Detection of Hysteroscopic Hysteromyoma in Real-Time Based on Deep Learning

Aihua Zhao¹,a, Jian Zhang²,b, Shixuan Wang³,c, Yan Wang⁴,d, Xin Zhu⁵,e, Wenfeng Shen⁶,f, Wenwen Wang⁷,g*

¹School of Computer Engineering and Science, Shanghai University, Shanghai, China
²School of Computer Engineering and Science, Shanghai University, Shanghai, China
³Department of Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China
⁴Department of Gynecology, Hubei Maternal and Child Health Hospital, Wuhan, Hubei, China
⁵University of Aizu, Aizu, Fukushima, Japan
⁶School of Computer Engineering and Science, Shanghai University, Shanghai, China
⁷Department of Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology, Wuhan, Hubei, China

aemail:aizhao@shu.edu.cn, bemail:zhangjian@shu.edu.cn,
cemail:shixuanwang@tjh.tjmu.edu.cn, demail:echowy12@163.com,
eemail:zhuxin@u-aizu.ac.jp, demail:wenfengshen@shu.edu.cn
* Corresponding Author: *email: wenwenwang@hust.edu.cn

Abstract: Hysteromyoma is the most common benign tumor in women. By the age of 50, 70% of women have one or more uterine fibroids, and about 30% of them have symptoms and need treatment [1]. In hysteroscopic surgery, doctors' inexperience and fatigue will reduce the accuracy of hysteromyoma diagnosis. In this paper, a hybrid model based on YOLOv3(YOLO) Network and DCGAN network(DCGAN) is proposed to detect hysteromyoma in real time to assist doctors in diagnosis and reduce subjective randomness. The real-time detection speed of the hybrid model reaches 25FPS, and the accuracy rate reaches 91.73%, which meets the requirements of clinical application and improves the diagnosis efficiency of hysteromyoma.

1. Introduction

Hysteromyoma, a benign tumor formed by hyperplasia of uterine smooth muscle, is a high incidence rate among women aged 30~50 years. Hysteromyoma account for 5%~10%[2] of the incidence of uterine diseases. Patients will show increased menstruation, long duration, etc... Hysteromyoma will have adverse effects on women's normal pregnancy[3]. With the development of minimally invasive technology, hysteroscopic hysteromyomectomy is widely used, which has the characteristics of non-invasive, can ensure the integrity of the uterus, will not damage the ovarian function of patients, the operation is favored by patients and their families in clinical[4]. However, hysteroscopic hysteromyoma detection needs to rely on the subjective judgment or pathological examination of doctors. Doctors' lack of experience and fatigue after a large number of operations will greatly reduce...
the accuracy of hysteromyoma diagnosis. Due to the complexity of the uterus and the large individual differences of hysteromyoma, it is difficult for doctors to diagnose, especially for patients with malignant tumor, the degradation of treatment plan caused by missed clinic directly delays the treatment opportunity of patients.

With the continuous development of deep learning, cross research with the medical field is also booming. Computer aided diagnosis system is gradually widely used in the field of uterine medical imaging. In 2010, N.Sriraam et al.\[5\] proposed the method of using wavelet packet transform and back propagation neural network to distinguish normal and hysteromyoma ultrasound images. The collected uterine images were adjusted to 256 × 256 pixels, and tested on 76 images (36 normal uterine images and 36 hysteromyoma images). The classification accuracy was 95.1%. Qiaodan Zhang and others designed DeepLab model and Mask R-CNN instance segmentation model respectively to realize the segmentation of hysteromyoma ultrasound image\[6\]. However in fact, most of the applications in the field of deep learning are based on the segmentation of CT images, MRI images and ultrasound images. In the hysteroscopic image detection of hysteromyoma, the application of deep learning in this field is still blank.

In order to solve the above problems, reduce the workload of doctors and meet the needs of accurate, fast and efficient clinical diagnosis of hysteromyoma, this study applies deep learning to hysteroscopic hysteromyoma detection. At present, the mainstream target detection algorithms, such as SSD network\[7\] and Fast R-CNN\[8\], have good results in object location and classification accuracy, but the detection speed is too slow to be applied to clinic. YOLOv3\[9\] network uses a single convolutional neural network model to achieve end-to-end target detection, with fast algorithm speed, which can realize real-time detection of hysteromyoma and meet the real-time requirements of clinical application. Because the number of clear hysteroscopic images of hysteromyoma in clinic is far less than the data set requirements when training the deep learning network, the capacity of the training set directly determines the performance of the deep learning network, so it is necessary to increase the number of images by means of data enhancement. Compared with traditional data enhancement methods, such as random rotation, random clipping, flipping, etc., deep learning can generate more and high-quality images. Among them, DCGAN\[10\], a deep convolution generation network, introduces convolution operation on the basis of GAN network, which makes the generated images more realistic, especially in the field of data enhancement.

In view of the good real-time detection ability of YOLOv3 network and the ability of DCGAN network to generate high-quality data, this paper proposes to mix DCGAN network and YOLOv3 neural network to form a hybrid model for real-time detection of hysteromyoma during hysteroscopic surgery, so as to assist doctors in diagnosis, reduce doctors' workload and improve the accuracy of hysteromyoma detection.

2. Design of DCGAN and YOLOv3 Hybrid Model

2.1. DCGAN Network

Deep neural network has a huge demand for sample size, but the medical image data of hysteromyoma is scarce, and due to dilation and other operations, the clear hysteroscopic images that can be used as training set are rare, which can not meet the demand of neural network for training data. Generating adversary network is one of the most promising methods of unsupervised learning on complex distribution in recent years. DCGAN network combines convolutional neural network with adversary network to produce high quality output.

Due to the particularity of invalid area in hysteroscopic image, as shown in Fig.1, the size of hysteromyoma image is adjusted to 250px × 250px to remove the black edge, so as to reduce the size of feature image and accelerate the convergence speed of the model. As shown in Fig.2, the discriminator extracts samples from the real hysteromyoma image, the convolution core size is set to 5 × 5, and the convolution step size is set to 2. When the convolution step size is 2, the size of the feature map can be reduced by half, so the dimension can be reduced by stringing the convolution step size of 2 instead of pooling. All layers of the discriminator use Leaky ReLU as the activation function. The generator samples from the prior distribution noise and uses ReLU as the activation function...
except Tanh function in the output layer. Fix the generator and train the discriminator to distinguish the real hysteromyoma image from the generated hysteromyoma image. After the discriminator is updated circularly, the parameters of the generator are updated with Adam algorithm and 0.0002 learning rate to reduce the gap between the generated image and the real image. After 600 epochs iterations and updates, high-quality hysteromyoma images are finally generated and labeled into the training set of YOLOv3 network.

Figure 1. Hysteroscopic image of hysteromyoma

Figure 2. DCGAN network architecture

The DCGAN network is used as the data model and mixed with the YOLOv3 network to solve the problem that the real-time detection effect of YOLOv3 network is not good due to the small amount of available hysteromyoma image data.

2.2. Hybrid Model Design

As the real-time detection speed of YOLOv3 network is fast, it can meet the speed requirements of clinical application. DCGAN network has good data enhancement ability and generating high-quality hysteromyoma image ability. This paper proposes a hybrid model based on YOLOv3 network and DCGAN network to assist hysteromyoma detection under hysteroscopic surgery. The model framework is shown in Fig.3.

Firstly, the discriminator in DCGAN network is used to sample the hysteromyoma image. The discriminator is trained to distinguish the real hysteromyoma image from the image generated by the generator, and update the parameters of the generator. Secondly, in order to improve the accuracy of the model, the number of convolution layers is increased, the first 52 convolution layers of Darknet-53 are used, the full connection layer is removed, and the convolution with step size of 2 is used instead of pooling for down sampling, so that the size of feature map is reduced by half, and the negative gradient effect caused by pooling is reduced. Due to the great individual differences of hysteromyoma, in order to achieve multi-scale detection accuracy, three different scale feature maps are used to detect different sizes of hysteromyoma. Finally, the training set and the uterine fibroids image generated by DCGAN network are sent to darknet-53 for feature extraction. After five times of down sampling and residual structure, DBL feature extraction and convolution layer, the first feature map with the scale of 15px × 15px is output. The 15px × 15px feature map after DBL is up sampled and added with the
result of next down sampling to get $30\times 30$ feature map. The feature map of $50\times 50$ is obtained by adding the result of the up sampling and the third down sampling. Three feature maps of different scales are output, and each scale is responsible for predicting different sizes of uterine fibroids.

After nearly 100000 iterations of training, the model with the minimum average loss was selected as the final model to detect hysteromyoma.

![Figure 3. Hybrid network architecture](image)

### 3. Experimental Design

#### 3.1. Dataset

The data supporting this study are provided by the Department of Gynecology, Tongji Hospital Affiliated to Tongji Medical College, Huazhong University of Science and Technology, and Hubei Maternal and Child Health Hospital. All data used in this study are reviewed by the hospital ethics committee.

During hysteroscopy, for each patient, the doctor will save 5 hysteromyoma images to form a case. The amount of image data is not enough to form a deep learning training set. In order to increase the image data set, 87 hysteroscopic videos of 45 patients collected from 2008 to 2019 were processed by MATLAB program. Based on the principle of ensuring the quality of image data, the videos were divided into TIFF format pictures and screened. Among them, the images that need to be excluded are:

1. The expert group of hysteroscopy cannot see the image of hysteromyoma;
2. The poor quality image with blur or reflection;
3. The hysteromyoma in the image is oversize or undersize.

Dataset A (including 2190 images), Dataset B (including 139 images) and Dataset C (including 8 hysteroscopic videos of 5 patients) are formed. The images in Dataset A are from 73 hysteroscopic videos of 35 patients, all provided by the Department of Gynecology, Tongji Hospital, Tongji Medical College, Huazhong University of Science and Technology. Each image contains at least one hysteromyoma. The images in Dataset B are from 14 videos of 10 patients, all of which are provided by Hubei Maternal and Child Health Hospital. Each image contains at least one hysteromyoma. In addition, 8 hysteroscopic videos of 5 patients are selected from the data provided by the Gynecology Department of Tongji Hospital and Hubei Maternal and Child Health Hospital to form video dataset C for video analysis to verify the real-time detection effect of the deep learning algorithm.

The labeling of image data A and B is completed by professional endoscopists. Labelme is used to label the image. Each hysteromyoma is marked with a rectangle, and the label is defined to form an XML format annotation file.

#### 3.2. Evaluation Criterion

In clinical trials, sensitivity index and specificity index are widely used as test methods. Here, image sensitivity and image specificity are taken as the evaluation criteria of the algorithm.

In the test phase, if the test frame of the mixed model is correctly located on the hysteromyoma, the test result is true positive (TP); no test frame on the hysteromyoma is false negative (FN). Therefore, the sensitivity of hysteromyoma detection is defined as
Sensitivity = \frac{TP}{TP + FN} \times 100\% \quad (1)

If the detection frame of the mixed model does not appear on the image without hysteromyoma, the detection result is counted as a true negative (TN); false positive (FP) is defined as the detection frame on the non hysteromyoma area. Therefore, the specificity of hysteromyoma detection is defined as

\text{Specificity} = \frac{TN}{TN + FP} \times 100\% \quad (2)

3.3. Comparative Experiment Design
Three groups of comparative experiments are designed.
(1) Dataset A is used as training set, and dataset B is used as test set to test the effect of single YOLOv3 network;
(2) The DCGAN network is added to generate data. Dataset A is used as the training set and dataset B is used as the test set to test the effect of the hybrid model;
(3) The dataset C is used as the test set to verify the real-time detection effect of the hybrid model.
Adjust the image of dataset A and dataset B to 250px × 250px, and input them into DCGAN network to generate new image. Output Height is set to 125 and epoch is set to 600.

3.4. Analysis of Experimental Results
The statistical results are shown in Table 1. The image dataset B is used as the test set of YOLOv3 network and hybrid model. Through statistics, the image sensitivity of hybrid model is 93.07\%, and the image specificity is 92.11\%, which is higher than that of single YOLOv3 network; the real time effect of hybrid model is tested with video dataset C, and the sensitivity and specificity are 84.21\% and 88.89\% respectively. In the test, the detection speed of the hybrid model reaches 25fps, which can meet the speed requirements for clinical application.

| Training Set | Model | Test Set | Total of Image | TP | FN | TN | FP | Sensitivity | Specificity |
|--------------|-------|----------|----------------|----|----|----|----|--------------|-------------|
| Dataset A | YOLOv3 | Dataset B | 2329 | 90 | 9 | 36 | 4 | 90.91% | 90.00% |
| \( \text{DCGAN \ generates \ data, \ Dataset \ a} \) | Hybrid model | Dataset B | 3239 | 94 | 7 | 35 | 3 | 93.07% | 92.11% |
| \( \text{DCGAN \ generates \ data, \ Dataset \ a} \) | Hybrid model | Dataset C | 3239 | 16 | 3 | 8 | 1 | 84.21% | 88.89% |

4. Conclusions
The hybrid model of DCGAN network and YOLOv3 network is applied to the clinical detection of hysteromyoma, which has good diagnostic accuracy and can assist doctors in clinical diagnosis.

The processing of dataset is related to the effect of deep learning algorithm. Increasing the number of valid dataset or removing ambiguous image annotation will improve the accuracy of deep learning algorithm. Therefore the next step is to process the reflective image and unclear image in the hysteromyoma image, so as to increase the effective data set and improve the accuracy of the algorithm. In addition, the detection and classification of uterine polyps, endometrial cancer, endometrial hyperplasia and other lesions will be studied.

Acknowledgments
A. The Shanghai Engineering Research Center of Intelligent Computing System (No. 19DZZ2252600) Development Program of China (NO. 2017YFB0701600), the Science and Technology Commission Shanghai Municipality (STCSM) (No. 19511121002)
B. The special fund for innovation and development of AI in Shanghai 2019-RGZN-01080, sponsored by Shanghai Municipal Commission of Economy and informatization

References

[1] Williams ARW. Uterine fibroids - what's new?. F1000Res. 2017;6:2109. Published 2017 Dec 7. doi:10.12688/f1000research.12172.1

[2] Shuang Qiu. Comparison of hysteroscopic and laparoscopic myomectomy in the treatment of type II uterine fibroids [J]. Systems medicine, 2020, 5 (23): 126-128

[3] Youjia Chen. Symptoms and nursing of hysteromyoma [n]. Public health news, 2020-12-09 (027).

[4] Shuping Wang, Qunying Zhou. Analysis of reproductive prognosis of patients with different types of submucosal myoma treated by hysteroscopic resection [J]. Journal of Wannan Medical College, 2019, 38(3): 260-263.

[5] N. Sriraam, D. Nithyashri, L. Vinodashri and P. M. Niranjan, "Detection of uterine fibroids using wavelet packet features with BPNN classifier," 2010 IEEE EMBS Conference on Biomedical Engineering and Sciences (IECBES), Kuala Lumpur, 2010, pp. 406-409.

[6] Qiaodan Zhang. Research on segmentation method of adenomyoma ultrasound image based on deep learning [D]. China University of Geosciences (Beijing), 2020.

[7] Liu W, Anguelov D, Erhan D, et al. SSD: Single Shot MultiBox Detector[C]// European Conference on Computer Vision. Springer, Cham, 2016.

[8] Ren S, He K, Girshick R, et al. Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 39(6).

[9] Redmon J, Divvala S, Girshick R, et al. You only look once: unified, real-time object detection [M]// 2016 IEEE Conference on Computer Vision and Pattern Recognition ( CVPR ) . Las Vegas, NV, USA: IEEE Computer Society, 2016

[10] dford A, Metz L, Chintala S. Unsupervised representation learning with deep convolutional generative adversarial networks [EB/OL]. arXiv, 2018-11-01.