



Article

Gastrointestinal Disease Classification in Endoscopic Images Using Attention-Guided Convolutional Neural Networks

Zenebe Markos Lonseko ^{1,2}, Prince Ebenezer Adjei ^{1,2,3}, Wenju Du ^{1,2}, Chengsi Luo ^{1,2}, Dingcan Hu ^{1,2}, Linlin Zhu ⁴, Tao Gan ³ and Nini Rao ^{1,2,*}

¹ Center for Informational Biology, University of Electronic Science and Technology of China, Chengdu 610054, China; zenebem4@gmail.com (Z.M.L.); prince.e.adjei@gmail.com (P.E.A.); aduwenju@163.com (W.D.); luosicheng945@gmail.com (C.L.); hudingcan@163.com (D.H.)

² School of Life Science and Technology, University of Electronic Science and Technology of China, Chengdu 610054, China

³ Department of Computer Engineering, Kwame Nkrumah University of Science and Technology, Kumasi AK-039-5028, Ghana; gantao111@aliyun.com

⁴ Digestive Endoscopic Center of West China Hospital, Sichuan University, Chengdu 610017, China; llz1823@163.com

* Correspondence: raonn@uestc.edu.cn

Abstract: Gastrointestinal (GI) diseases constitute a leading problem in the human digestive system. Consequently, several studies have explored automatic classification of GI diseases as a means of minimizing the burden on clinicians and improving patient outcomes, for both diagnostic and treatment purposes. The challenge in using deep learning-based (DL) approaches, specifically a convolutional neural network (CNN), is that spatial information is not fully utilized due to the inherent mechanism of CNNs. This paper proposes the application of spatial factors in improving classification performance. Specifically, we propose a deep CNN-based spatial attention mechanism for the classification of GI diseases, implemented with encoder–decoder layers. To overcome the data imbalance problem, we adapt data-augmentation techniques. A total of 12,147 multi-sited, multi-diseased GI images, drawn from publicly available and private sources, were used to validate the proposed approach. Furthermore, a five-fold cross-validation approach was adopted to minimize inconsistencies in intra- and inter-class variability and to ensure that results were robustly assessed. Our results, compared with other state-of-the-art models in terms of mean accuracy (ResNet50 = 90.28, GoogLeNet = 91.38, DenseNets = 91.60, and baseline = 92.84), demonstrated better outcomes (Precision = 92.8, Recall = 92.7, F1-score = 92.8, and Accuracy = 93.19). We also implemented t-distributed stochastic neighbor embedding (t-SNE) and confusion matrix analysis techniques for better visualization and performance validation. Overall, the results showed that the attention mechanism improved the automatic classification of multi-sited GI disease images. We validated clinical tests based on the proposed method by overcoming previous limitations, with the goal of improving automatic classification accuracy in future work.

Keywords: gastrointestinal disease; endoscopic image; lesion classification; computer-aided diagnosis; spatial attention; deep learning



Citation: Lonseko, Z.M.; Adjei, P.E.; Du, W.; Luo, C.; Hu, D.; Zhu, L.; Gan, T.; Rao, N. Gastrointestinal Disease Classification in Endoscopic Images Using Attention-Guided Convolutional Neural Networks. *Appl. Sci.* **2021**, *11*, 11136. <https://doi.org/10.3390/app112311136>

Academic Editor: Vito Renò

Received: 24 September 2021

Accepted: 11 November 2021

Published: 24 November 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 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

Gastrointestinal (GI) diseases are common in the human digestive system. Stomach cancer, esophageal cancer, and colorectal cancer are most common in terms of incidence and fatality [1,2]. Endoscopic examinations are thus vital to detect diseases and form the critical initial step for diagnosing GI tract diseases generally [3]. Such examinations also enhance the assessment of the clinical features of lesions to determine their severity and type and to arrive at proper diagnoses. Variations in the expertise of different clinicians could introduce errors in some cases, especially with respect to controversial aspects of

diagnostic images and videos from endoscopic examinations. Such inconsistency may lead to misdiagnoses and negative impact on patient care.

Automatic disease classification potentially addresses this problem by providing clinicians with objective and reliable identification of several GI endoscopic images, thereby minimizing the misdiagnosis rate, improving prognosis, and economizing clinicians' valuable time. Automatic GI disease classification thus remains an open area of research for attaining better lesion detection and classification accuracy [4].

Recently, artificial intelligence based on deep learning (DL) has demonstrated remarkable progress in classification, detection, fault diagnosis, and segmentation tasks [5–9]. However, a drawback to the potential of DL is that large amounts of data are required for optimum performance. In addition, difficulty in obtaining GI images is increased by patient privacy concerns and annotation costs [2,10,11]. In short, optimal use of the DL approach to automatic GI disease classification is constrained by data scarcity. Unlike classic machine learning-based classifiers such as support vector machines (SVMs) that are used for feature extraction [12], convolutional neural networks (CNNs) have shown better performance in feature extraction, making them state-of-the-art for DL applications [8]. Effective use of CNNs has improved image recognition- and classification-related tasks [8].

The foundation of CNN is the convolution operation that fuses spatial and channel-wise information within local receptive fields to extract information features, a process known as spatial encoding [13]. However, the vanilla CNN mechanism would allow the CNN operation to pay more attention to the pixel regions that play a decisive role in classifying input images while ignoring irrelevant information [14]. To improve the suggestive power of a CNN network, several recent studies have shown the benefit of enhancing the spatial encoding ability of CNN via spatial attention modules. Other studies have indicated promising CNN-based approaches for disease classifications [9,12,15–20]. In addition, spatial attention mechanisms have been studied and applied to related tasks [16,17,21–23].

To the best of our knowledge, GI-tract disease classification of multi-sited, multi-class diseases and artifacts, from GI endoscopic images using attention-guided CNNs, has not previously been applied. Multi-class disease and artifact classification and generalization are essential not only for diagnosis but also to avoid training biases. Our goal in this study, therefore, is to design, test, and validate an efficient spatial attention-guided CNN-based disease classification model that can be used in a computer-aided diagnosis (CAD) system for automatic GI disease classification. We examine a total of 10 classes: esophagitis, polyps, ulcerative colitis, early esophagus cancer, normal cecum, normal Z-line, normal pylorus, dyed-lifted polyps, dyed-resection-margins, and artifacts. Moreover, we adapt a t-distributed stochastic neighborhood embedding (t-SNE) visualization technique to better understand the data. [14].

Automatic classification of GI diseases using CAD can assist in the diagnosis of GI disease and in the efficient, effective, and safe removal of lesions [24–26]. Dimensionality reduction (DR) is a key step in feature extraction in images that often contain irrelevant information or labeling. When applied in this work, t-SNE was effective in enhancing data pattern visualization and feature extraction by preserving high intrinsic dimensionality [27,28]. Accordingly, this study proposes an efficient DL-based CAD system for the classification of multi-class diseases and artifacts in GI endoscopic images.

The main contributions of this paper can be summarized as follows:

- (1) We propose an efficient method that incorporates spatial attention CNN for classifying multi-class diseases and artifacts in GI endoscopic images.
- (2) We performed extensive experiments to validate the effectiveness of the proposed model. Moreover, we compared our results with recent related models and demonstrated better outcomes.
- (3) The proposed method demonstrated significant performance accuracy in GI disease classification by using spatial attention mechanisms and t-SNE.
- (4) The proposed GI disease classification method was validated for clinical applications and has great potential for medical communities.

In the next section, we describe the materials and method of our proposed approach. In Section 3, we address the experiments and results. In Section 4, we provide a detailed discussion of the results before concluding the paper in Section 5.

2. Materials and Methods

2.1. Materials

A total of 12,147 GI endoscopic images were obtained from hospital and public data sources and used to validate our proposed approach. The details of each dataset are described below:

2.1.1. Kvasir Multi-Class Dataset

Eight thousand multi-class GI endoscopy images (8 classes with 1000 images in each class) were collected from the Kvasir dataset v2 [2] and verified by experienced medical specialists. The datasets were collected using endoscopic equipment at Vestre Viken Health Trust in Norway. The pathological results included esophagitis (ESO), polyps (POL), ulcerative colitis (ULC), etc., while the anatomical landmarks included Z-line (ZLI), pylorus (PYL), cecum (CEC), etc. In addition to images related to removing lesions, such as a dyed and lifted polyp (DLP), the dyed resection margins (DRM) were used for the experiments. The dataset consisted of images with resolutions ranging from 720×576 up to 1920×1072 pixels. Some of the included GI images have artifacts indicating the configuration of the position of the endoscope inside the bowel, using an electromagnetic imaging system [2].

2.1.2. Endoscopy Artifact Detection Challenge Dataset

The endoscopy artifact detection (EAD2019) challenge dataset [4] provided 2147 images collected from six different centers. The images varied in resolution from 295×299 to 1920×1080 pixels in JPG format. Most of the images were used for endoscopic artifact detection (EAD) classification.

2.1.3. Gastrointestinal Endoscopy Dataset

We used 2000 early esophagus cancer (EEC) GI endoscopy images collected from 389 patients and verified by physicians from the Digestive Endoscopy Center (Lab) of the West China Hospital in Sichuan, China. The images were saved in the RGB channel in JPG format, with resolutions ranging from 512×512 to 764×572 pixels. Permission was obtained from the medical ethical review committee of the Electronic University of Science and Technology of China (UESTC) and West China Hospital, and the informed consent of patients was received. Samples of raw GI endoscopic images are depicted in Figure 1.

2.2. Methods

This subsection describes the DL-based method for using the attention mechanism in the GI disease classification of multi-class and multi-sited endoscopic images. In step 1, unnecessary background features, text, and other artifacts were removed from each sample dataset during pre-processing. Images were resized to 224×224 pixels for the experiments. As shown in Figure 1, the raw images that contained unnecessary background and text were removed during pre-processing, and the class balancing technique was applied with ECA-DDCNN [15]. Eighty percent of the data were used for training and 20% for testing. To avoid overfitting and imbalance, data augmentation techniques were used, such as transposing, flipping, rotating, random brightness, blurring, distortion, contrasting, and limited adaptive histogram equalization. After data augmentation, the total training data were comprised of forty thousand images. In step 2, the pre-processed data were utilized for the whole model's experiments. In step 3, disease detection and classification were implemented with attention mechanisms. In step 4, disease classification or lesion categorization of the proposed method was validated using 20% of the test data. Finally, in step 5, the output was evaluated quantitatively and qualitatively using appropriate evaluation metrics and visualization techniques as shown in Figure 2.

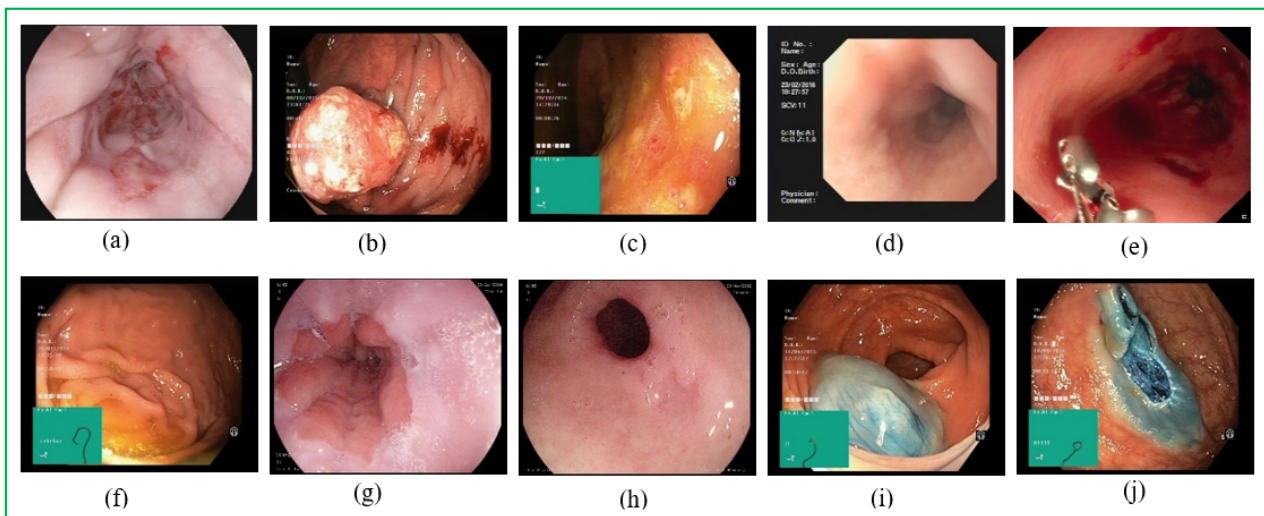


Figure 1. Raw input multi-class gastrointestinal endoscopy images. Respectively, (a–d) refers to pathological findings; (e) refers to artefacts and equipment; (f–h) refers to anatomical landmarks; and (i,j) refers to polyp removal images. (a) Esophagitis, which refers to an inflammation of the esophagus, visible as a break in the esophageal mucosa with the Z-line; (b) polyps are lesions within the bowel detectable as mucosal outgrowths; (c) ulcerative colitis refers to a chronic inflammatory disease affecting the large bowel; (d) early esophagus cancer; (e) images with artefacts and diagnostic equipment; (f) the cecum (the most proximal part of the large bowel); (g) The Z-line that marks the transition site between the esophagus and the stomach; (h) the pylorus (the area around the opening from the stomach into the first part of the small bowel, called the duodenum); (i) dyed-lifted-polyps; (j) dyed-resection-margins.

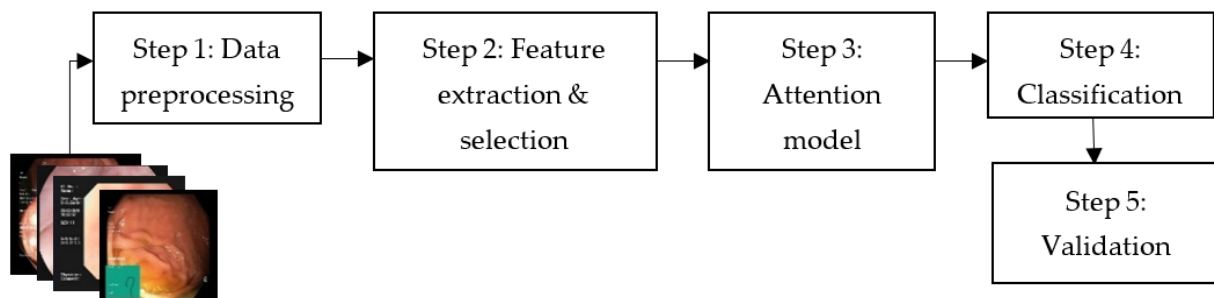


Figure 2. A framework of the proposed GI disease classification method.

In this work, we propose an attention network, a CNN, that adopts a mixed attention mechanism, as shown in Figure 3 (main network) and Figure 4 (sub-network). It comprises multiple attention layers/modules, which generate attention features by stacking layers to optimize the network [18]. The proposed model explores the performance of a robust classifier on CNN architecture. The spatial attention module compresses the channel and performs average and maximum pooling in the channel dimensions. The fully connected (FC) layers are then used to capture the non-linear cross-channel interaction, during which the complexity of the model is minimized by reducing dimensionality. The performance of the proposed approach was tested on the data used in training and validation. Attention features generated at stage 2 are shown in Figure 5.

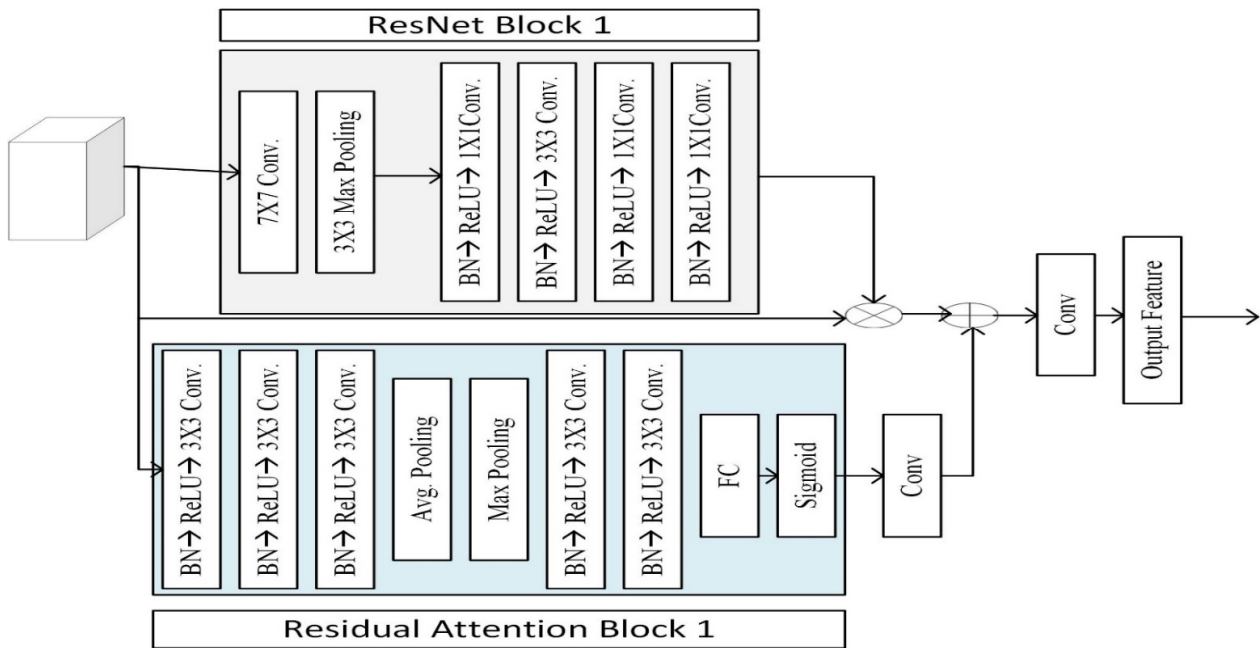


Figure 3. Overall spatial attention CNN architecture of the proposed method.

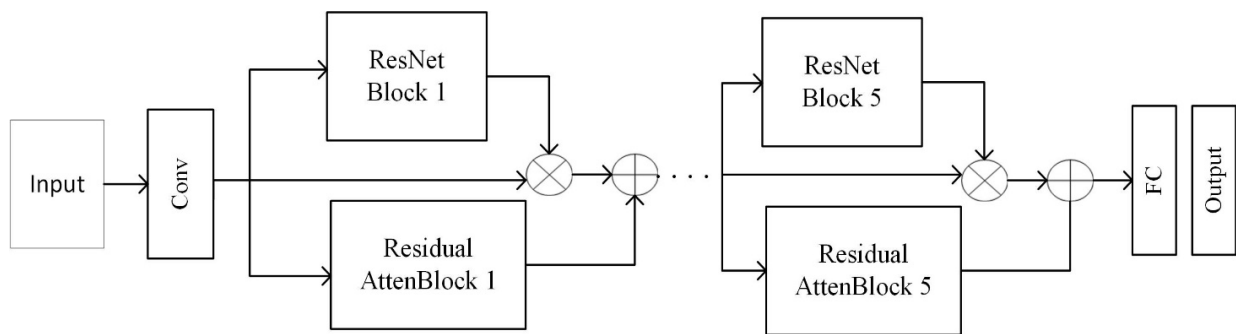


Figure 4. Attention residual learning block.

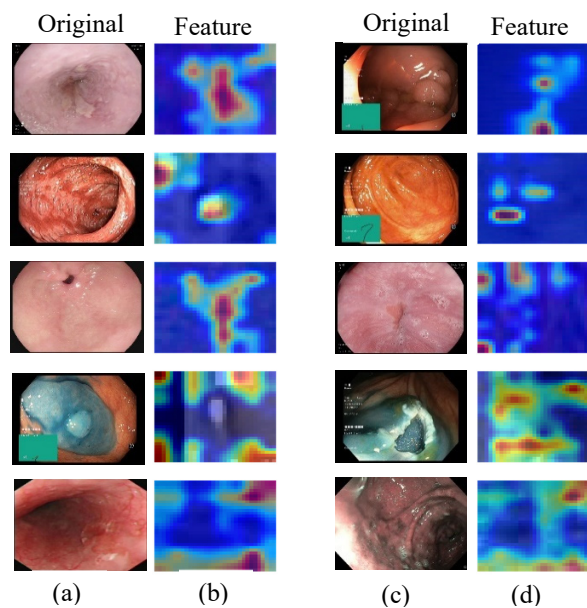


Figure 5. Sample heat map at stage 2 feature extraction of the proposed method. (a,c) demonstrate original images, class 0–9 respectively and (b,d) show attention features of all classes.

3. Experimental Setup

The proposed method explores the effect of a robust classifier on deep CNN architecture. Most common CNN architectures, such as GoogLeNet and ResNet, are used as a backbone [29]. The important features of each architecture were initially learned, followed by training with the three sets of GI endoscopic datasets. In the feature extraction procedure, the original architectures were not changed. The primary attention-based model was used for baseline network training. Stochastic gradient descent (SGD) was implemented to optimize the network with a 0.001 initial training rate. Total training epochs and batch size were 120 and 32, respectively. Input was resized to 128×128 pixels to crop the ROI. The programming language implemented in python 3.6.12 and Pytorch 1.7.1 DL library (<https://github.com/pytorch/pytorch> (accessed on 11 February 2021)) was used for the experiments. Experiments were implemented on a server-based Ubuntu 16.04.6 LTS. The system was equipped with four graphic processing units (GPUs) of NVIDIA GeForce RTX 2080Ti with 11 GB memory each.

We used several evaluation metrics to validate the classification performance of the proposed method, including accuracy (*Acc*), precision (*Pre*), recall (*Rec*), sensitivity (*Sen*), specificity (*Spe*), and F1-score (*F1*). *F1* measures a test accuracy by calculating the harmonic mean of *Pre* and *Rec* [2]. *Acc* specifies the ratio of GI-disease patients to non-diseased patients.

$$Acc = \frac{TP + TN}{TP + TN + FP + FN} \quad (1)$$

$$Pre = \frac{TP}{TP + FP} \quad (2)$$

$$Rec = \frac{TP}{TP + FN} \quad (3)$$

$$F1 = 2 \times \frac{Pre \times Rec}{Pre + Rec} \quad (4)$$

4. Results and Discussion

This section presents results and a discussion of the experimental results based on the DL approach. The DL approach has been shown to enhance the performance of GI disease classification tasks significantly. Table 1 depicts the model's performance and its parameters. The most common architectures [30–32] and baseline (ResNet50 = 90.28%, GoogLeNet = 91.38%, DenseNets = 91.60%, and baseline = 92.84%) demonstrated classification accuracy.

Table 1. Model performance complexity comparisons.

Models	Evaluation Metrics	
	Mean Accuracy (%)	Parameters (Million)
ResNet50 [30]	90.28	21.71
GoogLeNet [31]	91.38	5.61
DenseNets [32]	91.60	25.6
Baseline (Ours)	92.84	19.92

Table 2 reveals cross-validation results in different five-folds, with reference to the size of the test dataset and other parameters. In Table 3, we compared the proposed method with other related methods [10,23,29]. The mean accuracy for ECA-Net equaled 92.81; for DL-OCT, it was 90.60; for LSTM-CNN, it was 93.10; for Liu et al. it was 93; for our method, it was 93.19.

Table 2. Cross-validation statistical comparisons (mean values) on test datasets.

Folds	Evaluation Metrics			
	Precision	Recall	F1-Score	Accuracy
Fold1	91.8	91.8	91.7	92.16
Fold2	92.5	92.4	92.4	92.88
Fold3	92.4	92.5	92.6	92.91
Fold4	92.8	92.7	92.8	93.19
Fold5	92.8	92.6	92.7	93.12

Table 3. Statistical comparisons of related works (mean values).

Methods	Evaluation Metrics			
	Precision	Recall	F1-Score	Accuracy
ECA-Net [23]	92.4	92.4	92.2	92.81
DL-OCT [10]	90.1	90.05	90.1	90.60
LSMT-CNN [29]	92.8	92.8	92.6	93.10
Liu et al. [5]	92.7	92.6	92.7	93
Ours	92.8	92.7	92.8	93.19

Table 3 demonstrates a statistical comparison with other related methods. The comparison between the proposed method and other state-of-the-art methods was evaluated based on the same data and on standard evaluation metrics.

Figure 6 demonstrates statistical comparisons of the DL method's performance in the confusion matrix. The confusion matrix calculates the actual and predicted rates for each disease category. The results of all four methods presented on the confusion matrix indicate disease classification performance. The test data ranges from 300 to 400 in each class. Overall performance accuracies of the respective methods were ECA-Net = 92.81; DL-OCT = 90.60; LSTM + CNN = 93.10; and Ours = 93.19. The ECA-Net method [23] revealed the GI disease or artifacts classification performance, with the pylorus classification of 100% and the lowest classification in esophagitis (82%). The DL-OCT [10] method also demonstrated competitive classification accuracy, classifying pylorus at 98% and polyps at 81%. The LSMT-CNN [29] method exhibited competitive classification performance in pylorus (100%) and showed the lowest performance in the esophagitis class (82%). The proposed method classified pylorus at 100% and showed the lowest performance in the esophagitis disease classification (84%). Nearly all the methods classified pylorus best and misclassified the Z-line. Only a few data items were misclassified. Overall classification accuracy of all related methods was promising. The proposed method outperformed other methods significantly on GI disease or lesion classification by achieving a mean accuracy of 94.33%. Overall, the proposed method validated actual and predicted classifications better than other methods.

It is essential to determine the location (anatomical and physiological) of the digestive system. Figure 7 confirms disease classification results using t-distributed stochastic neighbor embedding (t-SNE) visualization techniques, which show the transformation of high-dimensional features into low dimensions by preserving key features of endoscopic images. This technique provides better identification of each class in terms of depth and color. The proposed method, including other related works [10,23,29] demonstrated t-SNE-based classification performance. Our method achieved better qualitative results when compared with other works.

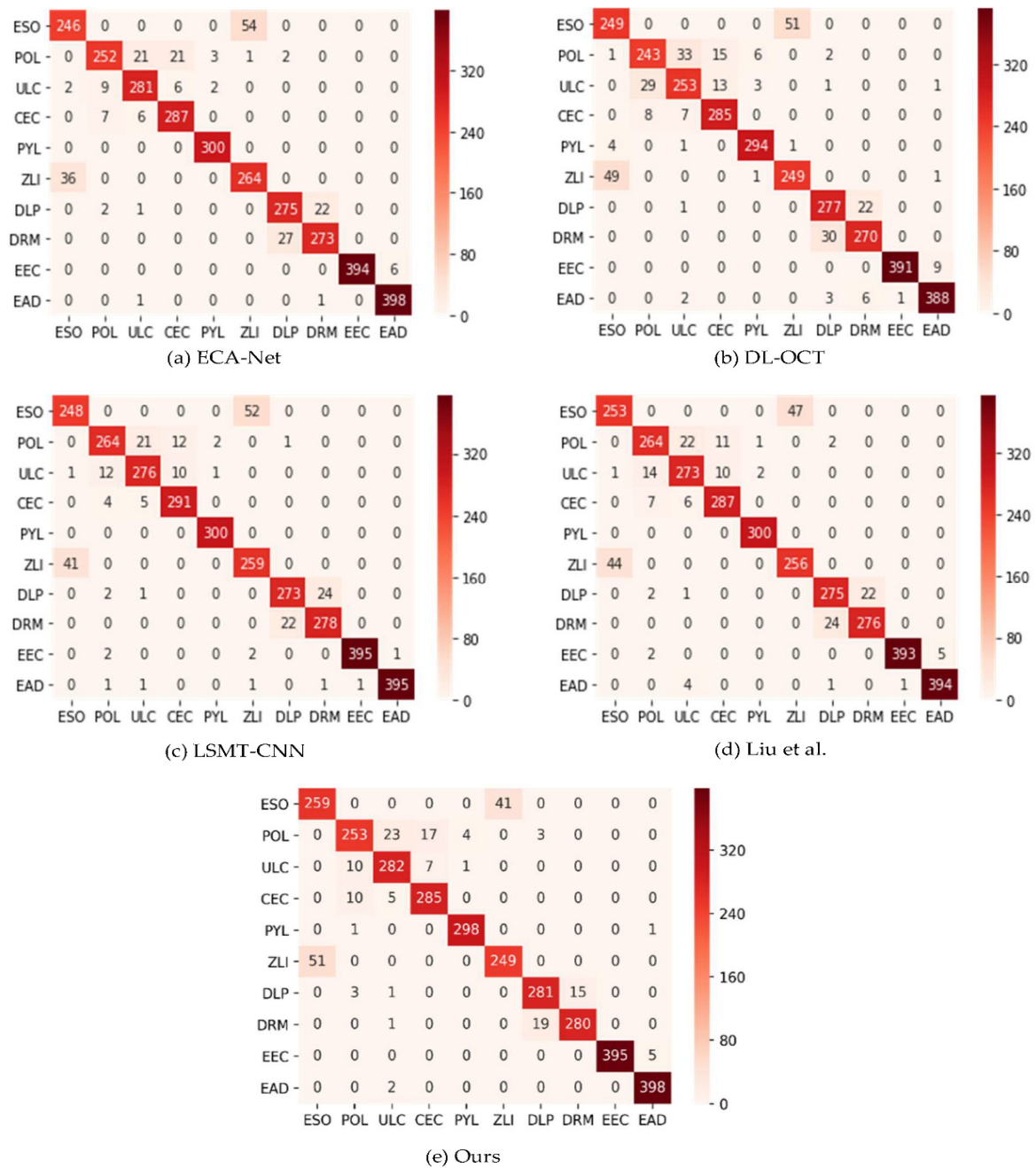


Figure 6. Confusion matrix of the GI disease classification on test datasets. The confusion matrix indicates the actual and predicted results of all ten classes: (a) The GI disease or artifacts classification performance of the ECA-Net [23], which classified pylorus better (100%), with the lowest classification (82%) achieved in esophagitis. (b) Classification accuracy of the DL-OCT [10] method’s performance. The method classified pylorus better (98%) and challenged the classification of polyps (81%). (c) Confirmed the competitive classification performance of the LSMT-CNN [29] method. The method better identified pylorus (100%) and showed the lowest classification results (82%) in the esophagitis class. (d) Liu et al. [5] demonstrated better classification in pylorus. (e) The proposed method, like other methods [23,29], classified pylorus better (100%), and the approach showed the lowest performance (84%) in the esophagitis disease classification.

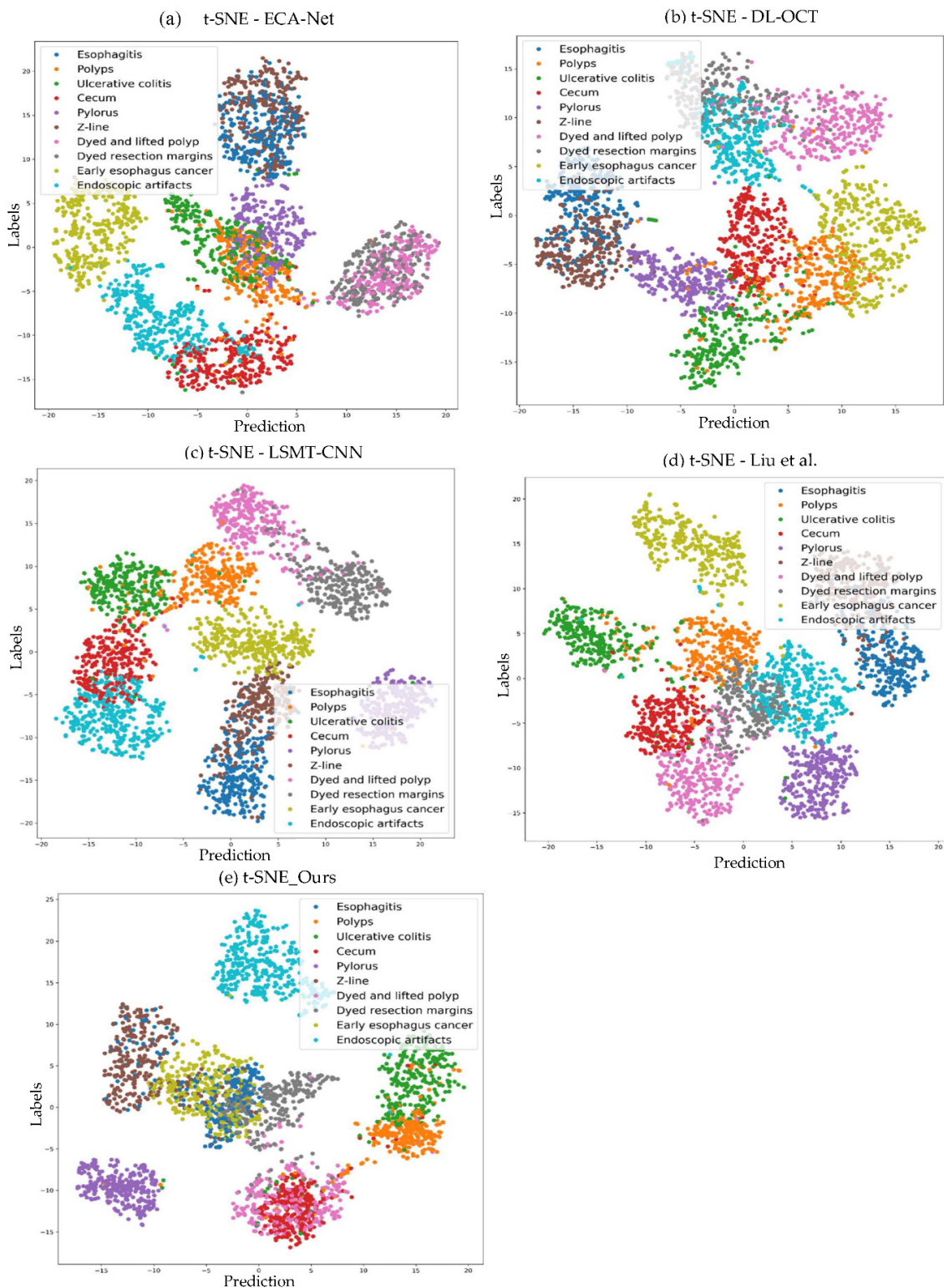


Figure 7. Validation of GI disease classification using the t-distributed stochastic neighbor embedding (t-SNE) visualization technique, which shows the transformation of high-dimensional features into low dimensions by preserving key features. The t-SNE technique provides better identification of each class in terms of depth-wise and color-wise approaches: (a) The ECA-Net [23] t-SNE method of classification performance and differentiates each class using multiple colors. (b) The DL-OCT [10] method's t-SNE-based classification of each GI lesion category. (c) The LSMT-CNN [29] method's t-SNE based classification results. (d) Liu et al. [5] t-SNE projection results. (e) The proposed method's t-SNE-based classification performance, which outperforms other related works significantly.

In this study, t-SNE was applied to preserve the local structure of the data by a non-linear dimension reduction (DR) approach. A performance analysis was implemented to demonstrate the performance difference in the methods [33]. The t-SNE visualization showed significant improvement in terms of feature extraction and DR for GI disease classification. It preserved complex non-linear data structures by maintaining the local similarity structure of the data [27]. The t-SNE technique could efficiently project complex 2D endoscopic datasets. High-dimensional space was preserved as much as possible. Our t-SNE-based disease classification exhibited better feature extraction performance than other methods.

The proposed method achieved better results in classifying GI diseases when compared with other recent related methods. However, the results of the proposed method come with some limitations. A fully supervised learning method needs a massive GI endoscopic image dataset, which is the method's main drawback. Another limitation of the proposed approach may be its lack of focus on pixel-wise classification and detection.

5. Conclusions

This work presents a deep CNN-based GI disease classification method by employing an efficient spatial attention mechanism. Three GI endoscopic image datasets with multi-classes were used to validate the proposed method. We validated model performance complexity and conducted five-fold cross-validation to confirm the proposed method. The experimental results show that our approach is more efficient when compared with other related methods. Accordingly, our proposed method has the potential to aid in the clinical diagnosis of various GI diseases.

In future works, we propose that special attention should be given to improving classification accuracy by combining DL techniques with clinical features and by exploring the stability of the proposed method. In addition, we intend to perform clinical testing of the proposed model in determining its value in assisting the diagnoses of GI diseases.

Author Contributions: Conceptualization, Z.M.L. and N.R.; methodology, Z.M.L., N.R. and P.E.A.; software, D.H.; validation, W.D. and T.G.; formal analysis, Z.M.L. and P.E.A.; investigation, C.L.; resources, N.R., T.G. and L.Z.; data curation, T.G., L.Z. and D.H.; writing—original draft preparation, Z.M.L.; writing—review and editing, Z.M.L., P.E.A. and N.R.; visualization, W.D.; supervision, N.R. and C.L.; project administration, N.R., and C.L.; funding acquisition, N.R. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the National Natural Science Foundation of China (Grant No. 61872405 and 61720106004), the Key R&D Program of Sichuan Province (Grant No. 2020YFS0243).

Institutional Review Board Statement: This study was conducted following the ethical standards of the institutional (IRB) and/or research committee and with the 1964 Helsinki declaration, and its latter amendments or comparable ethical standards.

Informed Consent Statement: All approaches were performed by following the regulations and relevant guidelines.

Data Availability Statement: In this study, we used two primary data sources (hospital and public) available in references. Except for the hospital dataset, which will be available upon requests made to a corresponding author due to data privacy restrictions, all the rest of the public datasets are available at <https://datasets.simula.no/kvasir/> (accessed on 10 March 2021); <https://polyp.grand-challenge.org/CVCClinicDB> (accessed on 10 March 2021); <https://doi.org/10.1007/s11548-013-0926-3> (accessed on 10 March 2021); <https://doi.org/10.17632/C7FJBXCGJ9.1> (accessed on 10 March 2021).

Acknowledgments: We acknowledge support by the National Natural Science Foundation of China (Grant No. 61872405 and 61720106004), the Key R&D Program of Sichuan Province (Grant No. 2020YFS0243).

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

References

1. Sung, H.; Ferlay, J.; Siegel, R.L.; Laversanne, M.; Soerjomataram, I.; Jemal, A.; Bray, F. Global Cancer Statistics 2020: GLOBOCAN Estimates of Incidence and Mortality Worldwide for 36 Cancers in 185 Countries. *CA. Cancer J. Clin.* **2021**, *71*, 209–249. [[CrossRef](#)]
2. Pogorelov, K.; Randel, K.R.; Griwodz, C.; Eskeland, S.L.; De Lange, T.; Johansen, D.; Spampinato, C.; Dang-Nguyen, D.T.; Lux, M.; Schmidt, P.T.; et al. Kvasir: A Multi-Class Image Dataset for Computer Aided Gastrointestinal Disease Detection. In Proceedings of the 8th ACM Multimedia Systems Conference, MMSys 2017, Taipei, Taiwan, 20–23 June 2017; pp. 164–169.
3. Muto, M.; Yao, K.; Kaise, M.; Kato, M.; Uedo, N.; Yagi, K.; Tajiri, H. Magnifying Endoscopy Simple Diagnostic Algorithm for Early Gastric Cancer (MESDA-G). *Dig. Endosc.* **2016**, *28*, 379–393. [[CrossRef](#)]
4. Ali, S.; Zhou, F.; Daul, C.; Braden, B.; Bailey, A.; Realdon, S.; East, J.; Wagnières, G.; Loschenov, V.; Grisan, E.; et al. Endoscopy Artifact Detection (EAD 2019) Challenge Dataset. *arXiv* **2019**, arXiv:1905.03209.
5. Liu, X.; Wang, C.; Bai, J.; Liao, G. Fine-Tuning Pre-Trained Convolutional Neural Networks for Gastric Precancerous Disease Classification on Magnification Narrow-Band Imaging Images. *Neurocomputing* **2020**, *392*, 253–267. [[CrossRef](#)]
6. Magalhaes, C.; Mendes, J.; Vardasca, R. Meta-Analysis and Systematic Review of the Application of Machine Learning Classifiers in Biomedical Applications of Infrared Thermography. *Appl. Sci.* **2021**, *11*, 842. [[CrossRef](#)]
7. Glowacz, A. Fault Diagnosis of Electric Impact Drills Using Thermal Imaging. *Measurement* **2021**, *171*, 108815. [[CrossRef](#)]
8. Takiyama, H.; Ozawa, T.; Ishihara, S.; Fujishiro, M.; Shichijo, S.; Nomura, S.; Miura, M.; Tada, T. Automatic Anatomical Classification of Esophagogastroduodenoscopy Images Using Deep Convolutional Neural Networks. *Sci. Rep.* **2018**, *8*, 7497. [[CrossRef](#)] [[PubMed](#)]
9. Abawatew, G.Y.; Belay, S.; Gedamu, K.; Assefa, M.; Ayalew, M.; Oluwasanmi, A.; Qin, Z. Attention Augmented Residual Network for Tomato Disease Detection and Classification. *Turk. J. Electr. Eng. Comput. Sci.* **2021**, *29* (Suppl. 1), 2869–2885.
10. Abdolmanafi, A.; Duong, L.; Dahdah, N.; Cheriet, F. Deep Feature Learning for Automatic Tissue Classification of Coronary Artery Using Optical Coherence Tomography. *Biomed. Opt. Express* **2017**, *8*, 1203–1220. [[CrossRef](#)] [[PubMed](#)]
11. Lonseko, Z.M.; Adjei, P.E.; Du, W.; Luo, C.; Wang, Y.; Hu, D.; Gan, T.; Rao, N. Semi-Supervised Gastrointestinal Lesion Segmentation Using Adversarial Learning. In Proceedings of the 2021 IEEE 3rd Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS), Tainan, Taiwan, 28–30 May 2021; IEEE: Piscataway, NJ, USA, 2021; pp. 63–66.
12. Liu, D.Y.; Gan, T.; Rao, N.N.; Xing, Y.W.; Zheng, J.; Li, S.; Luo, C.S.; Zhou, Z.J.; Wan, Y.L. Identification of Lesion Images from Gastrointestinal Endoscope Based on Feature Extraction of Combinational Methods with and without Learning Process. *Med. Image Anal.* **2016**, *32*, 281–294. [[CrossRef](#)]
13. Hu, J.; Shen, L.; Sun, G. Squeeze-and-Excitation Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 7132–7141.
14. Chen, Z.; Cao, M.; Ji, P.; Ma, F. Research on Crop Disease Classification Algorithm Based on Mixed Attention Mechanism. In *Journal of Physics: Conference Series*; IOP Publishing: Bristol, UK, 2021; Volume 1961, p. 12048.
15. Du, W.; Rao, N.; Dong, C.; Wang, Y.; Hu, D.; Zhu, L.; Zeng, B.; Gan, T. Automatic Classification of Esophageal Disease in Gastroscopic Images Using an Efficient Channel Attention Deep Dense Convolutional Neural Network. *Biomed. Opt. Express* **2021**, *12*, 3066. [[CrossRef](#)]
16. Ikenoyama, Y.; Hirasawa, T.; Ishioka, M.; Namikawa, K.; Yoshimizu, S.; Horiuchi, Y.; Ishiyama, A.; Yoshio, T.; Tsuchida, T.; Takeuchi, Y. Detecting Early Gastric Cancer: Comparison between the Diagnostic Ability of Convolutional Neural Networks and Endoscopists. *Dig. Endosc.* **2021**, *33*, 141–150. [[CrossRef](#)]
17. Zhu, Y.; Wang, Q.C.; Xu, M.D.; Zhang, Z.; Cheng, J.; Zhong, Y.S.; Zhang, Y.Q.; Chen, W.F.; Yao, L.Q.; Zhou, P.H.; et al. Application of Convolutional Neural Network in the Diagnosis of the Invasion Depth of Gastric Cancer Based on Conventional Endoscopy. *Gastrointest. Endosc.* **2019**, *89*, 806–815.e1. [[CrossRef](#)] [[PubMed](#)]
18. Wang, F.; Jiang, M.; Qian, C.; Yang, S.; Li, C.; Zhang, H.; Wang, X.; Tang, X. Residual Attention Network for Image Classification. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, 21–26 July 2017; Volume 2017, pp. 6450–6458.
19. Guan, Q.; Huang, Y.; Zhong, Z.; Zheng, Z.; Zheng, L.; Yang, Y. Thorax Disease Classification with Attention Guided Convolutional Neural Network. *Pattern Recognit. Lett.* **2020**, *131*, 38–45. [[CrossRef](#)]
20. Gessert, N.; Sentker, T.; Madesta, F.; Schmitz, R.; Kniep, H.; Baltruschat, I.; Werner, R.; Schlaefer, A. Skin Lesion Classification Using Cnns with Patch-Based Attention and Diagnosis-Guided Loss Weighting. *IEEE Trans. Biomed. Eng.* **2019**, *67*, 495–503. [[CrossRef](#)]
21. Chen, B.; Li, J.; Lu, G.; Zhang, D. Lesion Location Attention Guided Network for Multi-Label Thoracic Disease Classification in Chest X-Rays. *IEEE J. Biomed. Health Inform.* **2019**, *24*, 2016–2027. [[CrossRef](#)]
22. Tao, S.; Jiang, Y.; Cao, S.; Wu, C.; Ma, Z. Attention-Guided Network with Densely Connected Convolution for Skin Lesion Segmentation. *Sensors* **2021**, *21*, 3462. [[CrossRef](#)]
23. Wang, Q.; Wu, B.; Zhu, P.; Li, P.; Zuo, W.; Hu, Q. ECA-Net: Efficient Channel Attention for Deep Convolutional Neural Networks. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Seattle, WA, USA, 13–19 June 2020; pp. 11531–11539.

24. Hwang, J.H.; Jamidar, P.; Baig, K.R.K.K.; Leung, F.W.; Lightdale, J.R.; Maranki, J.L.; Okolo III, P.I.; Swanstrom, L.L.; Chak, A. GIE Editorial Board Top 10 Topics: Advances in GI Endoscopy in 2019. *Gastrointest. Endosc.* **2020**, *92*, 241–251. [[CrossRef](#)] [[PubMed](#)]
25. Mori, Y.; Kudo, S.; Mohmed, H.E.N.; Misawa, M.; Ogata, N.; Itoh, H.; Oda, M.; Mori, K. Artificial Intelligence and Upper Gastrointestinal Endoscopy: Current Status and Future Perspective. *Dig. Endosc.* **2019**, *31*, 378–388. [[CrossRef](#)]
26. Horie, Y.; Yoshio, T.; Aoyama, K.; Yoshimizu, S.; Horiuchi, Y.; Ishiyama, A.; Hirasawa, T.; Tsuchida, T.; Ozawa, T.; Ishihara, S.; et al. Diagnostic Outcomes of Esophageal Cancer by Artificial Intelligence Using Convolutional Neural Networks. *Gastrointest. Endosc.* **2019**, *89*, 25–32. [[CrossRef](#)]
27. Hajderanj, L.; Weheliye, I.; Chen, D. A New Supervised T-SNE with Dissimilarity Measure for Effective Data Visualization and Classification. In *ACM International Conference Proceeding Series*; ACM: New York, NY, USA, 2019; pp. 232–236.
28. Ghannam, R.B.; Techtmann, S.M. Machine Learning Applications in Microbial Ecology, Human Microbiome Studies, and Environmental Monitoring. *Comput. Struct. Biotechnol. J.* **2021**, *19*, 1092–1107. [[CrossRef](#)] [[PubMed](#)]
29. Öztürk, Ş.; Özkaya, U. Gastrointestinal Tract Classification Using Improved LSTM Based CNN. *Multimed. Tools Appl.* **2020**, *79*, 28825–28840. [[CrossRef](#)]
30. He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, Las Vegas, NV, USA, 26 June–1 July 2016; Volume 2016, pp. 770–778.
31. Ioffe, S.; Szegedy, C. Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. In Proceedings of the 32nd International Conference on Machine Learning, ICML 2015, Lille, France, 6–11 July 2015; Volume 1, pp. 448–456.
32. Huang, G.; Liu, Z.; Van Der Maaten, L.; Weinberger, K.Q. Densely Connected Convolutional Networks. In Proceedings of the 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017, Honolulu, HI, USA, 21–26 July 2017; Institute of Electrical and Electronics Engineers Inc.: Piscataway, NJ, USA, 2017; pp. 2261–2269.
33. Owais, M.; Arsalan, M.; Choi, J.; Mahmood, T.; Park, K.R. Artificial Intelligence-Based Classification of Multiple Gastrointestinal Diseases Using Endoscopy Videos for Clinical Diagnosis. *J. Clin. Med.* **2019**, *8*, 986. [[CrossRef](#)] [[PubMed](#)]