Multi-scale deep neural network for mitosis detection in breast cancer histological images

Boqian Wu[1], Binwen Fan[1*], Qiao Xiao[1], Tasleem Kausar[1], Wenfeng Wang[2*]

[1]Department of electronics and communication engineering,
Harbin Institute of Technology, Shenzhen,518000,China
[2]State Key Laboratory of Desert and Oasis Ecology, Xinjiang Institute of Ecology and Geography, Chinese Academy of Sciences, Urumqi, 830011, China

Abstract. Accurate assessment of the breast cancer deterioration degree plays a crucial role in making medical plan, and the important basis for degree assessment is the number of mitoses in a given area of the pathological image. We utilized deep multi-scale fused fully convolutional neural network (MFF-CNN) combing with conditional random felid (CRF) to detect mitoses in hematoxylin and eosin stained histology image. Analyze the characteristics of mitotic detection ----scale invariance and sparsity, as well as the difficulties ---- small amount of data , inconsistent image staining and sample class unbalanced. Based on this, mitotic detection model is designed. In this paper, a tissue-based staining equalization method is used, and to establish an effective training sample set, we select training samples by using CNN. A mitotic detection model fusing multi-level and multi-scale features and context information was designed, and the corresponding training strategy was made to reduce over-fitting. As preliminarily validated on the public 2014 ICPR MITOSIS data, our method achieves a better performance in term of detection accuracy than ever recorded for this dataset.

Keywords: Breast cancer, Mitosis detection, CNN, Stain-normalization, CRF, Multi-scale feature

1   Introduction

Mitosis counting refers to the number of dividing cell identified in a fixed number of high power fields (40x magnification in this experiment), which is laborious, subjective [1]. In hematoxylin and eosin (H&E) stained breast cancer sections, mitoses are discernible as hyperchromatic objects with dark color that lack clear nuclear membranes and have irregularity shape properties. In fact, mitosis is a complex process during which a cell nucleus undergoes four phase and exhibits highly variable appearance, moreover in most stages a mitotic nucleus looks like a non-mitotic nucleus shown in Fig.1 [2]. Therefore, the identification of mitosis may often suffer from disagreement between inter-observers. The computer-aid mitosis detector for breast cancer becomes a promising solution for these issues. Because it is a hard task to extract high-efficiency features manually, the performance of early studies based on

handcraft features were not so impressive [3-5]. In recent years, deep convolution neural networks (CNN) exhibits outstanding performance in classification [6][7][8], which can learn high-level feature representation from the raw dataset. Nowadays, more and more researchers begin to apply CNN model to mitosis detection.
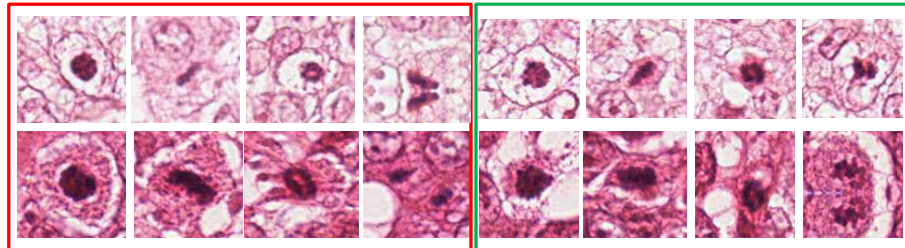

Fig. 1. The example of mitoses and non-mitoses (each column in red rectangle represents prophase, metaphase, anaphase and telophase of mitosis respectively, and non-mitoses with similarity appearance as mitosis were shown in the green rectangle)

The latest studies show that deep CNN with hierachical feature representation has made breakthroughs in mitosis detection [9], Table.1 lists the papers using deep learning in mitosis detection.

Haibo Wang et al. proposed an cascaded strategy combining the handcraft features and the features learned from deep CNN model together to generate more comprehensive features for mitosis detection [10]. This method was evaluated on the public ICPR12 mitosis dataset and yielded an F-measure of 0.7345 being the second best performance ever recorded in this dataset.

Ciresan et al. applied a DNN model to classify each pixel of the H&E stained mitosis images, consequently won the ICPR 2012 mitosis detection competition [11]. They proposed an approach to select relatively rare challenging non-mitosis samples based on the output of DNN model, which allowing the model to learn the significant differences between mitotic and non-mitotic nuclei efficiently. However, this training dataset includes 95.4% non-mitosis class samples and only 6.6% mitosis samples, so the detector tend to classify the mitosis as non-mitosis. In addition, the pixel-wise classifier is time-consuming, costing roughly 8 minutes per image.

Hao Chen et al. designed a faster mitosis detector by leveraging the FCN model, moreover, they proposed an innovative cascaded CNN model to detect mitosis [12]. The detection accuracy outperformed other methods by a large margin in 2014 ICPR MITOS-ATYPIA challenge.

However, above work didn't discuss the impact of the different level features from CNN on final accuracy.

Table 1. The overview of papers using deep learning in mitosis detection

reference	method
Cires ̧an et al. (2013)	Pixel-by-pixel classification based on CNN
Malon and Cosatto (2013)	Combines shape features with CNN
Wang et al. (2014)	Combines handcrafted features with CNN
Albarqouni et al. (2016)	Designs 'crowd sourcing' layer into CNN framework
Chen et al. (2016b)	Deep regression network
Janowczyk et al. (2016a)	Classifies tissue by using SAE and histogram matching based on tissue

The complete breast cancer metastasis probability assessment process is shown in Fig.2, including the interest region extraction steps, the staining normalization procedure, and the mitotic detection procedure. The area of interest has been given by medical professionals, and the goal of this paper is to design computer-aided system models for mitotic detection of breast cancer.
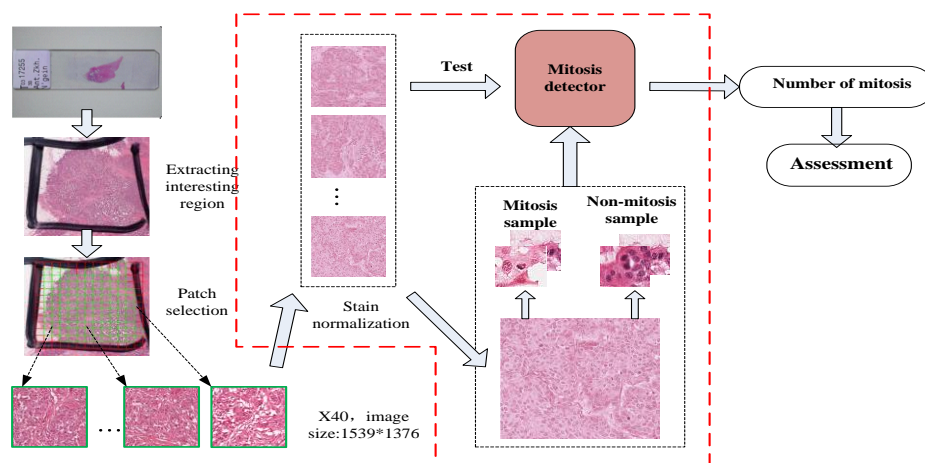
Fig.2. The process of breast cancer metastasis probability assessment process

We carry out the following research contents: choosing suitable image preprocessing method to solve the problem of stain inconsistency and establishing a high quality training sample set; designing the structure of mitosis detection model. We proposed an end-to-end MFF-CNN+CRF model to fuse the multiple level and multiple scale features and the contextual information. The establishments of the sub-paths would increase the number of parameters, therefore, the multi-step training strategy is proposed to reduce over-fitting. In stain normalization step, we used stacking automatic encoder to transform the image feature space and separating various tissue by the method of K-clustering and finally stain normalizing based on tissue. By using convolutional neural network as a sample filter, a representative high-quality training dataset is established, compared with the method of random sampling, the training effi-

ciency and detection accuracy of the model are improved. In summary, the main jobs of this study include:

- Proposing an end-to-end MFF-CNN+CRF model to fuse the multiple level and multiple scale features and the contextual information to improve the detection accuracy
- Completing the training of the proposed MFF-CNN+CRF model by a multi-step fine-tuning strategy.
- Normalizing the stain color based on the separate tissue
- leveraging convolutional neural network and the cascaded selection strategy to select representative non-mitosis training samples to build an effective dataset

Among above jobs, both the MFF-CNN+CRF model with corresponding training strategy are first time applied in the field of mitosis detection. The rest of this paper organized as follow. The section 2 will explain the new methodology in detailed. The section 3 will describe the experiment and exhibit the results. The section 4 will present the conclusion for this work.

2   METHODOLOGY

2.1   Stain-Normalization and training dataset building

The inconsistency of stain condition makes the appearances of H&E stained histology drastically different, so the classify performance was degraded [13][14]. For instance, many false mitosis may arise when the histopathology slide is over-stained. This experiment performed staining unmixing (separation of the hematoxylin and eosin stains) and normalized each hematoxylin and eosin image separately [16]. This method was based on what was described in [17]. First of all, we use the stacked auto-encoder to transfer the feature space of image, and then use K-clustering method to represent all kinds of tissues and separate the different kinds of tissues. Finally, we stain normalized each type of tissue separately. The whole process is shown in Fig.3. It was utilized not only to preprocess the training and testing image suffering from a poor staining.
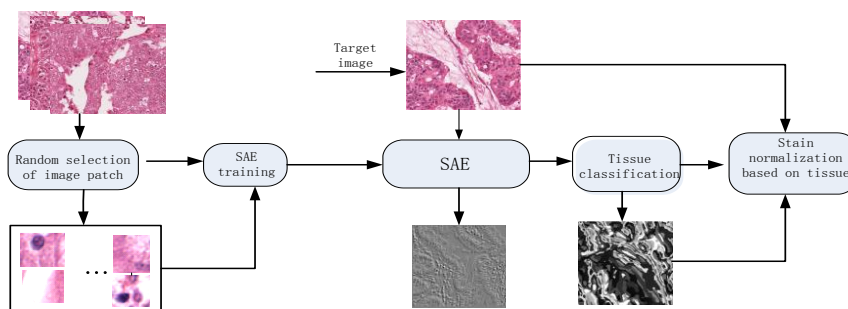

Fig.3. The process of stain normalization based on tissue

To build an effective dataset , we used g the CNN model as samples selector by leve-
raging the cascaded selection strategy., we extracted the non-mitosis patches by leve-
raging the output probability images of the previous stage. The values of probability
images represent the likelihood of mitosis and were used to the select the challenging
samples. And then, we augmented the mitosis training samples by rotations of
{0;45;90;135;180;215;270} degrees and translation.

2.2   Multi-scale Fused Fully Convolutional Network

The features from lower layers of CNN model responds to the general attributes such
as edge, color, texture and so on, while the higher layer features are more class-
specific and abstract [18][19][20]. Inspiring by the previous work which combined
the handcrafted features and the features learned from deep CNN [15], this study
proposed a sampler and efficient approach in which linked the output of conv3 and
conv4 layer to fc6 to combine the extracted different level features. The structure of
model is shown as Fig.4. To reduce over-fitting, we utilized the off-the-shelf CaffeNet
model for initialization and the training process of the whole FF-CNN model includ-
ing three steps, we trained the plain model without the two sub-paths firstly, later the
sub-path between conv4 and fc6_2 and the sub-path between conv3 and fc6_3 were
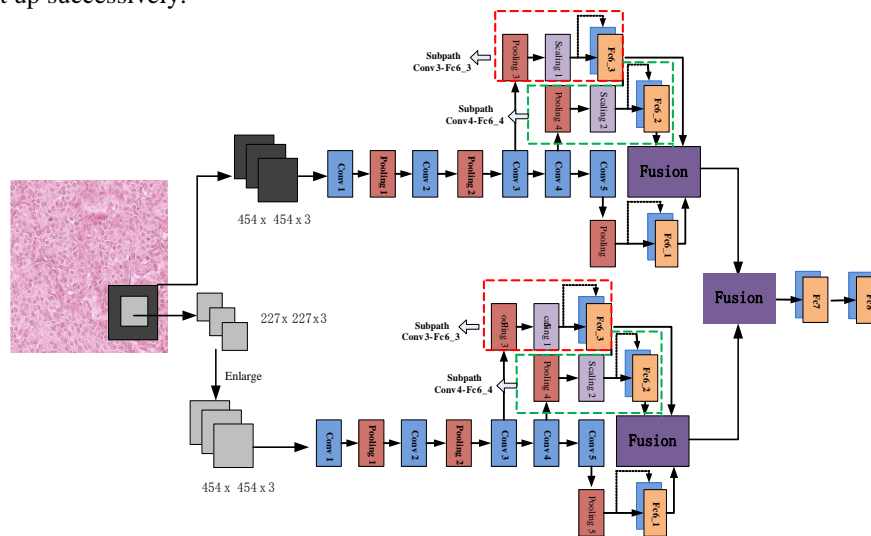set up successively.

Fig. 4**.** An overview of proposed Multi-scale FF-CNN model combining different level and
scale features together

The multi-scale FF-CNN model has two branches with different scale features and the
final class prediction for each pixel is the joint prediction from all two branches. More
contextual information could be captured by the larger scale branch, which helps
detect more real mitotic samples. Combining predictions achieve better mitotic detec-
tion. As we know, it will be more difficult to train a multi-scale FCN model, so we

modify the original loss function to make multi-scale learning features more recognizable, which can speed up training and improve performance. The loss function is as follows:

$$loss(\theta, D) = -\sum_k w_k \log p(Y = y^k \mid x^k, \theta) + \sum_{i=1}^{M} -\alpha_i \times (\sum_j w_j \log p(Y = y_i^j \mid x_i^j, \theta)) \quad (1)$$

Where M represents the number of the scale , here M=2; α_i represents the weight of the ith scale., x^j, y^j represents the value of the output and the real label respectively In order to focus on the discrimination characteristics of each scale, we will initialize larger α_i. Then we gradually lower α_i and focus on the joint training.

Table 2. The architecture parameter of final mitosis detector based on AlexNet FCN model

Layer	Feature maps	Filter size	Stride	padding	Learning rate
input	$1766 \times 1603 \times 3$	–	–	227×227	–
C1	$439 \times 399 \times 96$	11	4	–	10^{-5}
P1	$219 \times 199 \times 96$	3	2	–	–
C2	$219 \times 199 \times 256$	5	1	2×2	10^{-5}
P2	$109 \times 99 \times 256$	3	2	–	–
C3	$109 \times 99 \times 384$	3	1	1×1	10^{-3}
P3	$54 \times 49 \times 384$	3	2	–	–
C4	$109 \times 99 \times 384$	3	1	1×1	10^{-3}
P4	$54 \times 49 \times 384$	3	2	–	–
C5	$109 \times 99 \times 256$	3	1	1×1	10^{-3}
P5	$54 \times 49 \times 256$	3	2	–	–
C6_1	$49 \times 44 \times 4096$	6	1	–	10^{-3}
C6_2	$49 \times 44 \times 4096$	6	1	–	10^{-3}
C6_3	$49 \times 44 \times 4096$	6	1	–	10^{-3}
C7	$49 \times 44 \times 4096$	1	1	–	10^{-3}
C8	$49 \times 44 \times 2$	1	1	–	10^{-3}

Up-sample. Since the convolutional layer and pooling layers would reduce the data size., the size of output image is 49×44 not 1539×1376 as the original images. The ratio from the larger original image to the smaller output image is defined as "compress-ratio", in this case, $1539/49 \approx 32$ and $1376/44 \approx 32$ are the compress-ratio of the width and height respectively. Actually, the "compress-ratio" is roughly equal to the stride S of whole model . According to our compress-ratio, the value of displacement should be chosen as 32. In this up-sample approach, there will be $32 \times 32 = 1024$ operations to be implemented per image. The computation is redundant and time-consuming, finally, we set the sliding stride as 16 ,that is sliding number reduced to 4, ensuring a comparative accuracy at the same time. The method is explained in paper [21].

2.3   FF-CNN+CRF model

Conditional random field model (CRF) can predict pixel-level labels according to the context of the image and globally observe the image. In this paper, the CRF model connected to the FF-CNN model. The FF-CNN+CRF-RNN model is shown as Figure 5. The FF-CNN predicts the class probability to each pixel, and the CRF-RNN uses the predicted information as its input to globally optimize the segmentation results and they based on the intensity and location information of each pixel. Firstly, the whole slice image is taken as the input of FFCNN, and two probability images p^u , here $u = 1, 2$, $p^u = \left\{ p_{i,j}^u \mid i \in [1, 2, ..., w], j \in [1, 2, .., h] \right\}$ the images are the same size as the original image slice. $p_{i,j}^u$ represents the probability of a pixel predicting as mitosis.

Then, these tag prediction images are used as inputs to the CRF-RNN. Finally, the CRF-RNN obtains the globally optimized segmentation result of the original H&E image.
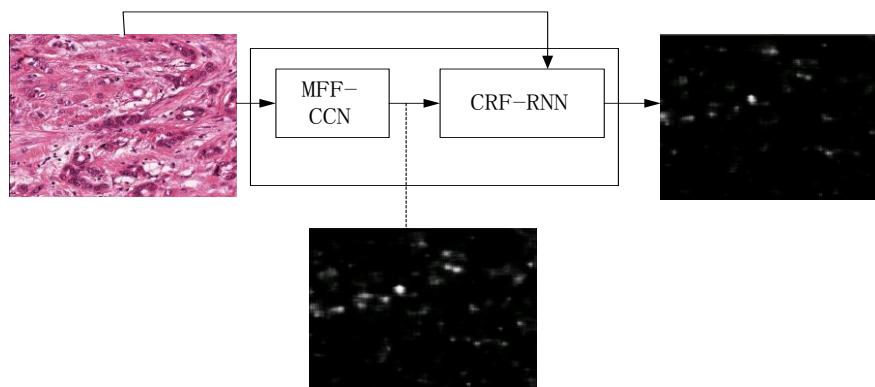

Fig. 5. The combination of MFF-CNN and CRF_RNN

The proposed FCNN+ CRF-RNN model is divided into three steps for training: (1) Training of MFF-CNN using image patch; (2) Training of CRF-RNN using image patch and fixing the parameters of MFF-CNN; and (3) Fine-tune the entire network with the whole image. In training steps 2 and 3, we first calculate the frequency weighted loss based on the current segmentation result and the real label, then the loss information is backpropagated to adjust the parameters of the FCNN and CRF-RNN synthesis networks. In training step 2, we fixed FCNN and adjusted the parameters in CRF-RNN. In training step 3, we set a small learning rate to fine-tune the parameters of the entire network. In our experiments, the initial learning rate was set to be, after training step 1, the learning rate was divided by 10 after every 20 epoches. In training steps 2 and 3, the learning rates are set to 10^{-8} and 10^{-10} respectively.

3   Experiment and result

3.1   Dataset

In this work, the performance of model was evaluated on 2014 ICPR MITOSIS data-set including 1,200 training images and 496 testing images. Due to different condition during the tissue acquisition process, the appearance of tissues are various, which makes the detection task more challenging. The spatial resolution of images acquired by the widely-used Apero-XT scanner is 0.25 $\mu m / pixel$ and the magnification is 40.

The stain-normalization method was leveraged to preprocess the images for the slice of A03, A07, A11, A14 , A15, A17 and A18, which made the edges of mitosis clearer and color consistent shown as Fig. 6, and to augment the training data for the slice of A04 and A05.

3.2   Evaluation and Performance

Evaluation was performed according to the ICPR 2014 content criteria, if the detected mitoses whose coordinates are closer than 30 pixels ($8 \mu m$) to true mitoses are defined as true positives (TP). Those detection locating without 32 pixels of true mitosis are counted as false positive (FP). The true mitoses which were not detected by the model are defined as false negative (FN). We compute the performance measure including precision: $P = \dfrac{N_{TP}}{N_{TP} + N_{FP}}$, recall: $R = \dfrac{N_{TP}}{N_{TP} + N_{FN}}$ and F-measure: $F = \dfrac{2 \times P \times R}{P + R}$,

Where N_{TP}, N_{FP}, N_{FN} refers to the number of true positive, false positive and false negative, respectively. We evaluated the performance of FF-CNN model and the one-way plain FCN model, FCN model, MFF-CNN model and MFF-CNN+CRF model and the effectiveness of frequency weighted loss function [20].

Table 3. Result of 2014 ICPR MITOSIS Datast

Method	TP	TN	FP	Precision	Recall	F-score
STRASBOURG	/	/	/	/	/	0.024
YILDIZ	/	/	/	/	/	0.167
MINES-CURIE-INSERM	/	/	/	/	/	0.235
CUHK	/	/	/	/	/	0.356
FCN	60	143	115	0.343	0.296	0.318
FFCN (without FC6_3)	62	141	117	0.346	0.305	0.324
FFCN	65	138	118	0.355	0.32	0.337
FFCN+FWLF	81	122	128	0.388	0.399	**0.393**
MFFCN+FWLF	92	111	135	0.405	0.453	**0.428**
MFFCN+FWLF+CRF	86	117	105	0.450	0.424	**0.437**

Detection Accuracy. The results of mitosis detection are showed in the Table.2 ("/" denotes that the results are not released). The MFF-CNN+CRF model yields the highest F-measure 0.437. In the experiment, we compared with MFF-CNN model and MFF-C+CRF model, we can clearly find that the number of misjudged mitosis greatly decreased from 135 to 105 shown in the Table.2. In Figure,8. The first column shows the H&E images, the second column shows the ground-truth image, the third column shows the test results of FF-CNN model, the fourth column is the test results of FF-CNN + CRF model, the fifth columns is the test results of multi-scale MFF-CNN + CRF model. The green circles in the third column of images shown the remove of misjudged mitosis samples by CRF-RNN model compared to the FF-CNN model. The yellow circles in the fourth column of images shown the adding of accurate mitosis samples by MFF-CNN model compared to the FF-CNN model. The experimental results in Fig.8. have shown that the CRF model resembles a "gate", which can filter out some of the wrong mitotic objects detected in the previous models. CRF model has added more contextual information, it can easily reduce the misjudgment mitoses, especially some isolated black noise points. The detection results for each slice are shown in the Fig.7. Even though the overall results is improved, the F-measure of A13 is still 0. Observing the H&E image of slice A13, their appearance are largely various and over-stained, most importantly there are only two mitoses within the 64 images.
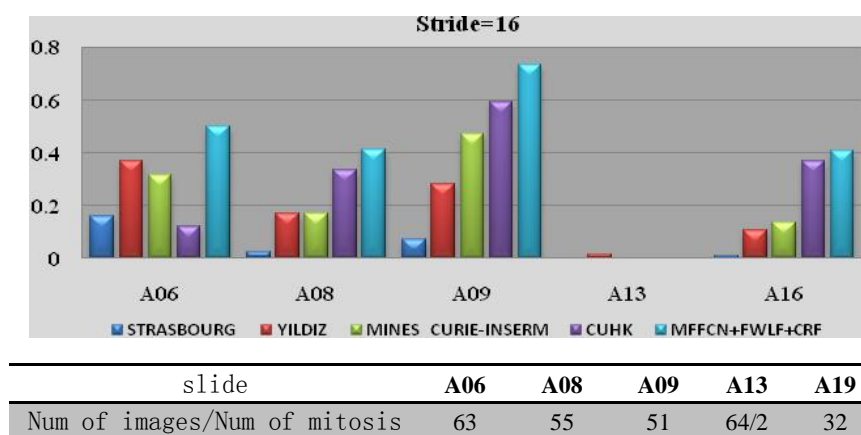

slide	A06	A08	A09	A13	A19
Num of images/Num of mitosis	63	55	51	64/2	32

Fig. 7. The performances on different slides of 2014 ICPR MITOSIS test data

4   Conclusion

Computer-assisted mitosis detector is meaningful to improve the efficiency and accuracy in cancer grading. In this paper, we proposed a Mutil-scale fused fully convolutional network combining the features from different level layers and multi-scale branches together and fine-tuned by a pre-trained caffe model. In addition, the MFF-

CNN model connect to a cascaded CRF-RNN model. We using the stain normalization based on tissue to reduce stain inconformity and training samples selection strategy to build the high quality training dataset . This work achieve a impressive accuracy rate. Future work will include developing an adaptive model for multiple dataset and applying unsupervised learning method in mitosis detection task .
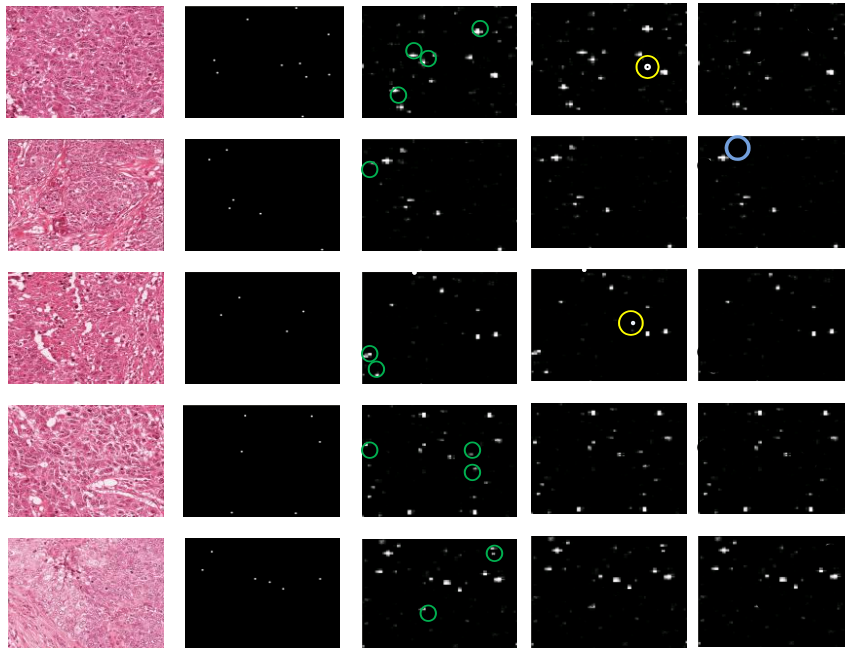

Fig. 8. The detection results by different models.

Acknowledgements

Reference

1. Roux L, Racoceanu D, Lomenie N, et al. Mitosis detection in breast cancer histological images An ICPR 2012 contest.[J]. Journal of Pathology Informatics, 2013, 4(1): 8-8.

2. Basavanhally A, Ganesan S, Feldman M, et al. Multi-Field-of-View Framework for Distinguishing Tumor Grade in ER+ Breast Cancer From Entire Histopathology Slides[J]. IEEE Transactions on Biomedical Engineering, 2013, 60(8): 2089-2099.

3. Irshad H. Automated mitosis detection in histopathology using morphological and multi-channel statistics features[J]. Journal of Pathology Informatics, 2013, 4(1): 10-10.

4. Irshad H, Jalali S, Roux L, et al. Automated mitosis detection using texture, SIFT features and HMAX biologically inspired approach[J]. Journal of Pathology Informatics, 2013, 4(2).

5. Huang C, Lee H C. Automated mitosis detection based on eXclusive Independent Component Analysis[C]. international conference on pattern recognition, 2012: 1856-1859.

6. Lecun, Yann, Y. Bengio, and G. Hinton. "Deep learning." *Nature*521.7553(2015):436-444.

7. Krizhevsky A, Sutskever I, Hinton G E, et al. ImageNet classification with deep convolutional neural networks[C]. neural information processing systems, 2012: 1097-1105.

8. Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]. computer vision and pattern recognition, 2015: 1-9.

9. Sainath T N, Mohamed A, Kingsbury B, et al. Deep convolutional neural networks for LVCSR[C]. international conference on acoustics, speech, and signal processing, 2013: 8614-8618.

10. Wang H. Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features.[J]. Journal of Medical Imaging, 2014, 1(3):1-8.

11. Dan C. Cireşan, et al. *Mitosis Detection in Breast Cancer Histology Images with Deep Neural Networks. Medical Image Computing and Computer-Assisted Intervention – MICCAI 2013*. Springer Berlin Heidelberg, 2013:411-418.

12. Chen H, Dou Q, Wang X, et al. Mitosis Detection in Breast Cancer Histology Images via Deep Cascaded Networks[C]//Thirtieth AAAI Conference on Artificial Intelligence. 2016

13. Kampffmeyer M, Salberg A B, Jenssen R. Semantic Segmentation of Small Objects and Modeling of Uncertainty in Urban Remote Sensing Images Using Deep Convolutional Neural Networks[C]// The IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016:680-688

14. Veta M, van Diest P J, Willems S M, et al. Assessment of algorithms for mitosis detection in breast cancer histopathology images.[J]. Medical Image Analysis, 2014, 20(1):237-248.

15. Chang, H., Loss, L. A., and Parvin, B., "Nuclear segmentation in h&e sections via multi-reference graph cut (mrgc)," in [ISBI'2012: Proceedings of the Sixth IEEE international conference on Symposium on Biomedical Imaging],(2012).

16. Veta M, Van Diest P J, Pluim J P, et al. Detecting mitotic figures in breast cancer histopathology images[J]. Proceedings of SPIE, 2013.

17. Jason Yosinski, Jeff Clune, Yoshua Bengio, and Hod Lipson, "How transferable are features in deep neural networks?," in Advances in Neural Information Processing Systems, 2014, pp.3320–3328

18. Janowczyk, A., Basavanhally, A., Madabhushi, A., 2016a. Stain normalization using sparse autoencoders (StaNoSA): Application to digital pathology. Comput Med Imaging Graph, in press.He K, Zhang X, Ren S, et al. Deep Residual Learning for Image Recognition[C]. computer vision and pattern recognition, 2015: 770-778.

19. Zhao L, Jia K. Multiscale CNNs for Brain Tumor Segmentation and Diagnosis[J]. Computational and Mathematical Methods in Medicine, 2016

20. Kampffmeyer M, Salberg A B, Jenssen R. Semantic Segmentation of Small Objects and Modeling of Uncertainty in Urban Remote Sensing Images Using Deep Convolutional Neural Networks[C]// The IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016:680-688.

21. Wu B, Kausar T, Xiao Q, et al. FF-CNN: An Efficient Deep Neural Network for Mitosis Detection in Breast Cancer Histological Images[M]//Proceedings of Medical Image Understanding and Analysis. 2017.