Hindawi

Research Article

The Application of Convolutional Neural Network Combined with Fuzzy Algorithm in Colorectal Endoscopy for Tumor Assessment

Te-Jen Su,[1,2] Feng-Chun Lee,[1,3] Cheuk-Kwan Sun,[4,5] Fu-Xiang Ke,[6] Shih-Ming Wang [ORCID],[7] and Ming-Chih Huang[8]

[1]*Department of Electronics, National Kaohsiung University of Science and Technology, Kaohsiung 807, Taiwan*
[2]*Graduate Institute of Clinical Medicine, Kaohsiung Medical University School of Dentistry, Kaohsiung 807, Taiwan*
[3]*Kaohsiung Change Gung Memorial Hospital, Kaohsiung 833, Taiwan*
[4]*Department of Emergency Medicine, E-Da Hospital, Kaohsiung, Taiwan*
[5]*School of Medicine for International Students, College of Medicine, I-Shou University, Kaohsiung, Taiwan*
[6]*Department of Electronics, National Kaohsiung University of Science and Technology, Kaohsiung 807, Taiwan*
[7]*Department of Computer Science and Information Engineering, Cheng Shiu University, Kaohsiung 833, Taiwan*
[8]*Department of Electronics, National Kaohsiung University of Science and Technology, Kaohsiung 807, Taiwan*

Correspondence should be addressed to Shih-Ming Wang; k1115@gcloud.csu.edu.tw

According to the Global Cancer Statistics 2020 published in the official journal of the American Cancer Society (ACS), colorectal cancer ranked 4th in incidence and 2nd in mortality, and the 2018 Cancer Registry Report of Taiwan Health Promotion Administration showed that colorectal cancer ranked 2nd in incidence and 3rd in mortality. With the rapid evolution of the times, the lifestyles of the people have shifted from what they used to be. In addition to uncontrollable factors such as family genetic disorders, diet, and bad habits, life stress may lead to an unhealthy body mass index (BMI), which, together with aging, increases the incidence of colorectal cancer. In this study, the convolutional neural network was used to assess the risk of tumor in the colon by colonoscopy. The endoscopic images of the colon, which were classified into three categories of healthy (normal), benign tumor, and malignant tumor, were adopted as training data. When this method is combined with the patient's physical data, the risk cancer can be calculated by the fuzzy algorithm. Based on the result of this study, the accuracy of the tumor profile by colonoscopy, that is, 81.6%, is more precise than that of colorectal cancer tumor analysis studies in the recent literature. The proposed method will help physicians in the diagnosis of colorectal cancer and treatment decisions.

1. Introduction

According to Global Cancer Statistics 2021 published in the official journal of the American Cancer Society (ACS), colorectal cancer is the 4th most common cancer and the 5th most common mortality, posing a serious threat to the health of the population [1, 2]. Another 2018 Cancer Registry Report by Health Promotion Administration reveals that colorectal cancer ranks 2nd in morbidity, 2nd in mortality, 1st in morbidity for men and 3rd for women, and 3rd in mortality for men and 4th for women [3].

In recent years, the incidence of colorectal cancer is on the rise year by year as a result of the changing living pattern, food culture, sedentary lifestyle, work environment, and other factors, even with the trend of increasing youthfulness. With the exclusion of family genetic history, the risk factors for colorectal cancer are associated with the poor habits of people's lives, irregular work habits, physical inactivity, work strain, lack of dietary control, and aging. It is also noted that in terms of the risk of colorectal cancer, ranging from age and genetic to environmental and lifestyle choices, factors such as obesity, low physical activity, active and passive

smoking, and high salt and red meat consumption are correlated to a higher risk of colorectal cancer [4–6]. Regarding the impact of age, a recent study demonstrated a steady yearly increase in the risk of young-onset colorectal cancer [7].

For early detection of colorectal cancer, carbohydrate antigen 19-9 (CA 19-9) and carcinoembryonic antigen (CEA) are commonly used biomarkers [8]. There is a strong correlation between CA 19-9 and CEA levels in colorectal cancer patients ($p = 0.001$ and $p < 0.0001$, respectively), both of which are important biomarkers in the progression of colorectal cancer [9]. Persistent smoking is known to alter the prognostic value of postoperative serum CEA levels in colorectal cancer patients because smoking can increase serum CEA levels independent of the disease status [10].

Another recent breakthrough in the diagnosis of colorectal cancer is deep neural network visualization [11–14]. An image analysis method based on deep learning can not only accurately classify different types of polyps in the whole slide image but also generate the main areas and features on slides through the model visualization method. This visualization method could significantly reduce the cognitive burden of clinicians [11]. In recent years, the convolutional neural network (CNN) model has been applied in the relevant medical literature. The accuracy of most model validation from previous studies [12–14] related to colorectal cancer falls between 75.1% and 83.9%. Based on the result of this study, the accuracy of the tumor profile by colonoscopy, that is, 81.6%, is more precise than those, 75.7% and 75.1% respectively, of colorectal cancer tumor analysis studies in references [12, 13]. It is slightly lower than that, 83.9%, in reference [14]. However, the source of image acquisition and research method are different. The cost is relatively high and is seldom employed. The proximity and complexity of the organs in the human body, the image resolution, size, and angle can affect the accuracy of identifying the targeted tissues and lesions using the training model.

The biggest difference between CNN and multilayer perceptron (MLP) lies in the additional convolution layer and pooling layer. These two layers enable CNN to have the capability in extracting details from image or speech features, instead of simply extracting data for calculation like other neural networks [15, 16].

The fuzzy theory was introduced by Lotfi [17]. Then, fuzzy logic of the concept of linguistic variables is proposed in reference [18]. Nowadays, fuzzy systems are applied in different fields, such as household appliances, industrial system control, and image recognition [19]. The research also pointed out that the application of professional fuzzy rules could help in detecting colorectal cancer and help doctors to easily identify diseases [20].

In this study, the convolutional neural network (CNN) was used for the training and learning of feature extraction from colonoscopy images. According to the health level, the training data were classified into healthiness, benign tumors, and malignant tumors. Colonoscopy images in the three categories were randomly selected as test data, and designated case images were used to assess the similarity of tumor profiles between the designated image and the image from test data.

The practical results of this study are summarized in the following four points:

(1) By analyzing the polyp profile in colonoscopy images, the results can be used as a reference for physicians to diagnose the symptoms as well as increase the detection efficiency and reduce the misdiagnosis rate.

(2) The patient's physical data are combined with the risk of tumor for assessment in the fuzzy system, which not only allows the patient to understand his current physical condition through data analysis but also enables the physician to make corresponding treatment decisions for the patient through the assessment results.

(3) After discussion on the results with clinicians, the accuracy of the raw data and assessment results are consistent with clinical analysis.

(4) The accuracy of the results of this study reaches 81.6%. Compared with the accuracy of colorectal cancer tumor analysis and research in the literature, the accuracy is better [12, 13], and the accuracy of the original data and evaluation results after discussion with clinicians is in line with clinical analysis.

2. Methodology

This study is based on the case data of colorectal cancer in a medical center in southern Taiwan. The colonoscopy images are trained and learned by convolutional neural networks. After completing the learning verification, the colonoscopy images of the designated patients will be tested and identified. Finally, the severity score and the patient-related information are fuzzy analyzed through a fuzzy algorithm, and the output is the risk of the patient's colorectal cancer, so that the doctor can diagnose colorectal cancer-related diseases Time aids.

3. Methods

Between January 2016 and December 2020, the medical records of the first 500 adults (i.e., age >18 years) of both genders undergoing first-time colonoscopy at a single referral center (i.e., Kaohsiung Chang Gung Memorial Hospital) regardless of indications were retrospectively reviewed. Exclusion criteria were as follows: (1) patients receiving previous colonoscopic examinations at other medical institutes, (2) those with normal colonoscopic findings, (3) those with a known history of benign or malignant colorectal diseases including familial polyposis and inflammatory bowel disease (i.e., Crohn's disease and ulcerative colitis), (4) those having received colorectal procedures (e.g., polypectomy and colorectal resection), (5) those without pathological analysis of colorectal specimens, and (6) those without complete information for the present study (e.g., body mass index and circulating CEA levels). Circulating CEA levels were determined in participants of annual physical checkups and those with positive stool occult blood test scheduled for colonoscopy.
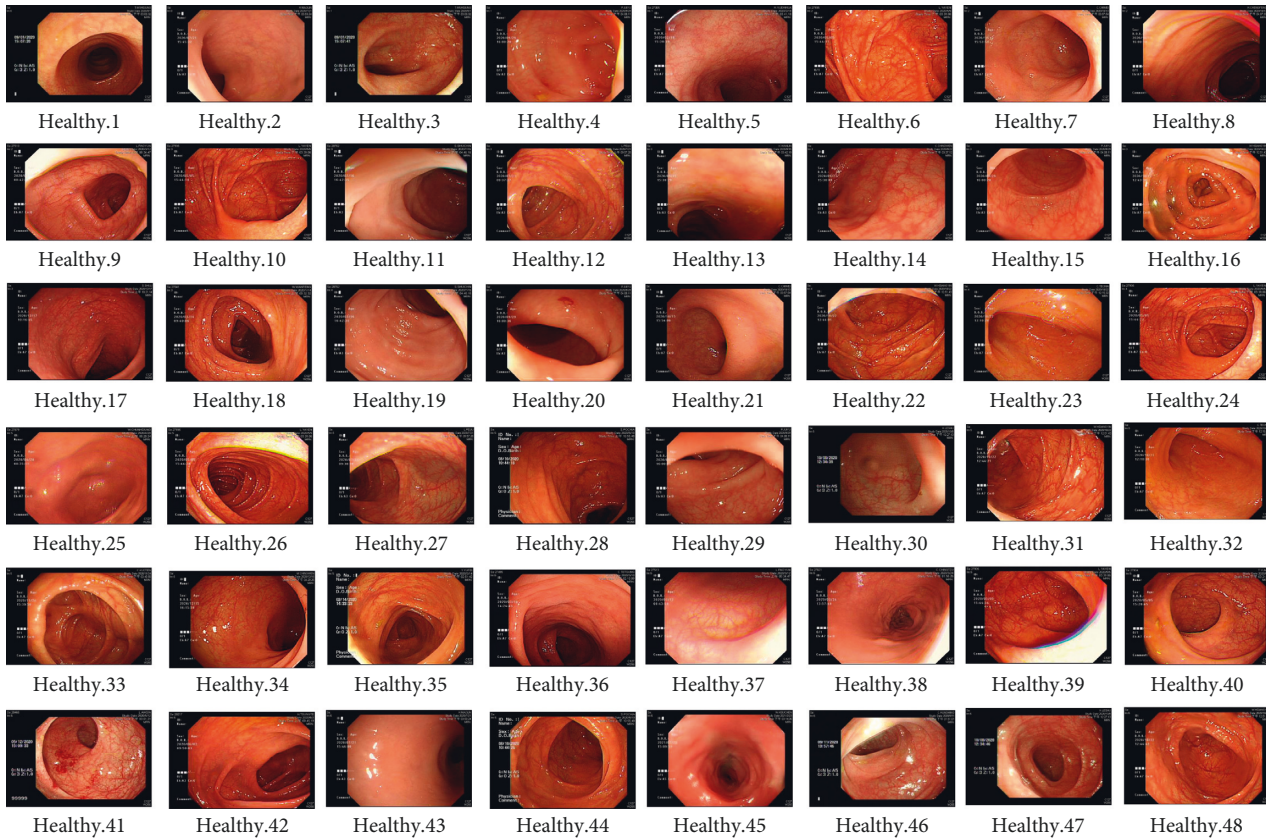
Healthy.1	Healthy.2	Healthy.3	Healthy.4	Healthy.5	Healthy.6	Healthy.7	Healthy.8
Healthy.9	Healthy.10	Healthy.11	Healthy.12	Healthy.13	Healthy.14	Healthy.15	Healthy.16
Healthy.17	Healthy.18	Healthy.19	Healthy.20	Healthy.21	Healthy.22	Healthy.23	Healthy.24
Healthy.25	Healthy.26	Healthy.27	Healthy.28	Healthy.29	Healthy.30	Healthy.31	Healthy.32
Healthy.33	Healthy.34	Healthy.35	Healthy.36	Healthy.37	Healthy.38	Healthy.39	Healthy.40
Healthy.41	Healthy.42	Healthy.43	Healthy.44	Healthy.45	Healthy.46	Healthy.47	Healthy.48

FIGURE 1: Healthy (normal).

4. Results

Of the 992 adult patients receiving colonoscopic examination within the study period, the medical records of the first 500 eligible for the current study were reviewed. The patient population comprised 275 males (55%) and 225 females (45%) with a mean age of 62.1 ± 11.8 (range, 31–85), a mean body mass index of 23.5 ± 3.7 (range, 17.5–31.6), and a mean circulating level of CEA 23.5 ± 280.6 (range, 0.5–1212.0).

In our routine practice, we take 12 images from a patient during a colonoscopic examination. Therefore, we had a total of 5950 images from 500 patients. All images were fed into an imaging analyzing software (Spyder 4.2.0) that divided the images into three categories, namely, normal image (Figure 1), benign (Figure 2), and malignant (Figure 3) tumors. There was no human handling or annotation of the images in the analytic process.

4.1. Research Environment. The colonoscopy images of 500 cases were divided into a training set, validation set, and test set. Among them, the colonoscopy images of 10 cases in the total data are taken as the test set, the rest is taken as the training set, and the images from the training set of 10% are used as the validation set of the training model for cross validation. Table 1 lists the environment and hardware configuration of this research.

4.2. System Software Design and Composition. Two kinds of software are used in the development of the system, one is Spyder, which is based on the development environment under the Python language. The open-source cross-platform scientific computing integrated development environment (IDE) of the Python language provides advanced code editing, interactive testing and debugging, computational science, data processing, and predictive analysis and supports multiple programming languages and operating systems. The second is Matlab, which is an interactive development environment based on algorithm development, data analysis, and numerical calculation.

4.3. Experimental Steps

(1) Obtain the image data of the colonoscopy through a medical center in the south, and search for the body-related data of the case based on the image data.

(2) Keep the required data, delete the unnecessary data such as blurred images, overexposed images, and unrecognizable shooting angles during the colonoscopy, and classify them into healthy (normal), benign, and malignant according to the type and appearance characteristics of polyps.

(3) From the three types of image data, after the training set and the test set are separated, the convolutional neural network is used to learn and train them, and the parameter values are adjusted to make the verification accuracy and loss value reach the expected value. Set goals.
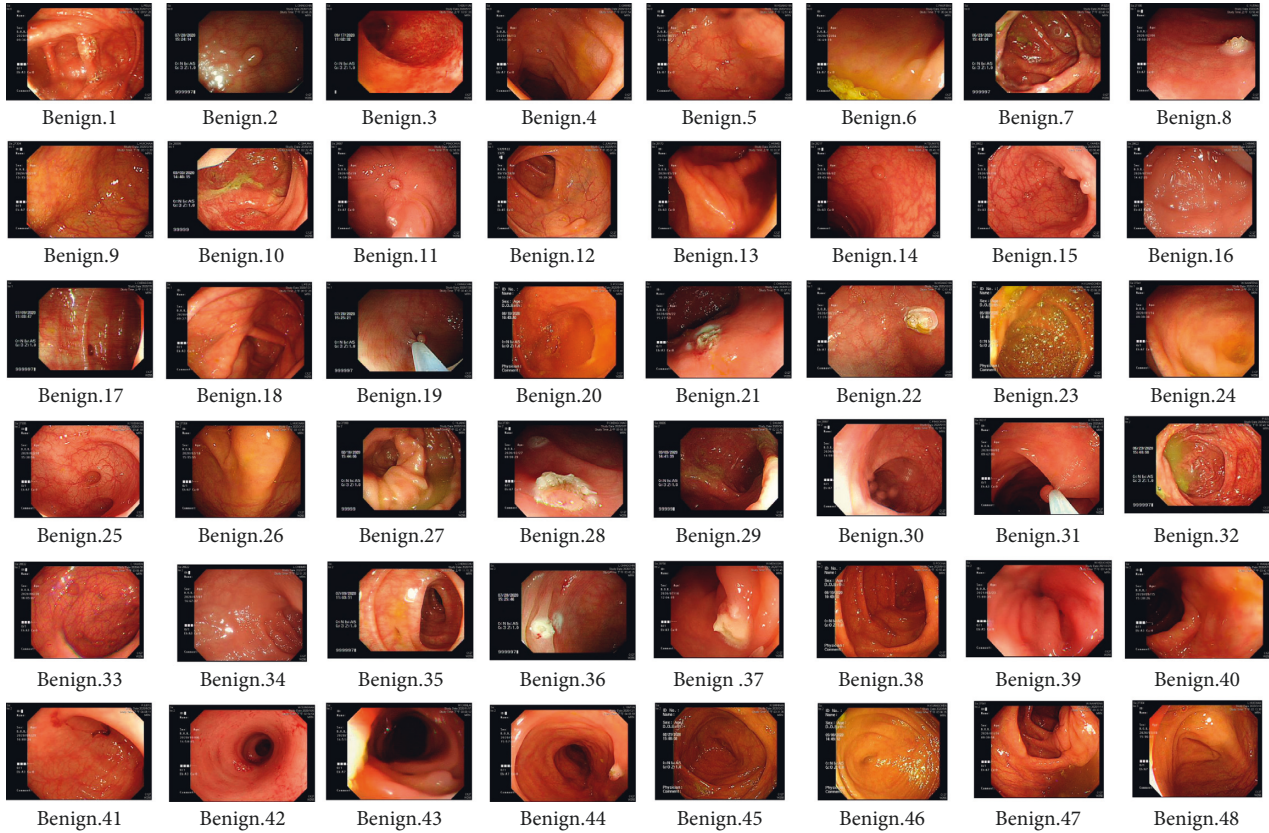
Figure 2: Benign tumor.

(4) After completing the learning and training, perform the result verification to check the corresponding similarity evaluation of healthy (normal), benign, and malignant in the test images. This value is also used as the severity percentage in the fuzzy input, which is called the tumor risk.

(5) Finally, input the age, BMI value, tumor risk, and carcinoembryonic antigen index in the body data of the corresponding case into the fuzzy algorithm to evaluate the risk of cancer, and display the result.

4.4. Data Preprocessing.

In the process of colonoscopy, the large intestine will be affected by factors such as the width of the intestine, bending of the intestine, the number of folds in the intestinal wall, the position and size of polyps, the shooting angle of the lens and whether it is accurately focused, and the overexposure or insufficient light source. The output image quality of the mirror inspection, coupled with the limited time of the inspection process, makes it inevitable that there will be poor quality and difficulty to identify image data in the screening results. Therefore, after filtering them, the difficult-to-identify or poor-quality images are deleted to improve the accuracy of the training model.

Since the human intestine is very long and the affected part only exists in a certain part of the general intestine, the results of colonoscopy screening may include normal images (labeled as "healthy") as well as those of polyps and malignant tumors. The colonoscopy images of these 500 patients are classified into three categories: "healthy" (Figure 1), "benign (Figure 2)," and "malignant" (Figure 3). This classification is only based on the appearance of polyps as a preliminary assessment, and the final judgment of the tumor profile must be approved by a professional physician. Screening and diagnosis are performed.

4.5. Convolutional Neural Network Model Architecture and Parameter Settings.

The neural network model used in this research is SmallerVGGNet, which is the simplified CNN model architecture of VGGNet [21], and the colonoscopy data of 500 cases were classified into healthy (normal), benign, and malignant categories. In order to avoid overfitting of CNN during training, the database was divided into a training set and a test set without duplication, and 5%–10% of the images in the training set were taken as the validation set, which was repeatable. The purpose was to observe the validation accuracy of the model after training and select the training model with the highest validation accuracy as the CNN model in this study for tumor risk assessment.

The CNN architecture of this study is based on the SmallerVGGNet neural network as a multiconvolutional deep learning classifier, which consists of 7 convolutional layers and 4 pooling layers with MaxPooling added after convolutional layers 1, 3, 5, and 7, respectively. The remaining model parameters are presented in Table 2, with the activation function being ReLU in the CNN training model, sigmoid in the multitag classifier, adam in the
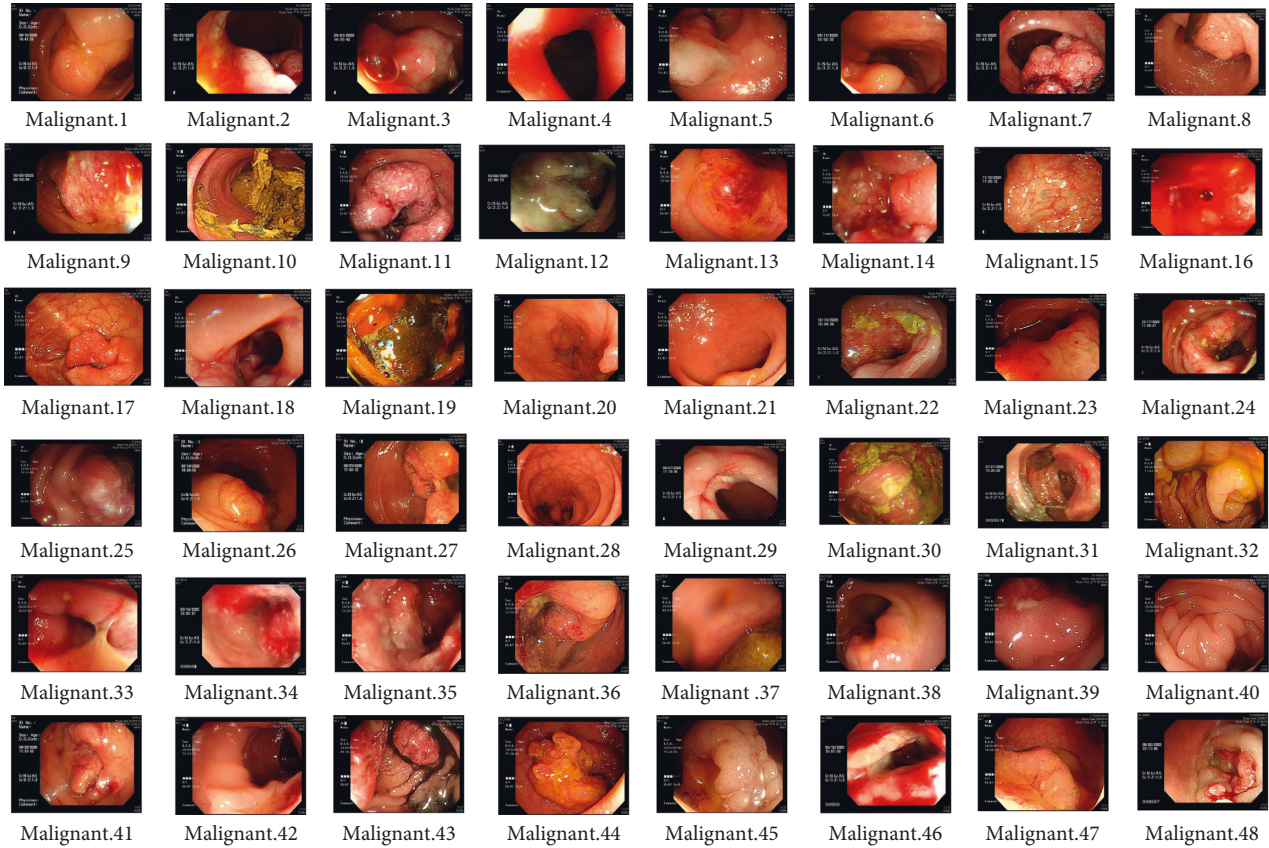
Malignant.1	Malignant.2	Malignant.3	Malignant.4	Malignant.5	Malignant.6	Malignant.7	Malignant.8
Malignant.9	Malignant.10	Malignant.11	Malignant.12	Malignant.13	Malignant.14	Malignant.15	Malignant.16
Malignant.17	Malignant.18	Malignant.19	Malignant.20	Malignant.21	Malignant.22	Malignant.23	Malignant.24
Malignant.25	Malignant.26	Malignant.27	Malignant.28	Malignant.29	Malignant.30	Malignant.31	Malignant.32
Malignant.33	Malignant.34	Malignant.35	Malignant.36	Malignant .37	Malignant.38	Malignant.39	Malignant.40
Malignant.41	Malignant.42	Malignant.43	Malignant.44	Malignant.45	Malignant.46	Malignant.47	Malignant.48

FIGURE 3: Malignant tumor.

TABLE 1: Experimental setup.

Test system	
Operating system	Windows 10 Professional (x64)
CPU	Intel Core i5-4570 (4C4T@3.2 GHz)
GPU	NVIDIA GeForce GT-710 2 GB
Memory	16 GB DDR3 1333 MHz
Development environment	Tensorflow-Keras (Spyder 4.2.0) Matlab R2020a
Program language	Python 3.6
Training set/test set	5950 piece/225 piece

TABLE 2: CNN model parameter settings.

Various parameters	
Layers	7 layers of convolutional layer and 4 layers of pooling layer
Convolutional kernel	3×3
MaxPooling	2×2
Activation function	Multilabel classification: sigmoid neural network training model: ReLU
Optimizer	Adam
Stride	1
Dropout	25%
Learning rate	0.001
Batch size	32
Epoch	200
Number of hidden layer neurons	1024

optimizer, stride of 1, dropout of 25%, the initial learning rate of 1e-3, batch size of 32, training iteration epoch of 200, and hidden layer neurons of 1024.

4.6. Fuzzy System Design.

The tumor risk estimated by the CNN training model can be combined with the patient's body-related data to derive the risk of colorectal cancer. Therefore, a fuzzy system was designed by establishing semantic variables of input and output, defining their membership functions, and formulating fuzzy rules, fuzzy inference, and defuzzification. As such, a fuzzy system for the risk of cancer was determined.

The tumor risk derived from the CNN model was combined with four input variables, including the corresponding age of the patients, BMI, and carcinoembryonic antigen. After fuzzification of the triangular and trapezoidal membership functions, the maximum-minimum (max-min) synthesis operator was used for the computation of the membership of the fuzzy set by the center-of-gravity method and the output was the risk of colorectal cancer.

4.6.1. Design of Fuzzy Parameters.

The fuzzy algorithm was applied to assess the risk of colorectal cancer, as shown in Figure 4, The tumor risk derived from the CNN model was combined with four input variables, including the corresponding age of patients, BMI, tumor risk, and
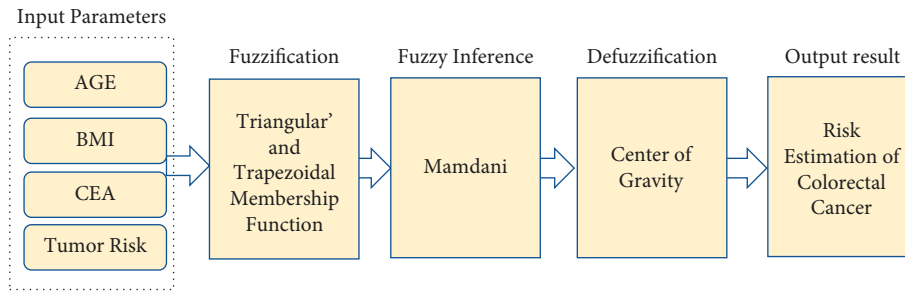
Input Parameters
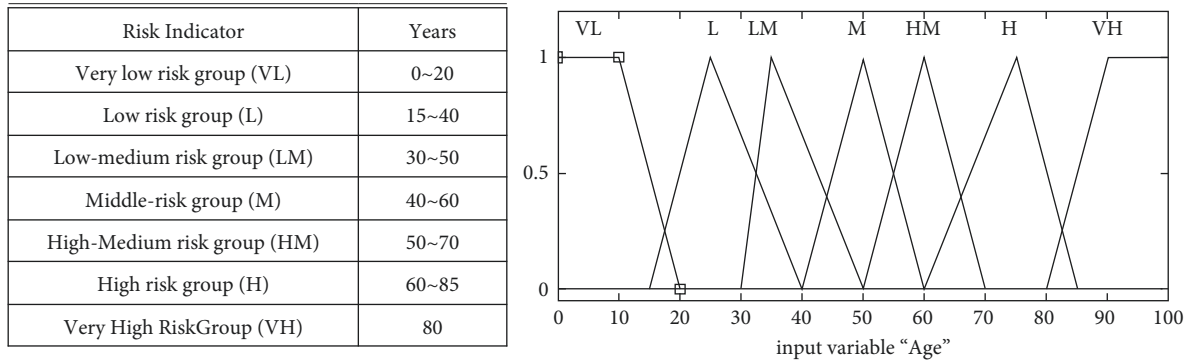

Figure 4: Schematic diagram of fuzzy evaluation.

Risk Indicator	Years
Very low risk group (VL)	0~20
Low risk group (L)	15~40
Low-medium risk group (LM)	30~50
Middle-risk group (M)	40~60
High-Medium risk group (HM)	50~70
High risk group (H)	60~85
Very High RiskGroup (VH)	80

Figure 5: Input attribution function design-age risk index.

Health Indicators	BMI
Underweight	0~18.5
Normal	16~25
Overweight	22~29
1 Degree Fat	25~32
2 Degree Fat	29~40
3 Degree Fat	35
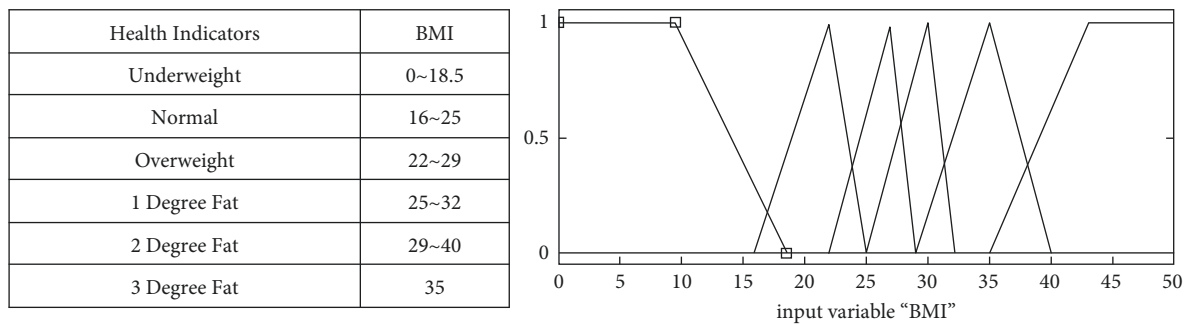
Figure 6: Input attribution function body mass index, BMI.

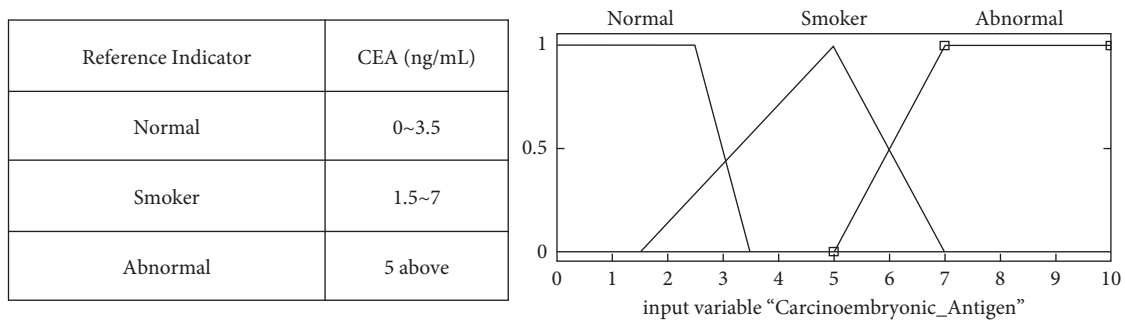| Reference Indicator	CEA (ng/mL)
Normal	0~3.5
Smoker	1.5~7
Abnormal	5 above

Figure 7: Input attribution function carcinoembryonic antigen, CEA.

carcinoembryonic antigen. After fuzzification of the triangular and trapezoidal membership functions, the maximum-minimum (max-min) synthesis operator was used for the computation of the membership of the fuzzy set by the center-of-gravity method and the output was the risk of colorectal cancer.

Severity Rating	Risk of Tumor (%)
Healthy(Normal)	0~40
Benign	25~75
Malignant	55~100
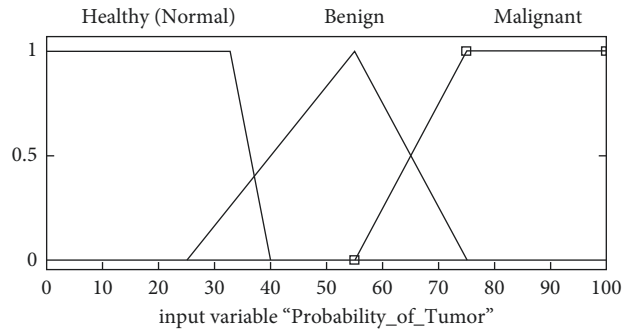
Figure 8: Input attribution function risk of tumor.

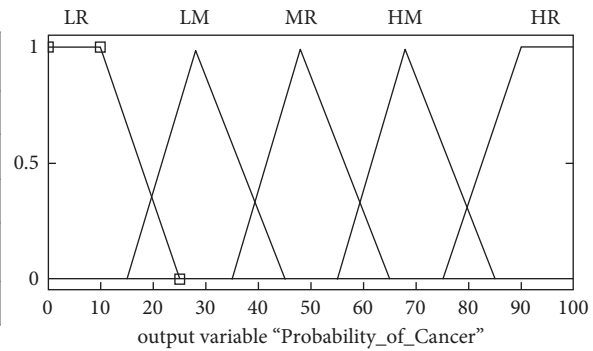| Risk Iindicator	Risk of Tumor (%)
Low-risk group(LR)	$0 < \times \leq 20$
Low-to-medium risk(LM)	$20 < \times \leq 40$
Medium risk group(MR)	$40 < \times \leq 60$
High-medium risk group(HM)	$60 < \times \leq 80$
High-risk group(HR)	$\times > 80$

Figure 9: Risk of cancer.

Table 3: Risk of tumor. (a) Healthy: age/BMI; (b) age/CEA; (c) CEA/BMI.

(a) Healthy: age/BMI						
BMI age (years)	0–20	15–40	30–50	40–60	50–70	60–85	Over 80
Underweight	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Normal weight	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Overweight	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Severe obesity	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Morbid obesity	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Super obesity	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
(b) Healthy: age/CEA						
CEA age (years)	0–20	15–40	30–50	40–60	50–70	60–85	Over 80
Normal	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Smoker	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
Abnormal	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk	Low risk
(c) Healthy: CEA/BMI		
BMI age	Normal	Smoker	Abnormal
Underweight	Low risk	Low risk	Low risk
Normal weight	Low risk	Low risk	Low risk
Overweight	Low risk	Low risk	Low risk
Severe obesity	Low risk	Low risk	Low risk
Morbid obesity	Low risk	Low risk	Low risk
Super obesity	Low risk	Low risk	Low risk

4.6.2. Establishment of Semantic Variables and Membership Functions. The fuzzy system has four input parameters and one output. Input parameters are age in Figure 5, BMI in Figure 6, carcinoembryonic antigen in Figure 7, and tumor risk in Figure 8, respectively, and the output is risk of colorectal cancer in Figure 9 The terms and the membership functions of each parameter are explained below.

Once the terms were established, the membership functions were defined based on the data from the pieces of literature. The graphs of membership function of this study were based on the triangular (Trimf) and trapezoidal (Trapmf) membership functions referenced to obtain better results and to facilitate the observation of the data in this study, while the range of age membership was determined

TABLE 4: Risk of tumor. (a) Benign: age/BMI; (b) age/CEA; (c) CEA/BMI.

(a) Benign: age/BMI

BMI age (years)	0–20	15–40	30–50	40–60	50–70	60–85	Over 80
Underweight	Low risk	Low risk	Low risk	Low-to-moderate risk	Moderate risk	Moderate risk	Moderate risk
Normal weight	Low risk	Low risk	Low risk	Low-to-moderate risk	Moderate risk	Moderate risk	Moderate risk
Overweight	Low risk	Low risk	Low-to-moderate risk	Moderate risk	Moderate risk	Moderate-to-high risk	Moderate-to-high risk
Severe obesity	Low risk	Low risk	Low to moderate risk	Moderate risk	Moderate-to-high risk	Moderate-to-high risk	High risk
Morbid obesity	Low risk	Low-to-moderate risk	Low-to-moderate risk	Moderate risk	Moderate-to-high risk	High risk	High risk
Super obesity	Low risk	Low-to-moderate risk	Low-to-moderate risk	Moderate-to-high risk	High risk	High risk	High risk

(b) Benign: age/CEA

CEA age (years)	0–20	15–40	30–50	40–60	50–70	60–85	Over 80
Normal	Low risk	Low risk	Low risk	Low-to-moderate risk	Moderate risk	Moderate risk	Moderate risk
Smoker	Low risk	Low risk	Low-to-moderate risk	Moderate risk	Moderate risk	Moderate-to-high risk	High risk
Abnormal	Low risk	Low-to-moderate risk	Low-to-moderate risk	Moderate risk	Moderate-to-high risk	High risk	High risk

(c) Benign: CEA/BMI

BMI age	Normal	Smoker	Abnormal
Underweight	Low risk	Low risk	Low risk
Normal weight	Low risk	Low risk	Low-to-moderate risk
Overweight	Low risk	Low-to-moderate risk	Moderate risk
Severe obesity	Low-to-moderate risk	Moderate risk	Moderate-to-high risk
Morbid obesity	Moderate risk	Moderate-to-high risk	High risk
Super obesity	Moderate-to-high risk	High risk	High risk

TABLE 5: Risk of tumor. (a) Malignant: age/BMI; (b) age/CEA; (c) CEA/BMI.

(a) Malignant: age/BMI

BMI age (years)	0–20	15–40	30–50	40–60	50–70	60–85	Over 80
Underweight	High risk	High risk	High risk	High risk	High risk	High risk	High risk
Normal weight	High risk	High risk	High risk	High risk	High risk	High risk	High risk
Overweight	High risk	High risk	High risk	High risk	High risk	High risk	High risk
Severe obesity	High risk	High risk	High risk	High risk	High risk	High risk	High risk
Morbid obesity	High risk	High risk	High risk	High risk	High risk	High risk	High risk
Super obesity	High risk	High risk	High risk	High risk	High risk	High risk	High risk

(b) Malignant: age/CEA

CEA age (years)	0–20	15–40	30–50	40–60	50–70	60–85	Over 80
Normal	High risk	High risk	High risk	High risk	High risk	High risk	High risk
Smoker	High risk	High risk	High risk	High risk	High risk	High risk	High risk
Abnormal	High risk	High risk	High risk	High risk	High risk	High risk	High risk

(c) Malignant: CEA/BMI

BMI age	Normal	Smoker	Abnormal
Underweight	High risk	High risk	High risk
Normal weight	High risk	High risk	High risk
Overweight	High risk	High risk	High risk
Severe obesity	High risk	High risk	High risk
Morbid obesity	High risk	High risk	High risk
Super obesity	High risk	High risk	High risk

from the statistics of reference [11], and BMI and CEA obtained from the information from the Ministry of Health and Welfare and major hospitals.

4.6.3. Establish Fuzzy Rule Base. After evaluating the polyp profile of colorectal endoscopy based on the CNN network model, combining the relevant risk factor parameters and the clinical experience of professional physicians, the corresponding results can be summarized, which is also used as a reference for the design of the fuzzy rule library. In the rule table of the fuzzy system in this study, there are 7 semantic variables in "age," 6 semantic variables in "BMI value," 3 semantic variables in "carcinoembryonic antigen index," and 3 in "tumor risk." There are semantic variables, so there are a total of 378 rules. Tables 3–5 list the comparison of the fuzzy rule base of healthy, benign, and malignant.

4.6.4. Fuzzy Inference and Defuzzification. In the fuzzy system of this study, the method of the center of gravity was utilized for the computation of the defuzzification. With the center of gravity method, the tumor risk and the three risk factors can be derived to assess the risk of cancer. Finally, the probabilistic assessment of the risk level allows the physician to know the current physical data of the patient to assist the physician in the diagnosis and treatment process, thus increasing the efficiency of diagnosis and reducing the rate of misdiagnosis. In the case of a patient aged 75, with a BMI of 25.5, a CEA of 10.55, and a tumor risk of 60.5%, the risk of cancer is calculated to be 75.3%, which corresponds to a risk assessment of the "moderate-to-high risk group."

5. Results

5.1. The Proposed CNN Training Model. Following the consummation of design of the neural network architecture,

the accuracy and loss function of the model were observed by varying the number of iterations and the ratio of images in the training and validation sets. The accuracy ranged from 0.6 to 0.65 for 50 and 75 iterations, indicating a poor training effect. Figure 10 shows the training results of the model with 100, 150, 175, and 200 iterations, and it can be therefore observed that when the ratio of the training set to the validation set was 9:1 for 200 iterations, the accuracy rate reached 81.6%, which was the model with the optimal training effect after multiple adjustments. Thus, it was chosen as the CNN training model for this study.

5.2. Analysis of the Risk of Tumor Detection by Colonoscopy. After the CNN training model was selected, the image data from the test set were classified and identified by a multi-convolutional classification Keras model. The results of the colonoscopy images in Figure 11 illustrate the percentage of the images in each of the three categories of healthy, benign, and malignant after assessment by the classification model, and the assessed probabilities were benign, 87.88%; malignant, 24.19%; healthy, 0.02%. Since the three categories of healthy, benign, and malignant were analyzed separately in the assessment process, the results were not 100% for the three categories combined; instead, the percentages of the three categories were assessed separately for each image. From the above assessment results, the risk of the polyp profile being a benign tumor was 87.88%.

5.3. Assessment of the Risk of Colorectal Cancer. There were four input parameters of the fuzzy system in this study, among which tumor risk was estimated by the CNN model, and the remaining age, BMI, and CEA indexes were the physical data of the patient corresponding to the colonoscopy images. The risk of colorectal cancer was measured by the fuzzy system, as shown in Figure 12 the result of
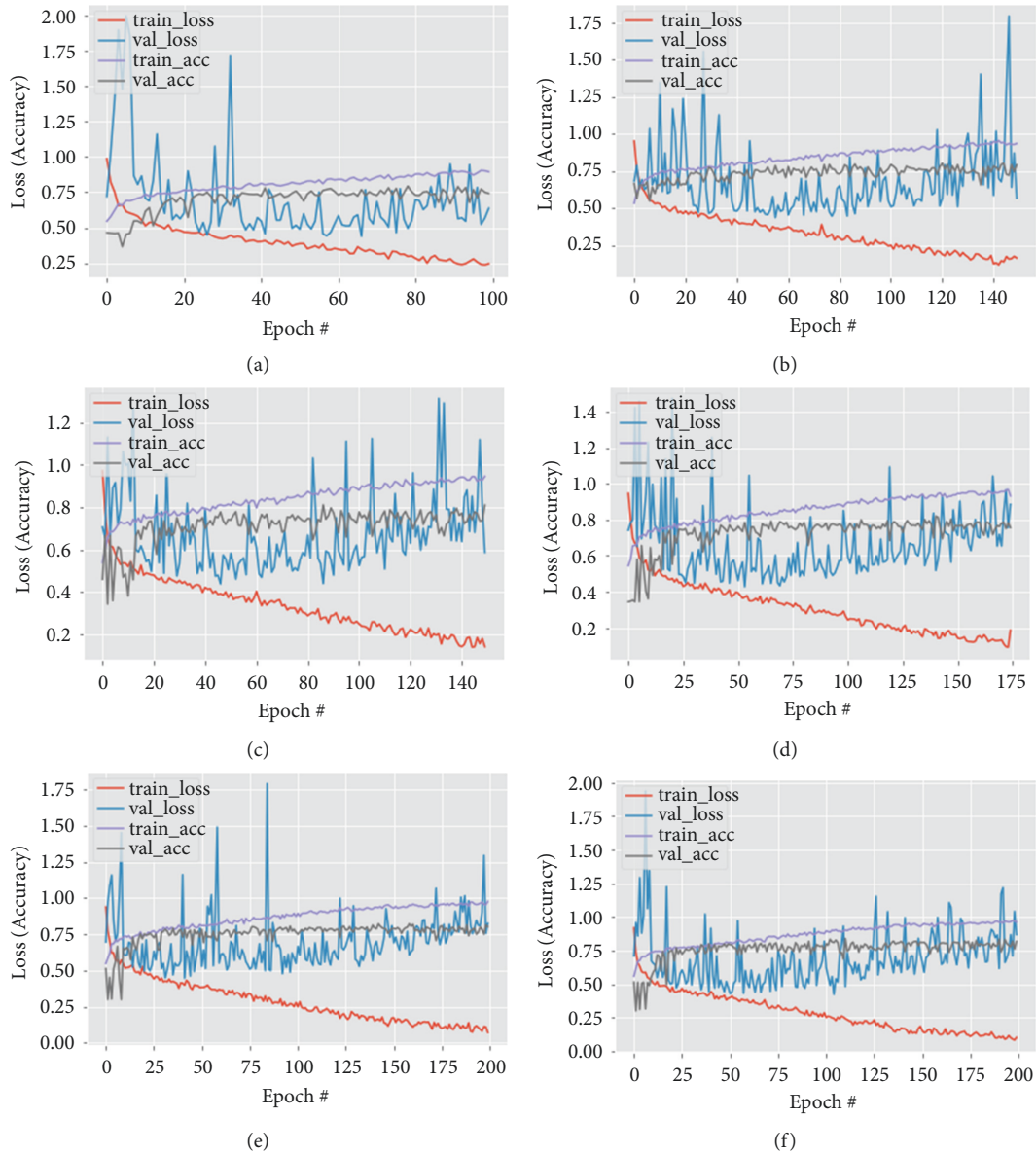
(a)

(b)

(c)

(d)

(e)

(f)

FIGURE 10: Epochs and validation accuracy difference comparison. (a) Epochs: 100, validation accuracy: 0.7424. (b) Epochs: 150, validation accuracy: 0.7918. (c) Epochs: 150, validation accuracy: 0.8108. (d) Epochs: 175, validation accuracy: 0.7555. (e) Epochs: 200, validation accuracy: 0.8027. (f) Epochs: 200, validation accuracy: 0.8161.
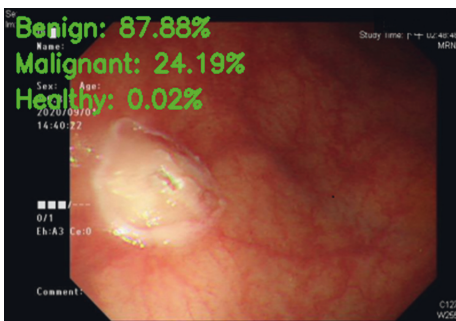

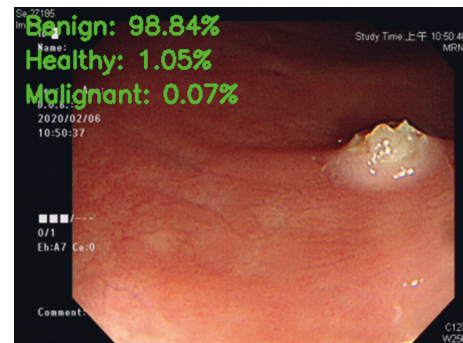FIGURE 11: Tumor risk analysis result: the risk of benign tumor is 87.88%.

FIGURE 12: Tumor risk assessment result: the risk of benign tumor is 98.84%.

colonoscopy image assessment was as follows: the risk of benign tumor is 98.84% and is 65.23% after the conversion of severity level, while the age of the patient was 44 years old, with a BMI of 31.6 and a CEA index of 2.01, and the risk of colorectal cancer was 60.6% after the assessment by the fuzzy system.

6. Discussion

The results of this study are intended for diagnostic purposes, not as a substitute for a physician's decision, thus there is no definite accuracy rate. Apart from discussing the results of the analysis of different tumor probabilities and physical conditions, the cases are also observed to see whether the accuracy rates are clinically relevant in various situations.

Case 1. Age = 56, BMI = 20.1, CEA = 5.15, PoT = 90.71%, and the fuzzy system assesses the risk of cancer as 91.7%, which corresponds to the interval of high-risk group. The risk factors in this case are all within normal values except for the age risk, which indicates that although this study adopts relatively stable risk factors as the experimental data, all risk factors can only function to the extent of increasing or decreasing the risk, and the actual physical condition has to be properly assessed by screening.

Case 2. Age = 61, BMI = 26.8, CEA = 1.09, PoT = 64.05%, and the fuzzy system assesses the risk of cancer as 60.2%, which corresponds to the interval of medium-high risk group. The possibility of cancer in this case is subject to assessment by the physician.

Case 3. Age = 47, BMI = 24.9, CEA = 434.78, PoT = 25.7%, and the fuzzy system assesses the risk of cancer as 10.1%, which corresponds to the interval of low-risk group. This case shows an unusually high CEA. Although CEA is only a reference value for risk factors, neither a high value means cancer, nor a low value means no cancer. As the human intestine is very long and the CEA in this case is significantly elevated above the abnormal range, this phenomenon suggests that the lesion may be located elsewhere in the intestine.

Taking into account the promising association between the imaging outcomes and the results of pathological analyses, the current study highlighted a time-efficient and noninvasive approach to the diagnosis of potential colorectal malignancies. The imaging tool may provide clinical guidance for clinicians to determine whether to proceed with high-risk procedures (e.g., polypectomy) or adopt a more conservative strategy, particularly in patients at high risk of complication (e.g., coagulopathy or impending colon perforation). Another advantage is the lack of requirement for specific software for operation. Nevertheless, the current study has its limitations. First, despite the involvement of up to 500 patients in the current study, the sample size is still relatively small to consolidate our findings. Second, although we included patient factors including age, CEA, and body mass index in our analysis, other risk factors for colorectal cancers such as dietary habits and family history were not taken into account. Further large-scale studies are warranted to validate the clinical application of the present imaging approach. Finally, the lack of real-time feedback is another limitation. Nevertheless, analyses of all images from a single patient can be completed within five minutes. In routine practice, such a short time could allow a clinician to decide the appropriate colonoscopic management strategy (i.e., invasive vs. conservative) based on the results of analysis when the patient is still under anesthesia or sedation.

The lack of real-time feedback is another limitation. Nevertheless, analyses of all images from a single patient can be completed within five minutes. In routine practice, such a short time could allow a clinician to decide the appropriate colonoscopic management strategy (i.e., invasive vs. conservative) based on the results of analysis when the patient is still under anesthesia or sedation.

Essentially the fuzzy rules based on the way-wise physicians' design and a large amount of learning information will reduce, but not completely avoid, the misdiagnosis rate.

7. Conclusion

In this study, a CNN training model is employed to analyze the assessment of tumor risk of healthy, benign, and malignant on colonoscopy images, and then the four parameters of tumor risk, age, BMI, and CEA are utilized to assess the risk of colorectal cancer by fuzzy algorithm, which assists physicians to effectively diagnose patients' symptoms through their current physical condition and data, thus reducing the misdiagnosis rate.

However, from the analysis and discussion of the experimental results, it is evident that despite the high priority of assessed tumor risk in the fuzzy system, as colonoscopy is the most direct way to screen for colorectal cancer, the consideration of risk factors has a certain degree of reference value for diagnostic signs, clinical analysis, postoperative follow-up, and prevention, in addition to the current physical condition.

Among the many colorectal cancer screening methods, colonoscopy is currently the most important and direct screening method. Compared with other methods, this method can directly observe the general situation of all tumors in the intestine. Compared with medical image computing together with MICCAI, CVC Colon DB, and ISIT-UMR, the Association for Computer-Aided Intervention used the image sequence data in the medical dataset as the training model of the deep convolutional neural network (DCNN) and trained two sets of settings Set-1 and DCNN. The accuracy rates of Set-2 are 75.71% and 79.78%, respectively [12]; based on computer-aided diagnosis (CAD), combined with convolutional neural network (CNN), through deep learning model after training, the polyp status analyzed by CAD in colonoscopy was used for verification test with CNN, and the accuracy rate of the research results was 75.1% [13].

The proximity and complexity of the organs in the human body, the image resolution, size, and angle can affect the accuracy of identifying the targeted tissues and lesions

using the training model. The biggest difference between CNN and multilayer perceptron (MLP) lies in the additional convolution layer and pooling layer. According to the Cancer Registration Annual Report of the National Health Administration of the Ministry of Health and Welfare of Taiwan, 17,302 people were initially diagnosed with colorectal cancer in 2019, of which 11,031 (63.8%) were colon cancer, and 6,271 (36.2%) were rectal, sigmoid junction, and anus [3]. Future research will be directed towards designing CNN network model to observe the difference in the diagnostic accuracy of colorectal cancer in different parts.

Data Availability

Access to data is restricted and the data are not freely available. Acceptable justifications for restricting access may include legal and ethical concerns, such as third-party rights, patient privacy, and commercial confidentiality.

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] S. Hyuna, F. Jacques, L. Rebecca et al., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *Ca - a Cancer Journal for Clinicians*, vol. 70, pp. 145–164, 2020.

[2] R. L. Siegel, K. D. Miller, and A. Jemal, "Cancer statistics, 2017," *CA: A Cancer Journal for Clinicians*, vol. 67, no. 1, pp. 7–30, 2017.

[3] *Cancer Registry Annual Report 2018*, pp. 1–5, Taiwan health promotion administration Ministry of health and Welfare, Taipei, Taiwan, 2020.

[4] L. Anna, R. Grzegorz, L. Tomasz, S. G. Aleksandra, and R. Sławomir, "Risk factors for the diagnosis of colorectal cancer," *Research Square*, vol. 29, pp. 1–15.

[5] A. H. Fatima and P. B. Robin, "Colorectal cancer epidemiology: incidence, mortality, survival, and risk factors," *Clinics in Colon and Rectal Surgery*, vol. 22, pp. 191–197, 2009.

[6] C. C. Kuan, C. L. Ko, H. C. Hong, C. C. Kung, L. W. Kuen, and C. S. Ling, "Path analysis of the impact of obesity on postoperative outcomes in colorectal cancer patients: a population-based study," *Journal of Clinical Medicine*, vol. 10, pp. 1–11, 2021.

[7] H. Alyssa, C. S. Eric, M. L. Jonathan et al., "Trends in the incidence of young-onset colorectal cancer with a focus on years approaching screening age: a population-based longitudinal study," *JNCI J Natl Cancer Inst*, vol. 113, pp. 863–868, 2021.

[8] A. Subki, A. ButtNS, and A. A. Alkahtani, "CEA and CA19-9 levels and KRAS mutation status as biomarkers for colorectal cancer," *Clinical Oncology*, vol. 6, pp. 1–8, 2021.

[9] L. Leilani, S. Silvia, W. Mathias, H. B. Doris, K. Marko, and L. Johannes, "Diagnostic and prognostic value of CEA and CA19-9 in colorectal cancer," *Diseases*, pp. 1–12, 2021.

[10] C. S. Huang, C. Y. Chen, L. K. Huang, W. S. Wang, and S. H. Yang, "Prognostic value of postoperative serum carcinoembryonic antigen levels in colorectal cancer patients who smoke," *PLoS One*, vol. 6, no. 5, pp. 1–14, 2020.

[11] K. Bruno, M. O. Andrea, P. M. Allen, M. N. Catherine, A. S. Matthew, and T. Lorenzo, "Looking under the hood: deep neural network visualization to interpret whole-slide image analysis outcomes for colorectal polyps," in *Proceedings of the 2017 IEEE Conference on Computer Vision and Pattern RecognitionWorkshops*, Honolulu, HI, USA, July 2017.

[12] P. Krushi, L. Kaidong, T. Ke et al., "A comparative study on polyp classification using convolutional neural networks," *PLoS One*, vol. 7, no. 30, pp. 1–16, 2020.

[13] K. Yoriaki, H. Hisashi, W. Tomohiro et al., "Computer-aided diagnosis based on convolutional neural network system for colorectal polyp classification: preliminary experience," *Oncology*, vol. 93, pp. 30–34, 2017.

[14] Y. Atsuo, N. Ryota, O. Keita, A. Tomonori, and K. Kazuhiko, "Automatic detection of colorectal neoplasia in wireless colon capsule endoscopic images using a deep convolutional neural network," *Endoscopy*, vol. 53, pp. 832–836, 2021.

[15] W. Liuli, Z. M. Liu, and Z. T. Huang, "Deep convolution network for direction of arrival estimation with sparse prior," *IEEE Signal Processing Letters*, vol. 26, pp. 1688–1692, 2019.

[16] D. D. Pukale, S. G. Bhirud, and V. D. Katkar, "Content based image retrieval using deep convolution eural network," *IEEE Xplore*, pp. 1–5, 2017.

[17] A. Z. Lotfi, "Fuzzy sets," *Information and Control*, vol. 8, pp. 338–353, 1965.

[18] A. Z. Lotfi, "Outline of a new approach to the analysis of complex systems and decision processes," *Man and Cybernetics*, vol. 3, pp. 28–44, 1973.

[19] S. K. Halgamuge and M. Glesner, "Neural networks in designing fuzzy systems for real world applications," *Fuzzy Sets and Systems*, vol. 65, no. 1, pp. 1–12, 1994.

[20] C. Tanjia, "Fuzzy logic based expert system for detecting colorectal cancer," *International Research Journal of Engineering and Technology*, vol. 05, pp. 389–393, 2018.

[21] S. Karen and Z. Andrew, "Very deep comvolutional networks for large-scale image recognition," in *Proceedings of the Conference Paper at International Conference on Learning Representations (ICLR)*, San Diego, CA, USA, May 2015.