

Research Article

Image Aided Recognition of Wireless Capsule Endoscope Based on the Neural Network

Bin Lu 

Department of Gastrointestinal Surgery, Affiliated Hospital of Shaoxing University (The Shaoxing Municipal Hospital), Shaoxing 312000, China

Correspondence should be addressed to Bin Lu; luchen131415@126.com

Received 19 January 2022; Accepted 22 March 2022; Published 7 April 2022

Academic Editor: Nima Jafari Navimipour

Copyright © 2022 Bin Lu. 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.

Wireless capsule endoscopy is an important method for diagnosing small bowel diseases, but it will collect thousands of endoscopy images that need to be diagnosed. The analysis of these images requires a huge workload and may cause manual reading errors. This article attempts to use neural networks instead of artificial endoscopic image analysis to assist doctors in diagnosing and treating endoscopic images. First, in image preprocessing, the image is converted from RGB color mode to lab color mode, texture features are extracted for network training, and finally, the accuracy of the algorithm is verified. After inputting the retained endoscopic image verification set into the neural network algorithm, the conclusion is that the accuracy of the neural network model constructed in this study is 97.69%, which can effectively distinguish normal, benign lesions, and malignant tumors. Experimental studies have proved that the neural network algorithm can effectively assist the endoscopist's diagnosis and improve the diagnosis efficiency. This research hopes to provide a reference for the application of neural network algorithms in the field of endoscopic images.

1. Introduction

Capsule endoscopy is a convenient, simple, and fast method of gastrointestinal examination, especially in the diagnosis of small bowel diseases. Small intestine capsule endoscopy technology is the starting point of capsule endoscopy. After more than ten years of development, capsule endoscopy has basically become an important inspection item for small bowel diseases. Compared with the traditional endoscope technology, the capsule endoscope can obtain the image of the entire digestive tract in real time during the whole process without feeling uncomfortable. The patient's entire digestive tract image can be displayed on the monitor; there are significant differences in the texture between the lesion area and the nonlesion area in the capsule endoscopic image, so the diagnosis of disease conditions by comparing textures in images has been widely used in clinical practice. Thousands of images will be generated during capsule endoscopic surgery. If the doctor reads and judges them one by one, it is easy to miss valuable information. Therefore, it is particularly important to find a computer-aided analysis method

that has a good feature extraction effect on the capsule endoscopic image. Existing studies have shown that it is feasible to identify various abnormalities in capsule endoscopic images through convolutional neural networks. Gomes developed an unsupervised homography evaluation method in the capsule endoscope framework and then applied it to the capsule positioning system. The network can evaluate the homography between two images [1]. Leenhardt has developed a computer-aided diagnostic tool to detect vasodilation. The improved algorithm based on the convolutional neural network has high diagnostic performance and can detect vasodilation in the static frame of small bowel capsule endoscopy [2]. Sainju proposed a supervised method to automatically detect the bleeding area in the capsule endoscope frame or image. During surgery, segmentation methods can be used to obtain regions from the image, and a well-trained neural network can identify data patterns generated by bleeding and nonbleeding regions [3]. Aoki researched whether a system based on convolutional neural networks can reduce the reading time of endoscopists and increase disease detection rates. The results show that the

convolutional neural network system designed by the team can reduce reading time, but it has no significant effect on the detection rate of lesions [4]. Li proposed a new computer-aided system for detecting bleeding areas in capsule endoscopic images. This scheme is very effective for detecting bleeding areas and can be used to distinguish between normal and bleeding areas in capsule endoscopic images [5]. Shahril studied the recognition performance of the bleeding area of the capsule endoscope image based on the deep convolutional neural network algorithm and proposed a preprocessing technology classification, which can distinguish between the normal area and the bleeding area by improving the accuracy of the capsule endoscope image. Experiments show that compared with capsule endoscopy images that do not use this enhancement technology, capsule endoscopy images that use this enhancement technology have a better classification effect [6]. Chen proposed a general depth framework of spatio-temporal cascade to understand the most common content in the entire gastrointestinal video. Compared with other methods, their proposed network framework can perform noise content detection and terrain segmentation at the same time, thereby reducing the number of images that need to be inspected and segmenting images of different lesion areas more accurately [7]. Yiftach has developed a deep learning algorithm that can automatically grade Crohn's disease during capsule endoscopy. The conclusion shows that the convolutional neural network has achieved high accuracy in detecting severe Crohn's disease. Convolutional neural network-assisted capsule endoscopy readings in Crohn's disease patients can potentially facilitate and improve the diagnosis and monitoring of these patients [8]. The image recognition and machine learning technology in artificial intelligence can effectively reduce the endoscopic image reading work of gastroenterologists, reduce the workload of reading physicians, quickly identify various suspected lesions, and improve the diagnostic efficiency of the capsule endoscope. There are two main purposes. The first is to reduce the video of capsule endoscopy to increase the speed of capsule endoscopy, and the second is to perform image detection for specific diseases. However, image features mainly include shape, color, and texture, and feature extraction and selection directly affect the performance of subsequent image classifiers. The scene of the capsule endoscope image is complex and changeable, and there are uncertain factors such as bubbles, peristalsis, and lighting changes. Therefore, effective and automatic filtering of redundant images is still a problem. Based on deep learning, this research hopes to establish an effective capsule endoscopic diagnosis system, and the neural network is trained for prediction after preprocessing by extracting texture features that can significantly show the lesion to improve the speed and accuracy of doctors' diagnosis and help doctors make better diagnoses.

2. Methodology

2.1. Capsule Endoscopy. Capsule endoscopy is one of the advanced methods to detect and diagnose human digestive

tract diseases. Compared with traditional inserting gastrointestinal endoscopes, capsule endoscopes have superior performance such as noninvasive, safe, and full-process detection [9–12]. The length of the capsule endoscope is about 20 mm and the diameter is about 10 mm. Its internal structure is shown in Figure 1. It is mainly composed of CMOS image sensor, lens, LED, ASIC transmitter, and power module.

The working principle of the capsule endoscope is shown in Figure 2. After the patient swallows the capsule endoscope, with the help of gravity and natural peristalsis of the gastrointestinal tract, the endoscope moves forward, passing through the mouth, esophagus, stomach, duodenum, jejunum, ileum, colon, and other parts, and finally discharged through the anus. During this process, the capsule endoscope will perform full range imaging (2–3 frames/sec) of the digestive tract it passes through, and the image will be stored in the data recorder carried by the patient through wireless transmission. According to statistics, the average residence time of the capsule endoscope in the digestive tract is about 8 hours, and thousands of color images can be collected during the entire process. Due to the characteristics of gastrointestinal peristalsis, there are often a large number of redundant images in the collected images, which have extremely high similarities. Doctors manually screen out images with lesions from these large numbers of images. Screening is labor intensive and inefficient, and it usually takes 2–3 hours to focus on carefully examining each frame. Since the endoscope has passed through different parts of the digestive tract, such as the stomach, duodenum, and colon, the color information, intestinal diameter, and movement state of each part are different, so the focus is not good during endoscopic shooting. The captured image scenes are also complex and changeable, with great differences in structure, color, and texture. There may be many uncertain factors, such as food residues, air bubbles, digestive juices, and blood. Therefore, how to quickly and automatically filter out redundant images in capsule endoscopy to improve the diagnosis efficiency of doctors is a hot issue in the field of neural network diagnosis technology and medical image processing.

2.2. Image Preprocessing. The RGB color model is a color superposition model based on human perception of colors, that is, three primary colors are added together in different ways to represent various colors. The RGB color model is device-oriented and is mainly used for the representation and display of images in the electronics industry. The laboratory color model is a color model based on physiological characteristics and has nothing to do with equipment. The color gamut of the lab color model is much larger than that of the RGB color model, CMYK color model, and human vision, and it makes up for the shortcomings of uneven color distribution in the RGB color model. Although lab color space is not often used, it has excellent results consistent with human visual analysis when analyzing certain images with specific color characteristics [13]. Usually, the original digital image we obtain is an RGB image, and the RGB image needs

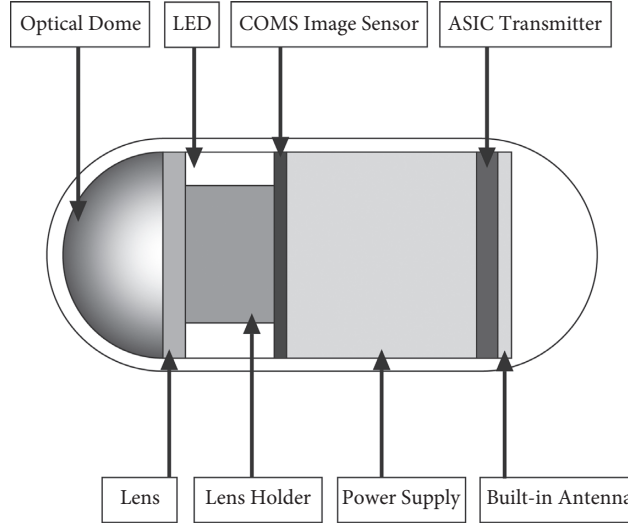


FIGURE 1: The structure of the capsule endoscope.

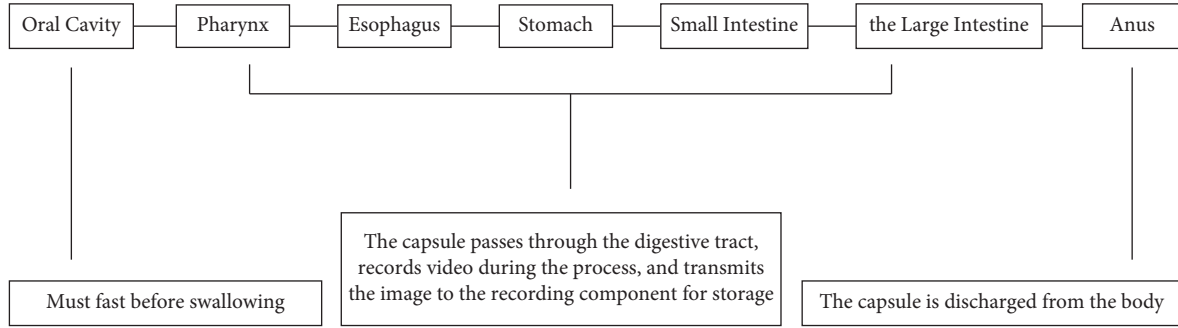


FIGURE 2: Schematic diagram of the capsule endoscope.

to be converted to XYZ color space first. Assuming the representation of the pixel in the RGB color space, first convert the RGB mode to the XYZ color space in the following way (XYZ is the coordinate system, X, Y, and Z are the 3 components of the target color space):

$$\begin{cases} R = \text{gamma}\left(\frac{r}{255.0}\right), \\ G = \text{gamma}\left(\frac{g}{255.0}\right), \\ B = \text{gamma}\left(\frac{b}{255.0}\right). \end{cases} \quad (1)$$

Among them, the gamma function is the gamma correction in the sRGB standard, which is usually defined as

$$\text{gamma}(x) = \begin{cases} \left(\frac{x + 0.055}{1.055}\right)^{2.4}, & \text{if } (x) > 0.04045, \\ \frac{x}{12.92}, & \text{others,} \end{cases} \quad (2)$$

$$\begin{cases} X = 0.4124 * R + 0.3576 * G + 0.1805 * B, \\ Y = 0.2126 * R + 0.7152 * G + 0.0722 * B, \\ Z = 0.0193 * R + 0.1192 * G + 0.9505 * B. \end{cases}$$

It can be seen from the above formula that when the X, Y, and Z components are very close to 1, the coefficients of R, G, and B are 0.950456, 1.0, and 1.088754, respectively. To map to the RGB mode in the same range, the formula for converting XYZ mode to lab mode is as follows:

$$\begin{cases} L = 116f\left(\frac{Y}{Y_n}\right) - 16, \\ a = 500\left[f\left(\frac{X}{X_n}\right) - f\left(\frac{Y}{Y_n}\right)\right], \\ b = 200\left[f\left(\frac{Y}{Y_n}\right) - f\left(\frac{Z}{Z_n}\right)\right]. \end{cases} \quad (3)$$

Among them, X_n , Y_n , and Z_n are usually set to 95.047, 100.0, and 108.883, and

$$f(x) = \begin{cases} t^{1/3}, & \text{if } (t) > \left(\frac{6}{29}\right)^3, \\ \frac{1}{3} * \left(\frac{29}{6}\right)^2 t, & \text{others.} \end{cases} \quad (4)$$

2.3. Extract Texture Features. There are significant differences in the texture between the lesion area and the non-lesion area in the capsule endoscopic image [14–16], so it is very important to extract texture features. In this research, Daubechies function is chosen as the wavelet basis function [17]. The formula can be expressed as

$$\begin{aligned} Q^i &= \{L_a^i, H_\beta^i\}, \\ i &= 1, 2, 3; \\ \alpha &= 1, 2, 3; \\ \beta &= 1, 2, \dots, 9. \end{aligned} \quad (5)$$

Among them, L represents the low frequency part of the image when it is decomposed in the horizontal and vertical directions; H represents the middle and high frequency part of the image; α represents the decomposition level; β represents the wavelet frequency band; and i represents the R , G , and B components of the image. Commonly used computer-aided analysis methods are mainly carried out in the low frequency band of the image. In this article, we choose the middle and high frequency bands to reconstruct the image and extract texture information. The selected subbands are represented as follows:

$$\begin{aligned} O^i &= \{H_\beta^i\}, \\ i &= 1, 2, 3; \\ \beta &= 1, 2, \dots, 9. \end{aligned} \quad (6)$$

Calculate the matrix $W_\theta T$, $T \in \{R, G, B\}$ of the RGB components of the converted image. The pixel pair (m, n) in the matrix means the distance in the image is d , the color scale is m and n , and the number of occurrences of two pixels in the direction θ . In applications, θ is usually selected as 0° , 45° , 90° , and 135° . It can reflect the distribution characteristics of brightness, it can also reflect the location distribution characteristics between pixels of the same or close to brightness, and it is also a second-order statistical feature of

image brightness changes. Then, normalize the obtained co-occurrence matrix, and write $W_\theta T(m, n)$ as the value of the pixel (m, n) into the normalized co-occurrence matrix, where $T \in \{R, G, B\}$, $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$. The expression is as follows:

$$\begin{aligned} E_T^\theta &= \sum_{m=1}^D \sum_{n=1}^D [w_T^\theta(m, n)]^2, \\ I_T^\theta &= \sum_{m=1}^D \sum_{n=1}^D (m-n)^2 w_T^\theta(m, n), \\ II_T^\theta &= - \sum_{m=1}^D \sum_{n=1}^D w_T^\theta(m, n) \log w_T^\theta(m, n), \\ A_T^\theta &= \frac{\sum_{m=1}^D \sum_{n=1}^D (m, n) w_T^\theta(m, n) - \mu_1^\theta \mu_2^\theta}{\sigma_1^\theta \sigma_2^\theta}, \end{aligned} \quad (7)$$

$$\mu_1^\theta = \sum_{m=1}^D m \sum_{n=1}^D w_T^\theta(m, n),$$

$$\mu_2^\theta = \sum_{n=1}^D n \sum_{m=1}^D w_T^\theta(m, n),$$

$$\sigma_1^\theta = \sum_{m=1}^D (m - \mu_1^\theta)^2 \sum_{n=1}^D w_T^\theta(m, n),$$

$$\sigma_2^\theta = \sum_{n=1}^D (n - \mu_2^\theta)^2 \sum_{m=1}^D w_T^\theta(m, n).$$

Among them, $E_\theta T$, $I_\theta T$, $II_\theta T$, and $A_\theta T$, respectively, represent the energy, contrast, entropy, and correlation of the co-occurrence matrix θ of each color component; D is the maximum color level of the image. In this study, $d = 1$ is used to construct the texture feature vector of the image based on the feature value calculated above, which can be given by the following formula.

$$Z_T = \left[\overline{E}_T^\theta, \widehat{E}_T^\theta, \overline{I}_T^\theta, \widehat{I}_T^\theta, \overline{II}_T^\theta, \widehat{II}_T^\theta, \overline{A}_T^\theta, \widehat{A}_T^\theta \right]. \quad (8)$$

Among them, $\overline{X}_T^\theta = \sum_{\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} X_T^\theta / 4$, $\widehat{X}_T^\theta = \sqrt{\sum_{\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}} (X_T^\theta - \overline{X}_T^\theta)^2 / 4}$; $X \in \{E, I, II, A\}$, $T \in \{R, G, B\}$, $\theta \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$. The 8-dimensional texture features of the RGB components obtained above are correspondingly added as the final extracted texture features:

$$F_{\text{texture}} = Z_R + Z_G + Z_B. \quad (9)$$

2.4. Network Training. This research is based on neural network algorithms to realize the recognition and segmentation of endoscopic images of digestive tract capsules. The support vector machine (SVM) is a statistical learning theory proposed and established by Vapnik et al. It is usually used for data classification and regression prediction. It has

many unique advantages in solving nonlinear, high-dimensional pattern, and small sample recognition problems [18–21]. The principle is as follows: find an optimal hyperplane that meets the classification requirements and maximize the distance between the edges on both sides of the hyperplane on the basis of ensuring the classification accuracy. Assuming that the number of samples in the training set is a and the training set is $\{(x_1, y_1), (x_2, y_2), \dots, (x_a, y_a)\}$, the SVM maps the input to the high-dimensional feature space Λ to $\phi(X)$, and the corresponding classification function is

$$f(x) = \text{sgn}(w \cdot \phi(x) + b). \quad (10)$$

In the formula, w and b represent the weight vector and the offset, respectively. According to the principle of structural risk minimization, formula (10) is transformed into

$$\begin{aligned} \min J &= \frac{1}{2} \|w\|^2 + C \sum_{c=1}^a (\xi_c^* + \xi_c), \\ \text{s.t.} \quad &\begin{cases} y_c - w \cdot \phi(x) - b \leq \varepsilon + \xi_c, \\ w \cdot \phi(x) + b - y_c \leq \varepsilon + \xi_c^*, \\ \xi_c, \xi_c^* \geq 0, c = 1, 2, \dots, a. \end{cases} \end{aligned} \quad (11)$$

In the formula, ξ_c and ξ_c^* are the upper and lower limits of the relaxation factor, respectively; C represents the penalty factor. By adjusting C , a balance can be achieved between training error and generalization ability. Introducing the Lagrangian multiplier becomes a convex quadratic optimization problem:

$$L(w, b, \xi, \xi^*, \lambda, \lambda^*, \gamma, \gamma^*) = \frac{1}{2} \|w\|^2 + C \sum_{c=1}^a (\xi_c + \xi_c^*) - \sum_{c=1}^a \lambda_c [\xi_c + \varepsilon - y_c + f(x_c)] - \sum_{c=1}^a \lambda_c^* [\xi_c^* + \varepsilon - y_c + f(x_c)] - \sum_{c=1}^a (\xi_c y_c - \xi_c^* y_c^*), \quad (12)$$

where λ_c and $\lambda^* c$ represent Lagrangian multipliers. In order to speed up the solution, formula (10) is transformed into a dual form:

$$\begin{aligned} \psi(\lambda, \lambda^*) &= -\frac{1}{2} \sum_{c,e=1}^a (\lambda_c - \lambda_c^*)(\lambda_e - \lambda_e^*) [\phi(x_c), \phi(y_e)] + \sum_{c=1}^a (\lambda_c - \lambda_c^*) y_c - \sum_{c=1}^a (\lambda_c - \lambda_c^*), \\ \text{s.t.} \quad &\begin{cases} w = \sum_{c,e=1}^a (\lambda_c - \lambda_c^*) x_c, \\ \sum_{c=1}^a (\lambda_c - \lambda_c^*) = 0, \\ 0 \leq \lambda_c, \lambda_c^* \leq C. \end{cases} \end{aligned} \quad (13)$$

Introduce the function $K(x_c, x)$ to replace the vector inner product $(\phi(x_c), \phi(x))$; then, the classification decision function of SVM is

$$f(x) = \text{sgn} \left(\sum_{c=1}^a (\lambda_c - \lambda_c^*) K(x_c, x) + b \right). \quad (14)$$

3. Results

The data used are 1317 patients who underwent capsule endoscopy in our hospital from September 2016 to September 2020, including 1438 males and 879 females, aged 13–86 years old, with an average age of 45.23 years. Finally, a total of 22785313 capsule endoscopic images were obtained. Excluding redundant images collected in various parts of the

digestive tract due to complex scenes, these images include food residues, bubbles, digestive juices, blood, and other uncertain factors that cause interference, and finally, 22565431 effective images are obtained. The images are divided into normal, inflammation, ulcers, polyps, lymphatic dilatation, hemorrhage, vascular disease, and bulge. A total of 1722499 images of the capsule endoscopy images of 300 patients were retained as a verification set to verify the performance of the model. A total of 5782 lesions were confirmed in 1317 patients. Then, according to the proportion and random selection method, each disease is divided into the training set and test set. Finally, there are 18881193 images in the training set and 3776238 images in the test set. The lesion samples in the image are all marked by an endoscopist with many years of hospital clinical experience. The data labeling work is divided into three steps: the

doctor first determines the patient's disease based on the previously recorded cases (such as polyps and ulcers) and then filters out the disease images from the capsule endoscopy video. Diseased images are provided by professionals. The doctor will label them one by one. Labels are divided into malignant tumors, benign lesions, and normal. Input the 3776238 test set images into the constructed convolutional neural network to verify the accuracy of the neural network. The convolutional neural network model constructed in this study can output malignant tumors, benign lesions, and normal results. Calculate the output results of the convolutional neural network model, and calculate the accuracy (Pr) of the model classification, the recall rate (Re) [22, 23], and the average accuracy (MA). The calculation is as follows:

$$\begin{aligned} \text{Pr} &= \frac{\text{TP}}{\text{TP} + \text{FP}}, \\ \text{Re} &= \frac{\text{TP}}{\text{TP} + \text{FN}}, \\ \text{MA} &= \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FN} + \text{TN} + \text{FP}}. \end{aligned} \quad (15)$$

According to this formula, it can be seen that increasing the number of true positive samples can increase the accuracy rate, while reducing the number of false negative samples can increase the recall rate. The new sample greatly increases the number of true positive samples and reduces the number of false negative samples. Therefore, the use of new samples can improve accuracy and recall.

4. Discussion

According to Table 1, after inputting 3776238 sample images of the test set into the convolutional neural network algorithm model, the accuracy, recall, and average accuracy of the algorithm model are 96.47%, 96.13%, and 97.69%, respectively. The total accuracy of this study reached 97.69%, which shows that the convolutional neural network constructed in this study can effectively help doctors perform endoscopic image recognition. Among them, the highest average accuracy is the normal type and the lowest is the benign lesion. The reason may be that normal endoscopic images have no pathological features, so this algorithm can more accurately identify normal endoscopic images. Among the types of benign lesions, it may be because there are more types of benign lesions, and there are more items to be compared with the algorithm, which leads to inaccurate output results of the algorithm. There is a significant difference in the average reading time of capsule endoscopic images for each patient between manual reading and assisted reading models based on convolutional neural networks. In this study, the RGB color model was converted to the lab color model, and the endoscopic images of these two color models were input into the validation set sample data for analysis and comparison. The result analysis is given in Table 2.

TABLE 1: Neural network output result index table.

Norm	Pr (%)	Re (%)	MA (%)
Malignant tumor	96.76	95.34	97.92
Benign lesions	95.18	94.69	96.34
Normal	97.46	98.37	98.81
Total	96.47	96.13	97.69

TABLE 2: Comparison table of the output results of the two color modes.

Color mode	Pr (%)	Re (%)	MA (%)
RGB	75.24	76.34	73.12
Lab	96.47	96.13	97.69

According to Table 2, it can be concluded that the results obtained by inputting the validation set samples of the RGB color model into the convolutional neural network are significantly lower than using the lab color model, so it can be judged that the laboratory color model can improve the accuracy of computer recognition. Although the convolutional neural network proposed in this study can effectively identify the endoscopic image of the digestive tract, in actual use, due to the particularity of the digestive tract environment, there will be many uncertainties in the working image of the capsule endoscope. For example, food residues, bubbles, digestive juices, and blood increase the difficulty of processing and analyzing redundant images. Therefore, it is necessary to filter out these useless images when processing images. The result of this research is the result of screening out these redundant images, so some errors may occur in practical applications.

5. Conclusion

This research proposes a convolutional neural network model algorithm that can automatically recognize and classify endoscopic images of digestive tract capsules. First, the redundant image is filtered to obtain a valid endoscopic image and the RGB color model of the image is converted to a lab color model, the image features are extracted and the training set image is used to train the convolutional neural network, and finally, input the validation set image into the network model to verify the performance of the model. Based on the above experiments, it can be concluded that the convolutional neural network constructed in this study can effectively distinguish and recognize the images of the digestive tract capsule endoscopy. The time it takes is much shorter than the doctor's diagnosis time and can quickly make an accurate diagnosis, so it can effectively assist the doctor in the diagnosis and treatment process. The convolutional neural network we built can screen and identify thousands of endoscopic images in a short time and then divide the identified endoscopic images into three types: malignant tumors, benign lesions, and normal. The convolutional neural network algorithm proposed in this research requires a large number of annotated image databases for training. Therefore, in the future research process, the

database must be continuously expanded, and a large amount of image data annotated by doctors need to be collected. Since there are a large number of redundant images in endoscopic images, it is necessary to develop an algorithm that can be automatically identified in future research to promote the classification of redundant images and effective images. In future research, we will try to integrate the advantages of other algorithms to make the constructed algorithm model more effective and more accurate to assist physicians in clinical diagnosis, thereby improving the cure rate of patients with gastrointestinal diseases.

Data Availability

The datasets and codes used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

References

- [1] S. Gomes, M. T. Valério, M. Salgado, H. P. Oliveira, and A. Cunha, "Unsupervised neural network for homography estimation in capsule endoscopy frames," *Procedia Computer Science*, vol. 164, pp. 602–609, 2019.
- [2] R. Leenhardt, P. Vasseur, C. Li et al., "A neural network algorithm for detection of GI angiectasia during small-bowel capsule endoscopy," *Gastrointestinal Endoscopy*, vol. 89, no. 1, pp. 189–194, 2019.
- [3] S. Sainju, F. M. Bui, and K. A. Wahid, "Automated bleeding detection in capsule endoscopy videos using statistical features and region growing," *Journal of Medical Systems*, vol. 38, no. 4, p. 25, 2014.
- [4] T. Aoki, A. Yamada, K. Aoyama et al., "Clinical usefulness of a deep learning-based system as the first screening on small-bowel capsule endoscopy reading," *Digestive Endoscopy*, vol. 32, no. 4, pp. 585–591, 2020.
- [5] B. Li and M. Q.-H. Meng, "Computer-aided detection of bleeding regions for capsule endoscopy images," *IEEE Transactions on Biomedical Engineering*, vol. 56, no. 4, pp. 1032–1039, 2009.
- [6] R. Shahril, A. Saito, and A. Shimizu, "Bleeding classification of enhanced wireless capsule endoscopy images using deep convolutional neural network," *Journal of Information Science and Engineering*, vol. 36, no. 1, pp. 91–108, 2020.
- [7] H. Chen, X. Wu, G. Tao, and Q. Peng, "Automatic content understanding with cascaded spatial-temporal deep framework for capsule endoscopy videos," *Neurocomputing*, vol. 229, no. 15, pp. 77–87, 2017.
- [8] B. Yiftach, A. Liran, and S. Shelly, "Ulcer severity grading in video-capsule images of Crohn's disease patients: an ordinal neural network solution," *Gastrointestinal Endoscopy*, vol. 93, no. 1, pp. 187–192, 2020.
- [9] C. A. Burke, J. Santisi, J. Church, and G. Levinthal, "The utility of capsule endoscopy small bowel surveillance in patients with polyposis," *American Journal of Gastroenterology*, vol. 100, no. 7, pp. 1498–1502, 2005.
- [10] B. S. Lewis and P. Swain, "Capsule endoscopy in the evaluation of patients with suspected small intestinal bleeding: results of a pilot study," *Gastrointestinal Endoscopy*, vol. 56, no. 3, pp. 349–353, 2002.
- [11] R. Caspari, M. Von Falkenhäusen, C. Krautmacher, H. Schild, J. Heller, and T. Sauerbruch, "Comparison of capsule endoscopy and magnetic resonance imaging for the detection of polyps of the small intestine in patients with familial adenomatous polyposis or with peutz-jeghers' syndrome," *Endoscopy*, vol. 36, no. 12, pp. 1054–1059, 2004.
- [12] V. P. Beltrán, P. Nos, and G. Bastida, "Evaluation of post-surgical recurrence in Crohn's disease: a new indication for capsule endoscopy," *Gastrointestinal Endoscopy*, vol. 66, no. 3, pp. 533–540, 2007.
- [13] L. He, X. Gao, W. Lu, X. Li, and D. Tao, "Image quality assessment based on S-CIELAB model," *Signal, Image and Video Processing*, vol. 5, no. 3, pp. 283–290, 2011.
- [14] K. Yao, M. Kato, and J. Fujisaki, "Techniques using the hemoglobin index of the gastric mucosa," *Endoscopy*, vol. 37, no. 5, pp. 479–486, 2005.
- [15] D. Dey, D. G. Gobbi, and K. Surry, "Mapping of endoscopic images to object surfaces via ray-traced texture mapping for image guidance in neurosurgery," *Proceedings of SPIE-The International Society for Optical Engineering*, vol. 3976, pp. 290–300, 2000.
- [16] L. P. Hariri, M. S. Alexandre, and D. Dmv, "Endoscopic optical coherence tomography and laser-induced fluorescence spectroscopy in a murine colon cancer model," *Lasers in Surgery and Medicine*, vol. 38, no. 4, pp. 305–313, 2010.
- [17] M. M. Panja and B. N. Mandal, "A note on one-point quadrature formula for Daubechies scale function with partial support," *Applied Mathematics and Computation*, vol. 218, no. 8, pp. 4147–4151, 2011.
- [18] C. J. C. Burges, "A tutorial on support vector machines for pattern recognition," *Data Mining and Knowledge Discovery*, vol. 2, no. 2, pp. 121–167, 1998.
- [19] N. Cristianini and J. Shawe Taylor, "An introduction to support vector machines and other kernel-based learning methods," *Computer Science, Pattern Recognition and Machine Learning, Computational Biology and Bioinformatics*, vol. 32, no. 1, pp. 1–28, 2001.
- [20] C. W. Chih-Wei Hsu and C. J. Chih-Jen Lin, "A comparison of methods for multiclass support vector machines," *IEEE Transactions on Neural Networks*, vol. 13, no. 2, pp. 415–425, 2002.
- [21] I. Guyon, J. Weston, and S. Barnhill, "Gene selection for cancer classification using support vector machines," *Machine Learning*, vol. 46, no. 1–3, pp. 389–422, 2002.
- [22] J. Grafman, A. Pascual-Leone, D. Alway, P. Nichelli, E. Gomez-Tortosa, and M. Hallett, "Induction of a recall deficit by rapid-rate transcranial magnetic stimulation," *NeuroReport*, vol. 5, no. 9, pp. 1157–1160, 1994.
- [23] H. L. Roediger, C. C. Stellon, and E. Tulving, "Inhibition from part-list cues and rate of recall," *Journal of Experimental Psychology: Human Learning & Memory*, vol. 3, no. 2, pp. 174–188, 1977.