Mitosis Detection in Breast Cancer Histology Images via Deep Cascaded Networks

Hao Chen[†], Qi Dou[†], Xi Wang[‡], Jing Qin[§], Pheng-Ann Heng[†,*]
[†] Department of Computer Science and Engineering, The Chinese University of Hong Kong
[‡] College of Computer Science, Sichuan University, China
[§] School of Medicine, Shenzhen University, China
[*] Shenzhen Institutes of Advanced Technology, Chinese Academy of Sciences, China

Abstract

The number of mitoses per tissue area gives an important aggressiveness indication of the invasive breast carcinoma. However, automatic mitosis detection in histology images remains a challenging problem. Traditional methods either employ hand-crafted features to discriminate mitoses from other cells or construct a pixel-wise classifier to label every pixel in a sliding window way. While the former suffers from the large shape variation of mitoses and the existence of many mimics with similar appearance, the slow speed of the later prohibits its use in clinical practice. In order to overcome these shortcomings, we propose a fast and accurate method to detect mitosis by designing a novel deep cascaded convolutional neural network, which is composed of two components. First, by leveraging the fully convolutional neural network, we propose a coarse retrieval model to identify and locate the candidates of mitosis while preserving a high sensitivity. Based on these candidates, a fine discrimination model utilizing knowledge transferred from cross-domain is developed to further single out mitoses from hard mimics. Our approach outperformed other methods by a large margin in 2014 ICPR MITOS-ATYPIA challenge in terms of detection accuracy. When compared with the state-of-the-art methods on the 2012 ICPR MITOSIS data (a smaller and less challenging dataset), our method achieved comparable or better results with a roughly 60 times faster speed.

Introduction

Breast cancer is the most common cancer among women and a major cause of death worldwide (Boyle, Levin, and others 2008). According to the Nottingham Grading System, which is recommended by the World Health Organization for breast cancer screening (Elston and Ellis 1991), three morphological features in histology sections, including tubule formation, nuclear pleomorphism and the number of mitotic figures, are critical for the assessment of breast cancer. Among them, the number of mitoses gives an important aggressiveness indicator of the invasive breast carcinoma and hence is of great significance in diagnosis and treatment. However, the manual annotation by histologists is time-consuming and subjective with limited reproducibility. To this end, the development of automatic detection methods is essential for improving the efficiency and reliability of pathological examination.

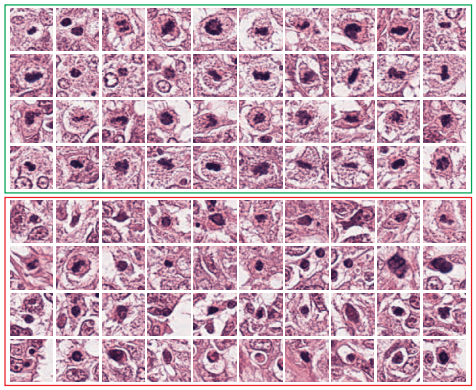

Figure 1: Example of mitoses and mimics (green rectangle encloses the true mitoses and red rectangle encloses the mimics that carry similar appearance).

However, automatic mitosis detection in breast histology images is a very challenging task for several reasons. First, the mitosis is characterized by a large variety of shape configurations, which are related to the high variation of biological structures, as shown in the green rectangle in Figure 1. In addition, the development of mitosis can be divided into four main phases: prophase, metaphase, anaphase and telophase. In different phases, the shape of nucleus is quite different, which further complicates the detection task. For instance, a mitotic cell has two distinct nuclei in the telophase, but they are not yet full individual cells. In this case, it must be counted as one single mitosis. Second, other cell types (e.g., apoptotic cells) usually carry similar morphological appearance with mitosis (as shown in the red rectangle in Figure 1), resulting in lots of false positives in the detection process. Third, the different conditions of histology image acquisition process, including sampling, cutting, staining, and digitalizing, increase the variabilities of mitosis appearance (Cruz-Roa et al. 2011; Veta et al. 2014). This is common when the tissue samples are acquired from different patients or at different time slots.

In recent years, several automatic methods have been developed to detect mitoses from breast histology images. Early studies employed hand-crafted features that capture specific characteristics of mitosis for automatic detection (Sommer et al. 2012; Khan, El-Daly, and Rajpoot 2012; Irshad 2013; Veta, van Diest, and Pluim 2013; Malon and Cosatto 2013; Wang et al. 2014; Tek 2013). However, these methods often suffer from the large shape variations of mitosis. In addition, the mimics with similar appearance are usually mistakenly recognized as mitoses by utilizing these features. In a recent study (Cireşan et al. 2013), researchers proposed to detect mitosis by applying deep convolutional neural networks (CNN), which can learn high-level feature representations in a data driven way, and achieved a higher detection accuracy than other methods. This method constructed a pixel-wise classifier based on CNN, which is quite computationally demanding and time-consuming. This prohibits its use in clinical practice. Another factor that may degrade the performance of the current CNN-based detection methods is the insufficiency of training samples, which may cause overfitting during the training process.

To overcome these shortcomings of previous methods, we propose a fast and accurate method to detect mitosis by designing a novel deep cascaded neural network (CasNN), which is composed of two models. First, by leveraging the fully convolutional network (FCN), we present a coarse retrieval model to identify and locate the candidates of mitosis. The proposed retrieval model can retrieve mitosis candidates efficiently on the whole sliding image while preserving a high sensitivity. Based on these candidates, a fine discrimination model is developed to further single out mitosis from hard mimics. As we reduce the search range from the whole image to the candidates only, the cascaded model can detect the mitoses in a common histology image within 0.5 seconds, around 60 times faster than the state-of-the-art method (Cireşan et al. 2013). To reduce the overfitting problem caused by the limited number of training samples and hence further improve the detection accuracy, the fine discrimination model transfers deep and rich feature hierarchies learned from a large number of cross-domain images, then fine tuned on the mitosis detection task. We evaluated our method on two public available datasets. Our method outperformed other competitors in 2014 ICPR MITOS-ATYPIA challenge by a large margin in terms of detection accuracy. When compared with the state-of-the-art methods on the 2012 ICPR MITOSIS data, which is much smaller and less challenging than the dataset in 2014 ICPR MITOS-ATYPIA challenge, our method achieved a comparable or better detection performance with a much faster speed.

Related Work

Studies for automatic detection of mitosis on hematoxylin and epsin (H&E) stained biopsies began thanks to the introduction of scanners for whole slide imaging (WSI) on glass slides. Previous studies applied domain-specific hand-crafted features to describe the morphological, statistical or textural characteristics of mitosis (Sommer et al. 2012; Khan, El-Daly, and Rajpoot 2012; Irshad 2013; Malon and Cosatto 2013; Wang et al. 2014; Tek 2013). Some of them combined two or more such features in order to improve the detection accuracy. For example, Sommer et al. (2012) constructed a pixel-wise classifier with shape and texture features to detect mitotic cells from histology images. Irshad et al. (2013) proposed a framework that included comprehensive analysis of statistics and morphological features for mitosis detection by employing a decision tree classifier. Tek (2013) classified mitotic and non-mitotic regions by using generic features and an ensemble of cascade adaboosts. However, these hand-crafted features require considerable efforts to design and validate. Furthermore, they cannot sufficiently represent the characteristics of mitoses with a large variation of shapes and textures, therefore resulting in a low detection accuracy.

Compared with methods based on hand-crafted features, the deep CNN with hierarchical feature representation learning has made breakthroughs in object recognition related problems (Krizhevsky, Sutskever, and Hinton 2012; Szegedy et al. 2014; Simonyan and Zisserman 2014; Chen et al. 2015c). As for mitosis detection, Ciresan et al. utilized a deep CNN as a pixel-wise classifier to detect mitosis and achieved the best performance at 2012 ICPR MITOSIS challenge with F1 score 78.2% (Cireşan et al. 2013) and 61.1% at 2013 MICCAI challenge (Veta et al. 2014), respectively. We consider this method as the state-of-the-art. However, the pixel-wise classifier of deep CNN is computationally demanding and time-consuming. Considering a single whole slide that consists of thousands of high-power fields (HPFs), it takes a long time to run across all sub-windows for detection, which may prohibit its use in clinical practice. In this regard, a novel model supporting fast and accurate mitosis detection simultaneously is demanded.

In recent years, several public datasets of breast cancer histology images have been available for algorithm assessment, e.g., the MITOSIS contest at ICPR 2012 (Roux et al. 2013), the AMIDA13 contest at MICCAI 2013 (Veta et al. 2014), and MITOS-ATYPIA challenge at ICPR 2014 with extensively enlarged data and partial pathological agreement considered. In our experiments, we conducted extensive experiments on the ICPR 2012 and 2014 challenge datasets.

Method

Figure 2 shows the architecture of the proposed method, which consists of two models of convolutional neural networks and they are combined in a cascaded manner. The first model can quickly retrieve the mitosis candidates while preserving a high sensitivity by taking the advantage of the fully convolutional network. We call it the *coarse retrieval model* N_c, which outputs a score map indicating the probability of mitosis candidates. The retrieved candidates are fed into the second model for further discrimination of mitoses and mimics with similar appearance. The second model is effectively constructed by transferring deep and rich feature hierarchies trained by deep convolutional neural networks on a large natural image dataset. We call it the *fine discrimination model* N_f, which is with higher capability of feature representation than CNN trained only on limited histopathological images, and hence can discriminate mitoses from hard mimics more precisely. Note that as the N_f performs only

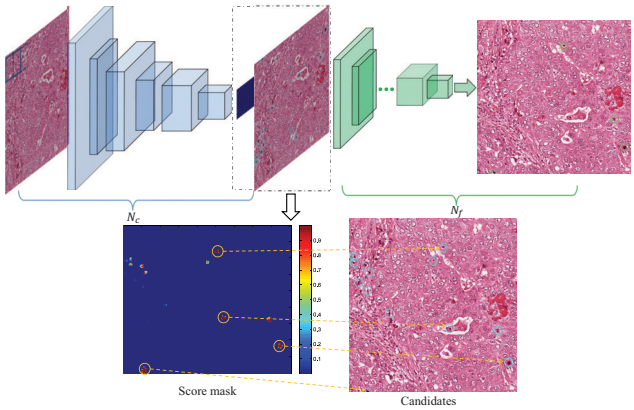

Figure 2: An overview of the proposed deep cascaded networks for fast and accurate mitosis detection.

on the candidates generated by N_c instead of the whole image, the detection process can be dramatically accelerated.

Coarse Retrieval Model

Considering that mitoses are sparsely distributed in histology images, a step of retrieving the regions of interest (ROI), i.e., mitosis candidates, could reduce the detection time dramatically, as the subsequent detection process could focus only on the candidates. Previous studies obtained the mitosis candidates relying on the pre-defined measurements on domain-specific morphological textures, color ratios or histogram distribution (Sommer et al. 2012; Irshad 2013; Malon and Cosatto 2013; Wang et al. 2014; Tek 2013). However, these methods were prone to losing mitoses, as these hand-crafted features could not accurately describe the complicated characteristics of mitoses. Recently, there were studies that explored the fast scanning approach with deep max-pooling convolutional neural networks for the mitosis detection (Giusti et al. 2013; Cireşan et al. 2013). Although these methods are more accurate than previous studies based on low-level features, they are computation-intensive and time-consuming in a pixel-wise classification way.

Different from previous methods, we utilize a fully convolutional network (FCN) (Ofer Matan and Denker 1991; Long, Shelhamer, and Darrell 2014) for fast retrieving the mitosis candidates. Traditional CNN contains the convolutional (C), sub-sampling, e.g., max-pooling (M), and fully-connected (FC) layers. Both C and M layers are translation invariant and can be operated on input of arbitrary size. However, the introduction of FC layers requires the input with a fixed size, as shown following:

$$h_j^l = \sigma(\boldsymbol{W}_j^l \boldsymbol{h}^{l-1} + b_j^l) \tag{1}$$

where $\boldsymbol{W}_j^l$ is the weight matrix connecting the neurons $\boldsymbol{h}^{l-1}$ in $(l-1)$th FC layer and jth index neuron h_j^l in the lth FC layer, b_j^l is the bias and $\sigma(\cdot)$ is the element-wise non-linear activation function. In fact, the fully connected layers are equivalent to the convolutional layers with kernel size 1×1:

$$\boldsymbol{h}_j^l = \sigma(\sum_{m=1}^{M} \boldsymbol{W}_{jm}^l \otimes \boldsymbol{h}_m^{l-1} + \boldsymbol{b}_j^l) \tag{2}$$

where $\otimes$ denotes the 2D spatial convolution, $\boldsymbol{W}_{jm}^l \in \mathbf{R}^{1\times 1}$ is the convolution kernel connected to jth feature map $\boldsymbol{h}_j^l$ and the mth feature map in the previous layer $\boldsymbol{h}^{l-1}$, and M is the total number of feature maps in $\boldsymbol{h}^{l-1}$. By employing Eq. (2), we can convert the fully connected layers into a fully convolutional fashion. Once the filters have been learned, they can be applied to the input image of arbitrary size.

The proposed coarse retrieval model has several advantages. First, since the FCN can take the whole image as input and generate the score mask with only one pass of forward propagation, it is capable of retrieving mitosis candidates efficiently on the whole sliding image instead of each pixel while preserving a high sensitivity. Second, it can also help to build a representative training database for the fine discrimination model N_f. Because mitoses rarely appeared in the whole HPF, non-mitosis training samples can be well represented by putting the false positives from N_c into the training samples of N_f. Thus the capability of the model N_f in distinguishing the mitoses from the hard mimics can be greatly enhanced by these false positives. The architecture of N_c can be seen in Table 1.

Table 1: The architecture of coarse retrieval model N_c

Layer	Feature maps	Kernel size	Stride
Input	94x94x3	-	-
C1	90x90x32	5	1
M1	30x30x32	3	3
C2	28x28x32	3	1
M2	14x14x32	2	2
C3	12x12x32	3	1
M3	6x6x32	2	2
FC4	100	-	-
FC5	2	-	-

Score Mask Generation The proposed coarse retrieval model is trained on the training samples with a fixed size input ($94 \times 94 \times 3$) by minimizing the cross entropy loss. Once the training is done, the N_c can be converted into a FCN model by Eq. (2). Then, the trained filters can be applied to scan the whole sliding image instead of employing the traditional patch-by-patch manner, which speeds up the detection process dramatically. Hence, the score mask indicating probability of mitosis candidates can be obtained after running through the converted coarse retrieval model only once. Each position of the output score mask corresponds to a specific region (size 94×94) in the original HPF image. Actually, it is equivalent to scanning the whole slide image with a fixed stride, which is determined largely by the stride of max-pooling layers. We will detail this in the following section.

Mitosis Candidates Localization Derived from the proposed model, the mitosis candidates can be located by mapping the index with higher scores on the score mask into the original coordinates of input image. Assuming non-overlapping region pooling, index mapping with convolution and max-pooling operations is formulated as:

$$\hat{x}_i = \frac{x_i - c_i}{s_i} + 1 \tag{3}$$

where c_i denotes the kernel size of convolutional or max-pooling layer, $\hat{x}_i$ is the position index after C or M operation on x_i, and s_i denotes the stride of convolutional or max-pooling layer. The original position index can be obtained by inverting above operations. For example, based on the network architecture shown in Table 1, for each position index $\hat{p}_s$ from the score mask, we can get the index p_o in the original image as following:

$$\begin{aligned} p_o = c_1 - 1 + s_{m_1}(c_2 - 1) + s_{m_1}s_{m_2}(c_3 - 1) \\ + s_{m_1}s_{m_2}s_{m_3}\hat{p}_s = p_0 + s\hat{p}_s \end{aligned} \tag{4}$$

where $p_0 = 22$ and $s = 12$ according to the architecture in Table 1. Thus we can retrieve the mitosis candidates with a sparse distribution based on the above index mapping. This is quite efficient when the detected objects are rarely distributed as in the case of mitosis detection. Despite with max-pooling layers, the probability maps can give quite dense predictions considering the equivalent stride 12 compared with the image size 2048×2048. Therefore, this approach can efficiently retrieve the candidates with a high sensitivity while reducing the computational workload. Subsequently, mitosis candidates are input into the deep discrimination model N_f for fine classification after local smoothing and non-max suppression.

Fine Discrimination Model

Knowledge Transfer across Domains The deep CNN with powerful feature representation achieved remarkable performance on recognition related tasks with large scale training resources available. However, limited datasets in medical applications, such as mitosis detection in breast histology images, increase the difficulties for training a powerful model to discriminate objectives from their mimics. The situation is further deteriorated when there exist lots of false positives with similar appearance. Although various transformations (e.g., translation, rotation, scaling, etc.) could be used to augment the training database, the training samples may be still insufficient to train a powerful model. Further improvement can be obtained with the help of transferring knowledge learned from related auxiliary tasks, where the training data can be easily acquired. Previous studies indicated that the filters (i.e., prior knowledge for recognition tasks) trained on large scale images of ImageNet (Russakovsky et al. 2014) could be transferred to different applications in other domains empirically (Jia et al. 2014; Yosinski et al. 2014; Chen et al. 2015b; Gupta, Ayhan, and Maida 2013; Chen et al. 2015a). Although medical and natural images are two different modalities and high level abstraction information is distinct, they do share statistical strength in low level details, e.g., orientation edges and junctions. Ideally, transfer learning from related data should somehow capture the salient factors of variation that explain the data, and benefit the recognition related target tasks.

Table 2: Number of HPFs/mitoses using Scanner Aperio-XT

Dataset (HPFs/mitoses)	ICPR 2012	ICPR 2014
Training data	35/226	1200/749
Testing data	15/100	496/NA

Therefore, we optimized the new medical task by employing an off-the-shelf model CaffeNet (Jia et al. 2014). The parameters of the previous layers (C1-C5) in the N_f were initialized by the pre-trained filters of CaffeNet model, which was trained on large scale images of ImageNet. This process can be considered as a pre-training phase of neural network with good initialization. By leveraging the transfer knowledge learned from large scale images, we fine tuned the N_f on the histology images by minimizing the following cross-entropy function:

$$\arg\min_{\boldsymbol{\theta}} \sum_{n=1}^{N}\sum_{k=1}^{K} -t_k \log p(y_k = 1|\boldsymbol{I}_n) + \lambda||\boldsymbol{W}||_2^2 \tag{5}$$

where $\boldsymbol{\theta} = \{\boldsymbol{W}, \boldsymbol{b}\}$ denotes the parameters of N_f, λ is the parameter for controlling the balance between the data loss term and the regularization term, $p(y_k = 1|\boldsymbol{I}_n)$ is the output probability for kth class given the input sub-window patch $\boldsymbol{I}_n$, t_k is the corresponding ground truth, K and N are the total number of classes and training samples, respectively. In the training process, dropout method (Hinton et al. 2012) was utilized to reduce the co-adaption of intermediate features.

Model Averaging In order to improve robustness, we trained multiple models of N_f for reducing the variance and improving the robustness capability (Geman, Bienenstock, and Doursat 1992). Three architectures with different number of neurons in three fully connected layers (i.e., FC6-FC8), 1024-256-2, 1024-512-2, 512-256-2, were trained, respectively. The sub-window sample was categorized as a mitosis when its averaged output posterior probability was above the threshold T (determined with cross-validation in our experiments), otherwise, categorized as non-mitosis.

Experiments

Materials and Preprocessing

The datasets were obtained from the 2012 and 2014 ICPR MITOSIS contests[1]. In our experiments, we evaluated our method on the HPF images acquired by the widely-used Apero-XT scanner. The numbers of HPF and mitoses in these two datasets are reported in Table 2. The centroids of mitoses were annotated by experienced pathologists. The ground truths of 2014 ICPR testing data were held out by the organizers for evaluation. For each dataset, we split training

[1]More details: http://ipal.cnrs.fr/ICPR2012/, http://mitos-atypia-14.grand-challenge.org/

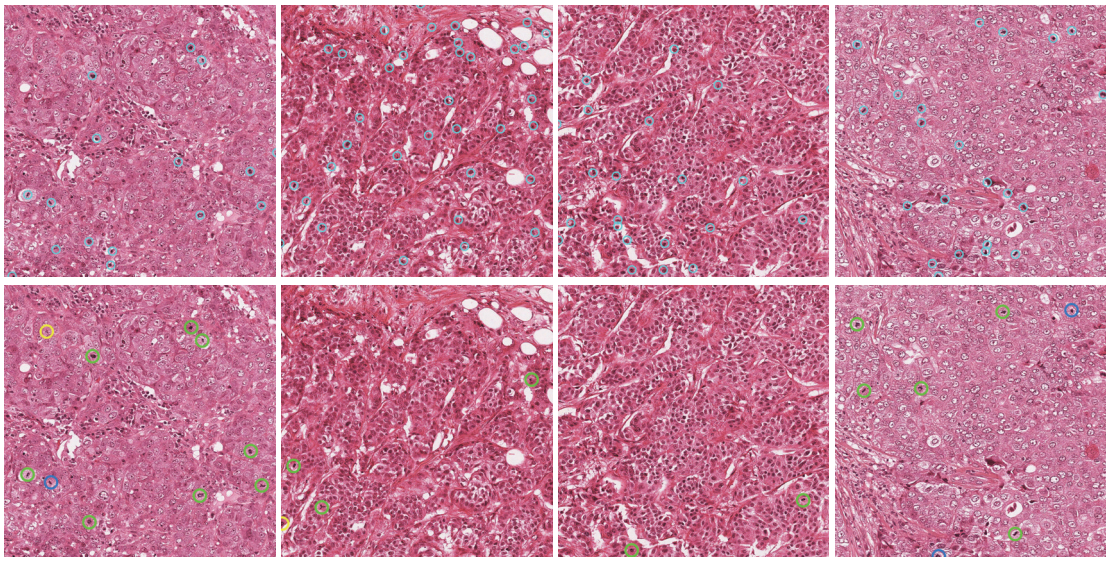

Figure 3: Detected results by our method on the testing data of 2012 ICPR MITOSIS contest. The first row shows the results of the coarse retrieval model and the second row shows the final detection results (cyan, yellow, blue and green circles denote the mitosis candidates, false negatives, false positives and true positives, respectively).

Table 3: Results of 2012 ICPR MITOSIS Dataset

Method	Precision	Recall	F1 score
UTRECHT (Veta et al. 2013)	0.511	0.680	0.584
NEC (Malon and Cosatto 2013)	0.747	0.590	0.659
SUTECH	0.699	0.720	0.709
IPAL (Irshad 2013)	0.698	0.740	0.718
DNN (Cireşan et al. 2013)	**0.886**	0.700	0.782
Coarse retrieval model N_c	0.211	**0.891**	0.342
RCasNN	0.720	0.713	0.716
CasNN(single)	0.738	0.753	0.745
CasNN(average)	0.804	**0.772**	**0.788**

data with ground truth into two sets for training and validation (about 1/7 of total training data), respectively. Patches extracted from mitotic regions were augmented by different transformations, including translation, rotation and flipping, for enlarging the training database.

Qualitative Evaluation

A score mask of coarse retrieval model and its corresponding retrieved candidates are shown in Figure 2. It is observed that large scores fired on the mitotic regions of score mask, while most of the non-mitotic regions have been suppressed as zeros, demonstrating the effectiveness of the coarse retrieval model with FCN. Four typical examples on the testing data of 2012 ICPR MITOSIS contest are shown in Figure 3. It is observed that the tissue appearance has large variations, which increases the difficulties for mitosis detection. Thanks to the advantages of FCN, the coarse retrieval model can efficiently retrieve the mitosis candidates while preserving most of true mitoses. Furthermore, our fine discrimination model can effectively get rid of most false positives from the candidates generated by the coarse retrieval model. There still exist a few false positives and false negatives in the final results. The false positives are mistakenly detected because the appearance of them is quite similar with the true mitosis, while the false negatives tend to be poorly stained in the tissue preparation process, which indicates that the consistence of tissue preparation is also important for accurate detection. Despite a few false results, our method could successfully detect most mitoses in these histology images.

Quantitative Evaluation and Comparison

According to the evaluation criteria of the MITOSIS challenges, a detection would be counted as a correct one if its Euclidian distance to a ground truth mitosis is less than $8\mu m$. All the detections that are not fallen within $8\mu m$ of a ground truth are counted as false positives. All the ground truths that do not have detections within $8\mu m$ are counted as false negatives. We employed the following evaluation measurements including recall: $R = N_{tp}/(N_{tp} + N_{fn})$, precision: $P = N_{tp}/(N_{tp} + N_{fp})$ and F_1 score: $F_1 = 2RP/(R + P)$, where N_{tp}, N_{fp} and N_{fn} are the number of true positives, false positives and false negatives, respectively. The ranking was made according to the overall F1-measure, where all the annotations were considered as a single dataset regardless to which patient they belong.

Evaluation on 2012 ICPR MITOSIS Dataset The 2012 ICPR MITOSIS dataset consists of 35 training images and 15 testing images. For the training process, we extracted a total of 29,100 mitosis samples after augmentation. In order to build a representative dataset for training the fine discrimination model, false positives from the coarse retrieval model were also employed. Note that the size of training data (to-

Table 4: Results of 2014 ICPR MITOSIS Dataset

Method	Precision	Recall	F1 score
STRASBOURG	-	-	0.024
YILDIZ	-	-	0.167
MINES-CURIE-INSERM	-	-	0.235
CUHK	-	-	0.356
RCasNN	0.360	0.424	0.389
CasNN(single)	0.411	0.478	0.442
CasNN(average)	**0.460**	**0.507**	**0.482**

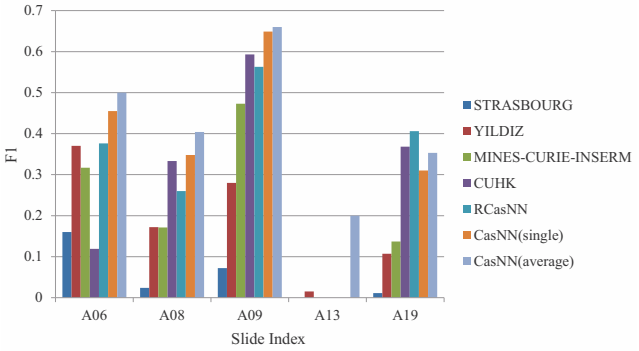

Figure 4: Results of different slides on 2014 ICPR MITOSIS testing data.

tal 394,275 training samples including 7.4% mitoses, 67.8% random selected negative samples, and 24.8% false positives from N_c) is smaller compared to the training dataset used in the state-of-the-art method (Cireşan et al. 2013) (1 million training instances with 6.6% as mitoses).

Although with less aggressive data augmentation, our method achieved a competitive performance on testing data with a much faster speed. The detailed results are reported in Table 3. Compared to the method with the best performance in 2012 ICPR contest (Cireşan et al. 2013), our method (CasNN) with model averaging achieved a comparable F1 score 0.788 and a higher recall 0.772. The single CasNN model outperformed the randomly initialized model (RCasNN) with respect to all the evaluation measurements, demonstrating that the knowledge transferred from deep and rich hierarchies can help to improve the performance. As the organizers from 2012 ICPR MITOSIS contest indicated that the dataset in this contest is by far too small for a good assessment of reliability and robustness of different algorithms (Roux et al. 2013), we further evaluated our method on the 2014 ICPR MITOSIS dataset, which is a much larger and more challenging dataset.

Evaluation on 2014 ICPR MITOSIS Dataset The dataset from 2014 ICPR is extensively expanded including 1200 training images and 496 testing images. One of the most difficult challenges in this dataset is the variability of tissue appearance, mostly resulted from the different conditions during the tissue acquisition process. As a result, the dataset is much more challenging than that in 2012 ICPR. The results of different methods are reported in Table 4 ("-" denotes that the results are not released). Our approach achieved the best performance with F1 score 0.482 outperforming other methods by a large margin. The performance of the single CasNN model outperformed the RCasNN model, demonstrating the efficacy of knowledge transfer strategy consistently. The per slide results of different methods are shown in Figure 4. It is observed that our method (CasNN) with model averaging achieved higher F1 score than other methods on most slides. Because the variation of tissue appearance in A13 slide is extremely large, the detection performance of different methods is very low. In comparison, our results achieved better performance than others.

Computation Time

In breast cancer diagnosis, a single whole slide usually consists of thousands of HPFs. Hence, the processing time of one HPF should be considered seriously in clinical applications (Veta et al. 2014). The superior advantage of the proposed cascaded framework is that it can reduce detection time dramatically while achieving a satisfactory accuracy. Our system was implemented with the mixed programming of MATLAB and C++. The coarse retrieval model took about 0.45 seconds to process per 4Mpixels HPF (size 2084×2084) and the fine discrimination model with 10 input variations cost about 0.49 seconds using a workstation with a 2.50 GHz Intel(R) Xeon(R) E5-2609 CPU and a NVIDIA GeForce GTX TITAN GPU. Totally, it took about 0.5 seconds for each input variation and was roughly 60 times faster than the state-of-the-art method (Cireşan et al. 2013), which took about 31 seconds with an optimized GPU implementation. Meanwhile, our approach achieved comparable detection accuracy to (Cireşan et al. 2013). This makes our approach possible for real-world clinical applications.

Conclusion

Automatic mitosis detection from breast cancer histology images can help to improve the efficiency and reliability of breast cancer screening and assessment. In this paper, we have proposed a novel deep cascaded network to tackle this challenging task by leveraging an efficient coarse retrieval model and a knowledge transferred fine discrimination model. Compared to the state-of-the-art methods, the proposed approach yielded better performance with a much faster speed, making it more practical in clinical applications. Future investigations include assessing the proposed method on more histology images and accelerating the algorithm with GPU optimization.

Acknowledgements

This work is supported by Hong Kong RGC General Research Fund (No. CUHK412513) and Shenzhen-Hong Kong Innovation Circle Funding Program (No. SGLH20131010151755080 and GHP/002/13SZ). The authors also gratefully thank Dr. Ludovic Roux for providing the data and helping the evaluation.

References

Boyle, P.; Levin, B.; et al. 2008. *World cancer report 2008.* IARC Press, International Agency for Research on Cancer.

Chen, H.; Dou, Q.; Ni, D.; Cheng, J.-Z.; Qin, J.; Li, S.; and Heng, P.-A. 2015a. Automatic fetal ultrasound standard plane detection using knowledge transferred recurrent neural networks. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*. Springer. 507–514.

Chen, H.; Ni, D.; Qin, J.; Li, S.; Yang, X.; Wang, T.; and Heng, P. 2015b. Standard plane localization in fetal ultrasound via domain transferred deep neural networks. *IEEE Journal of Biomedical and Health Informatics*.

Chen, H.; Shen, C.; Qin, J.; Ni, D.; Shi, L.; Cheng, J. C.; and Heng, P.-A. 2015c. Automatic localization and identification of vertebrae in spine ct via a joint learning model with deep neural networks. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015*. Springer. 515–522.

Cireşan, D. C.; Giusti, A.; Gambardella, L. M.; and Schmidhuber, J. 2013. Mitosis detection in breast cancer histology images with deep neural networks. In *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2013*. Springer. 411–418.

Cruz-Roa, A.; Díaz, G.; Romero, E.; and González, F. A. 2011. Automatic annotation of histopathological images using a latent topic model based on non-negative matrix factorization. *Journal of pathology informatics* 2.

Elston, C., and Ellis, I. 1991. Pathological prognostic factors in breast cancer. i. the value of histological grade in breast cancer: experience from a large study with long-term follow-up. *Histopathology* 19(5):403–410.

Geman, S.; Bienenstock, E.; and Doursat, R. 1992. Neural networks and the bias/variance dilemma. *Neural computation* 4(1):1–58.

Giusti, A.; Cireşan, D. C.; Masci, J.; Gambardella, L. M.; and Schmidhuber, J. 2013. Fast image scanning with deep max-pooling convolutional neural networks. *arXiv preprint arXiv:1302.1700*.

Gupta, A.; Ayhan, M.; and Maida, A. 2013. Natural image bases to represent neuroimaging data. In *Proceedings of the 30th International Conference on Machine Learning (ICML-13)*, 987–994.

Hinton, G. E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; and Salakhutdinov, R. R. 2012. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:1207.0580*.

Irshad, H. 2013. Automated mitosis detection in histopathology using morphological and multi-channel statistics features. *Journal of pathology informatics* 4.

Jia, Y.; Shelhamer, E.; Donahue, J.; Karayev, S.; Long, J.; Girshick, R.; Guadarrama, S.; and Darrell, T. 2014. Caffe: Convolutional architecture for fast feature embedding. *arXiv preprint arXiv:1408.5093*.

Khan, A. M.; El-Daly, H.; and Rajpoot, N. M. 2012. A gamma-gaussian mixture model for detection of mitotic cells in breast cancer histopathology images. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, 149–152. IEEE.

Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, 1097–1105.

Long, J.; Shelhamer, E.; and Darrell, T. 2014. Fully convolutional networks for semantic segmentation. *arXiv preprint arXiv:1411.4038*.

Malon, C. D., and Cosatto, E. 2013. Classification of mitotic figures with convolutional neural networks and seeded blob features. *Journal of pathology informatics* 4.

Ofer Matan, Christopher J. C. Burges, Y. L., and Denker, J. S. 1991. Multi-digit recognition using a space displacement neural network. In *Advances in Neural Information Processing Systems*, 488–495.

Roux, L.; Racoceanu, D.; Loménie, N.; Kulikova, M.; Irshad, H.; Klossa, J.; Capron, F.; Genestie, C.; Le Naour, G.; and Gurcan, M. N. 2013. Mitosis detection in breast cancer histological images An ICPR 2012 contest. *Journal of pathology informatics* 4.

Russakovsky, O.; Deng, J.; Su, H.; Krause, J.; Satheesh, S.; Ma, S.; Huang, Z.; Karpathy, A.; Khosla, A.; Bernstein, M.; et al. 2014. Imagenet large scale visual recognition challenge. *International Journal of Computer Vision* 1–42.

Simonyan, K., and Zisserman, A. 2014. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*.

Sommer, C.; Fiaschi, L.; Hamprecht, F. A.; and Gerlich, D. W. 2012. Learning-based mitotic cell detection in histopathological images. In *Pattern Recognition (ICPR), 2012 21st International Conference on*, 2306–2309. IEEE.

Szegedy, C.; Liu, W.; Jia, Y.; Sermanet, P.; Reed, S.; Anguelov, D.; Erhan, D.; Vanhoucke, V.; and Rabinovich, A. 2014. Going deeper with convolutions. *arXiv preprint arXiv:1409.4842*.

Tek, F. B. 2013. Mitosis detection using generic features and an ensemble of cascade adaboosts. *Journal of pathology informatics* 4.

Veta, M.; van Diest, P. J.; Willems, S. M.; Wang, H.; Madabhushi, A.; Cruz-Roa, A.; Gonzalez, F.; Larsen, A. B.; Vestergaard, J. S.; Dahl, A. B.; et al. 2014. Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Medical image analysis*.

Veta, M.; van Diest, P.; and Pluim, J. 2013. Detecting mitotic figures in breast cancer histopathology images. In *SPIE Medical Imaging*, 867607–867607. International Society for Optics and Photonics.

Wang, H.; Cruz-Roa, A.; Basavanhally, A.; Gilmore, H.; Shih, N.; Feldman, M.; Tomaszewski, J.; Gonzalez, F.; and Madabhushi, A. 2014. Cascaded ensemble of convolutional neural networks and handcrafted features for mitosis detection. In *SPIE Medical Imaging*, 90410B–90410B. International Society for Optics and Photonics.

Yosinski, J.; Clune, J.; Bengio, Y.; and Lipson, H. 2014. How transferable are features in deep neural networks? In *Advances in Neural Information Processing Systems*, 3320–3328.