

DESCRIPTION

TITLE OF INVENTION: LESION FEATURE RECOGNITION SYSTEM FOR GASTROINTESTINAL ENDOSCOPY

CROSS-REFERENCE TO RELATED APPLICATION

- 5 [0001]The present application claims the priority of Chinese Patent Application No. 202311040222.1, originally filed with the Chinese Patent Office on August 17, 2023 and entitled “System for Recognizing Gastrointestinal Endoscopic Lesion Features”, which is incorporated in its entirety herein by reference.

TECHNICAL FIELD

- 10 [0002]The present disclosure relates to the field of artificial intelligence, in particular to a system for recognizing gastrointestinal endoscopic lesion features.

BACKGROUND

- [0003]The application of artificial intelligence (AI) in clinical medicine has become a hot research topic. A primary approach involves constructing a multi-layered deep learning
15 model, collecting a large number of gastrointestinal endoscopic lesion images from hospitals, requiring endoscopists to annotate lesions represented in the images as training data, and allowing the AI to learn data features, to achieve the purpose of identifying some gastrointestinal diseases ultimately. Typically, it provides real-time assistance for an endoscopist during examinations through software prompts. When lesions are detected, a
20 prompt signal appears on a screen to remind the endoscopist to perform observation operations, and to provide a preliminary diagnostic result. Such technologies are being researched and applied across multiple areas including esophageal cancer, gastric cancer, and colorectal adenomatous polyps. They have demonstrated clinical value by enhancing the examination efficiency of endoscopists and reducing missed diagnoses. Beyond
25 localized lesions, AI-assisted diagnosis research has also made continuous progress in lesions such as chronic atrophic gastritis and ulcerative colitis. Although the AI has

already achieved medical-level proficiency in recognizing some static images of lesions, and even some technologies can diagnose the malignant or benign risk, in general, such applications still fall short of the maturity required for clinical use. With a high false positive rate, they remain distant from widespread adoption, with most achievements currently confined to scientific research and clinical trials.

[0004]At present, the research focus of mainstream AI technologies is primarily on recognizing and diagnosing endoscopic features for a single disease, and such technologies do not have the function of differentiating multiple diseases. To add diagnostic capabilities for a new disease, a new model must be built. After completing new sample training, the model operates in parallel with an original model to enable diagnosis. Clinically, dozens of diseases are commonly diagnosed through digestive endoscopy. To effectively assist endoscopists, the AI technologies must possess the diagnosing and identifying capabilities to simultaneously recognize dozens of diseases. This places higher demands on both computing power and accuracy for AI technologies utilizing single-disease models. It is challenging for the current hardware capabilities of small computers to simultaneously satisfy the requirements of low latency and versatility. Second, the AI technologies require a large number of samples as a prerequisite for model training. With the addition of diagnosis requirements for a new disease, corresponding costs for sample collection and annotation also increase. This also prevents such AI applications from rapidly expanding their diagnostic capabilities. Third, current technical solutions exhibit significant limitations in reaching diagnostic conclusions for diseases characterized by complex and variable features. Due to technical limitations, the more variable and complex the features of a single lesion are, the more challenging it becomes for the AI to reduce false positives and false negatives to an acceptable level.

[0005]In the field of endoscopic diagnosis of gastrointestinal diseases, various AI technologies and methods currently exist, such as deep learning, convolutional neural networks, support vector machines, and Bayesian networks. These technologies and methods can theoretically be used for endoscopic diagnosis of gastrointestinal diseases. However, they have limitations or insurmountable difficulties in achieving the goal of identifying various diseases. For example, 1) recognition of a single disease or lesion: some AI technologies may only be capable of recognizing a single disease or lesion, rather than simultaneously recognizing several different lesions and diseases. 2) High sample requirements: some AI technologies require an extensive amount of labeled data

and the expertise and experience of professional doctors in order to train high-quality models. 3) High model complexity: For some AI technologies, building a plurality of models or retraining models may increase model complexity and training time. 4) High misdiagnosis rate: Some AI technologies may fail to account for the interactions and similarities between different lesions, potentially leading to misdiagnosis issues. Consequently, to address the aforementioned technical problems, a novel solution is required. Only then can AI-assisted digestive endoscopy diagnosis technologies become widely applicable.

SUMMARY OF INVENTION

10 [0006]Technical problem to be solved:

[0007]Traditional AI models are usually trained and optimized for a single disease or a single lesion, requiring an extensive amount of labeled data and expertise and experience of professional doctors. Moreover, in practical application, interactions and similarities may exist between different lesions, and a traditional single-disease training method can not deal with these problems well. In addition, establishing a model for each disease separately increases complexity of the models, cost of sample processing, and training time, limiting efficiency and feasibility in practical applications.

[0008]In particular, an aspect of the present disclosure aims to solve the problems existing in the traditional AI models when they assist in endoscopic diagnosis of gastrointestinal diseases, and proposes a novel system. That is, features of different lesion types are analyzed. A plurality of feature points consistent with clinical diagnosis bases are extracted, and classified and summarized as a recognition target of an AI model. In this way, different lesions and diseases can be automatically recognized simultaneously, to avoid issues of establishing a plurality of models and retraining. Correct disease diagnosis can be obtained by performing secondary determination of a recognized feature combination. For clinical work where patient conditions are unpredictable and dozens of lesions require recognition, this method can enhance diagnostic accuracy and efficiency, providing endoscopists with a truly usable auxiliary diagnostic tool, and make up for shortcomings of single diagnostic capabilities, poor anti-interference, high computing power requirements, and a high misdiagnosis rate of a traditional AI-assisted endoscopic examination technology.

[0009]Technical solution:

[0010]A system for recognizing gastrointestinal endoscopic lesion features includes:

[0011]1) an image collection module;

[0012]2) a gastrointestinal endoscopic multi-level feature recognition and training module
5 including a first multi-task classification neural network configured to recognize n preset
different categories of gastrointestinal endoscopic feature groups and label values in the
gastrointestinal endoscopic feature groups, where $2 \leq n \leq 10$; and

[0013]3) a feature result collection module configured to collect recognition results
10 output by the gastrointestinal endoscopic multi-level feature recognition and training
module,

[0014]where the gastrointestinal endoscopic feature groups comprise a local lesion
feature group, and a label value of the local lesion feature group includes a local
depressed lesion, a local elevated lesion, or a local flat lesion; and

[0015]a downstream network of the local lesion feature group includes a local lesion
15 multi-level feature recognition and training module, the local lesion multi-level feature
recognition and training module includes a second multi-task classification neural
network configured to recognize m preset different categories of local lesion features and
complementary identification values corresponding to the categories, and $1 \leq m \leq 5$.

[0016]According to the other aspect of the present disclosure, an electronic apparatus is
20 provided. The electronic apparatus includes an electronic apparatus body and the above
system for recognizing gastrointestinal endoscopic lesion features. The system for
recognizing gastrointestinal endoscopic lesion features is mounted on the electronic
apparatus body.

[0017]The present disclosure has the beneficial effects as follows:

25 [0018]According to the present disclosure, a novel AI model based on feature point
classification and summarization is adopted. A plurality of different lesions and diseases
can be recognized simultaneously. The issues of establishing a plurality of models and
retraining are avoided. The accuracy and efficiency of endoscopic diagnosis of
gastrointestinal diseases are improved. Specifically:

30 [0019]1) The accuracy of diagnosis is improved: the AI model adopted by the present
disclosure can automatically recognize, classify, and summarize a plurality of different
lesions and diseases simultaneously, avoiding the issues of establishing a plurality of
models and retraining. Through the secondary determination of the recognized feature

combination, more accurate disease diagnosis results can be obtained, the misdiagnosis rate can be reduced, and the accuracy of diagnosis can be improved.

[0020]2) The efficiency of diagnosis is improved: the traditional AI models need to establish a model for each disease separately, increasing the complexity of the models, the cost of sample processing, and the training time, and limiting the efficiency and the feasibility of the AI models in practical applications. The AI model based on feature point classification and summarization adopted by the present disclosure can recognize a plurality of different lesions and diseases simultaneously. The issues of establishing a plurality of models and retraining are avoided. The efficiency of endoscopic diagnosis of gastrointestinal diseases is improved.

[0021]3) A workload of manual labeling is reduced: the AI model adopted by the present disclosure can automatically analyze, classify, and summarize features of different lesions and diseases, reducing the workload of manual labeling.

[0022]4) A clinical application value is improved: the AI model of the present disclosure can recognize a plurality of different lesions and diseases simultaneously, providing endoscopists with a truly usable auxiliary diagnostic tool. examination efficiency of the endoscopists can be improved. Missed diagnosis can be reduced. A value of AI in clinical application can be expanded. The AI technology application can be widely popularized.

BRIEF DESCRIPTION OF DRAWINGS

[0023]FIG. 1 is a schematic diagram of a system for recognizing gastrointestinal endoscopic lesion features constructed in an example of the present disclosure.

DETAILED DESCRIPTION OF THE EMBODIMENTS

[0024]A system for recognizing gastrointestinal endoscopic lesion features is disclosed in the present disclosure. Those skilled in the art can refer to contents herein to appropriately improve process parameters for implementation. It is specifically noted that all similar substitutions and modifications will be apparent to those skilled in the art and are deemed to be included in the present disclosure. It will be apparent to those skilled in the art that modifications or suitable variations and combinations of what is described herein can be

made to practice and apply the disclosed technology without departing from the content, spirit and scope of the present disclosure.

[0025]In the present disclosure, scientific and technical terms used herein have the meanings commonly understood by those skilled in the art, unless otherwise specified.

5 Throughout the specification and claims, unless explicitly indicated otherwise, the term “comprise,” or its variations of “include” or “contain,” and the like, will be understood to include stated elements or components but not to exclude other elements or components. The terms “such as,” “for example,” and the like are intended to refer to illustrative embodiments and are not intended to limit the scope of the present disclosure.

10 [0026]In order to enable those skilled in the art to better understand the technical solutions of the present disclosure, the present disclosure is described in further detail below in combination with specific examples.

[0027]As shown in FIG. 1, disclosed in the example is a system for recognizing gastrointestinal endoscopic lesion features of the present disclosure. The system includes:

15 [0028]S1. image collection module.

[0029]The image collection module is configured to obtain images stored in real time or in advance in a medium. The image collection module may include a video obtaining module and/or a picture obtaining module. In some embodiments of the present disclosure, the image collection module may use a video collection card to transmit
20 high-definition multimedia interface (HDMI), digital video interactive (DVI), standard data interface (SDI), S-Video, and other digital/analog video signals to a computer, read the video signals through opencv, and transcode the video signals into images in RGB image format frame by frame.

[0030]In some embodiments of the present disclosure, when the image collection module
25 obtains a video real-time image signal output by an endoscope apparatus, the image collection module performs interval sampling on the obtained frame-by-frame images, for example, with an interval sampling rate of 1 second/time, to obtain images to be recognized, and records frame numbers of the images to be recognized.

[0031]S2. gastrointestinal endoscopic multi-level feature recognition and training module.

30 [0032]The gastrointestinal endoscopic multi-level feature recognition and training module includes a first multi-task classification neural network configured to recognize n preset different categories of gastrointestinal endoscopic feature groups and label values in the gastrointestinal endoscopic feature groups, where $2 \leq n \leq 10$. The gastrointestinal

endoscopic feature groups include a local lesion feature group. In some embodiments of the present disclosure, n may be 2, 3, 4, 5, 6, 7, 8, 9, or 10. In an embodiment of the present disclosure, n=5. That is, the gastrointestinal endoscopic feature groups include five different categories, that is, a mucosal feature group, a vascular texture feature group, a hemorrhage feature group, a mucus feature group, and the local lesion feature group. In some other embodiments of the present disclosure, at least one of the mucosal feature group, the vascular texture feature group, the hemorrhage feature group, or the mucus feature group, or other relevant gastrointestinal endoscopic feature groups may be selected. Mucosal features, vascular texture features, hemorrhage features, and mucus features mentioned above belong to diffuse lesion features.

[0033]In some embodiments of the present disclosure, label values corresponding to the above different categories of gastrointestinal endoscopic feature groups may include that, for example, label values of the mucosal feature group include mucosal thinning, a linear erythematous mucosa, a patchy erythematous mucosa, a mucosal edema, an island-like mucosa, a residual glandular mucosa, a gastric mucosa, an intestinal mucosa, or an intestinal metaplasia; label values of the vascular texture feature group include obscured submucosal vessels, linear vessels, dendritic vessels, or spotty and patchy vessels; label values of the hemorrhage feature group include no hemorrhage, spotty and patchy hemorrhage, mild hemorrhage, or massive hemorrhage; label values of the mucus feature group include yellowish mucus, viscous mucus, bile reflux mucus, clear mucus, or a mucus lake; and a label value of the local lesion feature group includes a local depressed lesion, a local elevated lesion, or a local flat lesion.

[0034]To achieve an objective of the present disclosure, in some embodiments of the present disclosure, the first multi-task classification neural network includes one main feature extraction network configured to recognize the n preset different categories of gastrointestinal endoscopic feature groups, and n branch feature extraction networks located downstream of the main feature extraction network and configured to recognize the label values in the gastrointestinal endoscopic feature groups respectively. In an embodiment of the present disclosure, n=5. A method for constructing the first multi-task classification neural network includes:

[0035]Step A) at least three (for example, 4, 6, 8, 9, and 10) multi-branch parallel feature extraction units are used to form the main feature extraction network, and a first main feature extraction network feature map is obtained through operation.

[0036]Step B) at least three (for example, 4, 5, 6, and 8) multi-branch parallel feature extraction units are used to form an n-th branch feature extraction network; and the first main feature extraction network feature map is input, and finally an n-th branch feature extraction network feature map is obtained through multi-branch parallel convolution operation.

[0037]A method for designing a multi-branch parallel convolution operation unit includes:

[0038]a) a slicing unit is constructed and configured to divide an original input feature map into four groups to obtain a first slice feature group, a second slice feature group, a third slice feature group, and a fourth slice feature group.

[0039]b) one grouping feature extraction unit is constructed by three depthwise separable convolutions, and features are extracted from the first slice feature group, the second slice feature group, and the third slice feature group separately. The three depthwise separable convolutions have sizes of 3×3 , 1×13 , and 13×1 respectively, and are used for extracting spatial features, vertical channel features, and horizontal channel features. Finally, an extracted spatial feature map, a vertical channel feature map, a horizontal channel feature map, and the fourth slice feature group are re-concatenated, and a first output is obtained through one regularization layer.

[0040]c) a feature enhancement unit is constructed, where the feature enhancement unit is composed of one 1×1 convolution layer, an activation function layer, one 1×1 convolution layer, and one residual connection, a second output is obtained by performing series operation on a 1×1 convolution, an activation function, and an 1×1 convolution, and a third output is obtained by performing horizontal direction feature map concatenating on the original input feature map and the second output through the residual connection, that is, an output result of the multi-branch parallel convolution operation unit is obtained.

[0041]The 1×1 convolution layer is formed by conventional technical means in the art, and the activation function is GeLU.

[0042]Step C) a first prediction layer, a second prediction layer, a third prediction layer, a fourth prediction layer, and a fifth prediction layer are constructed by using a fully connected layer and a Sigmoid function as a combination unit. Specifically, a first branch feature map, a second branch feature map, a third branch feature map, a fourth branch feature map, and a fifth branch feature map are input to the fully connected layer, the network layer compresses the high-dimensional feature maps into a 1-dimensional vector,

and finally label recognition results of a first feature group, a second feature group, a third feature group, a fourth feature group, and a fifth feature group are obtained through Sigmoid function operation. A Sigmoid operation formula is:

$$p_k = \text{Sigmoid}(W_k^T \cdot F + b_k)$$

[0043]where p_k indicates a probability of solving for category k, W_k^T is a trainable hyperparameter for category k, and b_k is a trainable biased hyperparameter for category k.

[0044]In some embodiments of the present disclosure, training sets are input to the first multi-task classification neural network in batches for training.

[0045]A source of the training sets is: gastrointestinal picture data is collected according to a plurality of groups of preset multi-label feature results, feature labels are labeled one by one to obtain a first feature label data set, and the first feature label data set is split into a training set and a verification set according to a ratio of 8:2. The feature groups and the corresponding label values are shown in Table 1.

Table 1

Feature group name	Feature label value
Mucosal feature group	Mucosal thinning, a linear erythematous mucosa, a patchy erythematous mucosa, a mucosal edema, an island-like mucosa, a residual glandular mucosa, a gastric mucosa, an intestinal mucosa, and an intestinal metaplasia
Vascular texture feature group	Obscured submucosal vessels, linear vessels, dendritic vessels, and spotty and patchy vessels
Hemorrhage feature group	No hemorrhage, spotty and patchy hemorrhage, mild hemorrhage, and massive hemorrhage
Mucus feature group	Yellowish mucus, viscous mucus, bile reflux mucus, clear mucus, and a mucus lake
Lesion feature group	Local erosive lesion, local elevated lesion, diffuse lesion

[0046]In order to achieve an objective of the present disclosure, a method for training the first multi-task classification neural network includes:

[0047]Step D) an original label value output correspondingly by each of the gastrointestinal endoscopic feature groups is predicted, calculation is performed through a BCE cross-entropy loss function, an n-th loss function is obtained, and a total loss of the network is calculated in a weighted summation manner, where a specific formula is as

follows:

$$L_k = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1-y_i) \log(1-p_i)]$$

[0048]where k is a number of branches of the feature groups; L_k is a derivative of a loss for k feature branches; N is a number of pictures in a training batch; i is a corresponding category; y_i is a true label value of category i in the pictures, that is, y_i is 1 in a case where a picture belongs to category i, and y_i is 0 in a case where a picture does not belong to category i; and p_i is a prediction confidence coefficient of the picture for category i;

$$L = w_1 \times L_1 + w_2 \times L_2 + \dots + w_k \times L_k$$

[0049]where L is the total loss, and w is a weight of the loss function; and the total loss is obtained by summing.

[0050]In the present disclosure, an adjustment strategy for the weight of the loss function is stage attenuation. That is, fixed weights are [2.5, 2.5, 2.5, 2.5, 2.5] at the beginning of training, and are adjusted every 10 rounds. An adjustment basis is: in a case where $L_t < L_{(t-10)}$, $w_t = w_t * \theta$; and in a case where $L_t > L_{(t-10)}$, $w_t = w_t * (2-\theta)$. Further, θ is an attenuation coefficient, and $\theta=0.85$.

[0051]Step E) parameters of a neural network model are updated by using a backpropagation technology according to the total loss of the network, training is stopped in a case where a final loss function is stabilized to a preset value or reaches a preset maximum iteration number, and the first multi-task classification neural network trained in advance is obtained.

[0052]The first multi-task classification neural network constructed and trained is used for recognizing and predicting the gastrointestinal endoscopic lesion features. 1) The frame-by-frame images obtained by the image collection module are sampled at intervals with a sampling rate of 1 second/time, to obtain images to be recognized, and frame numbers of the images to be recognized are recorded. 2) Multi-group multi-label feature recognition is performed on the image to be recognized by using the first multi-task classification neural network constructed in advance to obtain a recognition result of a first mucosal feature group, a recognition result of a first vascular texture feature group, a recognition result of a first hemorrhage feature group, a recognition result of a first mucus feature group, and a recognition result of first local lesion feature group.

[0053]The recognition result of the first mucosal feature group, the recognition result of the first vascular texture feature group, the recognition result of the first hemorrhage feature group, the recognition result of the first mucus feature group, and the recognition result of the first local lesion feature group obtained above are input to a feature result collection module for appropriate purposes, which includes, but is not limited to, auxiliary diagnosis of gastrointestinal diseases.

[0054]To achieve an objective of the present disclosure, in some embodiments of the present disclosure, a downstream network of the local lesion feature group may further include a local lesion multi-level feature recognition and training module. The local lesion multi-level feature recognition and training module includes a second multi-task classification neural network configured to recognize m preset different categories of local lesion features and complementary identification values corresponding to the categories, and $1 \leq m \leq 5$. In some embodiments of the present disclosure, m may be 1, 2, 3, 4, or 5. In an embodiment of the present disclosure, $m=4$. That is, the local lesion features include four different categories, for example, a lesion morphology feature, a lesion area feature, a lesion surface feature, or a lesion border feature.

[0055]Further, S4. local lesion multi-level feature recognition and training module.

[0056]The local lesion multi-level feature recognition and training module includes a second multi-task classification neural network configured to recognize four preset different categories of local lesion features and complementary identification values corresponding to the categories. The second multi-task classification neural network may be obtained by using the same methods for constructing and training the first multi-task classification neural network as described above. In some embodiments of the present disclosure, a method for constructing the second multi-task classification neural network includes:

[0057]Step F) a lesion segmentation artificial intelligence neural network is constructed based on a DeepLabV3+ algorithm, and predicting images of the local lesion feature group, to obtain a first lesion segmentation map. Specifically,

[0058]i) Required data sets are constructed. A contour of a lesion is labeled. Targets include an elevated lesion target and an erosive lesion target. A first segmented lesion data set is obtained.

[0059]Second, multi-level labels of the lesion are labeled to construct a second label feature data set.

[0060]ii) A lesion segmentation model is trained according to the first segmented lesion data set to obtain a first lesion segmentation model. Input data is an image, and an output is a mask map containing only a segmented region.

[0061]Step G) a second multi-task classification neural network algorithm is use for predicting multi-level feature labels of the lesion. The first lesion segmentation map is input to obtain prediction results of a first lesion morphology feature group, a first lesion area feature group, a first lesion surface feature group, and a first lesion border feature group. Specifically,

[0062]i) The network is constructed by using the same method for constructing the first multi-task classification neural network as described above. Since only four feature groups are provided, four branches are set, and therefore subsequent four loss functions are provided.

[0063]ii) The second multi-task classification neural network algorithm constructed in advance is obtained through training, such that the second multi-task classification neural network algorithm has a lesion feature group label recognition capability. That is,

[0064]when recognition labels in a first lesion feature recognition group include "local erosion lesion" and "local elevated lesion," the second multi-task classification neural network algorithm constructed in advance is used for recognizing complementary identification values of the feature groups. A first complementary lesion morphology feature group, a first complementary lesion area feature group, a first complementary lesion surface feature group, and a first complementary lesion border feature group are obtained.

[0065]1) In the first complementary lesion morphology feature group, corresponding labels of "local erosive lesion" include: red and swollen, and not red and swollen; and corresponding labels of "local elevated lesion" include: flat, lateral, pedunculated, and irregular.

[0066]2) In the first complementary lesion area feature group, corresponding labels of "local erosive lesion" include: punctate, superficial and small, and deep and large; and corresponding labels of "local elevated lesion" include: multiple and solitary.

[0067]3) In the first complementary lesion surface feature group, corresponding labels of "local erosive lesion" include: no white coating, grayish white, yellowish white; and corresponding labels of "local elevated lesion" include: light white, light pink, and deep red.

[0068]4) In the first complementary lesion border feature group, corresponding labels of "local erosive lesion" include: smooth and burr; and corresponding labels of "local elevated lesion" include: smooth and rough.

[0069]Step H) a label conversion module is constructed, and the complementary
5 identification values obtained in step G) are fused with the label values of the local lesion feature group, to obtain a label value of a new local lesion feature group (second local lesion feature group).

[0070]A recognition result of the second local lesion feature group, together with the recognition result of the first mucosal feature group, the recognition result of the first
10 vascular texture feature group, the recognition result of the first hemorrhage feature group, and the recognition result of the first mucus feature group, may be input to the feature result collection module for appropriate purposes, which includes, but is not limited to, auxiliary diagnosis of gastrointestinal diseases.

[0071]To achieve an objective of the present disclosure, in some embodiments of the
15 present disclosure, the system for recognizing gastrointestinal endoscopic lesion features further includes a mucosal location recognition module configured to convert the label values in the mucosal feature group according to a mucosal location and output the recognition result to the feature result collection module. The location includes esophagus, gastric body, gastric fundus, gastric antrum, duodenum, cecum, ascending colon,
20 transverse colon, descending colon, sigmoid colon, or rectum.

[0072]Further, S3. mucosal location recognition module.

[0073]A method for constructing the mucosal location recognition module includes:

[0074]step I) a time-series segment classification model is used for classifying and
25 recognizing a location of a site in a current video/picture to obtain a location recognition result.

[0075]i) Video clips of any appropriate duration during digestive endoscopy are collected.

[0076]ii) Data is preprocessed, dimensionality is reduced and an image is scaled to an appropriate size by using an interval sampling technology according to a total number of frames in a collection duration.

30 [0077]iii) ResNet-3D is used for inputting and training the preprocessed data, and parameters are updated by a backpropagation technology in the process, to obtain a preset time-series segment classification model, such that the model has a position recognition capability.

[0078]step J) result conversion is performed on the first mucosal feature group in a feature recognition result of a first mucosal feature group image to be recognized after sampling and obtained by the gastrointestinal endoscopic multi-level feature recognition and training module. Specifically,

5 [0079]i) Images read frame by frame are accumulated to obtain a segment set reaching an appropriate duration, and the segment set is processed by using the preprocessing module.

[0080]ii) Location prediction is performed by using the preset time-series segment classification model to obtain a frame number range and a prediction result corresponding to the location.

10 [0081]iii) The first mucosal feature group corresponding to a frame number in the interval is reversely searched for according to the frame number range. In a case where corresponding locations of the interval of the frame number range are "gastric body," "gastric fundus," and "gastric antrum", and the first feature group in the interval contains a label of "intestinal mucosa", the label of "intestinal mucosa" is removed from the first
15 mucosal feature group, and a label of "intestinal metaplasia" is added.

[0082]iv) The label is added to the first mucosal feature group according to a recognition result corresponding to the location in the frame segment.

[0083]The result output by the mucosal location recognition module is a recognition result of the second mucosal feature group.

20 [0084]The recognition result of the second mucosal feature group, together with the recognition result of the first vascular texture feature group, the recognition result of the first hemorrhage feature group, the recognition result of the first mucus feature group, and the recognition result of the first local lesion feature group, may be input to the feature result collection module for appropriate purposes, which includes, but is not limited to,
25 auxiliary diagnosis of gastrointestinal diseases.

[0085]In some other embodiments of the present disclosure, the recognition results output by the gastrointestinal endoscopic multi-level feature recognition and training module may include: the recognition result of the first/second mucosal feature group, the recognition result of the first vascular texture feature group, the recognition result of the first hemorrhage feature group, the recognition result of the first mucus feature group, and
30 the recognition result of the first/second local lesion feature group. The recognition results are input to the feature result collection module for appropriate purposes, which includes, but is not limited to, auxiliary diagnosis of gastrointestinal diseases.

[0086]To achieve an objective of the present disclosure, in some embodiments of the present disclosure, the system for recognizing gastrointestinal endoscopic lesion features further includes a disease diagnosis decision making module configured to obtain a disease diagnosis result by using the recognition results.

5 [0087]Further, S5. disease diagnosis decision making module.

[0088]A method for constructing the disease diagnosis decision making module includes:

[0089]Step K) disease diagnosis labeling is performed on medical images in the first label data set, to obtain a first disease data set. Inputs of the data set are, for example, the recognition result of the second mucosal feature group, the recognition result of the first
10 vascular texture feature group, the recognition result of the first hemorrhage feature group, the recognition result of the first mucus feature group, and the recognition result of the second local lesion feature group (that is, label values). An output is a disease category. In some embodiments of the present disclosure, the disease category includes, but is not limited to, ulcerative colitis and atrophic gastritis. In some other embodiments of the
15 present disclosure, the output may alternatively be nature of a disease, including, for example, but not limited to, a local high-risk lesion, a local low-risk lesion, and a diffuse lesion and typing.

[0090]Step L) a preset disease diagnosis decision making algorithm is constructed. Specifically, a disease contained in the medical images is labeled based on an XGBOOST
20 algorithm, and data is trained iteratively batch by batch to obtain the disease diagnosis decision making algorithm constructed in advance.

[0091]Step M) the disease diagnosis decision making module may further include a preprocessing module configured to align the recognition result of the second mucosal feature group, the recognition result of the first vascular texture feature group, the
25 recognition result of the first hemorrhage feature group, the recognition result of the first mucus feature group, and the recognition result of the second local lesion feature group (that is, label values) to obtain an aggregate feature array.

[0092]Step N) the disease diagnosis decision making algorithm constructed in advance is used for performing inputting and algorithm inference on a feature set, to obtain the
30 disease diagnosis result.

[0093]The above descriptions are merely preferred embodiments of the present disclosure. It should be noted that those of ordinary skill in the art may further make several improvements and modifications without departing from the principle of the present

disclosure, but such improvements and modifications should be deemed as falling within the scope of protection of the present disclosure.

Claims

[Claim 1] A system for recognizing gastrointestinal endoscopic lesion features, comprising:

1) an image collection module;

5 2) a gastrointestinal endoscopic multi-level feature recognition and training module comprising a first multi-task classification neural network configured to recognize n preset different categories of gastrointestinal endoscopic feature groups and label values in the gastrointestinal endoscopic feature groups, wherein $1 \leq n \leq 10$; and

10 3) a feature result collection module configured to collect recognition results output by the gastrointestinal endoscopic multi-level feature recognition and training module,

wherein the gastrointestinal endoscopic feature groups comprise a local lesion feature group, and a label value of the local lesion feature group comprises a local depressed lesion, a local elevated lesion, or a local flat lesion; and

15 a downstream network of the local lesion feature group comprises a local lesion multi-level feature recognition and training module, the local lesion multi-level feature recognition and training module comprises a second multi-task classification neural network configured to recognize m preset different categories of local lesion features and complementary identification values corresponding to the categories, and $1 \leq m \leq 5$.

20 [Claim 2] The system for recognizing gastrointestinal endoscopic lesion features according to claim 1, wherein the image collection module comprises a video obtaining module and/or a picture obtaining module,

wherein the video obtaining module is configured to read a video signal output by an endoscope apparatus, to obtain video frame images; and

25 preferably, the video obtaining module performs interval sampling on the obtained frame-by-frame images to obtain images to be recognized, and records frame numbers of the images to be recognized.

30 [Claim 3] The system for recognizing gastrointestinal endoscopic lesion features according to claim 1, wherein the gastrointestinal endoscopic feature groups further comprise at least one of a mucosal feature group, a vascular texture feature group, a hemorrhage feature group, or a mucus feature group; and

preferably, label values of the mucosal feature group comprise mucosal thinning, a linear erythematous mucosa, a patchy erythematous mucosa, a mucosal edema, an

island-like mucosa, a residual glandular mucosa, a gastric mucosa, an intestinal mucosa, or an intestinal metaplasia; label values of the vascular texture feature group comprise obscured submucosal vessels, linear vessels, dendritic vessels, or spotty and patchy vessels; label values of the hemorrhage feature group comprise no hemorrhage, spotty and patchy hemorrhage, mild hemorrhage, or massive hemorrhage; and label values of the mucus feature group comprise yellowish mucus, viscous mucus, bile reflux mucus, clear mucus, or a mucus lake.

[Claim 4] The system for recognizing gastrointestinal endoscopic lesion features according to claim 1, wherein the first multi-task classification neural network comprises one main feature extraction network configured to recognize the n preset different categories of gastrointestinal endoscopic feature groups, and n branch feature extraction networks located downstream of the main feature extraction network and configured to recognize the label values in the gastrointestinal endoscopic feature groups respectively, and a method for constructing the first multi-task classification neural network comprises:

step A) using at least three multi-branch parallel feature extraction units to form the main feature extraction network, and obtaining a first main feature extraction network feature map through operation;

step B) using at least three multi-branch parallel feature extraction units to form an n-th branch feature extraction network respectively; and inputting the first main feature extraction network feature map, and finally obtaining an n-th branch feature extraction network feature map through multi-branch parallel convolution operation,

wherein a method for designing a multi-branch parallel convolution operation unit comprises:

a) constructing a slicing unit configured to divide an original input feature map into four groups to obtain a first slice feature group, a second slice feature group, a third slice feature group, and a fourth slice feature group;

b) constructing one grouping feature extraction unit by three depthwise separable convolutions, extracting features from the first slice feature group, the second slice feature group, and the third slice feature group separately, wherein the three depthwise separable convolutions have sizes of 3*3, 1*13, and 13*1 respectively, and are used for extracting spatial features, vertical channel features, and horizontal channel features, finally, re-concatenating an extracted spatial feature map, a vertical channel feature map, a horizontal channel feature map, and the fourth slice feature group, and obtaining a first

output through one regularization layer; and

c) constructing a feature enhancement unit, wherein the feature enhancement unit is composed of one 1*1 convolution layer, an activation function layer, one 1*1 convolution layer, and one residual connection, obtaining a second output by performing series operation on a 1*1 convolution, an activation function, and an 1*1 convolution, and obtaining a third output by performing horizontal direction feature map concatenating on the original input feature map and the second output through the residual connection, that is, obtaining an output result of the multi-branch parallel convolution operation unit; and

step C) using a fully connected layer and a Sigmoid function as a combination unit to construct an n-th prediction layer, and obtaining a label recognition result of an n-th gastrointestinal endoscopic feature group through Sigmoid function operation, wherein a Sigmoid operation formula is:

$$p_k = \text{Sigmoid}(W_k^T \cdot F + b_k)$$

wherein p_k indicates a probability of solving for category k, W_k^T is a trainable hyperparameter for category k, and b_k is a trainable biased hyperparameter for category k.

[Claim 5] The system for recognizing gastrointestinal endoscopic lesion features according to claim 1 or 4, wherein a method for training the first multi-task classification neural network comprises: inputting training sets to the first multi-task classification neural network in batches for training, and specifically comprises:

step D) predicting an original label value output correspondingly by each of the gastrointestinal endoscopic feature groups, performing calculation through a BCE cross-entropy loss function, obtaining an n-th loss function, and calculating a total loss of the network in a weighted summation manner, wherein a specific formula is as follows:

$$L_k = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1-y_i) \log(1-p_i)]$$

wherein k is a number of branches of the feature groups; L_k is a derivative of a loss for k feature branches; N is a number of pictures in a training batch; i is a corresponding category; y_i is a true label value of category i in the pictures, that is, y_i is 1 in a case where a picture belongs to category i, and y_i is 0 in a case where a picture does not belong to category i; and p_i is a prediction confidence coefficient of the picture for category i;

$$L=w_1\times L_1+w_2\times L_2+ \dots +w_k\times L_k$$

wherein L is the total loss, and w is a weight of the loss function; and the total loss is obtained by summing; and

step E) updating parameters of a neural network model by using a backpropagation technology according to the total loss of the network, stopping training in a case where a final loss function is stabilized to a preset value or reaches a preset maximum iteration number, and obtaining the first multi-task classification neural network trained in advance.

[Claim 6] The system for recognizing gastrointestinal endoscopic lesion features according to claim 3, further comprising a mucosal location recognition module configured to convert the label values in the mucosal feature group according to a mucosal location and output the recognition result to the feature result collection module;

wherein the location comprises esophagus, gastric body, gastric fundus, gastric antrum, duodenum, cecum, ascending colon, transverse colon, descending colon, sigmoid colon, or rectum.

[Claim 7] The system for recognizing gastrointestinal endoscopic lesion features according to claim 1, wherein the local lesion features comprise at least one of a lesion morphology feature, a lesion area feature, a lesion surface feature, or a lesion border feature.

[Claim 8] The system for recognizing gastrointestinal endoscopic lesion features according to claim 1 or 7, wherein a method for constructing the second multi-task classification neural network comprises:

step F) constructing a local lesion segmentation artificial intelligence neural network, wherein the local lesion segmentation artificial intelligence neural network is configured to classify and predict images input by the local lesion feature group, to obtain a first lesion segmentation map;

step G) constructing a local feature multi-label recognition network based on a segmented region, wherein the local feature multi-label recognition network is configured to recognize the m preset different categories of local lesion features and complementary identification values corresponding to the categories; and

step H) constructing a label conversion module, wherein the label conversion module is configured to fuse the complementary identification values obtained in step G) with the label value of the local lesion feature group, to obtain a new label value of the

local lesion feature group.

[Claim 9] The system for recognizing gastrointestinal endoscopic lesion features according to claim 1, further comprising a disease diagnosis decision making module configured to obtain a disease diagnosis result by using the recognition results.

- 5 [Claim 10] An electronic apparatus, comprising an electronic apparatus body and the system for recognizing gastrointestinal endoscopic lesion features according to any one of claims 1 to 9, wherein the system for recognizing gastrointestinal endoscopic lesion features is mounted on the electronic apparatus body.