College of Engineering Chengannur
Department of Computer Engineering
M. Tech. Computer Science (Image Processing)
03CS7903 Seminar II
Abstract of Proposed Seminar Topic

Ensemble of Instance Segmentation Models for Polyp Segmentation in Colonoscopy Images

07/MCS/2019 CHN19CSIP03 Krishnendu S.

September 7, 2020

Keywords: colorectal-cancer, polyp, colon-cancer, rectal-cancer, colonoscopy, polyp detection, object detection, neural network, MaskR-CNN, ResNet50, ResNet101, instance segmentation, image segmentation, semantic segmentation, ensemble model, histogram equalization, image matting model

Abstract

Colorectal cancer is one of the main causes of cancer death in worldwide.Colorectal cancer occurs in the colon or rectum.These cancers can also be called colon cancer or rectal cancer,depending on where they start.Most colorectal cancers start as a growth on the inner lining of the colon or rectum.These growth are called polyps.Some types of polyps can change into cancer overtime,but not all polyps become cancer.The chance of a polyp turning into cancer depends on the type of polyp it is.So early detection of polyp is key to ensure patient survival.Colonoscopy has been widely applied as a common practice to detect polyps and to detect changes or abnormalities in the large colon.However missing polyps in such procedure could happen and thus preventing early disease detection and treatment.Traditional manual screening is time consuming,operator dependent and error prone.Hence automated detection approach is highly demanded in clinical practice.Moreover polyp detection and segmentation is very challenging due to variations in polyp size,color,shape,texture and variations between polyps and hard mimics.In this proposed method,here apply a powerful object detection neural network 'Mask R-CNN' to detect and gives instance segmentation of polyps in colonoscopy images.Instance segmentation is a subtype of image segmentation which identifies each instance of each object within the image at the pixel level ie, associate a class label to each pixel similar to semantic segmentation except that it treats multiple objects of the same class as individual objects or separate entities.The proposed method also apply an ensemble method to combine the two Mask R-CNN models with different backbone structures(ResNet50 and ResNet101) to enhance the performance.

The overview of the proposed method is that,it consists of three components that are,data augmentation,two Mask R-CNN's with different backbone structures and the ensemble method.This is the first work to use Mask R-CNN for the task of polyp segmentation.

One of the challenges in training polyp segmentation model is the insufficient number of data for training,because access data is limited due to privacy concerns.Since the endoscopy procedures involving moving camera control,color calibration are not consistent, the appearance of endoscopy images significantly changes across different laboratories.The data augmentation steps bring endoscopy images into an extended space that can cover all their variances.By augmenting the training data can reduce the overfitting problem on training models.In this proposed model the applied augmentation methods are,vertical flipping,horizontal flipping,random rotation,random scaling,random shearing,random Guassian blurring,random contrast normalization,random brightness and random cropping and padding.

Mask R-CNN for instance segmentation is a flexible,small generic object instance segmentation method,which effi-

ciently detects object in an image,while simultaneously generating a high quality segmentation mask for each instance.Mask R-CNN is an extension of the faster R-CNN by adding a branch of predicting an object mask in parallel with the existing branch for bounding box recognition.The proposed frame work used two different backbone structures that are ResNet50 and Resnet101 for each Mask R-CNN.Using a ResNet backbone for feature extraction with Mask R-CNN gives excellent gains in both accuracy and speed.The Mask R-CNN is a two stage frame work.The first stage is a region proposed network(RPN).RPN is a new proposal generation network from faster R-CNN.The second stage has two parallel branches.The first one is the bounding box branch for detection.It contains bounding box regression and classification.The second one is the mask branch for segmentation.Mask R-CNN is simple to train and adds only a small overhead to faster R-CNN,running at 5fps.Traditional faster R-CNN has two outputs for each candidate object,a class label and a bounding box offset,to this here add a third branch that outputs the object mask.But the additional mask output is distinct from the class and box outputs.The key element of mask R-CNN including pixel to pixel alignment which is the main missing piece of faster R-CNN.

The Mask R-CNN with different backbone structures converge to different solutions,although it uses the same training data.For instance,Mask R-CNN with ResNet101 produced better segmentation results than Mask R-CNN with ResNet50 for some polyp images and viceversa.Based on this observation here use the ensemble method to combines two predictive mask by bitwise combination operation.Ensemble method combines the output of different models to reduce generalization error.This method can produce more accurate solution than a single model.The aggregate result of multiple model is always noisy than the individual models.This leads to model stability and robustness.The main causes of error in learning models are due to noise bias and variance.Ensemble method help to minimize these factors.

The training of these proposed model is divided into two phases that are,warmup phase and fully training phase.In warmup phase,here temprory freeze the backbone model and update only the rest part of the network via stochastic gradient descent(SGD) with momentum.In fully training phase,unfreeze the backbone model and update the entire network via SGD with momentum. The proposed model is evaluated on three public available datasets that are CVC-ClinicDB,ETIS-Larib,and CVC-ColonDB and trained on COCO dataset.

The main advantage of the proposed system is that the instance segmentation of Mask R-CNN network.The Mask R-CNN can recognize the polyp boundary as much as possible what other existing methods could not do.Moreover the RPN used here doesnot scan the actual image,instead the network scans the feature map making it much faster.The drawback of this system is that in Mask R-CNN the performance of the detection steps limits the performance of segmentation.

In future the proposed model can be improved by using more preprocessing method of image processing such as histogram equalization to improve image quality thereby reducing the false postive rate in detection.Then an image matting model can be used to seperate foreground,background and unknown regions.Furthermore in the exploration of Mask R-CNN can use other backbone structures.

References

[1] J.Kang and J.Gwak. Ensemble of Instance Segmentation Models for Polyp Segmentation in Colonoscopy Images. *IEEE Access*, 7:26440–26447, February 2019.

[2] L.Aabakken J.Bergsland Y.Shin, H.A.Qadir and I.Balasingham. Automatic Colon polyp Detection Using Region Based Deep CNN and Post Learning Approaches. *IEEE Access*, 6:40950–40962, August 2018.

[3] Y.Cui Y.Yuan X.Xing H.Seo L.Xing X.Jia, X.Mai and M.H.Meng. Automatic polyp Recognition in Colonoscopy Images Uing Deep Learning and Two-Stage Pyramidal Feature Prediction. *IEEE Transactions on Automatic Science and Engineering*, 17(3):1570–1584, July 2020.

[4] D.Wang Y.Cao X.Sun, P.Zhang and B.Liu. Colorectal Polyp Segmentation by U-net with Dilation Convolution. In *2019 18th IEEE International Conference on Machine Learning and Applications*, pages 851–858, May 2019.

[5] J.Herp A.Tashk and E.Nadimi. Fully Automatic Polyp Detection based on a Novel U-Net architecture and Morphological Post-Process. In *2019 International Conference on Control, Artificial intelligence, Robotics and Optimization*, pages 37–41, May 2019.

[6] Y.Bengio I.Goodfellow and A.Courville. *Deep Learning*. The MIT Press Cambridge,Massachusetts London,England, First edition, 2017.