



Research Article

Automated Classification and Segmentation in Colorectal Images Based on Self-Paced Transfer Network

Yao Yao,^{1,2} Shuiping Gou ,¹ Ru Tian,¹ Xiangrong Zhang,¹ and Shuixiang He ³

¹School of Artificial Intelligence, Xidian University, Xi'an, Shanxi 710071, China

²School of Information Engineering, Hangzhou Vocational and Technical College, Hangzhou, Zhejiang 310018, China

³Department of Gastroenterology, First Affiliated Hospital of Xi'an Jiaotong University, Xi'an, Shanxi 710071, China

Correspondence should be addressed to Shuiping Gou; shpgou@mail.xidian.edu.cn and Shuixiang He; hesx123@126.com

Received 13 November 2020; Revised 22 December 2020; Accepted 6 January 2021; Published 21 January 2021

Academic Editor: Zhiguo Zhou

Copyright © 2021 Yao Yao et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Colorectal imaging improves on diagnosis of colorectal diseases by providing colorectal images. Manual diagnosis of colorectal disease is labor-intensive and time-consuming. In this paper, we present a method for automatic colorectal disease classification and segmentation. Because of label unbalanced and difficult colorectal data, the classification based on self-paced transfer VGG network (STVGG) is proposed. ImageNet pretraining network parameters are transferred to VGG network with training colorectal data to acquire good initial network performance. And self-paced learning is used to optimize the network so that the classification performance of label unbalanced and difficult samples is improved. In order to assist the colonoscopist to accurately determine whether the polyp needs surgical resection, feature of trained STVGG model is shared to Unet segmentation network as the encoder part and to avoid repeat learning of polyp segmentation model. The experimental results on 3061 colorectal images illustrated that the proposed method obtained higher classification accuracy (96%) and segmentation performance compared with a few other methods. The polyp can be segmented accurately from around tissues by the proposed method. The segmentation results underpin the potential of deep learning methods for assisting colonoscopist in identifying polyps and enabling timely resection of these polyps at an early stage.

1. Introduction

The International Agency for Research on Cancer released research data on global cancer status in 2018. The report reported the incidence and mortality of 36 types of tumors in 185 countries around the world, comprehensively. The data showed that the incidence of colorectal cancer ranked third (10.2%), and the mortality rate ranked second (9.2%) [1]. As we known, the mortality rate of colorectal cancer can be reduced significantly by early removal of polyps [2] which can be found according to the early screening. Colorectal polyp, a benign disease, has specific imaging characteristics such as shape or surface structure and color [3]. Colorectal colonoscopy is the main method of diagnosing intestinal diseases. With a great number of colorectal images, the microscopic examination presents labor-intensive and time-consuming problems [4]. In addition, the pathological diagnosis of colonoscopy biopsy

samples is prone to deviations due to individual pathologists' experience and knowledge [5]. The accuracy of diagnosis depends on the experience of the microscopy doctor, and the difference in diagnosis accuracy between experienced doctors and less experienced doctors is greater than 10%. Therefore, it is necessary to distinguish polyps from normal tissue and tumor using colorectal optical images.

However, it is difficult for the diagnosis of colorectal optical images. Firstly, the low light and interference of liquid often result in poor imaging quality of colorectal images. Secondly, the edges of normal tissue and polyp types are blurred. It causes the classification accuracy of normal tissue, polyp, and tumor to be low. Thirdly, the individual differences of polyps are mainly manifested in shape, color, and surface contour for polyps. And colorectal polyps are heterogeneous resulted that the segmentation of polyp becomes challenging.

Previous research showed that deep learning has given good results in medical images processing, such as tumor detection, classification, segmentation, retrieval, and prediction, especially for diagnosis and treatment of the brain [6, 7], breast [8], lung [9, 10], gastric [11], prostate cancers [12, 13], and histopathology [14]. Meanwhile, endoscopy-assisted diagnosis has also made some progress using deep learning, especially in colorectal endoscopy. There are two types for colorectal image detection: pathology and optical colonoscopy images. Here is the introduction to the image diagnosis progress.

For the *pathology colorectal images*, the recent advancement of deep learning is adapted. Thakur et al. [15] reviewed the development of an AI system in CRC pathology image analysis using deep learning. Korbar et al. [16] proposed an automatic image-understanding method to help pathologists with histopathological characterization and diagnosis of colorectal polyps. Sena et al. [17] propose a deep learning approach to recognize four different stages of cancerous tissue development. Lizuka et al. [18] trained convolutional neural networks (CNNs) and recurrent neural networks (RNNs) on biopsy histopathology whole-slide images (WSIs) of stomach and colon. For the *optical colorectal images*, there are lots of researches on detection and segmentation of colorectal polyps. Some methods take into account time series: Urban et al. [19] used deep learning to localize and identify polyps in real time with 96% accuracy in screening colonoscopy. Klare et al. [20] proposed the APDS with which the colonoscopy system of the video stream is captured by a frame-grabber device in HD. Wang et al. [21] used real-time automatic detection system to increase colonoscopic polyp and adenoma detection rates; some methods take into account spatial information: Li et al. [22] used a fully convolutional neural network structure for segmenting colorectal polyps. Yang et al. [23] developed convolutional neural network (CNN) models which automatically categorized colorectal lesions into several stages ranging from nonneoplastic lesions to advanced CRC with conventional white-light colonoscopy images. Zhang et al. [24] developed a fully automatic algorithm to detect and classify hyperplastic and adenomatous colorectal polyps. Others are from the semantic information: Wickstrom and Kampffmeyer [25] proposed a novel method for estimating the uncertainty associated with important features in the input and demonstrated how interpretability and uncertainty can be modeled for semantic segmentation of colorectal polyps. The above colorectal image processing methods using deep learning have achieved good performance.

Based on the above analysis of colonic *pathology and optical* colorectal image literature, deep learning methods are proposed on detection or segmentation of colorectal polyps. However, unlike the recent research based on single task, our method takes into account multitask: colorectal image classification and polyp image segmentation. In the proposed STVGG, transfer learning and self-paced learning are used to solve the unbalance and the difficult sample learning. STVGG transfers ImageNet network parameters to VGG network and calculates the loss value of each image in the for-

ward propagation with the age parameter. In addition, the trained STVGG model of colorectal classification is shared to Unet segmentation model to deal with distinguishing polyp and normal tissues.

2. Materials and Methods

2.1. Data Acquisition and Preprocessing. A total of 50 patients were examined under colonoscopies, and images were collected from the anorectal department of a hospital in Shaanxi Province, China, under ethical approval. Three experienced endoscopists were invited to classify the normal tissue, polyp, and tumor, and the ground truth was acquired. The data preprocessing was as follows:

Firstly, data filtering: uncleaned or unclear colorectal images were removed. After image filtering, the set of endoscopic images consisted of 487, 1374, and 1200 images with normal tissue, polyp, and tumor, respectively, taken under either white light (WL) or narrow band imaging (NBI) endoscopy.

Secondly, dataset split: the data were divided into training set, validation set, and test set according to the ratio of 2:1:2.

Thirdly, data argumentation: the argumentation methods were rotation, flip, translation, and cropping. The training set and validation set were argumentized by four times

Finally, data resizing: the data was resized to $440 \times 440 \times 3$ to maintain the integrity of the intestinal wall.

2.2. Automatic Classification in Colorectal Endoscopy Based on STVGG. Because the performance of training network is poor by using colorectal images directly, Network pre-trained on ImageNet is introduced to obtain a good classification result. Meanwhile, polyp areas are more difficult to be classified than normal tissue and tumor. This paper introduces self-paced regularization items to assign different sample weights for training samples. Self-paced learning injects the difficulty metric into the optimization model and updates the model parameters based on the current sample ordering and the metric based on the learning effect. It obtains a new round of difficulty ordering of samples and finally achieves the purpose of adaptive sample ordering.

In our method, in order to fully use data of ImageNet, the parameters of C_1 and C_2 from pretrained VGG19 model on ImageNet are transferred to STVGG. And the practical colorectal images are used as training data to update other layer parameters of the STVGG model. The self-paced learning algorithm is introduced to STVGG for dealing with those difficult and unbalance samples to improve classification performance. The overview of STVGG classification method is shown in Figure 1, where C_i represents the i^{th} convolutional patch, F represents the fully connected layer, and G represents the global average pooling layer. In this study, the first fully connected layer F_6 of VGG19 is replaced with the global average pooling layer G_6 to reduce the amount of model parameters and prevent overfitting and to get the pretrained model with the parameters of C_1 and C_2 layers.

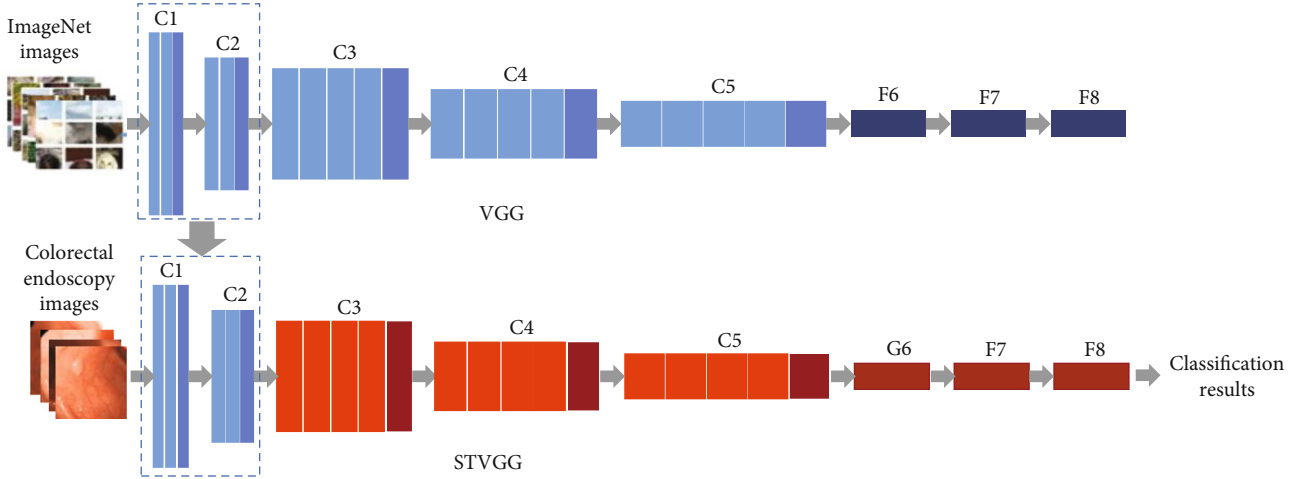


FIGURE 1: The overview of STVGG classification method.

The cross-entropy is selected as loss function $l(y_i, g(\mathbf{X}_i))$. Forward propagation is used to calculate the loss l in STVGG network:

$$l(y_i, g(\mathbf{X}_i)) = - \sum_{j=1}^c I\{y_i = j\} \log \left(\frac{e^{\omega_j^T z_i^l}}{\sum_{k=1}^c e^{\omega_k^T z_i^l}} \right). \quad (1)$$

In the Eq. (1), c is the number of disease categories in the colorectal endoscopy dataset, and ω_j^T is the weight parameter of the j^{th} output node. z_i^l represents the output vector in the last fully connected layer of the network. $I\{y_i = j\} \in \{0, 1\}$, when the predicted result of the sample is consistent with the label $I = 1$, otherwise $I = 0$.

Furthermore, self-paced learning is used to modify the STVGG network. The network objective function is rewritten as follows:

$$\min_{\omega, b} E_{\text{spl}}(\omega, b) = \min_{\omega, b} \left[\frac{1}{n} \sum_{i=1}^n v_i \cdot l(y_i, g(\mathbf{X}_i)) + f(v_i, \lambda) \right] \quad (2)$$

Parameters w and b are the weight and bias of the STVGG network, respectively, $\mathbf{v} = [v_1 \cdots v_i \cdots v_n]$ represents the weight of n samples, and $f(v_i, \lambda)$ is the binary self-paced regular term defined in Eq. (3).

$$f^H(v_i, \lambda) = -\lambda v_i; v_i * (l, \lambda) = \begin{cases} 1 & \text{if } l < \lambda, \\ 0 & \text{if } l \geq \lambda. \end{cases} \quad (3)$$

w , b , and \mathbf{v} of the STVGG network are optimized by iteration until the model converges to get a good classification network. Flowchart of STVGG algorithm is shown in Algorithm 1.

2.3. Automatic Segmentation in Polyp Image. After classification task is completed, the parameters $C_1 - C_5$ of the trained STVGG in the colorectal endoscopy classification task is shared to the segmentation task as the code part, while the

Unet network framework is used in colorectal endoscopy segmentation task. And the decoding part of the original Unet is also adjusted with the corresponding encode part. Compared with the original Unet, the channel number of downsampling in the last layer is not increased for the proposed model. The framework of our segmentation model is shown in Figure 2.

Each rectangular box corresponds to a multichannel feature map. The number on the left side of the rectangular box indicates the size of each channel of the feature map. The number at the top of the rectangle indicates the channel number in the feature map. The blue, red, green, and purple arrows indicate the convolution operation with a convolution kernel size of 3×3 , the max pooling with a stride of 2×2 , and the upsampling and the convolution with a convolution kernel size of 1×1 , respectively. The gray arrow indicates that the feature map of the encoding part is cropped and copied with the feature map of the decoding part.

2.4. Comparison with Other Methods. The selection of comparison methods is based on the baseline VGG model adding some training strategies, and the specific strategies are as follows:

(1) VGG19 with transfer learning strategies (VGG+TL)

The parameters $C_1 - C_2$ in VGG19 are transferred from ImageNet network to extract low-level features well shared with natural images in colorectal endoscopy images.

(2) VGG19 with the strategy of structure retention color normalization (VGG+SRCN)

The data are collected from different periods, different patients, and equipment in different periods. Therefore, SRCN strategy is used so that color features of processed image tend to be consistent and reduce intraclass differences.

(3) VGG19 with strategy of spatial pyramid pooling (VGG+SPP)

SPP [26] layer on the VGG19 network is adopted, and it obtains a fixed length feature vector to aggregate the features and avoid geometric distortion in feature maps.

Flowchart of STVGG algorithm.
 Input: Training set $\mathbf{D} = \{(\mathbf{x}_i, y_i), i = 1 \dots n\}$, \mathbf{x}_i represents the i^{th} training data, and y_i is the i^{th} data label.
 Initialization parameter: "age parameter" λ , a suitable initial value is given according to the approximate value range of the presample training error value; initialize the sample weight vector \mathbf{v} .
 Model training parameter settings: total number of training iterations epoch, minimum batch size for training and verification, initial learning rate during model training α , and decay rate of learning rate φ ; update increment of age parameters $k, k > 0$.
 a) Calculate network weights \mathbf{w} and bias \mathbf{b} by Eq. (2)
 b) Calculate and update loss function $l(y_i, g(\mathbf{X}_i))$
 c) Calculate self-paced regular term $f(v_i, \lambda)$ and update weight vector \mathbf{v}
 d) Calculate and update $\min_{\omega, b} E_{\text{spl}}(\omega, b)$
 e) Update age parameters λ and learning rate $\alpha, \lambda = \lambda + k, \alpha = \alpha \cdot \varphi, \varphi < 1$
 f) Repeat steps a to e until the number of iterations epoch = 0
 output: network weights \mathbf{w} and bias \mathbf{b}

ALGORITHM 1.

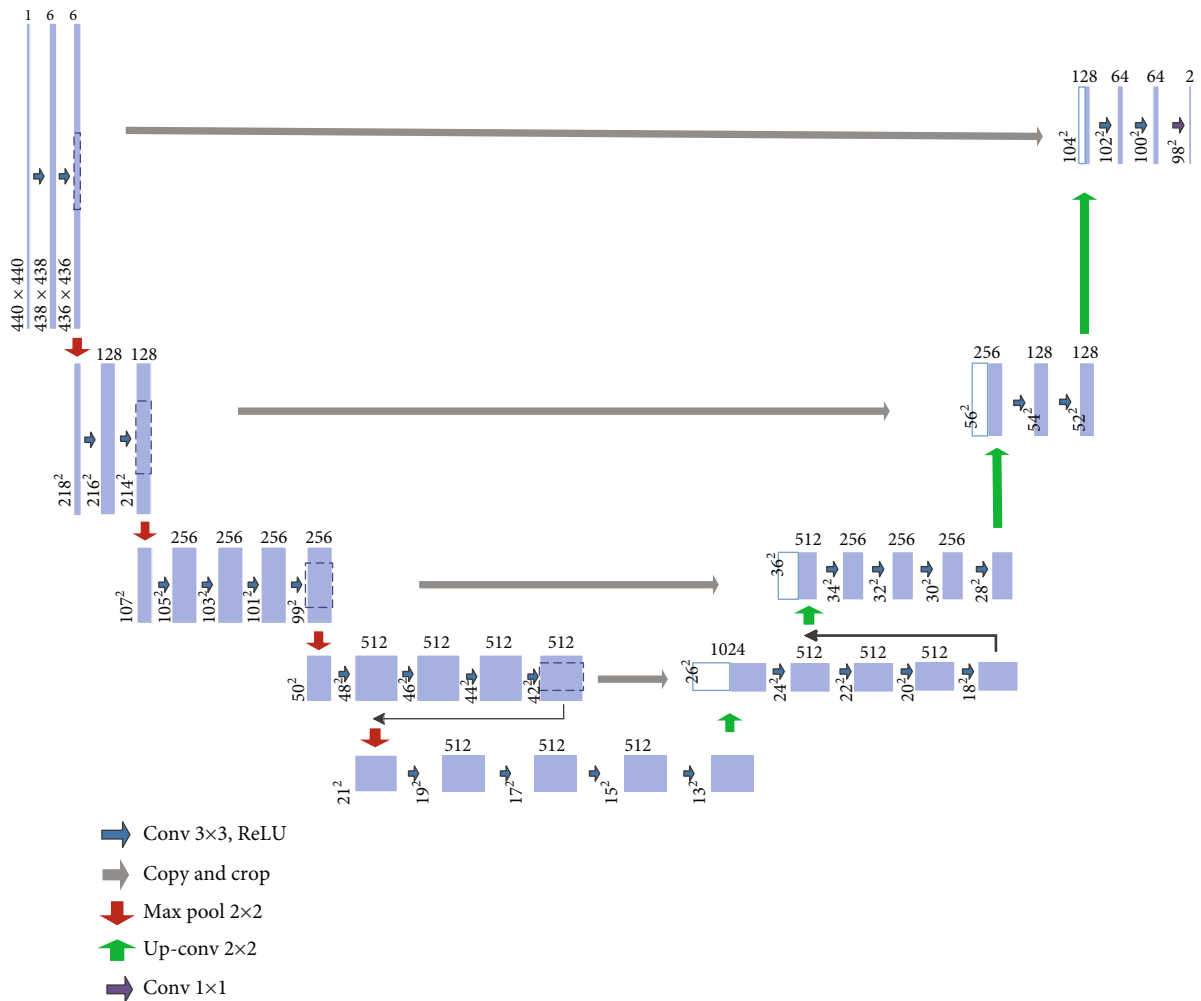


FIGURE 2: The framework of our segmentation model.

2.5. Evaluation of the Classification and Segmentation. The segmentation results are evaluated both visually and quantitatively, given the ground truth, our classification and segmentation results. The segmentation performance is eval-

uated by these evaluation metrics: accuracy, sensitivity, specificity, and dice similarity coefficient (DSC). We use TP, FP, TN, and FN to represent true positive, false positive, true negative, and false negative. And L_1 and L_2 represent the

TABLE 1: Classification accuracy obtained for different methods.

Category	VGG	VGG+SRCN	VGG+SPP	VGG+TL	STVGG
Tumor	0.98	0.98	0.94	0.96	0.98
Normal tissue	0.90	0.94	0.99	0.95	0.99
Polyp	0.7	0.84	0.91	0.89	0.95
Average accuracy	0.76	0.87	0.93	0.91	0.96

manual annotation and our method segmentation results, respectively, and these indexes are defined as follows:

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN}, \quad (4)$$

$$\text{Sensitivity} = \frac{|L_1 \cap L_2|}{|L_1|}, \quad (5)$$

$$\text{Specificity} = \frac{|\sim(L_1 | L_2)|}{|\sim L_1|}, \quad (6)$$

$$\text{DSC} = \frac{2 |L_1 \cap L_2|}{(|L_1| + |L_2|)}. \quad (7)$$

3. Experimental Results and Discussion

In this paper, the experimental environment was set in Python3.6.0, Tensorflow-GPU 1.7.0, Keras 2.1.3, SimpleITK 0.8.1, Nvidia Titan Pascal GPU (1080 Titan), and Cudnn V9.0. Verified by experiment, the colorectal image size of 440 was superior to 330 and 540. The average accuracy of the final test set was 0.03 and 0.01 higher than the latter two, respectively. The colorectal data with original image size range of 228 to 586 was resized to 440×440 .

Choosing optimization functions. Experiments showed that the average classification accuracy of SGD is 0.02 and 0.05 higher than RmSprop and Adam, respectively. Therefore, SGD was finally selected as our optimization function.

Selecting transfer layers. Experiments froze the parameters of the 4, 8, 12, and 16 layers, respectively. It showed that when the parameters of the first 4 layers were frozen, the best average classification accuracy was achieved.

Cross-entropy are chosen as loss function. The relevant parameters were as follows: learning rate is 0.0001, decay = $1e - 6$, and the parameter of Nesterov Momentum was set to 0.9. The batch size was 8. λ was initialized to 1.1, and the updating step was 0.05. As training began, λ became larger and the tolerance of difficult samples was greater.

3.1. Colorectal Endoscopy Image Classification. The colorectal endoscopy image classification accuracy is shown in Table 1. It can be seen from Table 1 that the polyp accuracy of VGG network was the lowest. The classification accuracies of VGG+SRCN on polyps and normal tissue are improved as it could decrease the intraclass differences. VGG+SPP also improved the classification accuracies of normal tissue and colorectal polyp but polyp accuracy was relatively low because the edges of normal tissue and polyp types are blurred. As an improvement strategy, the accuracy of

VGG+TL was also improved. But compared with strategies of SRVN and SPP, the effect is not significant. In this study, STVGG was proposed and the experimental results showed that the overall accuracies were greatly improved.

The main reason is that STVGG can classify difficult-to-classify samples, for example, the small inflammatory or hyperplastic polyps which are very similar to normal colonoscopy images, and the polyps with ulcers, large areas of bleeding, and reticulated polyps, which are closer to the characteristics of tumor. The STVGG method can significantly improve the accuracy of polyp under the condition of ensuring the classification accuracy of tumor and normal ones, and the method converges in about 10 generations of training.

3.2. Polyp Image Segmentation. In the above colorectal endoscopy image classification task, a relatively good classification result was obtained by STVGG model. Therefore, polyp segmentation was designed based on classification task. Doctors usually used polyp's images to make a decision whether surgical resection is required based on pathological diagnosis. Figure 3 shows five sequences of polyp images. (Ai) is the original image, (Bi) is ground truth, (Ci) is the segmentation result of Segnet network, (Di) is the segmentation result of Unet network, (Ei) is the segmentation result of TLVGG network, and (Fi) is the segmentation result of STVGG network.

Figure 3 indicates that the results of Segnet method are greatly affected by the surrounding environment, and the segmentation result is not good. The results of Unet method are more superior than those of Segnet in big target but the effect is not obvious. Compared to those methods, the results of TLVGG method made great progress especially in surrounding and small target, but the segmentation results of large targets are not ideal. Our method shows the best results, no matter it is segmentation of large and small objects or environmental interference.

Table 2 shows that the Segnet segmentation method does not segment the polyp in its complete shape. Using Unet for segmentation of polyps is accurate, but oversegmentation is also obvious. The Unet is more sensitive to light spots in the imaging process, and it is easy to treat the reflective part as a polyp.

The segmentation performances of the TLVGG network were obviously better than Segnet and Unet. The segmentation target contour was close to the real target, but there were still missed detections. The STVGG model worked best because ImageNet network parameters were transferred to VGG network to acquire good initial network. And self-paced learning was used to optimize the network so that the classification performance of label unbalanced samples was improved.

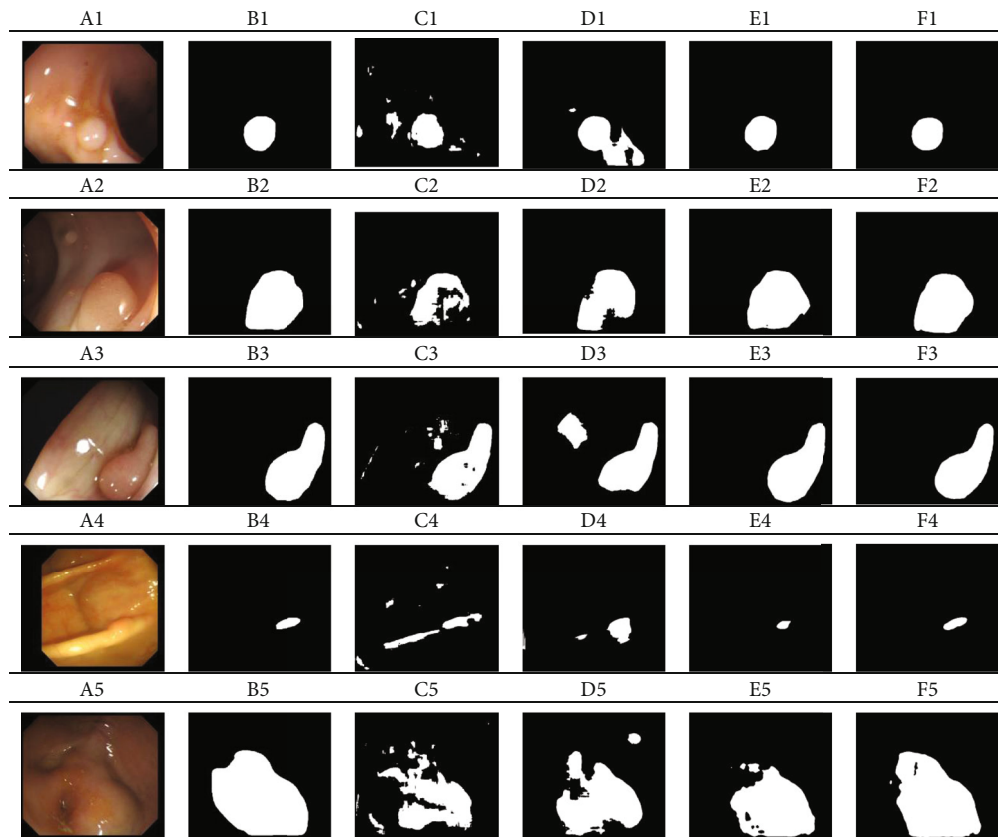


FIGURE 3: Segmentation in colorectal endoscopy images. (Ai) Original images, (Bi) ground truth, (Ci) Segnet, (Di) Unet, (Ei) TLVGG, and (Fi) ours.

TABLE 2: Segmentation indexes obtained from different methods.

	Segnet	Unet	TLVGG	STVGG
DSC	0.6210 ± 0.2370	0.6980 ± 0.3005	0.8267 ± 0.2066	0.8455 ± 0.2030
Sen	0.6916 ± 0.2677	0.7591 ± 0.3317	0.8222 ± 0.2462	0.8323 ± 0.2201
Spe	0.9766 ± 0.0180	0.9834 ± 0.0235	0.9933 ± 0.0095	0.9949 ± 0.0067

4. Conclusions

To address this issue of label unbalanced and difficult colorectal data, we presented an automatic processing pipeline for classification and segmentation based on colorectal images. STVGG network used transfer learning and self-paced learning in order to acquire good initial network and solve the problem of label unbalanced and difficult sample classification. And then STVGG network was shared as the encoding part of Unet as encoder of the segmentation task, and image segmentation task was achieved. The experimental illustrated that the proposed method obtained higher classification accuracy (96%) and segmentation performance compared with other a few methods. This proposed method may be applied to other image researches, such as stomach, ear, nose, and throat. Possible future improvements can be made in parameter adaptation.

Data Availability

The data used in the article are from the anorectal department of a hospital in Shaanxi Province, China, under ethical approval.

Conflicts of Interest

The authors declare that there is no conflict of interest regarding the publication of this paper.

Authors' Contributions

Yao Yao designed the algorithms and made the experiments. Shuiping Gou participated in guiding research plan. Ru Tian did the comparative experiment. Xiangrong Zhang polished the article. Shuixiang He provided the data and verified the clinical effect.

Acknowledgments

This study is supported by the Natural Science Foundation of Shaanxi Province under Grant Nos. 2019ZDLGY03-02-02 and 2017JM8026, Research Plan of Improving Public Scientific Quality in Shaanxi Province, No. E219360001, the Fundamental Research Funds for the Central Universities, No. JC2001, and Focus on developing of Shaanxi Province under Grant No. 2020GY-049. This is a research project of visiting engineers in colleges and universities of Zhejiang Province in 2020.

References

- [1] F. Bray, J. Ferlay, I. Soerjomataram, R. L. Siegel, L. A. Torre, and A. Jemal, "Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 68, no. 6, pp. 394–424, 2018.
- [2] A. G. Zauber, S. J. Winawer, M. J. O'Brien et al., "Colonoscopic polypectomy and long-term prevention of colorectal cancer deaths," *Obstetrical & Gynecological Survey*, vol. 67, pp. 687–696, 2012.
- [3] K. Søreide, B. S. Nedrebø, A. Reite, K. Thorsen, and H. Kørner, "Endoscopy, morphology, morphometry and molecular markers: predicting cancer risk in colorectal adenoma," *Expert Review of Molecular Diagnostics*, vol. 9, no. 2, pp. 125–137, 2014.
- [4] D. A. Joseph, R. G. S. Meester, and A. G. Zauber, "Colorectal cancer screening: estimated future colonoscopy need and current volume and capacity," *Cancer*, vol. 122, no. 16, pp. 2479–2486, 2016.
- [5] M. J. Van den Bent, "Interobserver variation of the histopathological diagnosis in clinical trials on glioma: a clinician's perspective," *Acta Neuropathologica*, vol. 120, no. 3, pp. 297–304, 2010.
- [6] D. G. Hewett, T. Kaltenbach, Y. Sano et al., "Validation of a simple classification system for endoscopic diagnosis of small colorectal polyps using narrow-band imaging," *Gastroenterology*, vol. 143, no. 3, pp. 599–607, 2012.
- [7] M. G. Ertosun and D. L. Rubin, "Automated grading of gliomas using deep learning in digital pathology images: a modular approach with ensemble of convolutional neural networks," in *In AMLA Annual Symposium Proceedings*, vol. 2015, pp. 1899–1908, American Medical Informatics Association, 2015.
- [8] B. E. Bejnordi, M. Veta, P. J. Van Diest et al., "Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer," *JAMA*, vol. 318, no. 22, pp. 2199–2210, 2017.
- [9] A. Teramoto, T. Tsukamoto, and Y. Kiriya, "Automated classification of lung cancer types from cytological images using deep convolutional neural networks," *Biomedical Research*, vol. 4067832, 2017.
- [10] J. Wang et al., "Notice of Violation of IEEE Publication Principles: Bag-of-Features Based Medical Image Retrieval via Multiple Assignment and Visual Words Weighting," *IEEE Transactions on Medical Imaging*, vol. 30, no. 11, pp. 1996–2011, 2011.
- [11] A. Meier, K. Nekolla, S. Earle et al., "End-to-end learning to predict survival in patients with gastric cancer using convolutional neural networks," *Annals Oncology*, vol. 29, 2018.
- [12] G. Litjens, C. I. Sánchez, N. Timofeeva et al., "Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis," *Scientific reports*, vol. 6, p. 26286, 2016.
- [13] H. Y. Chang, C. K. Jung, J. I. Woo et al., "Artificial intelligence in pathology," *Journal of pathology and translational medicine*, vol. 53, no. 1, pp. 1–12, 2019.
- [14] Y. Xu, Z. Jia, L. B. Wang et al., "Large scale tissue histopathology image classification, segmentation, and visualization via deep convolutional activation features," *BMC bioinformatics*, vol. 18, p. 281, 2017.
- [15] N. Thakur, H. Yoon, and Y. Chong, "Current trends of artificial intelligence for colorectal cancer pathology image analysis: a systematic review," *Cancers*, vol. 12, no. 7, p. 1884, 2020.
- [16] B. Korbar, A. M. Olofson, A. P. Miraflor et al., "Deep learning for classification of colorectal polyps on whole-slide images," *Journal of pathology informatics*, vol. 8, 2017.
- [17] P. Sena, R. Fioresi, F. Faglioni, L. Losi, G. Faglioni, and L. Roncucci, "Deep learning techniques for detecting preneoplastic and neoplastic lesions in human colorectal histological images," *Oncology Letters*, vol. 18, pp. 6101–6107, 2019.
- [18] O. Iizuka, F. Kanavati, K. Kato, M. Rambeau, K. Arihiro, and M. Tsuneki, "Deep learning models for histopathological classification of gastric and colonic epithelial tumours," *Scientific Reports*, vol. 10, p. 1504, 2020.
- [19] G. Urban, P. Tripathi, T. Alkayali et al., "Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy," *Gastroenterology*, vol. 155, pp. 1069–1078.e8, 2018.
- [20] P. Klare, C. Sander, M. Prinzen et al., "Automated polyp detection in the colorectum: a prospective study (with videos)," *Gastrointestinal Endoscopy*, vol. 89, pp. 576–582.e1, 2019.
- [21] P. Wang, T. M. Berzin, J. R. G. Brown et al., "Real-time automatic detection system increases colonoscopic polyp and adenoma detection rates: a prospective randomised controlled study," *Gut*, vol. 68, no. 10, pp. 1813–1819, 2019.
- [22] Q. Li, G. Yang, and Z. Chen, "Colorectal polyp segmentation using a fully convolutional neural network," in *2017 10th International Congress on Image and Signal Processing*, pp. 17–21, ShangHai, 2017.
- [23] Y. J. Yang, B.-J. Cho, M.-J. Lee et al., "Automated classification of colorectal neoplasms in white-light colonoscopy images via deep learning," *Journal of clinical medicine*, vol. 9, 2020.
- [24] R. Zhang, Y. Zheng, and T. W. C. Mak, "Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain," *IEEE Journal of biomedical and health informatics*, vol. 21, no. 1, pp. 41–47, 2017.
- [25] K. Wickstrøm, M. Kampffmeyer, and R. Jenssen, "Uncertainty and interpretability in convolutional neural networks for semantic segmentation of colorectal polyps," *Medical Image Analysis*, vol. 60, p. 101619, 2020.
- [26] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," *IEEE Transactions on Pattern Analysis & Machine Intelligence*, vol. 37, no. 9, pp. 1904–1916, 2014.