Answering Reviewer1

1）Material and method. The authors should justify why you select the Network type that you include in the study, as well as Python-3

ResNet was put forward in 2015 and won the first place in the classification task of the ImageNet competition, because it is "simple and practical". Later, many methods are based on ResNet50 or ResNet101 and are widely used in detection, segmentation, recognition and other fields. It makes a reference (X) for the input of each layer, learning to form residual functions, rather than learning some functions without reference (X). This residual function is easier to optimize and can greatly deepen the number of network layers. The extracted image features have strong robustness. ResNet50 is faster than ResNet100. Therefore, ResNet50 is selected as the skeleton network of semantic segmentation network in this paper.

Python 3 is a good deep learning programming language and supports multiple deep learning frameworks.

2）Must explain the acronyms you use for the evaluation index: PPV, NPV, MIOU, PA, and what they define, sensitivity, specificity, accuracy, precision...

It has been explained in the article.

3) Results: Table 2 should be explained in the results, not in the discussion section.

Adjusted in article.

4） Discussion: the results should be contrasted with previous articles, for example by Leenhardt et al, Articles related to other investigations of angiodysplasia and convolutional neural network have not been included in this section (Aoki, et al; Hosoe et al; Noya et al; Park et al, Tsubo i et al). Leenhardt et al applied segmentation in technology and has the highest level of lesion detection at present. The positive predictive value was 96%. However, the algorithm presented in this paper shows some advantages in the test set. Our accuracy rate is 98%.

5)Finally, highlight what this research contributes in relation to the previous ones.

Firstly, the segmentation network is introduced into the recognition of small intestinal vascular aberrations. Secondly, feature fusion mechanism is introduced into the semantic segmentation network to improve the recognition accuracy of the segmentation network. Finally, the model can detect small intestinal vascular malformation lesions with high accuracy and speed.

6） References. Are references 14 and 31 the same?

Modified in article.

Answering Reviewers 2

1)You have mentioned that ResNet was used for image classification while

using CE. So, how CNN & DL algorithm classification were made on the basis of annotation. How many experts did the annotation?

The semantic segmentation network includes an encoder and a decoder. The encoder extracts feature through convolution, and the decoder classifies each pixel value. A good classification network without SoftMax and full connection layer is a good feature extraction network. So ResNet-50 is selected as the skeleton network of the encoder in the semantic segmentation network of this paper. An experienced endoscopy team composed of three experts marked the annotation data.

The two doctors are the label personnel, and the one with the most experience is the arbitration.

2) Were the annotations, classifications and grading were confirmed by some other diagnostic modalities?

The labeling standard comes from a Chinese version of Capsule Endoscopy Atlas which the chief editor Houde Zhang is a famous Chinese endoscope expert. All doctors follow the specified standard pictures for labeling and confirmation.

3) What was the grader variance?

Vessels in the study were classified according to the pattern of discrimination and were not graded.

4) Role in identifying other lesions of the small intestine.

In this research, we studied the identification of small intestinal vascular

malformation. This method could be expanded to identify various lesions in the small intestine (including depressed and elevated lesions). And we are going to start the next step.

5) It's future applications and improvement in segment identification.

In the future, we will collect more data on small intestinal vascular malformation lesions to improve the accuracy of the model. In addition, continuous inter frame information of capsule endoscopy is added to the model to make the model more accurate in semantic segmentation.

6) How many images were ungradable / not classified and possible solutions?

The model proposed in this paper identified 85 lesions in the test set incorrectly. In the test set, 19 samples were identified without lesions. In practical application, it is required to identify manually and AI models together. The model proposed in this paper only plays an auxiliary role when doctors diagnose.