Using deep learning model based on residual network to detect pathologic type from cervical cancer MRI images

Kai Zhang[1, a], Jun Zhang[1], Xin Cai[2], Yalan Tang[2], Dan Lin[2], Mingxing Chen[2], Lan Li[2], Yajun Chen[1, b, *], Fan Xu[2, 3, 4]

[1]Department of Computer Science and Technology, School of China West Normal University, Nanchong, Sichuan 637000, PR China

[2]Department of Obstetrics and Gynecology, Nanchong Central Hospital, The Second Clinical Medical College, North Sichuan Medical College, Nanchong, Sichuan 637000, PR China

[3]Department of Obstetrics and Gynecology, The Affiliated Hospital of North Sichuan Medical College, Nanchong, Sichuan 637000, PR China

[4]Key Laboratory of Ultrasound Engineering in Medicine Co-Founded by Chongqing, Chongqing Medical University, Chongqing Collaborative Innovation Center for Minimally-invasive and Noninvasive Medicine, Chongqing, 400016, PR China

[a]137794494@qq.com, [b]scnccyj@163.com

*Corresponding author

Keywords: Artificial intelligence, Deep learning, Transfer learning, Computer-aided

Abstract: The detection of rare pathological types plays an important role in the clinical analysis of cervical cancer to ensure efficient treatment and improve recovery rate. Recently, with the rapid development of medical equipment, the diagnoses of pathological types become more accurate. Traditional approach is based on the experience of medical personnel or relevant physicians, some rare types are more likely to be neglected. In this study, we are concerned with the problem of develop a model based on deep learning to assist physicians in detecting pathological types of cervical cancer; Another important point, insufficient training data has always been a common limitation in the field of medical imaging. To solve this challenge, we propose a Resnet-based pre-training model to extract features of pathological images and classification. In particular, the introduction of migration learning not only reduces the scale of training data, but also effectively avoids over-fitting of deep models. We validate our model on a T2 sagittal image dataset of 641 patients with cervical cancer and compare it with the original algorithm. The experimental results show that the proposed model achieves effective performance in terms of cervical cancer pathologic detection.

1. Introduction

Cervical cancer is the fourth most common cancer affecting women worldwide, after breast, colorectal and lung cancer in the world [1]. The most common diagnosis of this disease is in the fifth decade of life, several years earlier than the median age of diagnosis of breast cancer, lung cancer and ovarian cancer, most of which is due to the lack of effective screening system [2]. Deaths from cervical cancer can be prevented with effective screening programmers, which can reduce morbidity and mortality.

The common pathological types of cervical cancer are squamous cell carcinoma (SCC) and adenocarcinoma (AC), accounting for more than 90%, among which adenocarcinoma mainly includes adenocarcinoma and adenosquamous carcinoma (ASC). With the increase of cervical cancer patients and the influence of doctors' own state, in the traditional artificial diagnosis, doctors often get the diagnosis result after screening for AC or SCC, but do not further rule out the possibility of ASC, which eventually leads to the misdiagnosis of AC or SCC in ASC patients.

Therefore, an effective screening method for the pathological classification of cervical cancer can play an important role in clinical treatment.

In this study, we constructed a dataset with clinical labels firstly. The labels are ASC, AC and SCC, which are defined on the clinical features. Secondly, a pre-trained DRN was retrained on the constructed dataset through a process called transfer learning. Thirdly, we have achieved classification and recognition of cervical cancer MRI images, and conducted comparative tests of several methods.

2. Method

In this section, we first introduce the dataset and image labeling. After that, we present a detailed description of the flowchart as shown in Fig. 1, and develop and evaluate our method.

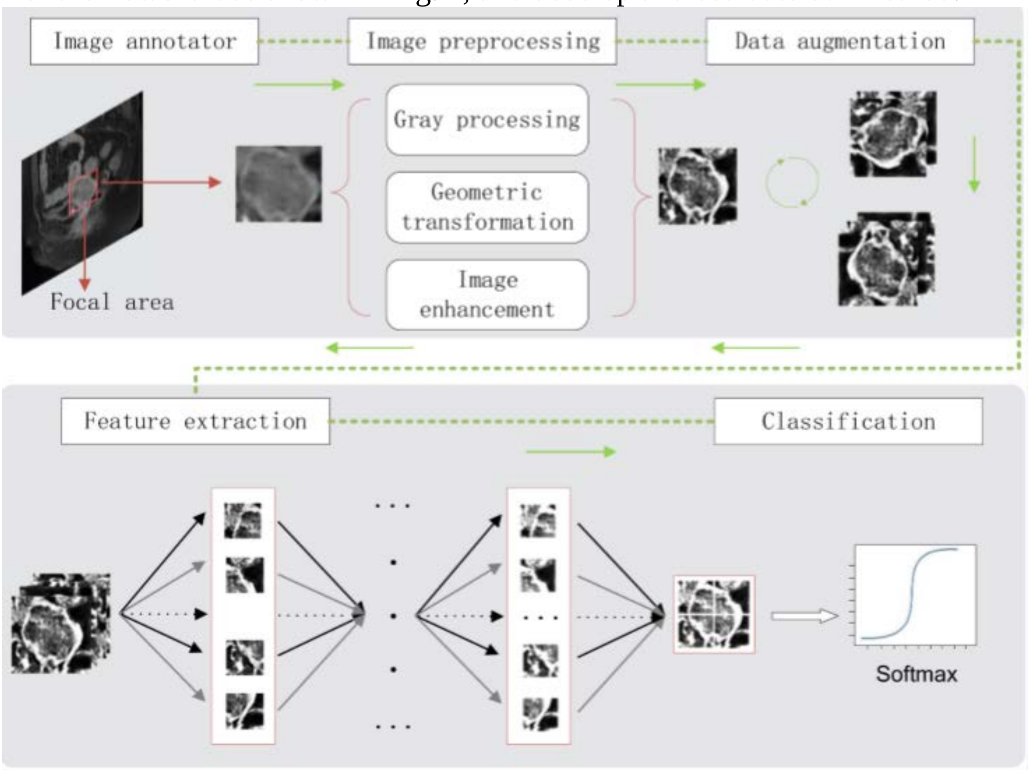

Figure 1. Flowchart of the proposed deep learning model integrating.

2.1 Dataset

We conducted experiments on a clinical dataset which was collected from the department of obstetrics and gynecology of the Nanchong central hospital. The lesion area in the case is marked by an expert with adequate clinical experience. Specifically, we validated our proposed model on the MR image of cervical cancer patients. For solving the problem of data unbalance, we used oversampling technology (SMOTE) [3] to expand the sample data.

Here is how to display a pop-up window from which to select and apply the AIP Proceedings template paragraph styles.

2.2 Data Preparation

All image data is exported from the original Medical Digital Imaging Communications (DICOM) format. This study only includes T2 sagittal sequences. First, the images were cropped so that only the part of the lesion area. Thereafter, all images were marked according to the pathological diagnosis. Lastly, the images were resized to a 100x100-pixel format to match the DRN requirements.

2.3 Method Description

In this study, due to the limited data samples, we used transfer learning to train a pre-trained DRN called ResNet [4]. ResNet is an extremely deep neural network and solves the degradation problem caused by increasing the network depth, so that the network performance can be improved by simply increasing the number of network layers. An in-depth introduction of the structure of ResNet is deviate from the topic of this paper, but a general description of the DRN is useful to its clarity.

3. Experimental Result

3.1 K-fold Cross Validation

K-fold cross-validation is often used for model tuning to find super-parameters that optimize the generalization performance of the model. In our study, we used 10-fold cross validation. The initial sample was divided into 10 sub-samples. 9 samples was randomly taken as a training data set each time, and the other 1 sample were used to test, and the proposed model was trained and tested 10 times.

3.2 Evaluation Metrics

To evaluate the performance of the proposed model, we considered four metrics such as the precision, the sensitivity, the specificity and the area under the curve (AUC). The precision is calculated by (1). In this experiment, the value was 0.923.

$$Precision = \frac{TP}{TP+FP} \tag{1}$$

The sensitivity is calculated by (6). In this experiment, the value was 0.915.

$$\text{Sensitivity} = \frac{TP}{TP+FN} \tag{2}$$

The specificity, also known as the recall, is calculated by (7). In this experiment, the value was 0.926.

$$\text{Specificity} = \frac{TN}{FP+TN} \tag{3}$$

The ROC area is a graphic representation of the classification performance of the model as shown in Fig. 2. In this experiment, the value was 0.93.

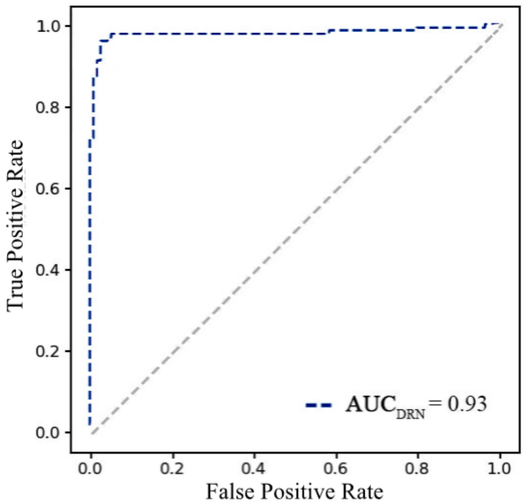

Figure 2. Operating characteristic curve (ROC) and AUC value of the model.

4. Discussion

To provide additional insight into the performance of the proposed model, we use traditional CNN and classic deep neural networks VGG to apply to our datasets. As shown in Table 1, Fig. 3, the deep neural network outperforms traditional CNN methods by a significant margin. The method of proposed in this paper performs relatively better than the method in [5] and achieves significantly better results that the method in [6]. In this task, our model has a precision of 0.923, a sensitivity of 0.915, and a specificity of 0.926. The area under the ROC curve was 93%.

Table 1. Compare with classical methods and other neural networks.

	Precision	**Sensitivity**	**Specificity**
CNN	0.677	0.704	0.685
VGG	0.849	0.825	0.838
Ours	0.923	0.915	0.926

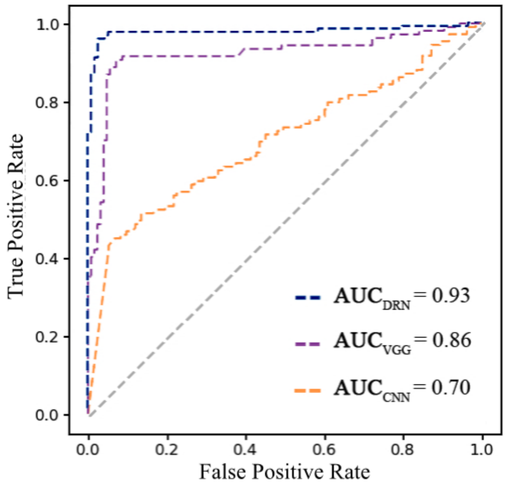

Figure 3. Operating characteristic curve (ROC) and AUC value of the three models.

The dataset for expert comparison were the MRI T2 sagittal images of cervical cancer captured in 2019 in Nanchong Central Hospital to compare the developed DRN's referral decisions with the decisions made by human experts. There are 641 images, and distributions of the SCC, AC, ASC. The confusion tables and error rates of the DRN models and the two experts are presented in Table 2, Table 3, Table 4. The DRN model outperformed the expert 1. Furthermore, the averagetime per case used by the DRN is much less than that of experts.

Table 2. The confusion table of the DRN

		Predicted condition		
		SCC	**AC**	**ASC**
True condition	SCC	467	12	8
	AC	10	96	10
	ASC	3	2	33

Table 3. The confusion table of the expert 1.

		Predicted condition		
		SCC	**AC**	**ASC**
True condition	SCC	477	0	10
	AC	5	101	10
	ASC	8	7	23

Table 4. The confusion table of the expert 2.

		Predicted condition		
		SCC	**AC**	**ASC**
True condition	SCC	480	1	6
	AC	1	108	7
	ASC	3	2	33

In designing our model, we improve the local area display effect caused of MRI images using the feature extraction techniques as proposed by Arturo et al [7]. We show that this is important in achieving a fairly balanced classification. We study augmenting our training dataset by sampling from classes of ASC. We find that data augmentation is also quite effective. The results showed that ResNet displayed a very high accuracy on T2 the imaging sequences in pathological type detection of cervical cancer, suggesting that the use of transfer learning was a suitable solution for the classification problem.

The limitations of our works lie on three aspects. Firstly, the number of cases is severely imbalanced. Secondly, preferably including more patients from various institutions, are needed to ensure the generalization performance in the clinical test. In future work, we plan to collect more actual clinical data in the future to alleviate the data proportion gap, and conduct more detailed processing and analysis of the collected data to make a more rigorous and closer to the standard data set. Thirdly, an unanswered question is whether the long-term clinical test will have other unknown effects on the doctor's diagnosis. More clinical testing is needed in the future development process.

5. Conclusions

In summary, we propose a DRN-based model for classification about MRI images of cervical cancer. The results reported in the present study suggest that DRN model could be a reliable tool for distinguishing cervical cancer from MR images.

Acknowledgments

This work was supported by the Bureau of Science and Technology Nanchong City Program (No: 14A0098, 16YFZJ0029, 18SXHZ0126); the Primary Health Development Research Center of Sichuan Province Program (No: SWFZ17-Y-25, SWFZ17-C-69); the Science and Technology Department of Sichuan province Program (No: 2015JY0056); the Nanchong Federation of Social Science Associations Program (No: NC2016B017).

References

[1] A. Jemal, M. M. Center, et al, *Global patterns of cancer incidence and mortality rates and trends* (Cancer Epidemiology and Prevention Biomarkers, 2010), pp: 1893–1907.

[2] R. Segal, K. Miller, A. Jemal, *Cancer statistics, 2018* (CA: a cancer journal for clinicians, 2018), pp. 7–30.

[3] N. V. Chawla, K. W. Bowyer, et al, *Smote: Synthetic minority over-sampling technique* (Journal of Artificial Intelligence Research, 2002), pp. 321–357.

[4] K. He, X. Zhang, S. Ren, et al, *Identity mappings in deep residual networks* (European conference on computer vision, Springer, 2016), pp. 630–645.

[5] K. Simonyan, A. Zisserman, *Very deep convolutional networks for large-scale image recognition*, arXiv preprint arXiv: 1409.1556.

[6] J. Dolz, K. Gopinath, Y. Jing, et al, *Hyperdense-net: A hyper-densely connected cnn for multi-modal image segmentation* (IEEE Transactions on Medical Imaging, 2018), pp. 99.

[7] A. Aquino, M. E. Gegundezarias, et al, *Detecting the optic disc boundary in digital fundus images using morphological, edge detection, and feature extraction techniques* (IEEE Transactions on Medical Imaging, 2010), pp. 1860–1869.