Claims

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

1) an image collection module;

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

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; 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

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 $1 \leq m \leq 5$.

[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

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.

[Claim 3] The system for recognizing gastrointestinal endoscopic lesion features according to claim 1, wherein 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; 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:

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

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;

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

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] 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 8, wherein the system for recognizing gastrointestinal endoscopic lesion features is mounted on the electronic apparatus body.