

Review

A Review of Application of Deep Learning in Endoscopic Image Processing

Zihan Nie ^{1,2} , Muhao Xu ^{1,2}, Zhiyong Wang ^{1,2}, Xiaoqi Lu ^{1,2} and Weiye Song ^{1,2,*} 

¹ School of Mechanical Engineering, Shandong University, Jinan 250061, China; 202414371@mail.sdu.edu.cn (Z.N.); ujnmxu@hotmail.com (M.X.); wzy2021@mail.sdu.edu.cn (Z.W.); 17604135762@163.com (X.L.)

² Key Laboratory of High Efficiency and Clean Mechanical Manufacture of Ministry of Education, Shandong University, Jinan 250061, China

* Correspondence: songweiye@sdu.edu.cn

Abstract: Deep learning, particularly convolutional neural networks (CNNs), has revolutionized endoscopic image processing, significantly enhancing the efficiency and accuracy of disease diagnosis through its exceptional ability to extract features and classify complex patterns. This technology automates medical image analysis, alleviating the workload of physicians and enabling a more focused and personalized approach to patient care. However, despite these remarkable achievements, there are still opportunities to further optimize deep learning models for endoscopic image analysis, including addressing limitations such as the requirement for large annotated datasets and the challenge of achieving higher diagnostic precision, particularly for rare or subtle pathologies. This review comprehensively examines the profound impact of deep learning on endoscopic image processing, highlighting its current strengths and limitations. It also explores potential future directions for research and development, outlining strategies to overcome existing challenges and facilitate the integration of deep learning into clinical practice. Ultimately, the goal is to contribute to the ongoing advancement of medical imaging technologies, leading to more accurate, personalized, and optimized medical care for patients.

Keywords: deep learning; endoscopy; image analysis; convolutional neural networks (CNNs)



Citation: Nie, Z.; Xu, M.; Wang, Z.; Lu, X.; Song, W. A Review of Application of Deep Learning in Endoscopic Image Processing. *J. Imaging* **2024**, *10*, 275. <https://doi.org/10.3390/jimaging10110275>

Academic Editor: Nikolaos Mitianoudis

Received: 28 September 2024

Revised: 24 October 2024

Accepted: 29 October 2024

Published: 1 November 2024



Copyright: © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

With the rapid development of technology, deep learning (DL) has demonstrated its transformative potential in various fields. From intelligent transportation systems (ITSs) to environmental monitoring, deep learning has not only improved the efficiency of traditional processing methods but also promoted the development of automation and intelligence [1,2]. These advances not only bring new opportunities to specific fields but also lay the foundation for the development of future cross-domain intelligent systems.

In modern medicine, endoscopy has become an increasingly vital tool, significantly enhancing the precision and efficiency of early disease diagnosis and treatment through direct visualization of internal organs [3]. Its applications in fields such as the digestive, respiratory, and urinary systems have made it a crucial method for both diagnosis and treatment. Endoscopic technology not only enables real-time observation of internal organs but also allows for the evaluation of lesions by imaging biological tissues, providing more accurate diagnostic information [4]. With the rapid advancements in endoscopy, the amount of data generated is also increasing, making efficient processing and analysis of this data essential for improving diagnostic accuracy and treatment outcomes. However, traditional image processing methods often suffer from slow speeds and limited accuracy, reducing their effectiveness in handling complex medical data [5].

To address these challenges, researchers have increasingly turned to deep learning technologies for endoscopic image analysis, significantly improving both efficiency and

accuracy. Convolutional neural networks (CNNs), in particular, have brought revolutionary improvements to image analysis by mimicking the brain's processing mechanisms. CNNs automatically learn and extract key features from images, enabling efficient and high-precision analysis [6]. This technology has enhanced capabilities in image segmentation, classification, and anomaly detection, reducing the workload for physicians. Computer-aided detection (CAD) and computer-aided diagnosis (CADx) can find or localize abnormal or suspicious areas in structural images, assisting doctors in making faster and more accurate decisions [7].

The quality of images is crucial to the success of deep learning-based analysis and diagnosis. In recent years, several studies have focused on enhancing the quality of endoscopic images using deep learning techniques such as image enhancement, noise removal, and resolution improvement. For instance, certain algorithms are able to create high-quality synthetic medical images or generate missing information, thereby improving diagnostic accuracy [8]. This not only enhances the quality of endoscopic examinations but also reduces the need for repeated procedures, lowering medical costs.

The advances in deep learning technology have also paved the way for personalized medicine. By utilizing deep learning models, researchers can develop more accurate diagnostic tools that provide tailored treatment recommendations based on a patient's specific condition. These innovations not only improve treatment outcomes but also optimize resource use [9]. As artificial intelligence continues to evolve, future endoscopic examinations are expected to become increasingly intelligent, not only through enhanced image analysis but also via data integration and comprehensive analysis, allowing for personalized medical services, health risk predictions, and tailored preventive measures [9–11].

In this paper, we review the application of deep learning technology in endoscopic image analysis. Since deep learning has made significant breakthroughs in image recognition, it has been widely adopted in medical imaging, particularly in feature extraction, disease diagnosis, and image enhancement for endoscopic images. We explore in detail the central role of CNNs in improving disease detection accuracy and processing efficiency, as well as how deep learning reduces physicians' workload through automated analysis, improving the detection and classification of lesions. Additionally, this paper reviews recent advances in deep learning for image quality enhancement, dynamic tracking, and three-dimensional reconstruction, discussing how these technologies can support clinical diagnosis and treatment. Through this analysis, we aim to demonstrate the potential of deep learning to further advance endoscopic technology and provide insights for future research directions.

2. Related Surveys

Several review papers delve into the intersection of artificial intelligence (AI), particularly deep learning, with gastrointestinal endoscopy and broader medical image analysis. Each provides valuable insights, although with differing scopes and focuses.

Shen et al. [7] provide a foundational overview of deep learning in medical imaging, including architectures such as CNNs and generative models. Alagappan et al. [10] adopt a forward-looking perspective, predicting the transformative impact of AI on gastrointestinal endoscopy. Choi et al. [8] examine the role of convolutional CNNs in endoscopic imaging, particularly for polyp detection and classification. Their review emphasizes the technical development and challenges related to CNN applications in diagnostic imaging. Pannala et al. [9] emphasize the real-time potential of AI, particularly through computer-aided detection (CADE) systems that aim to improve adenoma detection rates. Zhuang et al. [12] offer a focused examination of deep learning applications within digestive system imaging, analyzing recent advances.

Across these reviews, several common limitations emerge. None comprehensively tackle the full spectrum of challenges associated with deep learning integration in clinical workflows. Additionally, key technical aspects like super-resolution, noise reduction, and

cross-modality integration—essential for achieving higher diagnostic accuracy—are either mentioned briefly or overlooked. Many reviews also focus narrowly, either on future potentials or current applications, without sufficiently bridging the gap between recent research advancements and the practical strategies needed for clinical adoption. This leaves a need for more holistic perspectives that encompass both technical progress and real-world challenges. Table 1 shows the comparison of the reviews.

Table 1. Comparative overview of review papers.

Review Paper	Focus	Strengths	Gaps	How This Review Fills the Gap
Shen et al. (2017) [7]	Deep learning in medical imaging	Provides a comprehensive overview of deep learning models	Insufficient focus on real-time endoscopic applications	Addresses endoscopy-specific challenges like real-time processing
Alagappan et al. (2018) [10]	Future trends in AI for endoscopy	Predicts future AI impacts	Limited focus on current optimization strategies	Bridges research advancements with clinical implementation
Choi et al. (2020) [8]	CNN-based endoscopic imaging	Highlights polyp detection and technical evolution	Limited scope to segmentation and polyp detection	Addresses advanced optimizations like noise reduction
Pannala et al. (2020) [9]	Practical AI in endoscopy	Emphasizes real-time CAdE systems	Lacks discussion on super-resolution and denoising	Incorporates advanced image enhancement strategies
Zhuang et al. (2023) [12]	Digestive system image processing	Focused review of deep learning in digestive imaging	Limited cross-disciplinary scope	We integrate insights across broader endoscopic applications

3. Literature Review

3.1. Endoscopic Imaging Technology

Endoscopic imaging technology is a core tool in modern medical diagnosis, enabling doctors to directly visualize internal structures of the human body through minimally invasive or non-invasive methods. These technologies are crucial for early disease diagnosis, treatment monitoring, and surgical navigation. Below are five common endoscopic imaging techniques.

3.1.1. White Light Endoscope (WLE)

White light endoscopes (WLEs) are the most widely used endoscopes, utilizing traditional white light illumination to provide high-definition views of the surface structures within body cavities such as the digestive and respiratory tracts. The fundamental principles of white light endoscopy were first established through early studies, including those related to digestive tract diseases [13]. In recent years, WLE technology has continued to evolve, with advancements such as high-resolution endoscopes and video endoscopy systems, which enhance image resolution and detail. For example, modern high-definition WLEs can more accurately detect small lesions, such as intestinal metaplasia [14]. Additionally, combining WLEs with computer-aided diagnostic (CAD) systems that implement discriminant analysis of pit patterns, using quantitative features from training images, has improved the accuracy and efficiency of lesion detection [15].

3.1.2. Intravascular Ultrasound (IVUS)

Intravascular ultrasound (IVUS) involves inserting an ultrasound probe into blood vessels to obtain images of the vessel walls and surrounding tissues, and is widely used to evaluate atherosclerosis and guide endovascular procedures. This technology has been recognized by the Society for Cardiac Angiography and Interventions [16]. With ongoing technological advancements, IVUS now offers improved image quality and expanded applications. High-frequency and three-dimensional IVUS enhance the resolution and accuracy of plaque analysis. For instance, high-frequency IVUS provides finer details of vascular structures, helping doctors better assess plaque composition [17]. Moreover, combining IVUS with other imaging modalities, such as optical coherence tomography

(OCT), has further improved the evaluation of spontaneous coronary artery dissection and characterization of atherosclerotic plaques [18,19].

3.1.3. Endoscopic Ultrasound (EUS)

Endoscopic ultrasound (EUS) combines endoscopy and ultrasound to obtain high-resolution images from within the body. It is particularly useful for diagnosing diseases of the gastrointestinal tract and adjacent structures, such as pancreatitis or pancreatic cancer [20]. Recent advancements in EUS include elastography, where a transducer sends a shear wave through the pancreas and generates an elastogram by calculating the velocity of the wave passing through soft tissues [21].

3.1.4. Optical Coherence Tomography (OCT)

OCT uses light waves for cellular-level imaging, commonly applied in ophthalmology and cardiovascular diagnostics. OCT can generate microscopic three-dimensional images of tissue structures, aiding doctors in assessing pathological changes. First proposed by Huang et al. in 1991 [22], OCT has recently been adapted for endoscopic use. For example, wide-field OCT in endoscopy now provides high-resolution images over larger fields of view through improvements in optical design and data processing [23]. The latest advancements, such as optical coherence tomography angiography (OCT-A), show great potential for endoscopic applications [24].

3.1.5. Narrow-Band Imaging Endoscopy (NBI)

Narrow-band imaging (NBI) uses filtering technology to enhance endoscopic imaging by isolating specific blue and green wavelengths. This technology produces images with unique color distributions, aiding in the visualization of tissue structures [25]. NBI has made significant strides, especially in the early detection of gastrointestinal cancers [26].

Together, these endoscopic imaging technologies offer detailed internal views for various clinical scenarios, significantly improving the efficiency and accuracy of disease diagnosis and treatment. As technology continues to advance, these imaging techniques will play an even greater role in medical diagnostics.

3.2. Deep Learning Models

In medical image analysis, deep learning technologies, especially convolutional neural networks (CNNs)—have demonstrated tremendous potential. Deep learning mimics the human brain's ability to process complex data, such as images and sounds [27,28]. CNNs, with their multi-layer convolutional structures, automatically extract features from raw images. Convolutional layers use learnable filters (kernels) to capture local image features, while pooling layers reduce the size of feature maps to decrease computational complexity and retain essential information [29]. Non-linear activation functions (e.g., ReLU) further enhance the model's ability to recognize complex patterns [27].

In endoscopic image analysis, CNNs' automatic feature learning greatly improves the efficiency and accuracy of image processing. Studies have shown that CNNs can effectively perform tasks such as polyp detection, lesion classification, and region recognition. For example, a hybrid model combining CNNs with Swin Transformers was able to automatically segment polyps by learning from annotated endoscopic images, demonstrating high efficiency and accuracy when analyzing new images [30].

Additionally, deep learning has proven valuable in enhancing image quality, an important factor in improving the diagnostic value of endoscopic images. For example, using ImageNet-based technologies, researchers have improved image contrast and resolution, leading to better visualization of details and abnormalities [29]. EndoL2H, a model combining conditional adversarial networks with spatial attention, has been developed to achieve super-resolution endoscopic imaging, further enhancing diagnostic accuracy [31]. Deep learning-based denoising techniques have also shown excellent performance in reducing noise, advancing endoscopic image processing technology [32].

As CNN architectures and learning algorithms continue to be optimized, the accuracy and efficiency of endoscopic image analysis improve. For instance, deep residual networks (ResNet) have enhanced the training stability and performance of CNNs, yielding promising results in practical applications [33]. Multi-scale feature fusion and attention mechanisms have also contributed to better capturing subtle changes in endoscopic images [34].

Transfer learning techniques are another area of progress. By employing nucleus-centered image block sampling strategies and combining the classification results of multiple image blocks, researchers have improved robustness and accuracy in endoscopic image classification [35]. CNNs, therefore, play a pivotal role in improving image quality, detecting small lesions, and assisting doctors in making more accurate diagnoses [36,37].

In summary, the application of deep learning in endoscopic image analysis not only increases the automation level of image processing but also accelerates advancements in medical diagnostic technology. Continuous development and optimization of these technologies offer more accurate diagnostic support to clinicians, aiding early disease detection, treatment planning, and treatment monitoring [36].

3.3. Application of Deep Learning in Endoscopic Image Processing

To help readers quickly grasp the latest developments in deep learning in endoscopic image processing, Table 2 is an overview table showing the various application areas and their most popular models.

Table 2. Overview of Applications and Popular Models.

Application	Popular Models	Key Advances
Image segmentation	U-Net	Early cancer detection [38]
	SegNet	Polyp segmentation [39] Bleeding area monitoring [40]
Image Classification	ResNet-50	GERD classification [41]
	FCN	Nasopharyngeal malignancy [42] Early gastric cancer [43]
Image Enhancement	CNNs	Super-resolution [31]
	EndoL2H	Multi-image fusion [44] Low-light enhancement [45]
Dynamic Tracking and Analysis	SegNet	Polyp detection [46]
		Tumor classification [47] Surgical instrument tracking [48]
3D Reconstruction	SuperPoint	Colonoscopy reconstruction [49]
	Multi-view	Real-time 3D surface filling Tracking adaptive algorithms

3.3.1. Image Segmentation

Image segmentation is one of the core applications of deep learning in medical image analysis. Identifying and extracting specific regions or structures in images significantly improves the accuracy and efficiency of medical diagnosis. Deep learning models, especially CNNs, have been widely used in various medical image segmentation tasks, such as retinal layer segmentation, lung nodule detection, and cardiac image analysis. For example, a retinal segmentation method demonstrates the practicality of the model in processing actual medical images, especially in clinical environments where fast feedback is required, and efficiency improvement is crucial [50]. Another study proposed a CNN model optimized by multi-granularity visual features, which has made significant progress in improving the accuracy and speed of retinal layer segmentation [51]. The fusion of multi-granular features not only enhances the recognition ability of the model but also provides new perspectives for processing information at different levels. These medical image segmentation techniques provide valuable ideas and references for the segmentation of endoscopic images, laying the foundation for future work in this area.

In the field of gastrointestinal endoscopic image segmentation, deep learning technology has made substantial progress. In a review of the detection of precancerous lesions in upper gastrointestinal endoscopy, CNNs and transfer learning techniques were shown to be highly effective, with different models being evaluated for their efficacy in diagnosing precancerous lesions, thus highlighting the potential of deep learning in early cancer diagnosis [38]. This evaluation provides practical guidance for model selection in clinical practice.

The DeepPoly model uses a DoubleU-Net architecture which stacked one U-net network on another U-net network to perform polyp segmentation and classification. This approach efficiently utilizes semantic information to optimize the segmentation of different objects in medical images, achieving a mean dice-coefficient of 0.834 and 0.956 in segmentation for the Endotech challenge and Kvasir-SEG dataset [39]. This method, trained on a large-scale colonoscopy image dataset, demonstrates that the performance of deep learning models is closely tied to dataset size. Transfer learning alleviates the problem of insufficient data while maintaining high-precision detection effects. Using step-by-step transfer learning and five independent endoscopy datasets, a computer-assisted system was used to characterize tissues in NBI zoom images in Barrett's esophagus. The results overlapped with the expert ground truth, and the soft and sweet spot delineation scores reached 98% [52]. This result shows that transfer learning can maintain the detection accuracy of the model under limited data conditions, providing an effective solution for actual clinical applications.

Another example is a deep learning method based on UNet++ and VGG-16, used for accurately delineating the resection margins of early gastric cancer. By improving the accuracy testing method, which reconstructs the results on the endoscopic image and analyzes the distances between point pairs, this method significantly improved the delineation accuracy of the resection margin [53]. This progress reflects the importance of adaptability and flexibility in deep learning models, especially as application scenarios in medical imaging expand. The representative images of the resection extent are shown in Figure 1.

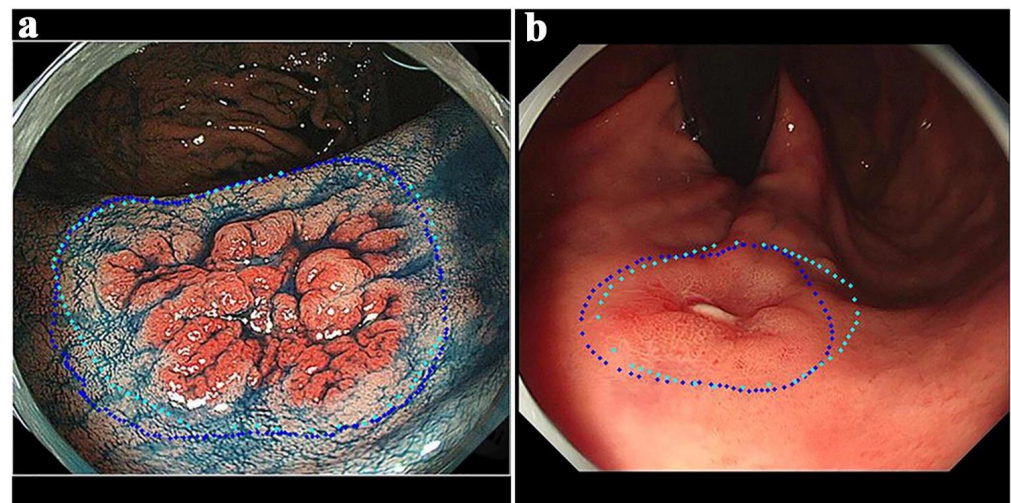


Figure 1. Representative images of the resection extent. (a) ENDOANGEL delineates the resection extent in the CE images. (b) ENDOANGEL delineates the resection extent in the WLE images. The dark dotted line is the predicted resection margin predicted. The light dotted line is the resection margin delineated by the expert (all images taken from [53]).

In addition, a semantic segmentation model based on SegNet, which is known for its high internal efficiency, has been applied for the detection of bleeding areas in capsule endoscopy images. Unlike traditional networks that output images in RGB, this model produces multiple channels for enhanced image interpretation. The training images were

labeled with three classes—bleeding, non-bleeding, and background—and the model significantly improved the detection accuracy of bleeding areas [40], demonstrating the potential of deep learning in processing complex medical images.

In respiratory endoscopic image analysis, a fully automatic vocal fold segmentation system based on deep learning maps three-channel RGB images into distance maps for training. This system analyzes vocal cord motion behaviors captured using flexible endoscopes with low-speed capability, segmenting the vocal cords and glottis areas in transnasal flexible laryngoscopy videos. The segmentation accuracy of this method reaches 0.9469, which is higher than 0.7358 of the encoder–decoder method [54]. This study reveals the significant potential of deep learning in deepening our understanding of complex physiological structures. Another study focused on the segmentation of airways and obstructive factors in endoscopic imaging, providing valuable insights for diagnosing and treating obstructive sleep apnea (OSA) [55]. However, the limited number of images in the test set may affect the generalizability of the model.

Though deep learning can perform image segmentation for various human endoscopy images, the occlusion of human tissues and organs by medical devices will affect the analysis and judgment when performing image analysis. For accurate segmentation of surgical instruments, DRR-Net, a dense-connected residual recurrent convolutional network, has been shown to perform well in instrument recognition within complex backgrounds [56]. This network adds Recurrent Neural Network (RNN) blocks and Adaptive Dilated Recurrent Convolutional Block (ADRCB) blocks to the traditional U-Net network backbone, allowing for time series analysis of video data. By combining data from previous frames and learning long-range context information, DRR-Net reflects the potential of deep learning to improve surgical safety and precision in real-world surgical scenarios. This feature suggests that the model can be integrated with a robotic system in the future to achieve fully autonomous surgical operations. Moreover, complex medical imaging environments can also pose challenges for image analysis. EndoUDA is a modality-independent deep learning segmentation method that leverages adaptive techniques and a joint loss function to improve generalization across different target domains, showing strong potential for wide-ranging applications in various endoscopic image modalities [57]. This research highlights deep learning's ability to find effective solutions across diverse medical imaging environments.

Artifact detection and segmentation are important challenges in deep learning applications. Artifacts are fundamental yet unavoidable problems in endoscopy, often complicating the detection of tissue abnormalities. Therefore, research on accelerating the identification of these categories and restoring frames is crucial [58]. An improved Cascade R-CNN model, combined with a feature pyramid network (FPN), has been employed for multi-class artifact detection. This combination improves the trade-off between mean average precision (mAP) and intersection over union (IoU), effectively enhancing artifact recognition and segmentation accuracy [59]. These studies highlight the importance of developing specialized techniques to handle artifacts, which in turn enhances the overall image quality and diagnostic level in endoscopy.

While expanding the types of segmentable images, it is also crucial to improve the integration of tasks. Multi-task deep learning models have also gained widespread use. A multi-task model that combines classification, image retrieval, and segmentation functions utilizes a mutual attention module to capture more diverse features during the segmentation task, significantly improving the accuracy and efficiency of endoscopic image analysis [60]. This type of multi-task learning approach not only enhances the overall performance of the model but also reduces the time and data required for training. Another model, DSI-Net, combines classification and segmentation tasks, significantly improving the multi-task processing capability of endoscopic images through deep collaborative interactive networks and attention mechanisms [61]. An example of DPENet segmentation results is shown in Figure 2.

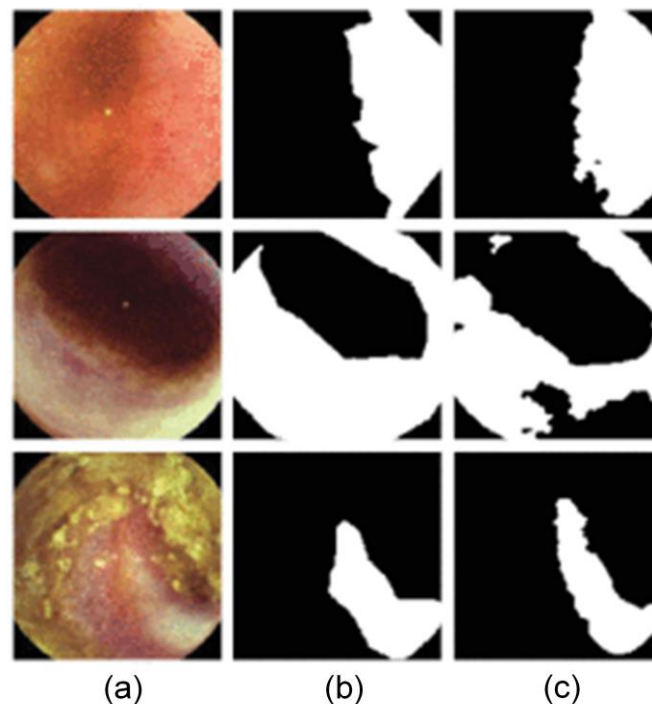


Figure 2. Example of DPENet segmentation results. White area: lesion. Black area: background. (a) The data collected by endoscopes (b) ground truth, (c) segmentation results of DSI-Net (all images taken from [61]).

In summary, these research findings demonstrate that the application of deep learning in endoscopic image segmentation not only improves the accuracy of lesion detection and segmentation but also drives the overall progress of endoscopic technology. The continuous development and application of these technologies provide robust support for clinical diagnosis and treatment, underlining the crucial role of deep learning in medical image analysis.

3.3.2. Image Classification

The application of deep learning techniques in endoscopic image classification has made significant progress. These techniques not only enhance the efficiency of diagnosis but also expand the scope of application for endoscopic technology. In the field of deep learning-based endoscopic image classification, CNNs have shown great potential [62].

For example, in the classification of gastroesophageal reflux disease (GERD), researchers designed a CNN architecture with high generalization ability, using a pre-trained model and applying dynamic data augmentation to the training set. This model classified GERD according to the Los Angeles classification (LA grade) system, achieving an accuracy of 89.3% [41]. The high generalization ability of the model was enhanced by the use of a data augmentation technique, a two-stage no-freezing fine-tuning policy, and an early stopping criterion technique [63].

Accurate detection of abnormalities is essential for effective diagnosis and treatment in medical imaging. A study utilized deep learning models consisting of a fully convolutional network based on the inception architecture to detect nasopharyngeal malignancies and verified the effectiveness and reliability of the model in real clinical settings, achieving an accuracy of 88.7% on the test set—7.5% higher than the expert in the prospective comparison phase [42]. Representative images of nasopharyngeal masses of different types are shown in Figure 3. The success of this study highlights the maturity of deep learning technology in the analysis of various endoscopic images and suggests that it may become a routine diagnostic tool in the future.

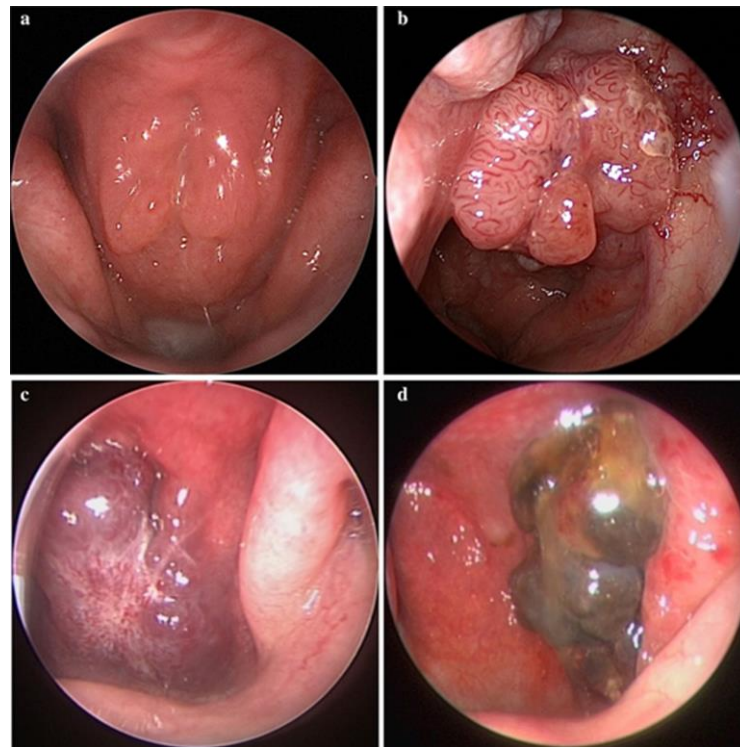


Figure 3. Representative images of nasopharyngeal masses. (a) Normal (adenoids hyperplasia); (b) nasopharyngeal carcinoma; (c) fibroangioma; (d) malignant melanoma (all images taken from [42]).

Another study developed a guided attention deep network based on ResNet-50 which combined a lightweight attention module and multi-scale feature extractor with U-Net. By extracting image features at different scales, the model significantly improved the recognition ability for early gastric cancer, achieving a classification accuracy of 92.2% [43]. However, this method's data processing range remains somewhat limited, and its recognition ability may be constrained when dealing with data from diverse regions. Ensemble learning can deal with this problem of poor robustness, another study found that the performance of pathological site classification could be enhanced through ensemble learning, a method using multiple network architectures and CNNs of different depths to effectively extract image texture features from input images. This technique not only compensates for the biases of individual models but also reduces error rates by integrating the prediction results of multiple models [64]. Building on this concept, Ghosh et al. (2023) further demonstrated that ensemble learning can successfully classify features that were incorrectly learned by individual base learners, improving accuracy to 95% [65].

Meanwhile, by adding a global spatial feature aggregation block to the ResNet-50 backbone, researchers were able to accurately aggregate spatial features, demonstrating excellent classification performance [66]. This improvement is particularly important for the positioning and analysis of complex tissues. In the future, it can be integrated with the smart capsule endoscopy system to achieve non-invasive detection and monitoring. A recent study by Mukhtorov et al. (2023) improved the transparency of endoscopic image classification through the development of interpretable deep learning models. This innovation is particularly important for enhancing clinicians' trust in AI-powered diagnostic tools [67].

In summary, the combination of different models and methods not only improves classification accuracy but also broadens the potential for clinical application in endoscopic diagnosis.

3.3.3. Image Enhancement

Image enhancement is a crucial application of deep learning in medical image processing, particularly in endoscopic image analysis. By improving the quality of images,

enhancement techniques enable doctors to identify lesions more accurately, thereby increasing diagnostic accuracy and efficiency. In recent years, significant progress has been made in this field through deep learning methods.

For colonoscopy images, the endoscopic image enhancement network (EIEN) based on Retinex theory has demonstrated successful illumination correction using a self-attention-guided multi-scale pyramid network. This approach stretches the green and blue channels in the reflection component and fuses the enhanced reflection component with the original image through weighting, thereby maintaining both image fidelity and the contrast of blood vessels and tissues [68]. However, this method relies on supervised training and requires a large amount of labeled data support. To solve the problem of insufficient labeled data, the application of unsupervised learning methods in image enhancement is increasing. Unsupervised deep learning methods self-learn image features to improve brightness and alleviate color distortion, providing important support for clinical decision-making [69]. Another deep unsupervised method for endoscopic image enhancement uses multi-image fusion technology to further improve image quality. This technique converts three derivative images generated from different enhancement methods into the HSI color space, then feeds the I channel image into the DerivedFuse network, which employs a novel reference-free quality metric as its loss function. The network accurately extracts and fuses features to enhance the intensity component of the original image. After enhancement, doctors can select images based on different parameter settings to suit their clinical needs [44], as shown in Figure 4. By fusing information from multiple images, this method significantly improves image detail and contrast, providing clearer visual information for doctors. The successful application of unsupervised learning highlights deep learning's potential in addressing the challenge of limited annotated medical data, particularly when data labeling is insufficient.

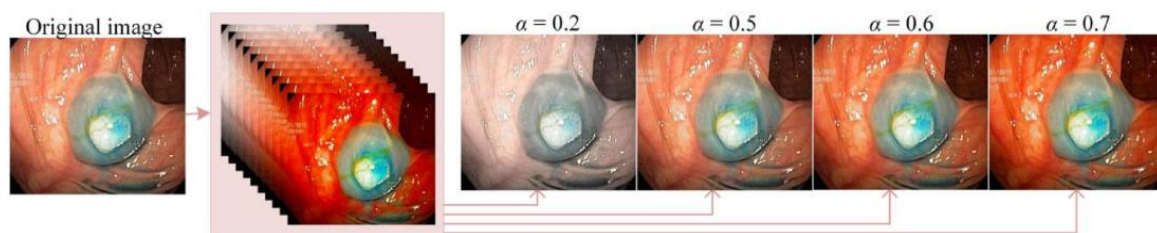


Figure 4. The visual effect of different α (parameter variable in adaptive adjustment function for changing saturation) values on image (all images taken from [44]).

Additionally, EndoL2H, a deep-learning-based super-resolution method for capsule endoscopy images, further exemplifies the importance of enhancing low-quality medical images through advanced techniques. This framework combines a conditional GAN with spatial attention to generate diagnostically relevant high-resolution images from low-resolution counterparts, ensuring the preservation of critical texture and details, even under extreme up-scaling factors. The system demonstrates excellent performance, with significant improvements in PSNR and SSIM metrics under high scaling, thereby supporting early diagnosis through enhanced visual clarity. Its application underscores the value of super-resolution in endoscopy, where limited resolution and image quality can hinder diagnostic performance [31].

Additionally, deep learning models have been applied to low-light environments, such as high-speed endoscopic videos, to improve image quality in complex conditions, the improvement is significant in terms of PSNR and SSIM scores, reaching 23.79 dB and 0.88, respectively [70]. The novel deep pyramid enhancement network (DPENet), derived from deep pyramid networks (DPNs), integrates global and local features at various scales through an image pyramid framework with three parallel branches. This method has proven effective in enhancing image contrast and details under low-light conditions [45]. Visual results from the low-light image enhancement (LIE) methods applied to the dataset

are shown in Figure 5. In further developments, automated frameworks have been introduced to streamline the restoration and enhancement of endoscopic images, ensuring fast processing while meeting the demands of rapid clinical diagnosis [71]. Electromagnetic interference noise removal is another key aspect of image enhancement, and researchers have developed deep learning algorithms to effectively reduce noise and enhance image clarity [72].

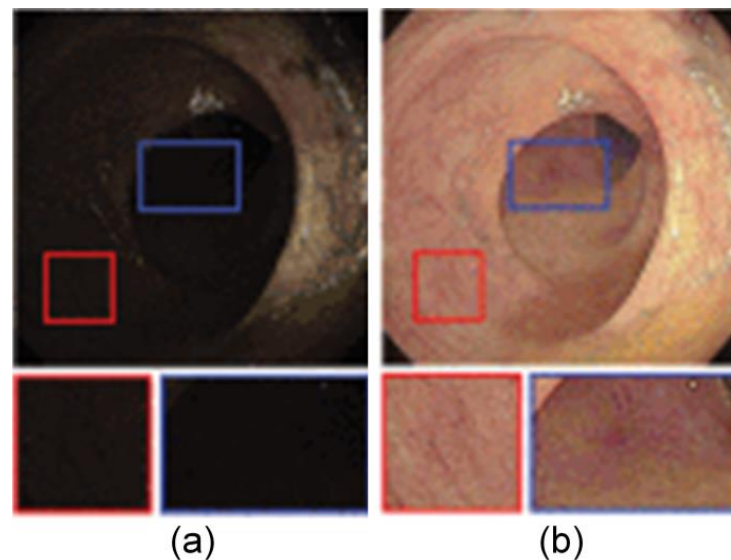


Figure 5. Visual results of LIE methods on the dataset. (a) Input image; (b) DPENet. In the blue and red boxes are the corresponding DOI fields (all images taken from [45]).

Finally, the application of deep learning to early gastric cancer detection has further demonstrated the practical value of image enhancement technology in clinical practice [73]. In the future, these models can be integrated into surgical navigation systems and telemedicine platforms to provide real-time and accurate diagnostic support and help doctors cope with complex lesion analysis tasks. By continuously optimizing model structures and algorithms, researchers are advancing the field of medical image analysis, broadening the future applications of deep learning in healthcare.

In summary, deep learning technology has made considerable progress in endoscopic image enhancement. By combining various methods and technologies, researchers are consistently improving image quality, diagnostic efficiency, and accuracy.

3.3.4. Dynamic Tracking and Analysis

Deep learning technology has shown tremendous potential for dynamic tracking and analysis in endoscopic imaging, particularly in real-time monitoring and precise lesion localization. As deep learning algorithms continue to evolve, researchers have significantly enhanced the efficiency and accuracy of endoscopic image processing by developing innovative models, offering more reliable support for clinical diagnosis and treatment.

In colonoscopy, real-time endoscopic image diagnosis support systems based on deep learning have proven effective in identifying and classifying colon lesions. These systems not only enhance physicians' work efficiency but also improve the detection rate of polyps and other lesions, highlighting the value of deep learning in dynamic endoscopic tracking [46,74]. For gastric tumor detection, a deep learning-based clinical decision support system (CDSS) was developed to automatically detect and classify tumors during real-time endoscopy, assisting in diagnosis and invasion depth prediction, with a lesion detection rate of 95.6% [47]. In the future, this system is expected to be further integrated into remote consultation and robot-assisted surgery systems to promote the popularization and development of high-quality medical services. Figure 6 presents representative identification results generated by the CDSS.



Figure 6. Representative examples of the clinical decision support system (CDSS). Green boxes show polyps detected by algorithm. (all images taken from [47]).

In capsule endoscopy, an automatic detection system has been developed that identifies abnormal lesions such as mucosal breaks, angioectasia, protruding lesions, and blood content in real-time [75]. By integrating Single Shot MultiBox Detector (SSD) and ResNet50, this system significantly improves the comprehensiveness of detection functions. Moreover, deep learning models have demonstrated great effectiveness in the automatic detection of early esophageal squamous cell carcinoma, with real-time monitoring and identification of abnormal esophageal lesions, thereby improving early diagnosis accuracy [76].

Regarding dynamic tracking of surgical instruments, researchers have developed deep learning-assisted robotic endoscope systems capable of tracking and locating surgical instruments in real-time, improving both the safety and precision of surgical procedures [48,77]. These advancements collectively promote the progress of endoscopic technology.

In conclusion, the application of deep learning in dynamic tracking and analysis of endoscopic images continues to evolve. By utilizing a range of deep learning models and optimized methods, the efficiency of real-time analysis has been significantly improved, offering valuable guidance for both diagnosis and surgical applications.

3.3.5. Three-Dimensional Reconstruction

Deep learning technology has shown great promise in the three-dimensional (3D) reconstruction of endoscopic images. Converting two-dimensional (2D) image data into 3D models allows doctors to gain a more intuitive understanding of internal structures, thereby supporting more accurate diagnoses and treatment planning and facilitating the development of assistive tools for practitioners [78]. In recent years, significant advancements in deep learning algorithms, particularly in CNNs and multi-view stereo vision technology, have led to remarkable progress in 3D reconstruction research for endoscopic imaging.

Advancements in tracking adaptive algorithms have enhanced the accuracy of detecting and matching feature points, optimizing 3D reconstruction with the improved SuperPoint algorithm by designing a new tracking loss. This technology improves both the speed and accuracy of endoscopic 3D reconstruction, providing doctors with real-time 3D image support, which increases the number of reconstruction points and improves stability [78]. However, the robustness of the algorithm in dealing with occlusions and missing areas still needs to be further improved.

To address this shortcoming, real-time colonoscopic 3D reconstruction technology has greatly enhanced the accuracy and efficiency of colonoscopy by dynamically reconstructing the 3D surface of the examined areas. This technology can effectively identify and fill in missing areas in the image, optimizing the doctor's field of view and improving operational precision [49]. This is particularly advantageous when examining complex or abnormal anatomical structures, where the ability to reconstruct in 3D provides irreplaceable benefits. The results of the reconstruction are shown in Figure 7.



Figure 7. Three-dimensional reconstruction for visualization of missing colonic surface (highlighted in black in the last image, 25% surface), small colon pouches that are occluded by ridges. The missing colonic surface is highlighted in black in the last image. (all images taken from [49]).

In summary, the application of deep learning in 3D reconstruction for endoscopic images is continuously evolving. The integration of multiple techniques has significantly improved both the accuracy and efficiency of these reconstructions. These technological advancements not only offer new perspectives for endoscopic image analysis but also provide more intuitive, real-time support for clinical applications.

4. Results

4.1. Summary of Findings

In this Results Section, we provide a comprehensive overview of the findings regarding the application of Deep Learning (DL) in enhancing endoscopic image processing. The results emphasize the positive impact of DL techniques across various applications, highlighting their effectiveness in improving diagnostic accuracy and operational efficiency. Table 3 shows the summary of findings on DL in endoscopic image processing.

Table 3. Summary of findings on DL in endoscopic image processing.

Findings	Description
Application Areas	DL is utilized in various areas, including classification, reconstruction, image enhancement, segmentation, and real-time pathology recognition across different types of endoscopic procedures, such as gastrointestinal and respiratory endoscopy.
Performance Metrics	Studies have shown that DL algorithms can enhance accuracy rates in lesion detection, particularly in identifying precancerous lesions and small tumors.
Impact on Workflow	DL systems have reduced the processing time for real-time analysis, allowing clinicians to make faster and more informed decisions during procedures.
Comparison with Traditional Methods	DL methods generally outperform traditional algorithms, such as manual review and conventional image processing techniques, in terms of speed and accuracy, particularly in polyp detection.
Quality of Image Output	Enhanced image quality through techniques like super-resolution has resulted in clearer visualization of details, aiding in more accurate diagnoses.

4.2. Strengths

In this section, we highlight the significant advantages and strengths of employing DL in endoscopic image processing, underscoring its contributions to modern medical diagnostics. Table 4 shows strengths of DL in endoscopic image processing.

Table 4. Strengths of DL in endoscopic image processing.

Strengths	Description
Enhanced Diagnostic Accuracy	DL techniques have been shown to significantly improve diagnostic accuracy, particularly in detecting early-stage malignancies and subtle anatomical abnormalities that might be overlooked by human observers.
Automation of Image Analysis	By automating repetitive tasks like image segmentation and classification, DL reduces the workload for clinicians, allowing them to focus on patient care and complex decision-making.
Real-time Processing Capabilities	DL systems facilitate real-time analysis, which is critical during endoscopic procedures, enabling immediate feedback and decision support to clinicians.
Adaptability and Continuous Learning	Many DL models can adapt to new data inputs and learn from ongoing clinical experiences, improving their accuracy and reliability over time.
Potential for Personalized Medicine	With ongoing advancements, DL technologies hold the promise of enabling personalized medicine by tailoring diagnostic and therapeutic approaches to individual patient profiles.

4.3. Limitations

This section outlines the challenges and constraints observed in the application of DL in endoscopic image processing. These limitations provide critical insights into areas that require further research and improvement. Table 5 shows limitations of DL in endoscopic image processing.

Table 5. Limitations of DL in endoscopic image processing.

Limitations	Description
Data Dependency	A major limitation is the high dependency on large, annotated datasets for training DL models. Many institutions face challenges in obtaining sufficiently large and diverse datasets, which may lead to biased model performance.
Generalization Issues	DL models may struggle to generalize across different endoscopic modalities and variations in imaging protocols, which can affect diagnostic accuracy when applied to new settings.
Interpretability	The “black box” nature of DL algorithms presents challenges in clinical settings, as clinicians often find it difficult to interpret the rationale behind AI-generated outputs, which is crucial for making informed medical decisions.
Computational Requirements	The computational resources needed for training and implementing DL models are substantial, often requiring advanced hardware that may not be readily available in all healthcare settings.

5. Discussion

The application of deep learning to endoscopic image processing represents a transformative advancement in medical diagnostics. Techniques based on convolutional neural networks have significantly improved the resolution, clarity, and diagnostic accuracy of endoscopic images. Through feature extraction, pattern recognition, and automated segmentation, deep learning models are surpassing the limitations of traditional image processing methods, offering faster and more reliable insights for clinical use.

Deep learning models, such as CNNs, excel at interpreting complex visual data, automatically identifying subtle patterns that may elude human perception. This capability is especially valuable for detecting early-stage malignancies or subtle abnormalities that could otherwise go unnoticed. The automation of detection and classification processes minimizes the risk of human error, leading to more accurate diagnoses and improved clinical outcomes. Furthermore, methods like EndoL2H leverage super-resolution techniques to enhance image quality, ensuring that even low-resolution endoscopic images retain diagnostically critical details [31]. Similarly, multi-image fusion approaches, as described by Huang et al. (2022), improve image brightness and color accuracy through unsupervised learning techniques, supporting enhanced visualization in clinical settings [64].

Despite these advancements, several challenges hinder the widespread adoption of deep learning in clinical practice. A major concern is the computational demand required to train and operate these models. Training deep networks requires substantial processing power and large-scale datasets, which may not always be available in medical institutions. This challenge is particularly relevant in healthcare settings where data scarcity is common, making it difficult to train models effectively using supervised learning approaches. As Hamid (2023) points out, data-centric AI approaches, which emphasize optimizing datasets alongside models, are critical for addressing these limitations [79].

Another challenge lies in ensuring that deep learning models maintain robustness across different clinical environments. Variability in endoscopic equipment, patient demographics, and imaging protocols can lead to domain shifts, which may negatively affect model performance. Developing models that are generalizable and capable of maintaining accuracy across diverse datasets remains a priority. Data-centric solutions, as emphasized by Hamid and Braun (2019), suggest that focusing on creating adaptable, high-quality datasets is essential for improving the scalability of AI models in healthcare [80].

Additionally, the “black box” nature of many deep learning models poses interpretability challenges. In medical diagnostics, transparency and accountability are essential, as clinicians need to understand the rationale behind AI-generated outputs to make informed decisions. While interpretability-focused models, such as those developed by Mukhtorov et al. (2023), provide some solutions, further research is needed to balance model complexity with explainability [61].

Despite these challenges, the potential of deep learning to alleviate the workload of clinicians is immense. Automating routine tasks allows healthcare providers to focus more on patient care, ultimately improving the efficiency of healthcare delivery. By integrating deep learning models into diagnostic workflows, clinicians can benefit from faster and more precise diagnoses, leading to better patient outcomes. As advancements continue, the field will need to prioritize both computational efficiency and transparency to ensure that deep learning tools are seamlessly integrated into clinical practice.

In conclusion, while the integration of deep learning into endoscopic image analysis holds great promise, addressing the challenges of computational demands, data scarcity, and interpretability will be crucial for its success. By adopting strategies that balance data-centric and model-centric approaches, researchers can develop more efficient and adaptable models that meet the demands of modern healthcare systems [79,80].

6. Conclusions

This paper provides a comprehensive review of deep learning applications in endoscopic image processing, covering advancements in image enhancement, segmentation,

classification, dynamic tracking, and 3D reconstruction. Key contributions include emphasizing the impact of super-resolution and multi-image fusion techniques, exploring unsupervised learning to address data scarcity, and highlighting model scalability across diverse clinical settings.

The analysis identifies future directions, including developing lightweight, interpretable models for efficient clinical use and exploring data-efficient techniques like domain adaptation to address limited annotated datasets. By aligning emerging AI trends with clinical needs, this work offers a roadmap for future research, aiming to enhance diagnostic accuracy, improve patient care, and drive the integration of deep learning into routine medical practice.

Author Contributions: Writing—original draft preparation, Z.N.; writing—review and editing, M.X., Z.W., X.L. and W.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Youth Project of Natural Science Foundation of Shandong Province, China (ZR2023QC262); the National Natural Science Foundation of China (62205181); the Natural Science Foundation of Shandong Province (ZR2022QF017); and the Shandong Province Outstanding Youth Science Fund Project (Overseas) (2023HWYQ-023).

Data Availability Statement: This is a review of published articles, and all the articles are publicly available.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Dilek, E.; Dener, M. Computer Vision Applications in Intelligent Transportation Systems: A Survey. *Sensors* **2023**, *23*, 2938. [\[CrossRef\]](#)
2. Boroujeni, S.P.H.; Razi, A. IC-GAN: An Improved Conditional Generative Adversarial Network for RGB-to-IR image translation with applications to forest fire monitoring. *Expert Syst. Appl.* **2024**, *238*, 121962. [\[CrossRef\]](#)
3. Valdastrì, P.; Simi, M.; Webster, R.J. Advanced Technologies for Gastrointestinal Endoscopy. *Annu. Rev. Biomed. Eng.* **2012**, *14*, 397–429. [\[CrossRef\]](#)
4. Kuipers, E.; Haringsma, J. Diagnostic and therapeutic endoscopy. *J. Surg. Oncol.* **2005**, *92*, 203–209. [\[CrossRef\]](#)
5. Rashed, B.M.; Popescu, N. Critical Analysis of the Current Medical Image-Based Processing Techniques for Automatic Disease Evaluation: Systematic Literature Review. *Sensors* **2022**, *22*, 7065. [\[CrossRef\]](#)
6. Esteva, A.; Robicquet, A.; Ramsundar, B.; Kuleshov, V.; DePristo, M.; Chou, K.; Cui, C.; Corrado, G.; Thrun, S.; Dean, J. A guide to deep learning in healthcare. *Nat. Med.* **2019**, *25*, 24–29. [\[CrossRef\]](#)
7. Shen, D.; Wu, G.; Suk, H.-I. Deep Learning in Medical Image Analysis. *Annu. Rev. Biomed. Eng.* **2017**, *19*, 221–248. [\[CrossRef\]](#)
8. Choi, J.; Shin, K.; Jung, J.; Bae, H.-J.; Kim, D.H.; Byeon, J.-S.; Kim, N. Convolutional Neural Network Technology in Endoscopic Imaging: Artificial Intelligence for Endoscopy. *Clin. Endosc.* **2020**, *53*, 117–126. [\[CrossRef\]](#)
9. Pannala, R.; Krishnan, K.; Melson, J.; Parsi, M.A.; Schulman, A.R.; Sullivan, S.; Trikudanathan, G.; Trindade, A.J.; Watson, R.R.; Maple, J.T.; et al. Artificial intelligence in gastrointestinal endoscopy. *VideoGIE* **2020**, *5*, 598–613. [\[CrossRef\]](#)
10. Alagappan, M.; Brown, J.R.G.; Mori, Y.; Berzin, T.M. Artificial intelligence in gastrointestinal endoscopy: The future is almost here. *World J. Gastrointest. Endosc.* **2018**, *10*, 239–249. [\[CrossRef\]](#)
11. Chahal, D.S.N.; Byrne, M.F. AI and Endoscopy: Future Perspectives. *Digit. Surg.* **2021**, *2021*, 319–338. [\[CrossRef\]](#)
12. Zhuang, H.; Zhang, J.; Liao, F. A systematic review on application of deep learning in digestive system image processing. *Vis. Comput.* **2023**, *39*, 2207–2222. [\[CrossRef\]](#)
13. Rex, D.K.; Helbig, C.C. High Yields of Small and Flat Adenomas with High-Definition Colonoscopes Using Either White Light or Narrow Band Imaging. *Gastroenterology* **2007**, *133*, 42–47. [\[CrossRef\]](#)
14. Panteris, V.; Nikolopoulou, S.; Lountou, A.; Triantafyllidis, J.K. Diagnostic capabilities of high-definition white light endoscopy for the diagnosis of gastric intestinal metaplasia and correlation with histologic and clinical data. *Eur. J. Gastroenterol. Hepatol.* **2014**, *26*, 594–601. [\[CrossRef\]](#)
15. Takemura, Y.; Yoshida, S.; Tanaka, S.; Onji, K.; Oka, S.; Tamaki, T.; Kaneda, K.; Yoshihara, M.; Chayama, K. Quantitative analysis and development of a computer-aided system for identification of regular pit patterns of colorectal lesions. *Gastrointest. Endosc.* **2010**, *72*, 1047–1051. [\[CrossRef\]](#)
16. Mintz, G.; Nissen, S.; Anderson, W.; Bailey, S.; Erbel, R.; Fitzgerald, P.; Pinto, F.; Rosenfield, K.; Siegel, R.; Tuzcu, E.; et al. American College of Cardiology Clinical Expert Consensus Document on Standards for Acquisition, Measurement and Reporting of Intravascular Ultrasound Studies (IVUS). *J. Am. Coll. Cardiol.* **2001**, *37*, 1478–1492. [\[CrossRef\]](#)
17. Potkin, B.N.; Bartorelli, A.L.; Gessert, J.M.; Neville, R.F.; Almagor, Y.; Roberts, W.C.; Leon, M.B. Coronary artery imaging with intravascular high-frequency ultrasound. *Circulation* **1990**, *81*, 1575–1585. [\[CrossRef\]](#)

18. Paulo, M.; Sandoval, J.; Lennie, V.; Dutary, J.; Medina, M.; Gonzalo, N.; Jimenez-Quevedo, P.; Escaned, J.; Bañuelos, C.; Hernandez, R.; et al. Combined Use of OCT and IVUS in Spontaneous Coronary Artery Dissection. *JACC Cardiovasc. Imaging* **2013**, *6*, 830–832. [\[CrossRef\]](#)
19. Li, X.; Li, J.; Jing, J.; Ma, T.; Liang, S.; Zhang, J.; Mohar, D.; Raney, A.; Mahon, S.; Brenner, M.; et al. Integrated IVUS-OCT Imaging for Atherosclerotic Plaque Characterization. *IEEE J. Sel. Top. Quantum Electron.* **2013**, *20*, 196–203. [\[CrossRef\]](#)
20. Gan, S.I.; Rajan, E.; Adler, D.G.; Baron, T.H.; Anderson, M.A.; Cash, B.D.; Davila, R.E.; Dominitz, J.A.; Harrison, M.E.; Ikenberry, S.O.; et al. Role of EUS. *Gastrointest. Endosc.* **2007**, *66*, 425–434. [\[CrossRef\]](#)
21. Yousaf, M.N.; Chaudhary, F.S.; Ehsan, A.; Suarez, A.L.; Muniraj, T.; Jamidar, P.; Aslanian, H.R.; Farrell, J.J. Endoscopic ultrasound (EUS) and the management of pancreatic cancer. *BMJ Open Gastroenterol.* **2020**, *7*, e000408. [\[CrossRef\]](#)
22. Huang, D.; Swanson, E.A.; Lin, C.P.; Schuman, J.S.; Stinson, W.G.; Chang, W.; Hee, M.R.; Flotte, T.; Gregory, K.; Puliafito, C.A.; et al. Optical Coherence Tomography. *Science* **1991**, *254*, 1178–1181. [\[CrossRef\]](#)
23. Kirsten, L.; Schindler, M.; Morgenstern, J.; Erkkilä, M.T.; Golde, J.; Walther, J.; Rottmann, P.; Kemper, M.; Bornitz, M.; Neudert, M.; et al. Endoscopic optical coherence tomography with wide field-of-view for the morphological and functional assessment of the human tympanic membrane. *J. Biomed. Opt.* **2019**, *24*, 031017. [\[CrossRef\]](#)
24. Ang, M.; Tan, A.C.S.; Cheung, C.M.G.; Keane, P.A.; Dolz-Marco, R.; Sng, C.C.A.; Schmetterer, L. Optical coherence tomography angiography: A review of current and future clinical applications. *Graefes Arch. Clin. Exp. Ophthalmol.* **2018**, *256*, 237–245. [\[CrossRef\]](#)
25. Gono, K. Narrow Band Imaging: Technology Basis and Research and Development History. *Clin. Endosc.* **2015**, *48*, 476–480. [\[CrossRef\]](#)
26. Barbeiro, S.; Libânio, D.; Castro, R.; Dinis-Ribeiro, M.; Pimentel-Nunes, P. Narrow-Band Imaging: Clinical Application in Gastrointestinal Endoscopy. *GE Port. J. Gastroenterol.* **2018**, *26*, 40–53. [\[CrossRef\]](#)
27. Kim, K.G. Book Review: Deep Learning. *Health Inform. Res.* **2016**, *22*, 351–354. [\[CrossRef\]](#)
28. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. *Nature* **2015**, *521*, 436–444. [\[CrossRef\]](#)
29. Krizhevsky, A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [\[CrossRef\]](#)
30. Park, K.-B.; Lee, J.Y. SwinE-Net: Hybrid deep learning approach to novel polyp segmentation using convolutional neural network and Swin Transformer. *J. Comput. Des. Eng.* **2022**, *9*, 616–632. [\[CrossRef\]](#)
31. Almalioglu, Y.; Ozyoruk, K.B.; Gokce, A.; Incetan, K.; Gokceler, G.I.; Simsek, M.A.; Ararat, K.; Chen, R.J.; Durr, N.J.; Mahmood, F.; et al. EndoL2H: Deep Super-Resolution for Capsule Endoscopy. *IEEE Trans. Med. Imaging* **2020**, *39*, 4297–4309. [\[CrossRef\]](#)
32. Kaur, A.; Kaur, G. A review on image enhancement with deep learning approach. *Accent. Trans. Image Process. Comput. Vis.* **2018**, *4*, 16–20. [\[CrossRef\]](#)
33. Shin, Y.; Qadir, H.A.; Aabakken, L.; Bergsland, J.; Balasingham, I. Automatic Colon Polyp Detection Using Region Based Deep CNN and Post Learning Approaches. *IEEE Access* **2018**, *6*, 40950–40962. [\[CrossRef\]](#)
34. Wang, S.; Cong, Y.; Zhu, H.; Chen, X.; Qu, L.; Fan, H.; Zhang, Q.; Liu, M. Multi-Scale Context-Guided Deep Network for Automated Lesion Segmentation with Endoscopy Images of Gastrointestinal Tract. *IEEE J. Biomed. Health Inform.* **2020**, *25*, 514–525. [\[CrossRef\]](#)
35. Zhang, L.; Lu, L.; Nogues, I.; Summers, R.M.; Liu, S.; Yao, J. DeepPap: Deep Convolutional Networks for Cervical Cell Classification. *IEEE J. Biomed. Health Inform.* **2017**, *21*, 1633–1643. [\[CrossRef\]](#)
36. Ronneberger, O.F.P.; Brox, T. U-net: Convolutional networks for biomedical image segmentation. In Proceedings of the Medical Image Computing and Computer-Assisted Intervention—MICCAI 2015: 18th International Conference, Munich, Germany, 5–9 October 2015; Proceedings, Part III 18. Springer International Publishing: Berlin/Heidelberg, Germany, 2015; pp. 234–241. [\[CrossRef\]](#)
37. Litjens, G.; Kooi, T.; Bejnordi, B.E.; Setio, A.A.A.; Ciompi, F.; Ghafoorian, M.; van der Laak, J.A.W.M.; van Ginneken, B.; Sánchez, C.I. A survey on deep learning in medical image analysis. *Med. Image Anal.* **2017**, *42*, 60–88. [\[CrossRef\]](#)
38. Yan, T.; Wong, P.K.; Qin, Y.-Y. Deep learning for diagnosis of precancerous lesions in upper gastrointestinal endoscopy: A review. *World J. Gastroenterol.* **2021**, *27*, 2531–2544. [\[CrossRef\]](#)
39. Hossain, S.; Rahman, M.; Syeed, M.M.; Uddin, M.F.; Hasan, M.; Hossain, A.; Ksibi, A.; Jamjoom, M.M.; Ullah, Z.; Samad, A. DeepPoly: Deep Learning-Based Polyps Segmentation and Classification for Autonomous Colonoscopy Examination. *IEEE Access* **2023**, *11*, 95889–95902. [\[CrossRef\]](#)
40. Ghosh, T.; Li, L.; Chakareski, J. Effective deep learning for semantic segmentation based bleeding zone detection in capsule endoscopy images. In Proceedings of the 2018 25th IEEE International Conference on Image Processing (ICIP), Athens, Greece, 7–10 October 2018; pp. 3034–3038. [\[CrossRef\]](#)
41. Lee, S.H.; Jang, B.I.; Kim, K.O.; Jeon, S.W.; Kwon, J.G.; Kim, E.Y.; Jung, J.T.; Park, K.S.; Cho, K.B.; Park, C.G.; et al. Endoscopic Experience Improves Interobserver Agreement in the Grading of Esophagitis by Los Angeles Classification: Conventional Endoscopy and Optimal Band Image System. *Gut Liver* **2014**, *8*, 154–159. [\[CrossRef\]](#)
42. Li, C.; Jing, B.; Ke, L.; Li, B.; Xia, W.; He, C.; Qian, C.; Zhao, C.; Mai, H.; Chen, M.; et al. Development and validation of an endoscopic images-based deep learning model for detection with nasopharyngeal malignancies. *Cancer Commun.* **2018**, *38*, 59. [\[CrossRef\]](#)

43. Ma, L.; Su, X.; Ma, L.; Gao, X.; Sun, M. Deep learning for classification and localization of early gastric cancer in endoscopic images. *Biomed. Signal Process. Control* **2023**, *79*, 104200. [\[CrossRef\]](#)
44. Huang, D.; Liu, J.; Zhou, S.; Tang, W. Deep unsupervised endoscopic image enhancement based on multi-image fusion. *Comput. Methods Programs Biomed.* **2022**, *221*, 106800. [\[CrossRef\]](#)
45. Yue, G.; Gao, J.; Cong, R.; Zhou, T.; Li, L.; Wang, T. Deep Pyramid Network for Low-Light Endoscopic Image Enhancement. *IEEE Trans. Circuits Syst. Video Technol.* **2023**, *34*, 3834–3845. [\[CrossRef\]](#)
46. Yamada, M.; Saito, Y.; Imaoka, H.; Saiko, M.; Yamada, S.; Kondo, H.; Takamaru, H.; Sakamoto, T.; Sese, J.; Kuchiba, A.; et al. Development of a real-time endoscopic image diagnosis support system using deep learning technology in colonoscopy. *Sci. Rep.* **2019**, *9*, 14465. [\[CrossRef\]](#)
47. Gong, E.J.; Bang, C.S.; Lee, J.J.; Baik, G.H.; Lim, H.; Jeong, J.H.; Choi, S.W.; Cho, J.; Kim, D.Y.; Bin Lee, K.; et al. Deep learning-based clinical decision support system for gastric neoplasms in real-time endoscopy: Development and validation study. *Endoscopy* **2023**, *55*, 701–708. [\[CrossRef\]](#)
48. Cheng, T.; Li, W.; Ng, W.Y.; Huang, Y.; Li, J.; Ng, S.H.; Chiu, P.W.Y.; Li, Z. Deep Learning Assisted Robotic Magnetic Anchored and Guided Endoscope for Real-Time Instrument Tracking. *IEEE Robot. Autom. Lett.* **2021**, *6*, 3979–3986. [\[CrossRef\]](#)
49. Ma, R.; Wang, R.; Pizer, S.; Rosenman, J.; McGill, S.K.; Frahm, J.M. Real-time 3D reconstruction of colonoscopic surfaces for determining missing regions. In Proceedings of the Medical Image Computing and Computer Assisted Intervention–MICCAI 2019: 22nd International Conference, Shenzhen, China, 13–17 October 2019; Proceedings, Part V 22. Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 573–582. [\[CrossRef\]](#)
50. He, X.; Song, W.; Wang, Y.; Poiesi, F.; Yi, J.; Desai, M.; Xu, Q.; Yang, K.; Wan, Y. Light-weight Retinal Layer Segmentation with Global Reasoning. *IEEE Trans. Instrum. Meas.* **2024**, *73*, 2520214. [\[CrossRef\]](#)
51. He, X.; Wang, Y.; Poiesi, F.; Song, W.; Xu, Q.; Feng, Z.; Wan, Y. Exploiting multi-granularity visual features for retinal layer segmentation in human eyes. *Front. Bioeng. Biotechnol.* **2023**, *11*, 1191803. [\[CrossRef\]](#)
52. de Groof, A.J.; Struyvenberg, M.R.; van der Putten, J.; van der Sommen, F.; Fockens, K.N.; Curvers, W.L.; Zinger, S.; Pouw, R.E.; Coran, E.; Baldaque-Silva, F.; et al. Deep-Learning System Detects Neoplasia in Patients with Barrett’s Esophagus with Higher Accuracy Than Endoscopists in a Multistep Training and Validation Study with Benchmarking. *Gastroenterology* **2020**, *158*, 915–929.e4. [\[CrossRef\]](#)
53. An, P.; Yang, D.; Wang, J.; Wu, L.; Zhou, J.; Zeng, Z.; Huang, X.; Xiao, Y.; Hu, S.; Chen, Y.; et al. A deep learning method for delineating early gastric cancer resection margin under chromoendoscopy and white light endoscopy. *Gastric Cancer* **2020**, *23*, 884–892. [\[CrossRef\]](#)
54. Hamad, A.; Haney, M.; Lever, T.E.; Bunyak, F. Automated segmentation of the vocal folds in laryngeal endoscopy videos using deep convolutional regression networks. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops, Long Beach, CA, USA, 16–17 June 2019.
55. Kang, Y.H.; Kim, J.Y.; Kim, Y.J.; Kim, S.H.; Kim, K.G.; Rhee, C.-S. Airway and Airway Obstruction Site Segmentation Study Using U-Net with Drug-Induced Sleep Endoscopy Images. *J. Imaging Inform. Med.* **2024**. [\[CrossRef\]](#)
56. Yang, L.; Gu, Y.; Bian, G.; Liu, Y. DRR-Net: A Dense-Connected Residual Recurrent Convolutional Network for Surgical Instrument Segmentation from Endoscopic Images. *IEEE Trans. Med. Robot. Bionics* **2022**, *4*, 696–707. [\[CrossRef\]](#)
57. Celik, N.; Ali, S.; Gupta, S.; Braden, B.; Rittscher, J. Endouda: A modality independent segmentation approach for endoscopy imaging. In Proceedings of the 24th International Conference, Strasbourg, France, 27 September–1 October 2021; Springer International Publishing: Cham, Switzerland, 2021; pp. 303–312. [\[CrossRef\]](#)
58. Ali, S.; Zhou, F.; Braden, B.; Bailey, A.; Yang, S.; Cheng, G.; Zhang, P.; Li, X.; Kayser, M.; Rittscher, J.; et al. An objective comparison of detection and segmentation algorithms for artefacts in clinical endoscopy. *Sci. Rep.* **2020**, *10*, 2748. [\[CrossRef\]](#)
59. Yang, S.; Cheng, G. Endoscopic artefact detection and segmentation with deep convolutional neural network. In Proceedings of the 1st International Workshop and Challenge on Computer Vision in Endoscopy, Nice, France, 13 April 2019.
60. Yu, X.; Tang, S.; Cheang, C.F.; Yu, H.H.; Choi, I.C. Multi-Task Model for Esophageal Lesion Analysis Using Endoscopic Images: Classification with Image Retrieval and Segmentation with Attention. *Sensors* **2021**, *22*, 283. [\[CrossRef\]](#)
61. Zhu, M.; Chen, Z.; Yuan, Y. DSI-Net: Deep Synergistic Interaction Network for Joint Classification and Segmentation with Endoscope Images. *IEEE Trans. Med. Imaging* **2021**, *40*, 3315–3325. [\[CrossRef\]](#)
62. Jia, X.; Xing, X.; Yuan, Y.; Meng, M.Q.H. Deep learning-driven models for endoscopic image analysis. In *Advances in Artificial Intelligence, Computation, and Data Science: For Medicine and Life Science*; Springer: Berlin/Heidelberg, Germany, 2021; pp. 271–300. [\[CrossRef\]](#)
63. Wang, C.-C.; Chiu, Y.-C.; Chen, W.-L.; Yang, T.-W.; Tsai, M.-C.; Tseng, M.-H. A Deep Learning Model for Classification of Endoscopic Gastroesophageal Reflux Disease. *Int. J. Environ. Res. Public Health* **2021**, *18*, 2428. [\[CrossRef\]](#)
64. Nguyen, D.T.; Lee, M.B.; Pham, T.D.; Batchuluun, G.; Arsalan, M.; Park, K.R. Enhanced Image-Based Endoscopic Pathological Site Classification Using an Ensemble of Deep Learning Models. *Sensors* **2020**, *20*, 5982. [\[CrossRef\]](#)
65. Gunasekaran, H.; Ramalakshmi, K.; Swaminathan, D.K.; Mazzara, M. GIT-Net: An Ensemble Deep Learning-Based GI Tract Classification of Endoscopic Images. *Bioengineering* **2023**, *10*, 809. [\[CrossRef\]](#)
66. Muruganantham, P.; Balakrishnan, S.M. Attention Aware Deep Learning Model for Wireless Capsule Endoscopy Lesion Classification and Localization. *J. Med. Biol. Eng.* **2022**, *42*, 157–168. [\[CrossRef\]](#)

67. Mukhtorov, D.; Rakhmonova, M.; Muksimova, S.; Cho, Y.-I. Endoscopic Image Classification Based on Explainable Deep Learning. *Sensors* **2023**, *23*, 3176. [\[CrossRef\]](#)
68. An, Z.; Xu, C.; Qian, K.; Han, J.; Tan, W.; Wang, D.; Fang, Q. EIEN: Endoscopic Image Enhancement Network Based on Retinex Theory. *Sensors* **2022**, *22*, 5464. [\[CrossRef\]](#)
69. Yue, G.; Gao, J.; Duan, L.; Du, J.; Yan, W.; Wang, S.; Wang, T. Colorectal endoscopic image enhancement via unsupervised deep learning. *Multimedia Tools Appl.* **2023**, 1–23. [\[CrossRef\]](#)
70. Gómez, P.; Semmler, M.; Schützenberger, A.; Bohr, C.; Döllinger, M. Low-light image enhancement of high-speed endoscopic videos using a convolutional neural network. *Med. Biol. Eng. Comput.* **2019**, *57*, 1451–1463. [\[CrossRef\]](#)
71. Asif, M.; Chen, L.; Song, H.; Yang, J.; Frangi, A.F. An automatic framework for endoscopic image restoration and enhancement. *Appl. Intell.* **2020**, *51*, 1959–1971. [\[CrossRef\]](#)
72. Gulenko, O.; Yang, H.; Kim, K.; Youm, J.Y.; Kim, M.; Kim, Y.; Jung, W.; Yang, J.-M. Deep-Learning-Based Algorithm for the Removal of Electromagnetic Interference Noise in Photoacoustic Endoscopic Image Processing. *Sensors* **2022**, *22*, 3961. [\[CrossRef\]](#) [\[PubMed\]](#)
73. Xu, M.; Zhou, W.; Wu, L.; Zhang, J.; Wang, J.; Mu, G.; Huang, X.; Li, Y.; Yuan, J.; Zeng, Z.; et al. Artificial intelligence in the diagnosis of gastric precancerous conditions by image-enhanced endoscopy: A multicenter, diagnostic study (with video). *Gastrointest. Endosc.* **2021**, *94*, 540–548.e4. [\[CrossRef\]](#)
74. Lee, J.Y.; Jeong, J.; Song, E.M.; Ha, C.; Lee, H.J.; Koo, J.E.; Yang, D.-H.; Kim, N.; Byeon, J.-S. Real-time detection of colon polyps during colonoscopy using deep learning: Systematic validation with four independent datasets. *Sci. Rep.* **2020**, *10*, 8379. [\[CrossRef\]](#)
75. Aoki, T.; Yamada, A.; Kato, Y.; Saito, H.; Tsuboi, A.; Nakada, A.; Niikura, R.; Fujishiro, M.; Oka, S.; Ishihara, S.; et al. Automatic detection of various abnormalities in capsule endoscopy videos by a deep learning-based system: A multicenter study. *Gastrointest. Endosc.* **2021**, *93*, 165–173.e1. [\[CrossRef\]](#)
76. Guo, L.; Xiao, X.; Wu, C.; Zeng, X.; Zhang, Y.; Du, J.; Bai, S.; Xie, J.; Zhang, Z.; Li, Y.; et al. Real-time automated diagnosis of precancerous lesions and early esophageal squamous cell carcinoma using a deep learning model (with videos). *Gastrointest. Endosc.* **2020**, *91*, 41–51. [\[CrossRef\]](#) [\[PubMed\]](#)
77. Lee, D.; Yu, H.W.; Kwon, H.; Kong, H.-J.; Lee, K.E.; Kim, H.C. Evaluation of Surgical Skills during Robotic Surgery by Deep Learning-Based Multiple Surgical Instrument Tracking in Training and Actual Operations. *J. Clin. Med.* **2020**, *9*, 1964. [\[CrossRef\]](#) [\[PubMed\]](#)
78. Barbed, O.L.; Montiel, J.M.; Fua, P.; Murillo, A.C. Tracking adaptation to improve superpoint for 3d reconstruction in endoscopy. In Proceedings of the International Conference on Medical Image Computing and Computer-Assisted Intervention, Vancouver, BC, Canada, 8–12 October 2023; Springer Nature: Cham, Switzerland, 2023; pp. 583–593. [\[CrossRef\]](#)
79. Hamid, O.H. Data-Centric and Model-Centric AI: Twin Drivers of Compact and Robust Industry 4.0 Solutions. *Appl. Sci.* **2023**, *13*, 2753. [\[CrossRef\]](#)
80. Hamid, O.H.; Braun, J. Reinforcement learning and attractor neural network models of associative learning. In Proceedings of the Computational Intelligence: 9th International Joint Conference, IJCCI 2017, Funchal-Madeira, Portugal, 1–3 November 2017; Revised Selected Papers. Springer International Publishing: Berlin/Heidelberg, Germany, 2019; pp. 327–349. [\[CrossRef\]](#)

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.