figure 2. Flowchart of transfer learning.

that leverages knowledge and model parameters learned from one task to improve the performance of another related task. By employing transfer learning, the learning performance of the target task can be enhanced, especially when there is limited or insufficient labeled data available. This approach utilizes the knowledge and experience gained from the source task to aid the learning process of the target task. The core principle of transfer learning is based on the observation that there are shared underlying features and knowledge across different tasks and domains. These underlying features and knowledge can be acquired through pre-training models on the source task. By employing transfer learning, these learned features and knowledge can be applied to the target task, thereby accelerating the learning process and improving its performance.

Transfer learning consists of several key steps, including pre-training, feature extraction, and fine-tuning. This learning method has extensive applications in various domains such as object recognition, object detection, natural language processing, and medical image analysis. It addresses the challenge of limited data and constrained computational resources in the target task by utilizing the learned features and knowledge from the source task. Transfer learning has widespread applications across different fields and provides powerful tools and methods for addressing practical problems. The process flowchart is illustrated in **Figure 2**.

Data augmentation

In image processing, data augmentation is often an indispensable part aimed at improving image quality, increasing dataset diversity, and enhancing the generalization capability of the models. The principle of data augmentation involves applying a series of transformation operations to the original image, such as geometric transformations (translation, rotation, scaling), color transformations (brightness adjustment, contrast enhancement, color distortion), and adding image noise, so as to generate new training samples. By introducing these transformations, uncertainties and diversities present in the real world can be simulated, making the model more robust.

With the widespread application of deep learning, unsupervised data augmentation has emerged as a novel approach. It involves enhancing the training dataset by synthesizing data automatically. Unlike traditional data augmentation methods, unsupervised data augmentation does not rely on any external supervision signals. Instead, it utilizes the features and statistical properties of its data to generate new training samples. Data augmentation finds extensive applications in various domains, in-

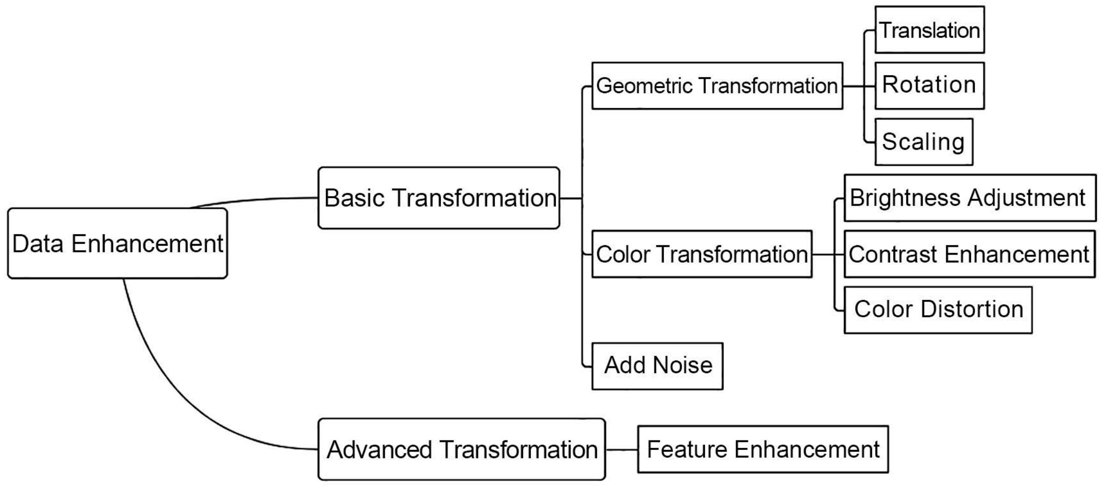

Figure 3. Data augmentation flowchart.

Table 2. High-quality images and corresponding image processing techniques

Characteristics of high-quality images	Corresponding image processing techniques
High resolution	Image sharpening and denoising
Good contrast	Histogram equalization and contrast enhancement
Rich details and textures	Local enhancement and texture enhancement
No blurring or motion blur	Image deblurring and motion blur correction
Uniform illumination	Image illumination equalization and background illumination correction
No noise or artifacts	Image denoising and artifact removal
Moving object detection	Moving object tracking and segmentation

cluding image classification, object detection, semantic segmentation, and anomaly detection, among others. The process flowchart of data augmentation is illustrated in **Figure 3**.

Image processing techniques in endoscopes

Medical image processing is an important branch of image processing that applies image processing techniques to the analysis and recognition of medical image data. Medical image data include computed tomography scans, MRI scans, X-rays, ultrasound images, and others. These image data contain rich information that can be utilized for disease diagnosis and treatment. Medical image processing techniques can enhance, segment, register, and classify these image data, extracting useful information to assist doctors in diagnosis and treatment decision-making. Medical image processing techniques can be applied in tumor detection, organ segmentation, lesion identification, surgical navigation, and more. The application of medical image processing techniques can improve diagnostic accuracy and treatment effectiveness while reducing healthcare resource waste and patient trauma. The characteristics and targeted image processing of high-quality medical images are illustrated in **Table 2**, as depicted in this academic paper.

Denoising of endoscopic images

After studying the principles of image noise, researchers discovered that the noise frequencies exhibit regular patterns. Based on the distribution characteristics and the nature of the images themselves, numerous traditional image denoising methods have emerged. One of the conventional approaches used in medical denoising is based on wavelet transform, which has also been applied in endoscopy. Gopi et al. proposed a method based on dual-tree complex wavelet transform using bivariate shrinkage to reduce noise in intestinal images [14]. In subsequent years, Patil et al. combined wavelet transform with singular value decomposition to further reduce noise [15]. The experimental results demonstrated that this approach outperformed the use of wavelet denoising alone, and the denoised images were advantageous for further processing tasks such as image segmentation and recognition. Furthermore, due to the use of small-sized sensors for imaging in endoscopy, images are often transmitted as high-speed serial signals or analog signals, which can introduce striped noise. Qu et al. proposed a real-time denoising method for removing stripe noise, where the Fourier transform of each row of the noisy image was implemented on a field-programmable gate array [16]. This method can identify frequency domain and suppress the stripe noise spectrum. The designed endoscope was applied in urological systems, effectively meeting the clinical requirements for pathological diagnosis. Another common type of noise in endoscopic images is salt-and-pepper noise. Ning et al. employed traditional median filtering and adaptive median filtering methods to evaluate their denoising effectiveness based on the peak signal-to-noise ratio and the degree of noise degradation in the original images [17]. The study analyzed the influence of filtering noise density on the quality of denoised images. Their results demonstrated that both methods effectively eliminated salt-and-pepper noise in computed tomography and MRI images while preserving image edges and details.

Most traditional medical image processing methods are developed based on the characteristics of the images themselves, statistical features of the noise, and distribution patterns. Generally, these methods can be classified into two categories: spatial domain-based and transform domain-based approaches. Traditional filters often process the entire noisy image without considering the specific characteristics of the noise. As a result, they fail to properly handle local textures, leading to the blurring of image details. Therefore, researchers have been striving to develop denoising algorithms that can effectively preserve local detail features while maintaining good edge preservation.

In recent years, with the development of machine learning, there has been an emergence of model-based and learning-based methods for processing endoscopic images. Machine learning techniques can be categorized into dictionary learning methods and deep learning methods. Magdalena et al. proposed a medical image processing method based on RGB compression [18]. This approach uses sparsity averaging with reweighted analysis to convert the image into a low-dimensional sparse representation. Compression and decompression algorithms are then employed to compress and reconstruct images. This method shows potential applications in wireless capsule endoscope image transmission and storage. In addition to noise caused by devices, uneven illumination also affects physicians' observation and diagnosis of lesions. The emergence of CNNs has opened up new possibilities in image processing. However, there are few CNN-based blind image-denoising models. Zhang et al. proposed an endoscopic image denoising algorithm based on spatial U-Net [19]. This algorithm utilizes spatial attention mechanisms to extract and enhance important information in the image, reducing the influence of noise and artifacts. The U-Net structure can effectively restore the details and texture information of endoscopic images while avoiding excessive image smoothing. Izadi et al. introduced the deep image prior method, which uses a randomly initialized neural network and white Gaussian noise as input to denoise real images through thousands of iterations [20]. However, this method requires manual termination of the iteration process, making it computationally intensive. Zou et al. proposed a block-based blind denoising method [21]. They also designed a two-stage CNN model to process each block. The first stage extracts image features using convolution, and the second stage performs blind denoising on the extracted features. Compared to Lempitsky's experiments, this approach improves the deep image prior method by resolving the challenge of determining the iteration's terminate point in the absence of specific noise models and prior information [22]. Ahmed et al. used a CNN approach, where a denoising CNN learned the noise distribution in endoscopic images, thereby removing noise and enhancing clarity and contrast, and then the enhanced images were classified to aid physicians in disease diagnosis [23]. **Table 3** summarizes some common denoising methods.

Image deblurring

Capsule endoscopy is a type of endoscopy that is particularly prone to motion blur. Factors such as camera movement during the procedure and its relatively small field of view contribute to the reduced quality of capsule endoscopy images. Consequently, these images exhibit motion blur, overexposure, and underexposure, all of which hinder subsequent texture analysis and lesion identification. Traditional deblurring methods primarily focus on image blur models and utilize known blur kernels for deconvolution operations to obtain sharp images, resulting in predominantly non-blind deblurring methods. Single-image deblurring methods can estimate either a uniform blur kernel or a spatially varying blur kernel for image deconvolution. Cho et al. argued that applying image deconvolution to video deblurring is problematic since some deconvolution artifacts may be further amplified by temporal factors and

Table 3. Common image denoising methods

Method	Keywords	Advantages	Disadvantages	Effect comparison
BM3D	Image denoising, block matching	Suitable for different types of noise	High computational complexity and poor performance for signal-to-noise ratio images	Compared with other methods, it performs better under the Gaussian noise effect, but can be generally applied for other types of noise.
NLM	Image denoising, non-local means	Better preservation of image details and edge information, with a certain degree of versatility	High computational complexity and poor performance for images with high noise levels	It performs better with Gaussian noise and pretzel noise, but maintains an average performance for other types of noise.
CNN	Image denoising, CNN	Suitable for various types of noise; good denoising effect	Requires a large amount of annotated data and high computational complexity.	Compared to traditional methods, deep learning-based methods have better denoising performance.
DnCNN	Deep learning, CNN	Fast training speed, good denoising effect, suitable for different types of noise	Requires a large amount of data for training and may be over-fitted.	Better result enhancement compared to traditional methods.
FFDNet	Grayscale images, noise removal	Better adaptability to different types of noise and better noise removal effect	Training takes a long time and requires a large amount of noisy data for training.	Compared to DnCNN, it performs better in the processing of different types of noise.
WGAN	Deep learning, GANs	Capable of generating more realistic images	Longer training time	Compared with traditional methods, WGAN performs better in noise removal and can reduce image distortion.

Note: BM3D, block-matching 3D; NLM, non-local mean; CNN, convolutional neural network; GAN, generative adversarial network; DnCNN, denoising convulsive neural network; FFDNet, fast and flexible denoising convolutional neural network; WGAN, Wasserstein generative adversarial network.

exaggerated in the video [24]. In endoscopic surgery, physicians need to diagnose and treat patients by observing their organs or tissues through the endoscope. However, the vibrations or movements of the endoscope during the surgical process often result in video blur, which affects the observation and diagnosis. Therefore, Peng et al. proposed a framework that can restore blurry frames by synthesizing image details from neighboring sharp frames [25]. They also introduced a non-parametric grid-based motion model to align sharp frames with blurry frames. By directly comparing blurry patches with the nearest matching sharp patches in the endoscopic images, the effectiveness of the algorithm was demonstrated, outputting clearer endoscopic images. Additionally, during endoscopic examinations, the use of biopsy tools can introduce various disturbances to endoscopic images, such as haze, noise, oversaturation, illumination disruptions, and fogging caused by changes of body temperature. Therefore, image defogging is also necessary to improve the visibility. Tchak et al. proposed a chromatic-based defogging method that simulates the formation process of haze in images [26]. They combined an adaptive dark channel prior approach with histogram equalization to eliminate foggy artifacts, restore the image, and enhance contrast and brightness. This approach improves image quality in endoscopic surgery without using additional suction devices. Wang et al. presented a method using recursive shadow suppression filters to reduce bubble frames through detecting bubbles in capsule endoscopy videos, thereby enhancing the visualization of gastrointestinal lesions [27].

Traditional deblurring methods, mostly based on fixed blur kernels, are unable to remove non-uniform motion blur. Moreover, these traditional algorithms often require significant processing time, making them impractical for real-world applications. In recent years, CNNs have been widely utilized in various image restoration tasks due to their efficiency and high function approximation capabilities. Commonly used network architectures include CNNs and GANs. In 2017, Nah et al. introduced a multi-scale deblurring algorithm based on CNNs [28]. By downscaling the images from low-resolution to high-resolution, this research achieved the

Table 4. Common image deblurring methods

Method	Keywords	Advantages	Disadvantages	Effect comparison
Wiener filters	Statistic, filtering	Simple and computationally efficient	Not applicable to complex blur and noise, limited effect	Weak effect, poor effect on complex blurs
Total variation	Edge-holding, regularization	Preserves image detail, suitable for complex blurs	Insufficient smooth details and long processing time	Better retention of image details, but may result in overly strong smoothing effects
Multi-scale fuzzy processing	Image pyramids, multi-scale	Capable of handling different levels of blur, with adjustable effects	High computational complexity, may lead to over-enhancement	Better results, can adapt to different levels of blur
CNN-based approach	Deep learning, CNNs	Capable of learning complex image features to provide better image detail recovery	Requires a large number of annotated image pairs for training and longer training time	Good recovery of image details and textures
DebulrGAN	Deep learning, end-to-end training	Direct mapping from blurred images to clear images	Requires a large number of annotated image pairs for training and longer training time	Ability to recover a clear image directly from a blurred image

Note: CNN, convolutional neural network; GAN, generative adversarial network.

removal of various types of blurs. Building upon this work, Shi et al. proposed a gradient-guided GAN that employed the Res2net structure as the backbone, consisting of image and gradient branches [29]. They designed a lightweight pre-processing network to correct excessive blur and improve training efficiency. Their results demonstrated significant improvements in visual restoration with no noticeable artifacts or structural deformations. However, the computational cost of this network is high, leading to significant training overhead. Lin et al. also presented an end-to-end encoder-decoder network architecture to remove blood blur [30]. They incorporated residual learning and cascaded learning methods to alleviate the insufficient clarity and lack of details when reconstructing issue images affected by blood occlusion. The aim was to enhance the visualization quality of endoscopic images impacted by blood contamination, so as to improve the image reading experience of physicians, aiding in more accurate diagnosis. The Vector Quantized-Variational AutoEncoder (VQ-VAE) model was selected as the backbone network, and a GAN was employed to train a dataset for generating blood blur, effectively reducing the cost of dataset collection. Deblurring is often combined with super-resolution enhancement algorithms. Yang et al. proposed an end-to-end endoscopic image deblurring and super-resolution algorithm, which consists of three components: a deblurring network, a super-resolution network, and a feature fusion network [31]. This approach significantly reduces computational costs during neural network training, overcoming issues related to insufficient training data and slow processing speeds. It enables real-time deep imprinting of endoscopic imaging. Additionally, the deblurring algorithm improves image clarity, and the super-resolution algorithm enhances the restoration of image details, which is crucial for observing subtle lesions or structural features. **Table 4** summarizes some common deblurring methods.

Image enhancement

Medical image enhancement techniques play a crucial role in medical image processing and serve as important preprocessing step for identification and detection algorithms in medical image analysis. Traditional medical image enhancement algorithms can be mainly categorized into two types: spatial domain-based and frequency domain-based methods, with most studies focusing on spatial domain-based approaches. Hai et al. proposed an image enhancement method based on contrast limited adaptive histogram equalization for enhancing stereoscopic endoscopic 3D images [32]. This method involves dividing the image into nearly equal-sized non-overlapping regions to obtain histogram clip limits. Then, the histogram of each region is redistributed and equalized through histogram equalization. A reference image is determined by searching conditions, and global histogram matching is employed to correct color discrepancies between the two stereoscopic views. This method effectively enhances image details, eliminates color biases, and addresses issues such as low contrast and blurry details in stereoscopic endoscopic 3D images, which may affect accurate lesion identification and assessment by physicians. It improves the visual quality and discernibility of the images. However, the method has a higher computational complexity and slightly larger computational cost. Sato et al. introduced a texture and color enhancement imaging technique to enhance the visual quality and discernibility of endoscopic images [33]. This method utilizes texture enhancement algorithms to enhance visual and structural information in the images, thereby making the lesion areas more clearly visible. In addition to noise caused by devices, uneven illumination can also affect the observation and diagnosis of lesions. Zhang et al. proposed an endoscopic image enhancement method based on an improved weighted guided filter [34]. By introducing parameters such as gradient threshold, luminance threshold, and blur coefficient, the method can enhance contrast in images containing lesions, thereby improving the detection rate of lesions. Wu et al. presented a synthesis algorithm based on Retinex and Pseudo-HDR [35]. The former corrects uneven illumination issues and improves contrast and brightness in images. The latter is used for synthesizing images with a wide dynamic range. This algorithm demonstrates good practicality and can be easily integrated and applied in actual endoscopic scenarios.

In the field of medical image processing, deep neural network technology has garnered significant interest from researchers due to its remarkable effectiveness. In recent years, researchers have focused on applying deep learning networks to enhance the quality of endoscopic images. To improve the visual quality and discernibility of endoscopic images, An et al. proposed a Retinex-Net network based on the Retinex theory, which employs image decomposition techniques to decompose the original endoscopic image into different frequency components, such as low-frequency and high-frequency components [36]. By enhancing the different components, the algorithm can enhance the contrast and details of the image and reduce the impact of image blur. Moreover, the reflection from the endoscope's mirror surface can also affect visual perception. Shen et al. proposed a content-aware mirror reflection suppression approach based on adaptive image restoration and neural networks, which effectively mitigates the influence of mirror reflections on the images [37]. Gomez et al. introduced a method for low-light image enhancement of high-speed endoscopic videos using CNNs [38]. In clinical examinations, inadequate illumination is often encountered in endoscopic imaging due to anatomical restrictions of the throat and technical limitations of recording devices. This method helps clinicians and researchers deal with images that frequently suffer from insufficient lighting, making clinical analysis more manageable. The lack of medical ground truth images for training has been a major bottleneck in deep learning-based medical imaging processing, mainly due to the difficulty in collecting original datasets. Unsupervised learning methods eliminate the need for ground truth medical images during training. Huang et al. studied a deep unsupervised endoscopic image enhancement method based on multi-image fusion [39]. By employing unsupervised mapping and deep learning networks, high-quality endoscopic images can be obtained from images with poor lighting conditions, low contrast, and color deviations without the need for real data, thus reducing the workload of data collection. Li et al. introduced a novel physics-based semi-supervised learning framework for endoscopic image enhancement [40]. They integrated a previous physical imaging defect model with the CycleGAN framework to address issues such as image haze, uneven lighting, and color deviations. This approach facilitates detail preservation and color restoration, achieving outstanding performance. Notably, the paper proposed a transfer learning network for generating datasets, which significantly improves network performance on small training datasets. **Table 5** provides a summary of some common image enhancement methods.

Image segmentation

Endoscopic image segmentation is one of the key tasks in the field of medical image processing. It plays a crucial role in endoscopic diagnosis and treatment. The purpose of image segmentation is to accurately separate different tissue structures, lesion areas, or regions of interest within the endoscopic images. Accurate image segmentation can assist physicians in precisely locating and identifying abnormalities, as well as facilitating accurate diagnosis, treatment planning, and surgical guidance. Due to factors such as lighting variations, noise, and diverse morphological appearances commonly present in endoscopic images, researchers have proposed numerous methods over the past few decades to address these challenges.

Early endoscopic image segmentation methods

Table 5. Common image enhancement methods

Method	Keyword	Advantage	Disadvantage	Effect comparison
Histogram equalization	Histogram, contrast enhancement	Enhance contrast by redistributing image pixel values	May cause excessive global contrast enhancement or distortion	Can enhance the contrast and details of an image, but may produce excessive enhancement or distortion
Fourier transform	Fourier spectrum, filtering	Enhance the edges and details of images by frequency domain filtering, capable of filtering out noise and blur	May cause spectral distortion or ringing phenomena, and may have limited effect on complex images and noise	Can enhance image edges and details, but may produce spectral distortion and ringing
Spatial domain filtering	Filtering, spatial domain processing	Specific image features can be processed by filters that can enhance image details and textures, and suppress noise	Parameters need to be adjusted for different image types and noise types, which may result in blurring or distortion of image details	Filtering and enhancement can be performed for different features, but the effect is limited in more complex cases.
Based on GANs	GAN, generators, discriminators	Can generate realistic images and maintain authenticity; can learn image features adaptively to provide better images	Requires large amount of training data and computational resources, with long training time and possible pattern collapse or unstable generation	Provides high-level image enhancement and restoration effects to produce realistic images
Based on CNN	CNN, super resolution	Capable of reconstructing low-resolution images into high-resolution images, can restore details and clarity for a wide range of image types and resolutions	Longer training time, requires large amount of training data, may have limited effect for complex images	Ability to reconstruct low-resolution images into high-resolution images and restore image details

Note: GAN, generative adversarial network; CNN, convolutional neural network.

primarily relied on traditional image processing techniques such as thresholding, edge detection, and morphological operations. The datasets used for traditional segmentation tasks were manually annotated by physicians. Liu et al. proposed a hue-texture embedded region model for narrow-band imaging-enhanced endoscopic image segmentation, which adjusts the hue component of the image to highlight the characteristic colors of structures such as vessels and lesions [41]. By combining color information and texture features, this approach achieves better segmentation of structures and lesion areas within the image. In the field of gastroenterology, segmentation of endoscopic images for detecting digestive ulcers has emerged as a novel domain in medical image processing. Rajivegandhi et al. introduced a collaborative and efficient algorithm for digestive ulcer detection based on watershed transformation, thresholding, and morphological operators [42]. This algorithm successfully detects ulcers, opening a new avenue for advancing gastroscopic image processing techniques and clinical applications.

With the rapid development of computer vision and machine learning, significant progress has been made in endoscopic segmentation. Machine learning and deep learning-based methods are gradually being proposed for handling complex endoscopic images. Endoscopic segmentation is primarily applied to disease detection and lesion segmentation. Jha et al. proposed a real-time polyp segmentation method based on NanoNet for both capsule endoscopy videos and colonoscopy videos, greatly improving detection efficiency [43]. Laves et al. presented a CNN-based semantic segmentation method and dataset for laryngoscopic images, utilizing four methods (SegNet, UNet, ENet, and ErfNet) and evaluating their accuracy

Table 6. Common image segmentation methods

Method	Keywords	Advantages	Disadvantages	Effect comparison
Segmentation based on thresholds	Threshold, pixel classification	Simple, intuitive, and easy to implement for situations where there is an obvious difference in pixel values between the apparent target and the background	Sensitive to light changes, noise and other factors, requiring the pre-setting of set suitable thresholds	Suitable for image segmentation where there is an obvious difference in pixel values between the apparent target and the background
Segmentation based on regional growth	Regional growth	Able to automatically identify continuous regions, good segmentation for regions with high similarity	Poor segmentation for regions with blurred target boundaries or low similarity, sensitive to light changes and noise	Suitable for region segmentation tasks with a high degree of similarity
Segmentation based on edge detection	Edge detection, contour extraction	Ability to accurately extract the outline of the target, good for images with clear edges	Poor image segmentation for blurred or complex backgrounds, sensitive to noise	Suitable for image segmentation where accurate target contours need to be extracted
FCN	CNNs	Ability to learn semantic information of images, better for targets of different scales	May be less effective in the case of fine targets and blurred edges	Provide better semantic segmentation results. Applicable to overall target segmentation
U-net	CNN, encoder, decoder	With encoding and decoding structure, capable of capturing multi-scale contextual information, good re-edge positioning and detail retention	Possible problems for datasets with unbalanced categories	Provide better target boundary segmentation and detail retention, suitable for medical image segmentation
Mask R-CNN	Target detection, instance segmentation	Combines target detection and segmentation tasks to accurately locate and segment multiple target instances, works well for large variations in target shape and size	High computational complexity, requiring more computational resources, may be less effective for complex backgrounds and overlapping targets	Provides accurate target instance segmentation results for multi-target localization and segmentation

Note: FCN, fully convolutional network; CNN, convolutional neural network.

using intersection over the union [44]. Lwasa et al. proposed a deep learning-based method for automatic segmentation of pancreatic tumors in contrast-enhanced endoscopic ultrasound image videos, which has the potential to become a powerful tool for early diagnosis and treatment of pancreatic tumors [45]. However, these methods require manual annotation of datasets, which is time-consuming and labor-intensive. Subsequently, unsupervised learning has gained attention, particularly in the context of endoscopic surgery. Sasmal et al. proposed an unsupervised image segmentation method for segmenting polyps in endoscopic images [46]. Mahmood et al. introduced an adversarial training-based reverse domain adaptation method, which uses unlabeled synthetic data to transform real data into a synthetically similar representation while preserving clinically relevant diagnostic features through regularization [47]. They generated a large dataset consisting of synthetically generated endoscopic images with annotated ground truth from normal colonoscopy images. **Table 6** summarizes some common segmentation methods.

Discussions and challenges

Currently, the integration of endoscopy and image processing technology has become a hot research topic in the medical field. An increasing number of researchers are exploring how to apply AI image processing techniques to the processing and analysis of endoscopic images. In existing studies, endoscopy combined with AI is mainly applied to lesion identification and classification, surgical assistance, training and education, pathology prediction, intelligent diagnosis, and therapy.

Although deep learning-based models have continued to dominate medical image processing, there are still numerous challenges associated with deep modeling. These challenges limit the application and implementation of new methods in clinical practice, but also provide researchers with new opportunities. The following challenges are proposed:

Data scarcity and annotation: Creating annotated datasets is a difficult and time-consuming task. Therefore, it is crucial to design deep learning models that can effectively utilize both labeled and unlabeled data to generate training datasets. Additionally, due to privacy and ethical concerns, it is challenging to collect large-scale medical data. Hence, developing data augmentation models that extract key information from limited samples to form extensive "realistic" datasets is essential.

Real-time requirements: In clinical practice, real-time processing and analysis of endoscopic images are critical for prompt decision-making and treatment. However, deep learning methods often require long training and inference times, which cannot meet the real-time processing demands of clinical endoscopy. Therefore, a challenge lies in how to improve the speed of deep learning models while ensuring the accuracy of the analysis.

Data security: Endoscopic images contain sensitive patient information, making data privacy and security crucial considerations during model training.

Perspectives

This paper primarily presents the clinical application of image processing techniques in endoscopic imaging. Tracheal intubation is a primary method for airway management during general anesthesia and is also one of the emergency measures for critically ill patients. Direct laryngoscopy or video laryngoscopy is commonly used for oral/nasal tracheal intubation, with the latter offering a clearer field of view and higher intubation efficiency compared to the former. However, in certain special cases such as patients undergoing oral and maxillofacial surgery or emergency intubation, the presence of saliva, blood, or sputum in the oral pharynx can contaminate the video laryngoscope lens during the intubation process, obstructing the view of the glottis and posing safety risks during anesthesia. In this study, we propose to utilize deep learning-based image denoising techniques to address the issue of image blurring. By capturing video images of patients undergoing video laryngoscopy for tracheal intubation and generating a dataset of software-simulated noisy images, the aim is to develop a deep learning algorithm for denoising these images. The effectiveness of the program should be validated using real blurry images, with the ultimate goal of assisting anesthesiologists in tracheal intubation procedures, improving intubation efficiency, and reducing complications. The deep learning model will be built based on the Transformer network to restore the blur caused by blood, secretions, and other contaminants in endoscopic images. Additionally, a novel network is also being explored to address the problem of laryngeal structure recognition. Due to the complexity of the laryngeal structure, it is crucial to employ a segmentation approach to accurately delineate the major laryngeal regions during the intubation process. This approach aims to enhance the accuracy and speed of intubation by facilitating the identification of laryngeal structures, particularly for less experienced clinicians.

References

[1] Talo M, Yildirim O, Baloglu UB, et al. Convolutional neural networks for multi-class brain disease detection using MRI images. Comput Med Imaging Graph 2019;78:101673.

[2] Gupta N, Gupta D, Khanna A, et al. Evolutionary algorithms for automatic lung disease detection. Measurement 2019;140:590-608.

[3] Xue Y, Xu T, Zhang H, et al. SegAN: Adversarial Network with Multi-scale L(1) Loss for Medical Image Segmentation. Neuroinformatics 2018;16(3-4):383-392.

[4] Schlegl T, Seeböck P, Waldstein SM, et al. Unsupervised Anomaly Detection with Generative Adversarial Networks to Guide Marker Discovery. Inf Process Med Imaging 2017;146-157.

[5] Hu K, Jin J, Zheng F, et al. Overview of behavior recognition based on deep learning. Artif Intell Rev 2023;56(3):1833-1865.

[6] Keysers D, Deselaers T, Gollan C, et al. Deformation models for image recognition. IEEE Trans Pattern Anal Mach Intell 2007;29(8):1422-1435.

[7] Liu J, Sun J, Wang S. Pattern Recognition: An overview. Int J Comput Sci Network Secur 2006;6(6):57-61.

[8] Neumann M, King D, Beltagy I, et al. ScispaCy: Fast and Robust Models for Biomedical Natural Language Processing. ArXiv 2019;abs/1902.07669.

[9] Otter DW, Medina JR, Kalita JK. A Survey of the Usages of Deep Learning for Natural Language Processing. IEEE Trans Neural Netw Learn Syst 2021;32(2):604-624.

[10] Qiu X, Sun T, Xu Y, et al. Pre-trained models for natural language processing: A survey. Sci China Technol Sci 2020;63(10):1872-1897.

[11] Afouras T, Chung JS, Senior A, et al. Deep

Audio-Visual Speech Recognition. IEEE Trans Pattern Anal Mach Intell 2022;44(12):8717-8727.
[12] Han W, Zhang Z, Zhang Y, et al. ContextNet: Improving Convolutional Neural Networks for Automatic Speech Recognition with Global Context. ArXiv 2020;abs/2005.03191.
[13] Li Z, Liu F, Yang W, et al. A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. IEEE Trans Neural Netw Learn Syst 2022;33(12):6999-7019.
[14] Gopi VP, Palanisamy P, Niwas SI. Capsule Endoscopic Colour Image Denoising Using Complex Wavelet Transform. Wireless Networks and Computational Intelligence 2012;220-229.
[15] Patil R, Bhosale S. Medical Image Denoising using Wavelet Transform and Singular Value Decomposition. Comput Sci 2021.
[16] Qu J, Jin K, Wang M, et al. Real-Time Stripe Noise Removal Method for Endoscope Image. 2021 IEEE 4th International Conference on Electronics Technology (ICET) 2021;865-870.
[17] Ning Cy, Liu Sf, Qu M. Research on removing noise in medical image based on median filter method. 2009 IEEE International Symposium on IT in Medicine & Education 2009;384-388.
[18] Magdalena R, Rahim T, Pratama IPAE, et al. RGB-Based Compressed Medical Imaging Using Sparsity Averaging Reweighted Analysis for Wireless Capsule Endoscopy Images. IEEE Access 2021;9:147091-147101.
[19] Zhang K, Zuo W, Chen Y, et al. Beyond a Gaussian Denoiser: Residual Learning of Deep CNN for Image Denoising. IEEE Trans Image Process 2017;26(7):3142-3155.
[20] Izadi S, Mirikharaji Z, Zhao M, et al. WhiteNNer-Blind Image Denoising via Noise Whiteness Priors. 2019 IEEE/CVF International Conference on Computer Vision Workshop (ICCVW) 2019;476-484.
[21] Zou S, Long M, Wang X, et al. A CNN-Based Blind Denoising Method for Endoscopic Images. 2019 IEEE Biomedical Circuits and Systems Conference (BioCAS) 2019;1-4.
[22] Ulyanov D, Vedaldi A, Lempitsky V. Deep Image Prior. International Journal of Comput Vision 2020;128(7):1867-1888.
[23] Ahmed A. Classification of Gastrointestinal Images Based on Transfer Learning and Denoising Convolutional Neural Networks. Proceedings of International Conference on Data Science and Applications 2022;631-639.
[24] Cho S, Wang J, Lee S. Video deblurring for hand-held cameras using patch-based synthesis. ACM Trans Graph 2012;31(4):Article 64.
[25] Peng L, Liu S, Xie D, et al. Endoscopic video deblurring via synthesis. 2017 IEEE Visual Communications and Image Processing (VCIP) 2017;1-4.
[26] Tchaka K, Pawar VM, Stoyanov D. Chromaticity based smoke removal in endoscopic images. Med Imaging 2017.
[27] Wang Q, Pan N, Xiong W, et al. Reduction of bubble-like frames using a RSS filter in wireless capsule endoscopy video. Opt Laser Technol 2019;110:152-157.
[28] Nah S, Kim TH, Lee KM. Deep Multi-scale Convolutional Neural Network for Dynamic Scene Deblurring. 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2017;257-265.
[29] Huang Z, Shi Z. Image Caption Combined with GAN Training Method. Intelligent Information Processing X 2020;310-316.
[30] Lin J, Jiang M, Pang Y, et al. A desmoking algorithm for endoscopic images based on improved U-Net model. Concurr Comput 2021;33(22):e6320.
[31] Yang X, Chen Y, Tao R, et al. Endoscopic Image Deblurring and Super-Resolution Reconstruction Based on Deep Learning. 2020 International Conference on Artificial Intelligence and Computer Engineering (ICAICE) 2020;168-172.
[32] Hai Y, Li L, Gu J. Image enhancement based on contrast limited adaptive histogram equalization for 3D images of stereoscopic endoscopy. 2015 IEEE International Conference on Information and Automation 2015;668-672.
[33] Sato T. TXI: Texture and Color Enhancement Imaging for Endoscopic Image Enhancement. J Healthc Eng 2021;2021:5518948.
[34] Zhang G, Lin J, Cao E-H, et al. A Medical Endoscope Image Enhancement Method Based on Improved Weighted Guided Filtering. Mathematics 2022.
[35] Wu JCH, Lin GS, Hsu HT, et al. Quality enhancement based on retinex and pseudo-HDR synthesis algorithms for endoscopic images. 2013 Visual Communications and Image Processing (VCIP) 2013;1-5.
[36] An Z, Xu C, Qian K, et al. EIEN: Endoscopic Image Enhancement Network Based on Retinex Theory. Sensors (Basel, Switzerland) 2022;22(14):5464.
[37] Shen DF, Guo JJ, Lin GS, et al. Content-aware specular reflection suppression based on adaptive image inpainting and neural network for endoscopic images. Comput Methods Programs Biomed 2020;192:105414.
[38] Gómez P, Semmler M, Schützenberger A, et al. Low-light image enhancement of high-speed endoscopic videos using a convolutional neural network. Med Biol Eng Comput 2019;57(7):1451-1463.
[39] Huang D, Liu J, Zhou S, et al. Deep unsupervised endoscopic image enhancement based

on multi-image fusion. Comput Methods Programs Biomed 2022;221:106800.

[40] Li W, Fan J, Li Y, et al. Endoscopy image enhancement method by generalized imaging defect models based adversarial training. Phys Med Biol 2022;67(9):095016.

[41] Liu X, Wang C, Bai J, et al. Hue-texture-embedded region-based model for magnifying endoscopy with narrow-band imaging image segmentation based on visual features. Comput Methods Programs Biomed 2017;145:53-66.

[42] Rajivegandhi C, Shree NSD, Khan S, et al. Detection of peptic ulcers based on thresholding and watershed segmentation. 2015 3rd International Conference on Signal Processing, Communication and Networking (ICSCN) 2015;1-5.

[43] Jha D, Tomar NK, Ali S, et al. NanoNet: Real-Time Polyp Segmentation in Video Capsule Endoscopy and Colonoscopy. 2021 IEEE 34th International Symposium on Computer-Based Medical Systems (CBMS) 2021;37-43.

[44] Laves MH, Bicker J, Kahrs LA, et al. A dataset of laryngeal endoscopic images with comparative study on convolution neural network-based semantic segmentation. Int J Comput Assist Radiol Surg 2019;14(3):483-492.

[45] Iwasa Y, Iwashita T, Takeuchi Y, et al. Automatic Segmentation of Pancreatic Tumors Using Deep Learning on a Video Image of Contrast-Enhanced Endoscopic Ultrasound. J Clin Med 2021;10(16).

[46] Sasmal P, Bhuyan MK, Dutta S, et al. An unsupervised approach of colonic polyp segmentation using adaptive markov random fields. Pattern Recognit Lett 2022;154:7-15.

[47] Mahmood F, Chen R, Durr NJ. Unsupervised Reverse Domain Adaptation for Synthetic Medical Images via Adversarial Training. IEEE Trans Med Imaging 2018;37(12):2572-2581.