

Improved YOLO v3 Network for Basal Cell Carcinomas and Bowen's Disease Detection

Jiantao Zhang^a, Yun Huang^a, Xiaobo Zhang^a, Yan Xue^b, Xinling Bi^{b,*}, Zhuo Chen^{c,*}

^a*School of Mechatronic Engineering and Automation, Shanghai University, Shanghai, China*

^b*Department of Dermatology, Changhai Hospital, Naval Medical University, Shanghai, China*

^c*Department of Dermatology, Shanghai Children's Medical Center, Shanghai Jiao Tong University School of Medicine, Shanghai, China*

The pathologists' workload is very heavy because they have to identify and diagnose a large number of histological images in daily work. So, there is an urgent need to develop a method for detecting and identifying skin cancer. Basal cell carcinomas (BCC) and Bowen's disease (squamous cell carcinoma in situ) are common malignant skin cancers. In this paper, we focus on the study of YOLO detection model in object detection, and use the YOLOv3 model to detect and identify BCC and Bowen disease. The results show that the model can effectively detect most lesions. Through the structural analysis of the YOLOv3 model algorithm, it is found that the number of convolutional feature maps in some layers' changes drastically. So, a new convolutional layer is added to gradually reduce the number of feature maps, and then an improved YOLOv3 algorithm is proposed. The improved algorithm is used to detect the lesions of the BCC and the Bowen's disease. The results show that the improved model algorithm has better detection performance. Under the condition of threshold value of 0.7, the recognition accuracy of BCC and Bowen's disease is 91.3% and 90.9% respectively.

Keywords: Object detection, YOLO, Digital pathology, Basal cell carcinoma, Bowen's disease

1. INTRODUCTION

Skin is the largest organ of the human body. Because of its direct exposure to air and sunlight, and some factors such as heredity, sebum secretion and lifestyle, the incidence rate of skin cancer is increasing in recent years. And it has become the most common type of cancer [1]. Many skin cancers are developed from early skin diseases. A large part of these skin diseases is due to the lack of early detection and treatment, resulting in the deterioration of cancer. Studies have shown that skin disease is one of the most common human diseases. Skin diseases are often accompanied by pain, itching and appearance damage, which bring inconvenience to patients.

In particular, skin cancer has become an important problem in the field of world health.

According to a statistical study in 2012, there are no less than 5 million cases of non-melanoma skin cancer diagnosed in the United States, including BCC and Bowen's disease [2]. Every year, the number of new cases of skin cancer exceeds that of prostate cancer, lung cancer, breast cancer and colon cancer [3]. The study also shows that one in five Americans has skin cancer in their lifetime.

BCC is the most common non melanoma skin cancer [4]. In China, 65% ~ 75% of skin tumors are BCC [5]. BCC is a malignant tumor originated from basal cells of epidermis or hair follicle epithelium. It often occurs on hairy skin around ears, nose and eyes. It usually presents local invasive growth with little metastasis [6]. The growth rate of cancer cells is slow. After resection, radiotherapy and chemotherapy are not needed. In recent years, the incidence rate of BCC has been increasing year by year

* Author to whom correspondence should be addressed.
E-mails: bixinling@163.com (Xinling Bi),
kishiyo0000@aliyun.com (Zhuo Chen)

[7,8]. Skin exposure to ultraviolet is the main cause of BCC. Other susceptible factors include radiotherapy, arsenic poisoning, contact tar and HPV infection.

Squamous cell carcinoma (SCC) is a kind of malignant tumor, which usually occurs in the epithelium of epidermis and its appendages. In addition, some parts, such as bronchus and bladder, can form SCC through squamous metaplasia, although they are not covered by squamous epithelium [9]. It can destroy the surrounding tissues and even metastasize to lymph nodes, and finally lead to life-threatening. The incidence rate of SCC is next to that of BCC. Bowen's disease is an early manifestation of SCC, also known as cutaneous SCC in situ, and the atypical cells only involve the epidermis. So, the early diagnosis of Bowen's disease is important due to its better prognosis.

Clinical practice shows that if the skin cancer is diagnosed at an early stage, then it can be easily treated. Pathological biopsy is one of the important methods in the diagnosis of skin cancer. Pathological cell images can be observed and comprehensively evaluated on the basis of cytology, but it is a heavy task to identify a large number of pathological images. Traditional diagnostic methods mainly rely on experienced doctors to detect and judge the abnormal cells from many microscopic cell images [10]. The workload of recognizing so many pathological images is heavy, which can easily lead to low accuracy and even misdiagnosis. In addition, the shortage of the medical resources cannot meet the growing medical needs of people, and the medical development of many countries is relatively backward. In these cases, it is necessary to develop an auxiliary tool to help identify skin diseases.

In recent years, with the rapid development of computing science and technology, the object detection technology based on deep convolution neural network has been widely used in various fields. The convolutional neural network can extract high-level features of image through the multi-layer structure. Due to the advantages of weight sharing mechanism, local feature acceptability and spatiotemporal multi sampling, the network model has good distortion invariance.

In the process of computer aided diagnosis for BCC and Bowen's disease, the core index is accuracy and speed. On the basis of accuracy reaching or exceeding human level, the processing speed should not be lower than human level. If the algorithm can complete real-time detection, it can play a better role in the clinical diagnosis scene. The object detection task includes regression task

and classification task. The prediction of object position in the image belongs to regression task and whether the selected region contains the tested object belongs to classification task. Before 2014, object detection in images mainly used traditional artificial design to extract texture and color in order to match objects. These algorithms have poor adaptability and robustness when the scale, angle and illumination of the object change. In 2014, RCNN (Regions with Convolutional Neural Networks) was first used in object detection task, and the good results were achieved. RCNN uses deep network in the task of object detection, and realizes object detection by generating candidate regions [11]. With the excellent performance of convolutional neural network in computer detection task, more and more researchers use convolutional neural network to design detection model. Object detection based on convolutional neural network is divided into RCNN neural network which classifies first and then locates, and YOLO (You Only Look Once) neural network which classifies and locates at the same time [12]. RCNN series deep learning network improves the accuracy and the robustness compared with the traditional hand selected object detection network. But the detection of RCNN series deep learning network is divided into two steps: finding the suggestion box and classifying and locating the suggestion box. So, the detection speed of the network in practical application is slow and the actual detection efficiency is reduced. It cannot meet the requirements of efficient detection. The algorithm of YOLO model can complete the task of location and classification at the same time. Its speed is much faster than RCNN series network.

This paper studies the efficient algorithm of YOLO model. YOLOv3 model was used to detect BCC and Bowen's disease. By analyzing the structure of YOLOv3 model algorithm, it is found that the number of convolution characteristic graphs in some layers change dramatically, which may lead to bad training results. Therefore, an improved YOLOv3 algorithm is proposed. The improved algorithm is used to detect the lesions of BCC and Bowen's disease, and the results are compared and discussed with the previous detection results.

This paper is organized as follows. Section 2 introduces the data used in the experiment. Section 3 presents the model of the YOLO. According to the deficiency of YOLOv3 model, improves the network based on it. In section 4, the YOLOv3 model and the improved YOLOv3 model are used to identify BCC and Bowen's disease. Section 5 compares and summarizes the

experimental results.

2. EXPERIMENTAL DATA

Histopathological features are the most reliable diagnostic criteria for cellular carcinogenesis. So far, there is no large-scale public database of BCC and Bowen's disease markers. The images of BCC and Bowen's disease on dermnet and dermquest skin community websites are insufficient for deep learning training. Therefore, it is necessary to establish the database of BCC and Bowen's disease. In this case, through cooperation with Shanghai Changhai Hospital, 800 histopathological images of

normal skin, BCC and Bowen's disease were collected from Changhai Hospital. All the pathological images were selected by professional dermatologists. The image resolution of the whole original slide is $0.252\text{ }\mu\text{m}$ per pixel. The 1064×728 image was taken at 100 times apparent magnification. In order to ensure that the size of the input vector is fixed, the size of the input image can be adjusted to 416×416 by batch program. Image samples of normal skin, BCC and Bowen's disease are shown in the Fig. 1, the Fig. 2 and the Fig. 3. At the same time, the lesion tissue is marked by a professional dermatologist, and the lesion is labeled by covering the lesion area with a rectangle, as

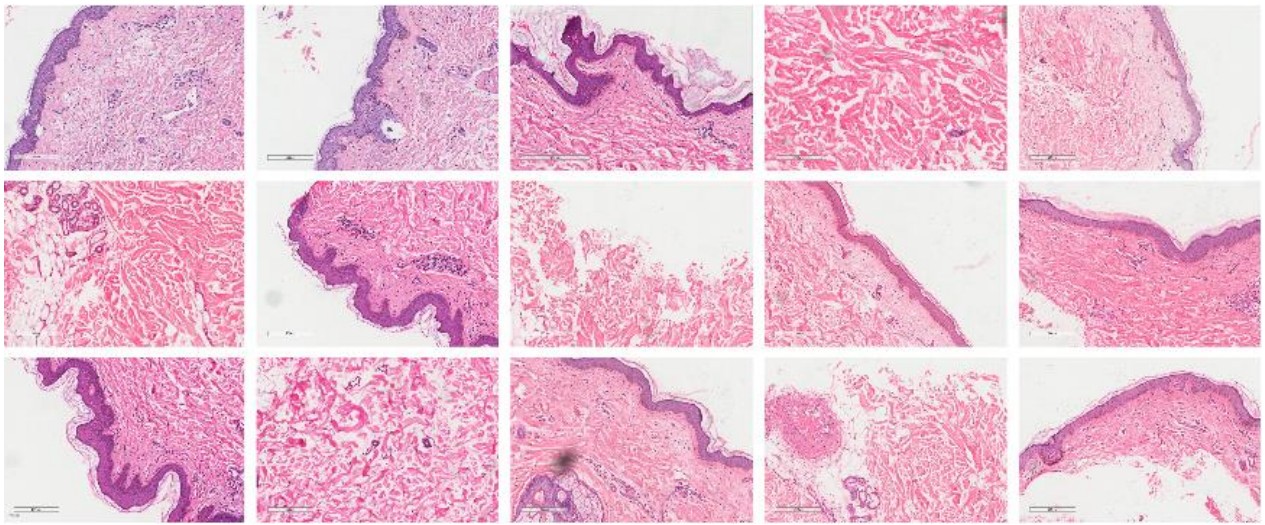


Fig. 1 Histologic image samples of normal skin

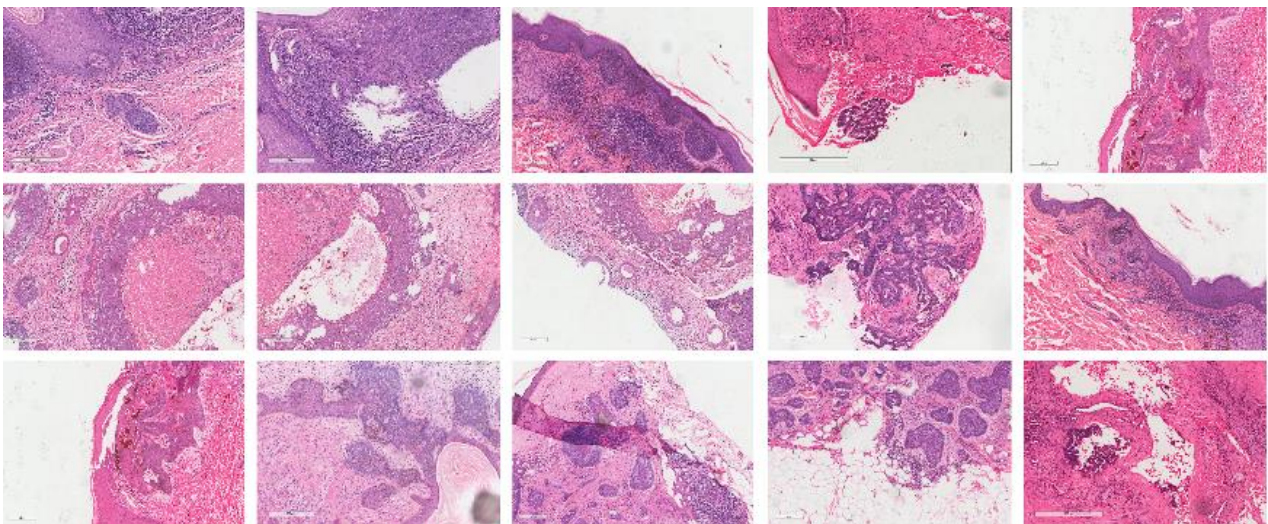


Fig. 2 Histologic image samples of BCC

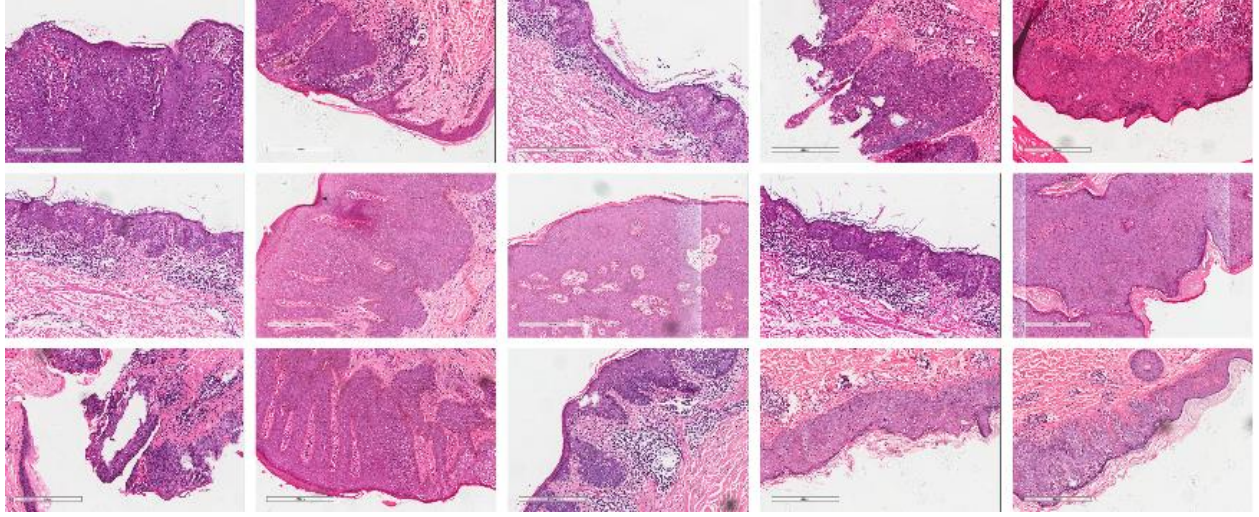


Fig. 3 Histologic image samples of Bowen's disease (SCC in situ)

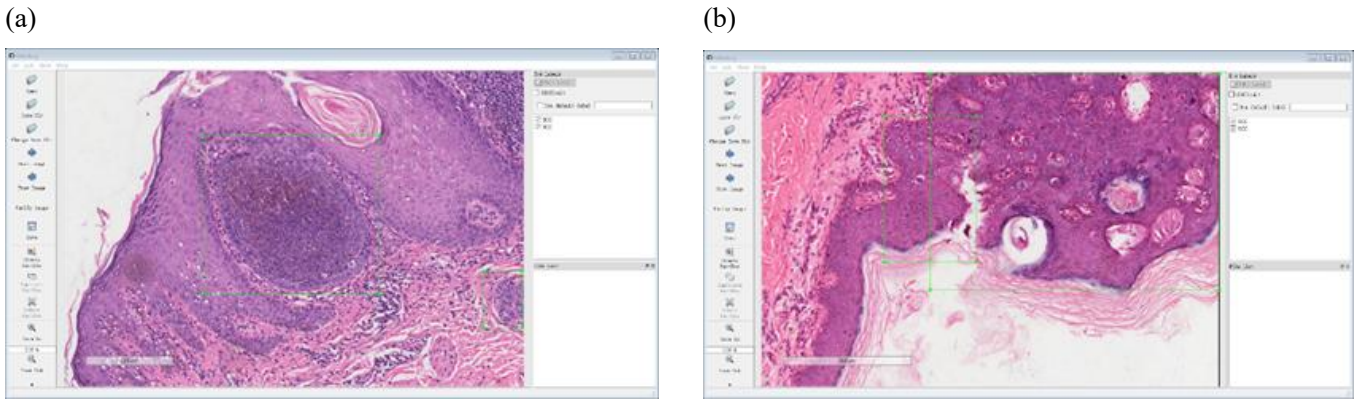


Fig. 4 Labeling of histopathological pictures of skin cancer. (a) labeling of BCC; (b) labeling of Bowen's disease

shown in the Fig. 4 (a) for BCC, and the Fig. 4 (b) for Bowen's disease. The vertex coordinates of the rectangle and the original image data form the image database of the lesion tissue with labels. In the detection task, pathological images are used for training, verification and testing. 600 images of BCC are used for labeling training, and 200 images are used for test. The data distribution of Bowen's disease was the same as that of BCC.

3. METHODS

3.1. YOLOv3

YOLO algorithm mainly transforms object detection, that is, category prediction and object location, into regression task, and uses a deep convolution neural network to simultaneously predict the category and boundary position of the object, so as to realize end-to-end object detection in a real sense. This method completes the two tasks at one time, so the speed is very fast [12]. The end-to-end training mode proposed by YOLO algorithm is different from Fast-RCNN [13] which uses Selective

Search algorithm to generate suggestion box, and Faster-RCNN [14] which uses special RPN network to generate regional suggestion box. YOLO divides the whole image into several regions, and all detection is based on the region, especially the traditional network detection framework divides classification and location into two tasks. For the network model, if you input an image, the YOLO algorithm will record the whole image just as if you only look at the image. This is the origin of the name of the YOLO network.

YOLOv1 uses 24 convolution layers and two fully connected layers. The network uses GoogLeNet's network structure design for reference. YOLOv1 uses 1×1 and 3×3 convolution layers. The smaller receptive field is to avoid too much information loss, reduce the number of parameters, and improve the nonlinear expression ability and generalization ability of the model through multi-layer structure.

Compared with SSD [15] (Single Shot Multi Box Detector) network and Faster-RCNN and other leading networks at that time, YOLOv1 has the problem of

insufficient accuracy, so YOLOv2 improves the positioning accuracy on the premise of ensuring the classification accuracy [16]. Compared with YOLOv1, YOLOv2 draws on the anchor boxes in the Faster-RCNN network.

As shown in Fig. 5, there are 5 anchor boxes generated by the YOLOv2 algorithm. The blue box represents the real box, the yellow boxes represent the anchor boxes, and the red box is the target center position. By using fixed anchor boxes to predict the Bounding box, the number of prediction boxes in each graph is increased from 98 to 845, which improves the accuracy and recall rate. At the same time, each anchor box has an independent classification vector (that is, the length of the vector is 5 and the number of classification categories).

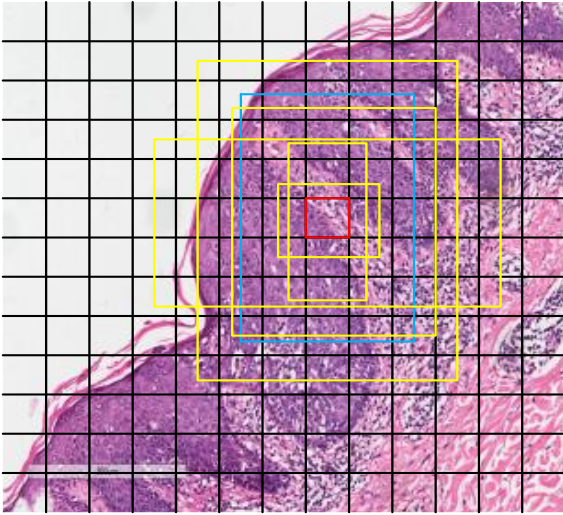


Fig. 5 In the YOLOv2 algorithm, each grid generates 5 anchor boxes

YOLOv3 [17], is equivalent to that of SSD in accuracy, but about three times faster than SSD in speed [18]. This result shows that under the premise of achieving the same performance, YOLOv3 can save nearly three times the time. Compared with YOLOv2, YOLOv3 mainly achieves performance improvement through the following points:

(1) The bounding box of YOLOv3 is better than that of YOLOv2. Using multi-scale prediction, the feature maps of 52×52 , 26×26 and 13×13 are obtained by 8-, 16- and 32-times down sampling on the input image of 416×416 . The width and height parameters of anchor box are still generated by K-means clustering, then nine sizes of anchor box are obtained. The nine sizes of anchor box are assigned to the three scales according to their sizes. The anchor box predicts three rectangles for each scale.

YOLOv3 allocates different sizes of anchors on different feature maps, which reflects the prediction characteristics of network multi granularity feature map and improves the small object prediction performance. The design of this idea also refers to the FPN network (feature pyramid networks for object detection) [19].

(2) YOLOv3 uses cross-entropy loss as the loss of category prediction, and uses logistic regression to predict the score of each category. Compared with YOLOv2, the cross-entropy loss function used by YOLOv3 is more helpful to improve the accuracy of system training.

(3) The first 52 layers of Darknet-53 (without full connection layer) are adopted in YOLOv3, and the structure is shown in Table 1. The network is a full convolution layer network, which uses a large number of residual connections and convolution steps to realize down-sampling operation.

Table 1 Structure of Darknet-53 model

	Type	Filters	Size/Stride	Output
	convolution	32	3×3	256×256
	convolution	64	$3 \times 3/2$	128×128
1×	convolution	32	1×1	
	convolution	64	3×3	
	residual			128×128
	convolution	128	$3 \times 3/2$	64×64
2×	convolution	64	1×1	
	convolution	128	3×3	
	residual			64×64
	convolution	256	$3 \times 3/2$	32×32
8×	convolution	128	1×1	
	convolution	256	3×3	
	residual			32×32
	convolution	512	$3 \times 3/2$	16×16
8×	convolution	256	1×1	
	convolution	512	3×3	
	residual			16×16
	convolution	1024	$3 \times 3/2$	8×8
4×	convolution	512	1×1	
	convolution	1024	3×3	
	residual			8×8
	avgpool		Global	
	connected		1000	
	Softmax			

3.2. The improved YOLOv3

The variant of darknet-53 is used in YOLOv3. In the detection task, the YOLOv3 algorithm provides a 106-layer fully convolutional underlying architecture. It uses three different scales for prediction [17]. The 85th layer and the 97th layer in the network are up-sampled twice (the feature map with higher layer has smaller resolution due to more convolution times). The feature maps of the result after downsampling are combined with the feature maps of lower layers. The core idea of the algorithm is to combine the high-level feature maps with the low-level feature maps, which can preserve the low-level features well and improve the final detection accuracy. The YOLOv3 network uses convolution kernel with the size of 1×1 and 3×3 to extract image features, and uses shortcut connections similar to the residual network. The YOLOv3 algorithm uses 18 feature maps with a size of 13×13 on the 81st layer, 18 feature maps with a size of 26×26 on the 93rd layer, and 18 feature maps with a size of 52×52 on the 105th layer. There are three network layers used for prediction: the 82nd, 94th, and 106th layers. By analyzing the feature extraction layer of the YOLOv3 model, it is found that the number of convolution kernels have intervals such as 32 to 64, 64 to 128 and 128 to 256 before 35 layers. Starting from the 37th layer, due to the reduction of the resolution of the feature map, the number of convolution kernels has increased to 512, with a convolution kernel interval of 256 to 512. In the 38th to 60th layers, 256 convolution kernels, 512 convolution kernels and residuals are regularly used as a unit of repetition. In the 62nd layer, the repetition of 1024 and 512 convolutional layers is used. In the 81st layer, the 1024 13×13 feature maps of 80 layers are convolved into 18 13×13 convolutions directly through 18 1×1 convolution kernels. Feature maps, similarly, in the 93rd layers, 512 26×26 feature maps in 92 layers are convolved into 18 26×26 feature maps through 18 1×1 convolution kernels. Considering that the number of feature maps reduced drastically, especially when the 1024 feature maps are directly convolved into 18 feature maps on the 80th layer, it may lead to poor model training results or limited training effects. It can be found from the Inception model research that the input dimension is usually reduced before large-scale convolution. It is the basic neural network design criteria [20]. Therefore, the network structure of YOLOv3 has been improved, and the improved network

model structure is shown in Table 2. The structure of other unshown layers remains unchanged. The improved YOLOv3 model has the same structure as the original YOLOv3 model in the first 80 layers. The difference is that

Table 2 Structure of improved YOLOv3 model

Layer	Type	Filters	Size/Stride	Output
75	convolution	512	1×1	13×13
76	convolution	1024	$3\times3/1$	13×13
77	convolution	512	1×1	13×13
78	convolution	1024	$3\times3/1$	13×13
79	convolution	512	1×1	13×13
80	convolution	1024	$3\times3/1$	13×13
81	convolution	512	1×1	13×13
82	convolution	256	1×1	13×13
83	convolution	18	1×1	13×13
84	detection			
85	route 79			
86	convolution	256	1×1	13×13
87	upsample	$2\times$		26×26
88	route 87 61			
89	convolution	256	1×1	26×26
90	convolution	512	$3\times3/1$	26×26
91	convolution	256	1×1	26×26
92	convolution	512	$3\times3/1$	26×26
93	convolution	256	1×1	26×26
94	convolution	512	$3\times3/1$	26×26
95	convolution	256	1×1	26×26
96	convolution	18	1×1	26×26
97	detection			
98	route 93			
99	convolution	128	1×1	26×26
100	upsample	$2\times$		52×52
101	route 97 36			
102	convolution	128	1×1	52×52
103	convolution	256	$3\times3/1$	52×52
104	convolution	128	1×1	52×52
105	convolution	256	$3\times3/1$	52×52
106	convolution	128	1×1	52×52
107	convolution	256	$3\times3/1$	52×52
108	convolution	18	1×1	52×52
109	detection			

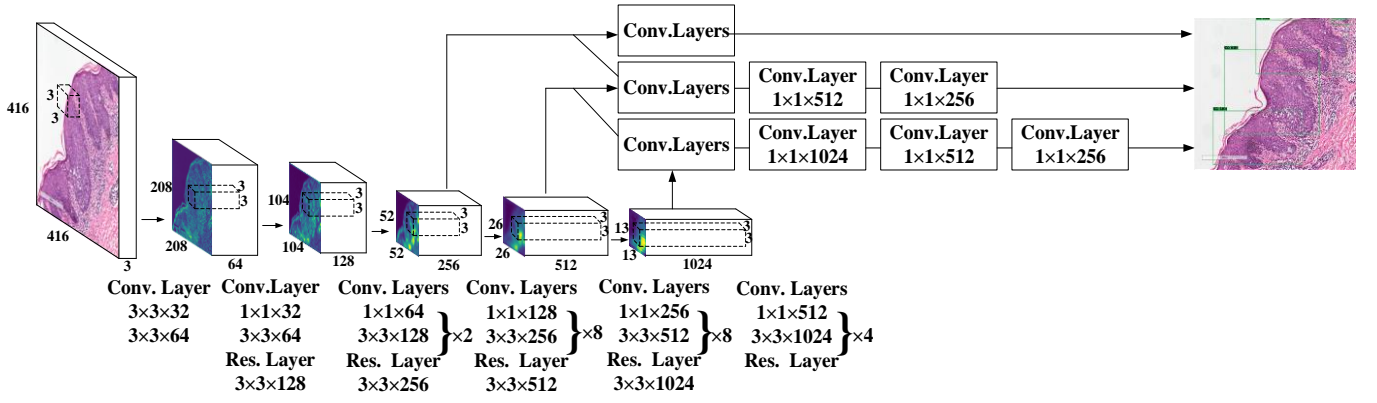


Fig. 6 Schematic diagram of detection using improved YOLOv3 model

the improved YOLOv3 model adds a new convolution layer after 80 layers. The convolution core is 1×1 , and the number of convolution cores is 512 and 256 respectively. Because the number of feature maps is gradually reduced, the number of feature maps is halved in layers 81 and 82. The purpose of using 1×1 convolution kernel is to connect the feature graph and change the mapping mode of the feature graph. Then, 18 feature graphs are generated to make the first prediction through the idea of multi-layer network prediction. In the same way, after 94 layers in the model, a 1×1 convolution kernel is added to connect the previous layer of feature maps. The number of the feature maps is halved by 95 layers, and the final model layer number is 109th layer. In addition, the number of layers of feature map resolution is changed corresponding to the route layer, such as 84 layers, 87 layers, and 97 layers. The Table 2 shows the detail structure of the improved YOLOv3 model. Fig. 6 is a schematic diagram of using the improved YOLOv3 model to detect the pathological picture of Bowen's disease.

4. EXPERIMENTS

4.1. Evaluation Standard

The clinical histopathological images of BCC and Bowen's disease are labeled. The labeling tool is labelImg. YOLOv3 is selected as the deep convolution neural network model, and the framework is Darknet. The initial parameters of the model are pre-trained on the coco (Microsoft common objects in context) data set. For BCC and Bowen's disease, the subsets of training and testing contain 600 and 200 images, respectively. The accuracy

and recall rate are used to measure the accuracy [21]. The calculation formula of accuracy is shown as:

$$Precision = \frac{TP}{TP + FP} \quad (1)$$

where, TP indicates the number of all true lesions in the detected lesions; FP indicates the number of incorrectly detected lesions in the detected lesions. Therefore, the meaning of precision expression is the probability of correct detection in the number of detected lesions.

The calculation formula of recall rate is shown as:

$$Recall = \frac{TP}{TP + FN} \quad (2)$$

where FN represents the number of lesions that have not been correctly detected. The meaning of the recall rate is the ratio of the detected correct lesions to the all actual lesions. That is the probability that the real lesions are correctly detected. In addition, the Precision-Recall curve (hereinafter referred to as PR curve) and the area under the curve (average accuracy, referred to as AP) are used to comprehensively evaluate model performance [21].

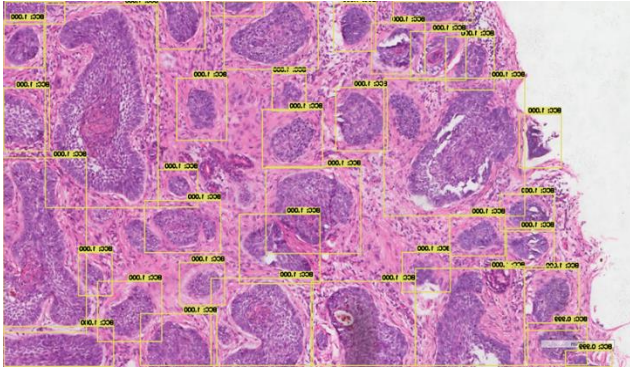
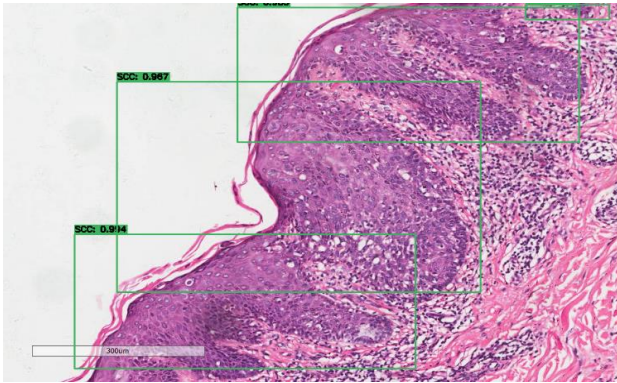
4.2. Experimental Results

Fig. 7 and Fig. 8 show the test results of BCC and Bowen's disease, respectively. It can be seen from the test results that the YOLOv3 model has a good detection effect in the histopathological images of BCC and Bowen's disease. In addition to a small number of missed and false detection of the lesion areas, most of the lesion areas are accurately detected. When the thresholds of the model are selected as 0.4, 0.5, 0.6 and 0.7, the precision and recall values of the model detection are shown in Table 3. The threshold represents the coincidence degree between the

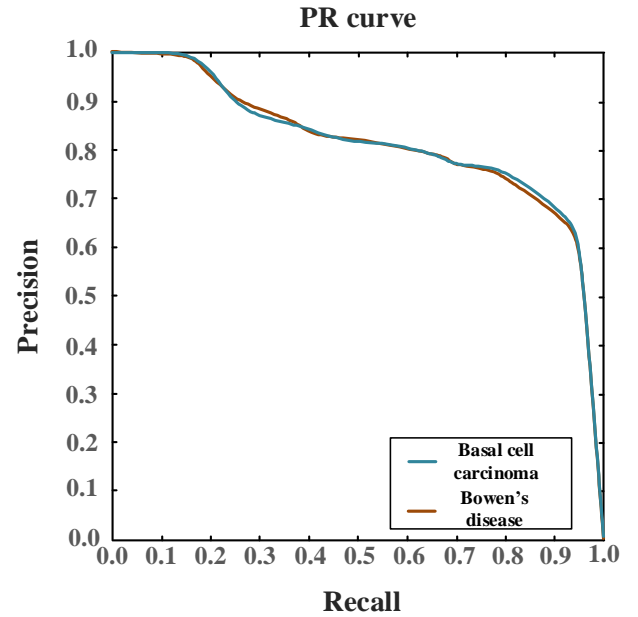
Table 3 Performance of YOLOv3 model in the detection of BCC and Bowen's diseases

		Threshold	Recall	Precision	Threshold	Recall	Precision	
BCC		0.4	0.557	0.810		0.4	0.667	0.786
		0.5	0.469	0.811	Bowen's	0.5	0.529	0.818
		0.6	0.365	0.852	disease	0.6	0.424	0.824
		0.7	0.302	0.864		0.7	0.364	0.857

prediction rectangle and the actual area. The PR curve of the model under the corresponding thresholds of BCC and Bowen's disease is shown in the Fig. 9.

**Fig. 7** Schematic diagram of test results for BCC**Fig. 8** Schematic diagram of test results for Bowen's disease

It can be seen from Table 3 that the recall rate and the accuracy of the model in predicting two kinds of cancer cells are different under different thresholds. The model with lower threshold also has higher recall rate and lower accuracy. It is also because more pathological regions will be identified in the low threshold state, which will improve the recall rate. At the same time, due to the lower threshold, there are many false positives in the identified regions. It reduces the accuracy. In the Fig. 9, both BCC and Bowen's disease have high AP values, which represent the area surrounded by the curve and coordinate axis. Therefore, the

**Fig. 9** PR curve of YOLOv3 model in BCC and Bowen's disease detection experiment

model based on YOLOv3 training has good detection performance in the detection task of BCC and Bowen's disease.

The experimental data processing and the measurement standard of the improved YOLOv3 model are consistent with that of original YOLOv3. When the model threshold is selected as 0.4, 0.5, 0.6, and 0.7, the Precision and Recall values detected by the model are shown in Table 4. The PR curve of the model under the corresponding threshold for BCC and Bowen disease is shown in the Fig. 10.

It can be found from the Table 4 that the improved YOLOv3 model has higher accuracy in the detection of BCC and Bowen's disease under the same threshold. The PR curve of Bowen's disease has higher recognition accuracy under low recall rate. With the increase of recall rate, the recognition accuracy of the model decreases gradually, but still remains at a high level. It can be seen from the Fig. 10 that the improved model has a higher AP value on the PR curve in both BCC and Bowen's disease.

The area enclosed by the curve and coordinate axis is larger. Through the above experimental results, it can be found that the performance of the improved YOLOv3 algorithm is improved when compared with the original YOLOv3 algorithm in the detection task of BCC and Bowen's disease under a given threshold.

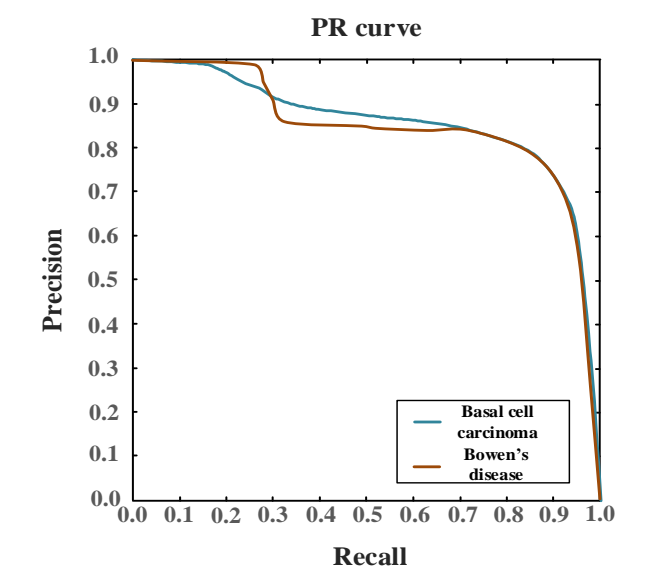


Fig. 10 PR curve of the improved YOLOv3 model in the detection experiment of BCC and Bowen’s disease

4. CONCLUSIONS

Skin disease is very common and thousands of people all over the world suffer from skin diseases every year. BCC and Bowen’s disease are the most common skin tumors. Its new cases exceed those of prostate cancer, lung

cancer, and breast cancer. Tissue biopsy is one of the important methods for the diagnosis of the skin cancer. It conducts comprehensive evaluation and diagnosis based on cytology by observing histological images. It is an arduous task to manually identify many pathological images, and it is easy to have problems with low assessment accuracy and even misdiagnosis. This paper uses the YOLOv3 model algorithm for detecting the lesions in BCC and Bowen’s disease. It is found that most lesion areas can be effectively detected. After in-depth research on the structure of the YOLOv3 model algorithm, it is found that the number of convolutional feature maps changed drastically. It may lead to bad training results. Therefore, the new convolutional layer is added and the number of feature maps is gradually reduced. An improved YOLOv3 algorithm is proposed. And the improved algorithm is used to detect the lesions of BCC and Bowen's disease. The results show that the improved model algorithm has better detection performance. Under the condition of threshold value of 0.7, the recognition accuracy of BCC and Bowen's disease is 91.3% and 90.9% respectively.

In the future research, it is necessary to combine clinical images with skin histological images to improve the learning efficiency of the model. In addition, convolutional neural network has advantages in the detection of skin cancer clinical images and skin pathology images. The development techniques of the object detection algorithm will be applied to detect more and more skin cancer images with higher accuracy.

Table 4 Performance of the improved YOLOv3 model in the detection of BCC and Bowen’s disease							
	Threshold	Recall	Precision		Threshold	Recall	Precision
BCC	0.4	0.635	0.870	Bowen’s disease	0.4	0.636	0.840
	0.5	0.540	0.872		0.5	0.485	0.842
	0.6	0.400	0.893		0.6	0.333	0.846
	0.7	0.328	0.913		0.7	0.303	0.909

Acknowledgments

This research was supported by Shanghai Pujiang Program (No. 18PJD053), Shanghai Natural Science Foundation of China (No. 18ZR1414300), Shanghai Science and Technology Innovation Program (No.

1839190090).

Declarations

Conflict of interest All the Authors of this article have no conflict of interest.

References

- [1] Torre, L. A., Sauer, A. M. G., Chen, Jr. M. S., et al. (2016). Cancer statistics for Asian Americans, Native Hawaiians, and Pacific Islanders, 2016: Concerging incidence in males and females. *CA: a cancer journal for clinicians*, 66(3):182-202.
- [2] Rogers, H. W., Weinstock, M. A., Feldman, S. R., Coldiron, B. M. (2015). Incidence estimate of nonmelanoma skin cancer (keratinocyte carcinomas) in the US population. *JAMA dermatology*, 151(10): 1081-1086.
- [3] Siegel, R. L., Miller, K. D., Fuchs, H., Jemal, A. (2015). Cancer statistics. *CA: a cancer journal for clinicians*, 65(1): 5-29.
- [4] Rubin, A. I., Chen, E. H., Ratner, D. (2015). Basal-cell carcinoma. *New England Journal of Medicine*, 353(21): 2262-2269.
- [5] Muzic, J. G., Schmitt, A. R., Wright, A. C., et al. (2017). Incidence and trends of basal cell carcinoma and cutaneous squamous cell carcinoma: a population-based study in Olmsted County, Minnesota, 2000 to 2010. *Mayo Clinic Proceedings*, 890-898.
- [6] Friedrich, R. E., Giese, M., Li, L., et al. (2015). Diagnosis, treatment and follow-up control in 124 patients with basal cell carcinoma of the maxillofacial region treated from 1992 to 1997. *Anticancer research*, 25(3A): 1693-1697.
- [7] Peris, K., Fargnoli, M. C., Garbe, C., Kaufmann, R., Bastholt, L., et al. (2019). Diagnosis and treatment of basal cell carcinoma: European consensus-based interdisciplinary guidelines. *European journal of cancer*, 118: 10-34.
- [8] Hrebikova, H., Beznoska, P., Bezrouk, A., Chvatalova, J., et al. (2001). Decellularized Skeletal Muscle: A Promising Biologic Scaffold for Tissue Engineering. *Journal of Biomaterials and Tissue Engineering*, 7(6): 491-498.
- [9] Alam, M., Ratner, D. (2001). Cutaneous squamous-cell carcinoma. *New England Journal of Medicine*, 344(13): 975-983.
- [10] Rastrelli, M., Tropea, S., Rossi, C. R., Alaibac, M. (2014). Melanoma: epidemiology, risk factors, pathogenesis, diagnosis and classification. *In vivo*, 28(6): 1005-1011.
- [11] Girshick, R., Donahue, J., Darrell, T., et al. (2014). Rich feature hierarchies for accurate object detection and semantic segmentation. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 580-587.
- [12] Redmon, J., Divvala, S., Girshick, R., et al. (2016). You only look once: Unified, real-time object detection. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 779-788.
- [13] Girshick, R. (2015). Fast R-CNN. *Proceedings of the IEEE international conference on computer vision*, 1440-1448.
- [14] Ren, S., He, K., Girshick, R., et al. (2017). Faster R-CNN: Towards real-time object detection with region proposal networks. *IEEE Trans. Pattern Anal. Machine Intelligence*, 39: 1137-1149.
- [15] Liu, W., Anguelov, D., Erhan, D., et al. (2016). SSD: Single shot multibox detector. *European conference on computer vision*, 21-37.
- [16] Redmon, J., Farhadi, A. (2017). YOLO9000: better, faster, stronger. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 7263-7271.
- [17] Redmon, J., Farhadi, A. (2018). YOLOv3: An incremental improvement. *arXiv preprint arXiv:1804.02767*.
- [18] Wu, H., Gu, X. (2015). Towards dropout training for convolutional neural networks. *Neural Networks*, 71:1-10.
- [19] Lin, T. Y., Dollár, P., Girshick, R., et al. (2017). Feature pyramid networks for object detection. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2117-2125.
- [20] Christian, S., Vincent, V., Sergey, I., Jonathon, S., Zbigniew, W. (2016). Rethinking the inception architecture for computer vision. *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2818-2826.
- [21] Flach, P., Kull, M. (2015). Precision-recall-gain curves: PR analysis done right. *Advances in neural information processing systems*, 838-846.