Artificial intelligence: Future for Polyp Diagnosis!

Xue Zhao, Wenbin Jiang, Tiexin Yu, Jiali Zhou
24520210157080, 24520210157070, 23020211153984, 24520211154666

Abstract

Colorectal polyp is an independent risk factor of colorectal cancer, as well as a common disease in children with serious effects on normal growth and development. Colonoscopy is the gold standard for the detection and diagnosis of colonic polyps, while the diagnosis quality various due to operator's diagnosis experience and their visual fatigue. Therefore, it is important to improve the diagnosis accuracy of colonic polyp from the colonoscopy. An artificial intelligence model which can identify colorectal polyp patients at early stage would be useful for earlier intervention. However, the full connection layer in the existed deep learning models not only brings a large number of parameters, but also easily makes these networks fall into over fitting, resulting in weak network generalization ability. In this study, we collected two children colonic polyp RGB image datasets from public database, CP-CHILD-A dataset and CP-CHILD-B, as training sets and testing sets respectively. By using global average pooling, we proposed the improved approaches: ResNets-GAP, which has much fewer parameters than origin ResNets. The accuracy, sensitivity and specificity of the model in the testing dataset are all above 95%. The experimental results show that the proposed approach has good effect on the automatic detection of colonic polyps.

Introduction

Colorectal cancer (CRC) is one of the most common cancer with high incidence rate and mortality in the world (Miller et al. 2021). Colonoscopy and polypectomy are the standard nursing methods for early diagnosis and prevention of colorectal cancer. Studies shown that they can significantly reduce the incidence rate and mortality of colorectal cancer (Atkin et al. 2017). Colonic polyp is a common disease in children, which seriously affects the normal growth and development of children. Colorectal polyp is an independent risk factor for colorectal cancer, especially the polyps with large diameter, large number and atypical hyperplasia. If colonic polyps cannot be treated early, it is more probably to develop to CRC. Therefore, early detection and treatment of colonic polyps effectively reduce the incidence rate of colon cancer (Johnson and Dachman 2000).

Colonoscopy is the gold standard for the detection and diagnosis of colonic polyps. The results of colonoscopy are easily affected by the operator's experience. Factors such as lack of experience and visual fatigue directly affect the accuracy of diagnosis (Azer 2019). Studies show that the missed detection rate of colonoscopy for colon polyps and colon tumors is as high as 25% (Zhao et al. 2019). Therefore, it is very important to ensure the quality of colonoscopy and improve the accuracy of colonoscopy diagnosis.

Colonoscopy CAD has been a hot spot in Artificial Intelligence research (Badgeley et al. 2019; Lo et al. 2018). In 2003, in the preliminary study of detecting polyps in white light colonoscopy images, wavelet transform was used as an image classifier (Karkanis et al. 2003; Maroulis et al. 2003). Since then, there have been more CAD applications based on specific databases (Wang et al. 2015; Tajbakhsh, Gurudu, and Liang 2015), but the number of images in these databases is limited, mostly less than 20. In 2015, Tyler Berzin and his team first proposed the concept of using artificial intelligence technology to assist endoscopic diagnosis (Mahmud et al. 2015). For a long time, researchers used low-level simple features to detect polyps through machine learning model. The feature acquisition method is usually artificially designed and extracted information such as boundary (shape), texture, intensity, color and spatio-temporal features, which is difficult to ensure the high accuracy of polyp intelligent recognition (Bernal et al. 2017; Zhu, Fan, and Liang 2010).

With the huge improvement of computer technology and data volume, the field of artificial intelligence, including Machine Learning (ML) and Deep Learning (DL), has ushered in unimaginable progress in recent years. The medical service industry is one of the earliest fields to apply artificial intelligence system on a large scale (Shortliffe et al. 1973), which has achieved encouraging and continuous improvement results. Computer Aided Diagnosis (CAD) systems can diagnose various diseases with high accuracy, even surpassing human experts nowadays. The image analysis systems based on artificial intelligence has also been widely used in radiology, pathology, ophthalmology, heart disease, skin disease and other professional fields (Farkhadov, Eliseev, and Petukhova 2020).

Especially for DL algorithms, such as Convolutional Neural Network (CNN) (LeCun et al. 1998), have achieved great success in image recognition, image segmentation (Wang et al. 2019), language understanding and other fields. With the development of deep learning algorithm and the significant improvement of computer computing power, the CAD of colonoscopy is becoming a reality (Mori et al. 2017). Several DL models, especially CNN, have been used for lesion detection of colonoscopy, including polyp classification, detection and segmentation, polyp tracking and improving the accurate diagnosis rate (Azer 2019; Badgeley et al. 2019; Lo et al. 2018).

With the development and optimization of deep learning, the CAD model of gastrointestinal endoscopy based on deep learning has emerged(Mori et al. 2017). Zhu y et al. (Zhu et al. 2019) used CNN to determine the depth of invasion of gastric cancer and confirmed that its recognition ability is better than that of endoscopists. Park et al. (Park and Sargent 2016) used CNN to automatically extract diagnostic features from colonoscopy images. Tajbakhsh et al. (Tajbakhsh, Gurudu, and Liang 2015) proposed a new polyp detection method based on convolutional neural network to make polyp location more accurate. Urban et al. (Urban et al. 2018) applied CNN system to colonoscopy images, and the cross validate accuracy of CNN in identifying polyps was 96.4%.

For image classification, the classical CNN network inputs the feature map generated by the last convolution layer into the fully connected layer, and then connects with the softmax logistic regression layer(Wang et al. 2020a). However, the full connection layer not only brings a large number of parameters, but also easily makes the network fall into over fitting, resulting in weak network generalization ability. Lin et al. (Lin, Chen, and Yan 2013) first proposed the method of global average pooling (GAP). Different from the traditional FC layer, the gap layer applies global average pooling to the whole feature mapping, so that each feature mapping can only get one output, which reduces the number of parameters, simplifies the network, avoids over fitting, and makes the classification task highly interpretable.

Therefore, we propose to optimize the global average pooling (GAP) of the classical CNN models DenseNet and ResNet to obtain DenseNet-GAP and ResNet-GAP. We hope to get a better colon polyp CAD detection model based on colonoscopy by comparing the optimized model with the classical model.

Methods

Different from the traditional detection method, we apply the image classification method of DL network to the colonic polyp detection. Due to the variety of colonic polyps in morphology and the complexity of intestinal environment, it is necessary to choose the CNN model with a high degree of non-linear. The classical ResNets shows excellent recognition ability on the labeled colonic polyp datasets. Based on the aforementioned network, we propose a new network structure by using GAP, which names ResNets-GAP. The improved networks have outstanding classification performance, and also greatly reduce the number of parameters compared with the original networks.

In the task of colonic polyp recognition based on DL network, the medical images are directly fed into the trained network model, and the diagnosis results are given by the system, which can give doctors a clear prompt. Doctors can judge more carefully according to the results. Such assistance can effectively help endoscopists reduce missed diagnosis and improve the accuracy of colonic polyp diagnosis. The processing framework is shown in Fig. 1.

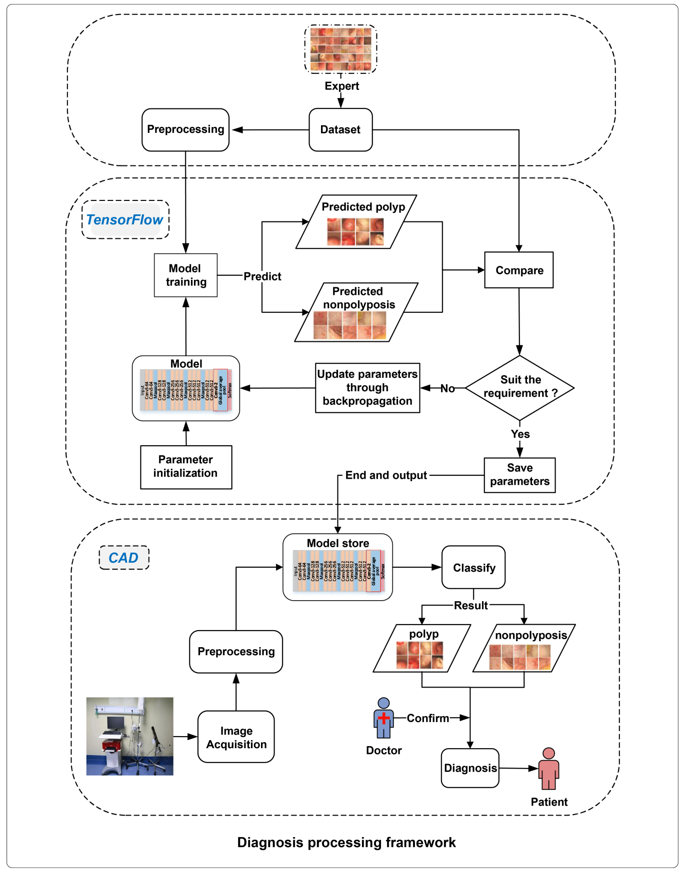
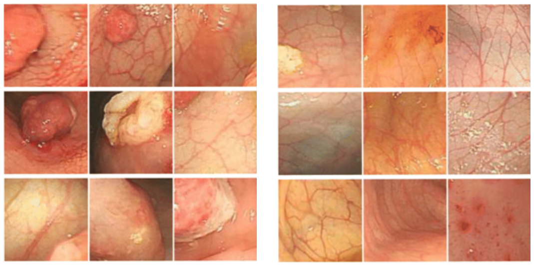


Figure 1: Diagnosis processing framework

Datasets

At present, there are very few public datasets related to colonic polyps and few images in the available datasets of colonic polyps. We collected two colonoscopy RGB images datasets of children from the study by Wang et al. (Wang et al. 2020b), CP-CHILD-A dataset and CP-CHILD-B dataset. As the training sets, CP-CHILD-A dataset contains 800 colonic polyp images and 6200 normal or other pathological images, as shown in Fig. 2(a) and 2(b), respectively. There are 300 colonic polyp images and 300 normal images in the testing sets, CP-CHILD-B, which uses to verify the generalization ability of the model.

Colonic polyp detection is difficult for the following reasons: (1) Other colonic lesions, such as inflammatory bowel disease, ulcerative colitis pictures, etc., may cause bleeding, follicles, etc. their pictures look like polyp pictures. (2)



(a) Colonic polyp (b) Non-polyp

Figure 2: Images in CP-CHILD-A

Many of the polyps in the picture do not appear in the field completely, and some of them even only appear in the corner of the picture. (3) Light and shooting angle also affect the imaging quality.

The large dataset CHILD-A was divided into 90% training set and 10% validation set by random sampling. Then, data augmentation methods, including random horizontal and vertical rotation, brightness change and channel shift, were accomplished with ImageDataGenerator in training dataset, which was expanded into 7543 images (5580 to 6530 for non-polyp and 620 to 1013 for polyp). The details of the datasets are shown in the Table .

Dataset	Training	Validation	Testing
None-polyp	6530	620	300
Polyp	1013	80	300

Table 1: Data distribution in three datasets

ResNets

Deeper networks often have better feature extraction ability and performance. But with the increase of network depth, the degradation problem will appear, i.e., with the increase of training times, the accuracy will reach saturation or even decline. The residual network model proposed by He et al. (He et al. 2016). effectively solves this problem. In ResNets, identity mapping is applied to optimize the residual part in order to better highlight the changes. Compared with the general convolution, the residual block uses skip connection to realize identity mapping and make it bypass the nonlinear transformation. ResNets are stacked by residual blocks, so the network is easier to optimize and its depth can be greatly increased, both of which can improve the recognition accuracy. The residual networks used in our experiments is ResNet50.

Net-GAP

For image classification, the feature maps generated by the last convolutional layer in CNN are usually fed into a fully connected layer, and finally connected with soft-max logistic regression layer (Wang et al. 2020a). However, the fully connected layer not only brings a huge number of parameters, but also makes the network easily fall into over-fitting, which leads to the weak generalization ability of the network. Lin et al. (Lin, Chen, and Yan 2013). proposed the global average pooling (GAP) method for the first time. Different from the traditional FC layer, GAP layer applies global average pooling to the whole feature map, so that each feature map only gets one output. GAP greatly reduces the number of parameters, thus greatly simplifying the network and avoiding over fitting. By using GAP, each feature map has only one output feature. This one-to-one correspondence mode of feature map and category strengthens the relationship between the credibility of feature map and concept (category), and makes the classification task highly understandable. Based on the above idea, we combine ResNet50 with GAP in this paper. By replacing the FC layers with GAP, the new network structures, named Net-GAP, where the "Net" is the original neural networks' model before replacement, is obtained.

Implementation details

Our Computer-aided Diagnosis system (CADs) for intestinal polyp utilized an improved ResNet50 as classification network. The training from scratch ResNet50 framework was constructed using TensorFlow that GAP layer was proposed between the last convolution and softmax layer substituting the average pooling and fully connected layers. The batch size was set up to 16 and the ResNet50 was trained with 30 epochs. The Adam method is adopted to minimize the crossEntropy loss function with the initial learning rate of 0.0005. Moreover, the input size of images was unified to 256×256.

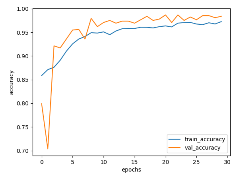
Evaluation criteria

As the evaluation criteria used by most medical image classification models, we use accuracy, sensitivity and specificity as the evaluation criteria. Accuracy refers to the ratio of all correct classification results in the classification model to the total input pictures. Sensitivity, also known as Recall, is the proportion that the model predicts the polyps correctly to all the results of the polyp samples. Specificity is the proportion that the model predicts the non-polyps correctly to all the results of the negative samples.

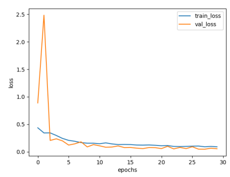
Results

We used the network models introduced in "Method" to conduct experiments on CP-CHILD-A. The results of training were shown in Fig. 3. The CADs achieved great performance with the accuracy and loss of 0.9728 and 0.0890 in training dataset that 0.9843 and 0.0535 in validation dataset, respectively (Fig. 3(a)- 3(b)). It was demonstrated in the confusion matrix that only 6 samples are classified as false negative (FN) and 5 samplesas false positive (FP) (Fig. 3(c)). Furthermore, the other evaluation index including precision, recall and specificity were analyzed in Table .

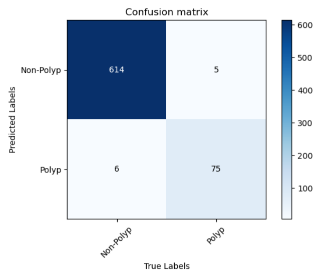
Subsequently, we test our CADs for polyp diagnosis with the CHILD-B dataset containing 300 images for both two categories. As a result, 15 polyp images were wrongly classified as non-polyp and 13 non-polyp images classified as polyp (Fig. 4(a)). The examples of these wrong predicted



(a) Accuracy of Training and Validation



(b) Loss of Training and Validation



(c) Confusion Matrix of Validation

Figure 3: The performance of CADs in training process: (a) The accuracy curve of training and validation dataset. (b) The loss curve of training and validation dataset. (c) Confusion matrix of training process.

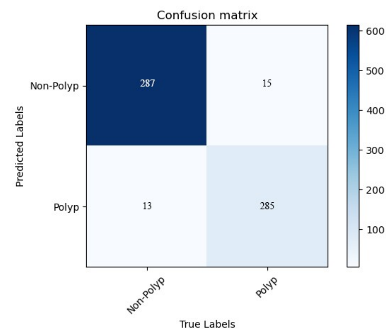
	Precision	Recall	Specificity
None-polyp	0.9920	0.9900	0.9380
Polyp	0.9260	0.9380	0.9900

Table 2: Evaluation index of the training CADs

images were shown (Fig. 4(b)- 4(c)). The reason why the model would wrongly classify we conclude as: (1) The features in high dimension are not extracted completely. (2) The structure of the model is limited. The model we used is ResNet50, the parameters may not enough to extract more features. (3) The images of test set are different from training set in image brightness, color etc., so may confuse the model. Nevertheless, in general, the sensitivity and specificity of the model was 0.9567 and 0.9500, respectively, demonstrating good classification performance and potential for clinical application of our CADs for polyp diagnosis.

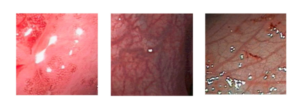
Conclusions

The combination of medical image processing and neural network is a new branch and industry hotspot in the field of digital medicine, and its application in the medical field has received extensive attention. Colonoscopy is easily limited by the operator s experience, and the factors such as inexperience and visual fatigue will directly affect the accuracy of diagnosis. In the experiments, we combine the classical CNN models ResNets with global average pooling, and propose the new network structures ResNets-GAP. The accuracy, sensitivity and specificity of the model in an indenpendent dataset are all above 95%. It demonstrates that the proposed approach has good effect on the automatic detection of colonic polyps.



(a) Confusion Matrix of Test

(b) False Negative Samples



(c) False Positive Samples

Figure 4: The performance of CADs in testing process: (a) Confusion matrix of testing process. (b) Examples of false negative samples. (c) Examples of false positive samples.

References

Atkin, W.; Wooldrage, K.; Brenner, A.; Martin, J.; Shah, U.; Perera, S.; Lucas, F.; Brown, J. P.; Kralj-Hans, I.; Greliak, P.; et al. 2017. Adenoma surveillance and colorectal cancer incidence: a retrospective, multicentre, cohort study. *The Lancet Oncology* 18(6):823–834.

Azer, S. A. 2019. Challenges facing the detection of colonic polyps: what can deep learning do? *Medicina* 55(8):473.

Badgeley, M. A.; Liu, M.; Glicksberg, B. S.; Shervey, M.; Zech, J.; Shameer, K.; Lehar, J.; Oermann, E. K.; McConnell, M. V.; Snyder, T. M.; et al. 2019. Candi: an r package and shiny app for annotating radiographs and evaluating computer-aided diagnosis. *Bioinformatics* 35(9):1610–1612.

Bernal, J.; Tajkbaksh, N.; Sanchez, F. J.; Matuszewski, B. J.; Chen, H.; Yu, L.; Angermann, Q.; Romain, O.; Rustad, B.; Balasingham, I.; et al. 2017. Comparative validation of polyp detection methods in video colonoscopy: results from the miccai 2015 endoscopic vision challenge. *IEEE transactions on medical imaging* 36(6):1231–1249.

Farkhadov, M.; Eliseev, A.; and Petukhova, N. 2020. Explained artificial intelligence helps to integrate artificial and human intelligence into medical diagnostic systems: Analytical review of publications. In *2020 IEEE 14th International Conference on Application of Information and Communication Technologies (AICT)*, 1–4. IEEE.

He, K.; Zhang, X.; Ren, S.; and Sun, J. 2016. Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, 770–778.

Johnson, C. D., and Dachman, A. H. 2000. Ct colonography: The next colon screening examination? *Radiology* 216(2):331–341.

Karkanis, S. A.; Iakovidis, D. K.; Maroulis, D. E.; Karras, D. A.; and Tzivras, M. 2003. Computer-aided tumor detection in endoscopic video using color wavelet features. *IEEE transactions on information technology in biomedicine* 7(3):141–152.

LeCun, Y.; Bottou, L.; Bengio, Y.; and Haffner, P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* 86(11):2278–2324.

Lin, M.; Chen, Q.; and Yan, S. 2013. Network in network. *arXiv preprint arXiv:1312.4400*.

Lo, S. B.; Freedman, M. T.; Gillis, L. B.; White, C. S.; and Mun, S. K. 2018. Journal club: computer-aided detection of lung nodules on ct with a computerized pulmonary vessel suppressed function. *American Journal of Roentgenology* 210(3):480–488.

Mahmud, N.; Cohen, J.; Tsourides, K.; and Berzin, T. M. 2015. Computer vision and augmented reality in gastrointestinal endoscopy. *Gastroenterology report* 3(3):179–184.

Maroulis, D. E.; Iakovidis, D. K.; Karkanis, S. A.; and Karras, D. A. 2003. Cold: a versatile detection system for colorectal lesions in endoscopy video-frames. *Computer Methods and Programs in Biomedicine* 70(2):151–166.

Miller, K. D.; Ortiz, A. P.; Pinheiro, P. S.; Bandi, P.; Minihan, A.; Fuchs, H. E.; Martinez Tyson, D.; Tortolero-Luna, G.; Fedewa, S. A.; Jemal, A. M.; et al. 2021. Cancer statistics for the us hispanic/latino population, 2021. *CA: A Cancer Journal for Clinicians*.

Mori, Y.; Kudo, S.-e.; Berzin, T. M.; Misawa, M.; and Takeda, K. 2017. Computer-aided diagnosis for colonoscopy. *Endoscopy* 49(08):813–819.

Park, S. Y., and Sargent, D. 2016. Colonoscopic polyp detection using convolutional neural networks. In *Medical Imaging 2016: Computer-Aided Diagnosis*, volume 9785, 978528. International Society for Optics and Photonics.

Shortliffe, E. H.; Axline, S. G.; Buchanan, B. G.; Merigan, T. C.; and Cohen, S. N. 1973. An artificial intelligence program to advise physicians regarding antimicrobial therapy. *Computers and Biomedical Research* 6(6):544–560.

Tajbakhsh, N.; Gurudu, S. R.; and Liang, J. 2015. Automated polyp detection in colonoscopy videos using shape and context information. *IEEE transactions on medical imaging* 35(2):630–644.

Urban, G.; Tripathi, P.; Alkayali, T.; Mittal, M.; Jalali, F.; Karnes, W.; and Baldi, P. 2018. Deep learning localizes and identifies polyps in real time with 96% accuracy in screening colonoscopy. *Gastroenterology* 155(4):1069–1078.

Wang, H.; Liang, Z.; Li, L. C.; Han, H.; Song, B.; Pickhardt, P. J.; Barish, M. A.; and Lascarides, C. E. 2015. An adaptive paradigm for computer-aided detection of colonic polyps. *Physics in Medicine & Biology* 60(18):7207.

Wang, W.; Yang, Y.; Wang, X.; Wang, W.; and Li, J. 2019. Development of convolutional neural network and its application in image classification: a survey. *Optical Engineering* 58(4):040901.

Wang, W.; Li, Y.; Zou, T.; Wang, X.; You, J.; and Luo, Y. 2020a. A novel image classification approach via densemobilenet models. *Mobile Information Systems* 2020.

Wang, W.; Tian, J.; Zhang, C.; Luo, Y.; Wang, X.; and Li, J. 2020b. An improved deep learning approach and its applications on colonic polyp images detection. *BMC Medical Imaging* 20(1):1–14.

Zhao, S.; Wang, S.; Pan, P.; Xia, T.; Chang, X.; Yang, X.; Guo, L.; Meng, Q.; Yang, F.; Qian, W.; et al. 2019. Magnitude, risk factors, and factors associated with adenoma miss rate of tandem colonoscopy: a systematic review and meta-analysis. *Gastroenterology* 156(6):1661–1674.

Zhu, Y.; Wang, Q.-C.; Xu, M.-D.; Zhang, Z.; Cheng, J.; Zhong, Y.-S.; Zhang, Y.-Q.; Chen, W.-F.; Yao, L.-Q.; Zhou, P.-H.; et al. 2019. Application of convolutional neural network in the diagnosis of the invasion depth of gastric cancer based on conventional endoscopy. *Gastrointestinal endoscopy* 89(4):806–815.

Zhu, H.; Fan, Y.; and Liang, Z. 2010. Improved curvature estimation for shape analysis in computer-aided detection of colonic polyps. In *International MICCAI Workshop on Computational Challenges and Clinical Opportunities in Virtual Colonoscopy and Abdominal Imaging*, 9–14. Springer.