

Efficient and Accurate Mitosis Detection

A Lightweight RCNN Approach

Yuguang Li², Ezgi Mercan¹, Stevan Knezevitch³, Joann G. Elmore⁴, and Linda G. Shapiro^{1,2}

¹Paul G. Allen School of Computer Science and Engineering, University of Washington, Seattle, WA, USA

²Department of Electrical Engineering, University of Washington, Seattle, WA, USA

³Pathology Associates, Clovis, CA, USA

⁴Department of Medicine, University of Washington, Seattle, WA, USA

ylee3@uw.edu, ezgi@cs.washington.edu, stevanrk@gmail.com, jelmore@uw.edu, shapiro@cs.washington.edu

Keywords: Breast Pathology, Automated Mitosis Detection, Convolutional Neural Networks, Histopathological Image Analysis

Abstract: The analysis of breast cancer images includes the detection of *mitotic figures* whose counting is important in the grading of invasive breast cancer. Mitotic figures are difficult to find in the very large whole slide images, as they may look only slightly different from normal nuclei. In the last few years, several convolutional neural network (CNN) systems have been developed for mitosis detection that are able to beat conventional, feature-based approaches. However, these networks contain many layers and many neurons per layer, so both training and actual classification require powerful computers with GPUs. In this paper, we describe a new lightweight region-based CNN methodology we have developed that is able to run on standard machines with only a CPU and can achieve accuracy measures that are almost as good as the best CNN-based system so far in a fraction of the time, when both are run on CPUs. Our system, which includes a feature-based region extractor plus two CNN stages, is tested on the ICPR 2012 and ICPR 2014 datasets, and results are given for accuracy and timing.

1 INTRODUCTION

A mitotic figure is a cell undergoing mitosis (division). In these actively dividing cells the chromosomes are visible by light microscopy. Instead of a nucleus, the chromosomes are visible as tangled, dark-staining threads. Counting of the mitotic figures is often used clinically as an indicator of tumor aggression (Medri et al., 2003). In clinical practice, the mitotic count is performed manually by a pathologist by carefully examining Hematoxylin and Eosin (H&E) stained tissue slides at high magnification using a microscope. This process is cumbersome and may contribute to inter-pathologist and intra-pathologist variation in tumor diagnosis of up to 20% (Yadav et al., 2012) (Baak et al., 2009) (Meyer et al., 2005) (Robbins et al., 1995) (Malon et al., 2012). The automation of this process could reduce time and cost and improve the comparability of results obtained from different labs (Malon et al., 2012). Mitosis detection in histopathological analysis is labor intensive. Mitotic figures evolve over a continuum spanning four distinct phases during which a

cell nucleus undergoes various transformations. Each phase is associated with a unique shape and texture. Scanned images from a single slide may not show all mitotic figures on the plane of focus, making their recognition more difficult due to areas being out of focus. Different regions may also mimic a mitotic figure, requiring a trained pathologist to differentiate between them. A low density of mitoses in histological images makes this work more labor intensive for a pathologist. In addition, differences in staining and tissue artifacts (e.g., tissue folding, tears, etc.) complicate this task.

Because of its importance in determining the severity of the cancer, the development of automated mitosis detection has become an active area of research with the goal of developing decision support systems to help pathologists. Contests have been held to encourage research in this topic, including the 2012 International Conference on Pattern Recognition (ICPR12) Mitosis Detection Contest (Roux et al., 2013), the Assessment of Mitosis Detection Algorithms 2013 Challenge (AMIDA13) (Veta et al., 2015) and the 2014 ICPR Mitosis Detection Chal-

lenge (MITOS-ATYPIA-14) (Roux, 2014). Over the years, the focuses have migrated from inter-group mitosis detection (one set of histopathology images provided for both training and testing in the ICPR12 challenge) to different-group mitosis detection (one set of histopathology images provided for initial training and testing and a separate validation set withheld for final testing in the ICPR14 challenge). Meanwhile, compared with the ICPR12 datasets, more realistic cases have been included in the ICPR14 datasets, including cases with inconsistent staining, bad lighting, tissue folding and inclusion of regions that look similar to texture of mitotic figures.

Multiple mitosis detection systems have been developed for these competitions and afterwards, using the data provided by the contests. In all cases, the top competitors have been deep neural networks. Motivated by these and by the object recognition work of Girshick (Girshick et al., 2016) that uses region-based convolutional neural networks to increase the efficiency of deep-neural-net-based machine learning, we have developed a region-based convolutional neural network (RCNN) system for mitosis detection in breast cancer images. Our system consists of three stages: 1) a feature-based random forest classifier that locates regions of interest in an image, 2) a candidate extractor stage that inputs the regions of interest from stage 1 to a CNN and filters out those that are not likely to be mitoses, and 3) a final predictor stage that inputs regions that remain after stage 2 and performs a scanning operation starting at their centers, using the same CNN to look for strong evidence of a mitotic figure. We have tested our system on both the ICPR12 and ICPR14 contest data sets and will show that it is much faster than the best reported system (Chen et al., 2016) while achieving accuracy measures (recall, precision, F-measure) that are almost as good. Our system would be more suitable for use in practice, since it can be run by practitioners in a medical environment without the use of GPUs.

2 RELATED LITERATURE

Because of the above contests, there have been a number of systems developed for mitotic figure detection in breast cancer images. Sertel et al. (Sertel et al., 2009) developed a computer-aided system based on pixel-level likelihood functions and 2-step component-based thresholding for mitotic counting in digitized images of neuroblastoma tissue. Roullier et al. (Roullier et al., 2010) proposed a multi-resolution unsupervised clustering method driven by domain-specific knowledge. In the 2012 ICPR contest, a va-

riety of approaches were developed and proven effective for the task (Roux et al., 2013). Irshad et al. (Irshad et al., 2013) proposed the framework of segmenting nuclei and finding mitotic regions among them, using selected block-wise color and texture features (e.g. co-occurrence features and run-length features) from the segmented area. Ciresan et al. (Ciresan et al., 2012) (Ciresan et al., 2013), whose work originally inspired our own, proposed an approach to sample from the original histological images and trained two separate multi-column deep learning neural networks. The same model has also been proven accurate in detecting mitoses from the AMIDA13 challenge. Their multi-column neural network (Ciresan et al., 2012) is an automated model to generate the optimized image descriptors and classify the input image patches with massive image training samples. The image descriptors generated from trained DNNs were proven to be helpful in object detection. Simo-Serra (Simo-Serra et al., 2005), Irshad (Irshad et al., 2013) and Wang (Wang et al., 2014) developed two different methods that merge DNN image descriptors and handcrafted features and improve the detection accuracy.

In microscopic images, mitoses are not frequently observed. Background tissues and mitotic regions have very distinguishable differences in color and texture. Differentiating between these two regions is usually as hard as detecting nuclei from the images. Convolutional neural networks (CNNs) for high-accuracy classification, however, are much more computationally expensive compared with regular nuclei detection algorithms. As a result, it is more computationally efficient to apply pre-processing to quickly eliminate most of the non-mitotic regions. A high-accuracy convolutional neural network can then be applied only to the difficult regions. Chen and Hao (Chen et al., 2016) proposed a two-stage mitosis detection pipeline, which improves the performance of both speed and accuracy. The pipeline first used a coarse retrieval model, a three-layer end-to-end Fully Convolutional Network (FCN), to segment mitosis candidates. It was followed by a fine discrimination model, a CaffeNet (Jia et al., 2014), to classify the selected patches. This model improves the F-measure of mitosis detection from the dataset of MITOS-ATYPIA-14 (Roux, 2014) by 13%. The pipeline takes around 0.5 seconds for each input image of 1000×1000 pixels with GPU and 31 seconds with an optimized CPU implementation. Wu proposed a fused fully-connected convolution neural net approach where the features from different layers are fused that outperforms the winner of the ICPR2014 mitosis detection challenge (Wu et al., 2017).

Compared with deep learning neural networks, pixel-wise classification with handcrafted image features is still cheaper and simpler for object segmentation. The FCN models introduced in previous research were initially proposed to handle image segmentation of more complicated subjects, for example 20 classes including vehicles and humans from ImageNet. These models, therefore, are not efficient in selecting mitotic candidates. On the other hand, small-sized multi-column neural networks proved to be very useful in detecting mitoses (Ciresan et al., 2013). Larger-scale neural nets such as CaffeNet (Chen et al., 2016) might potentially cause overfitting considering the complexity of the mitosis detection task. Smaller neural network models, therefore, have their advantages in higher test accuracy, lower training requirements and higher testing speed. In this paper, we develop a lighter-weight model and compare its results to those of Chen and Hao, who did not compete in the ICPR 2014 contest but beat the winners of that contest in 2016.

3 METHODOLOGY

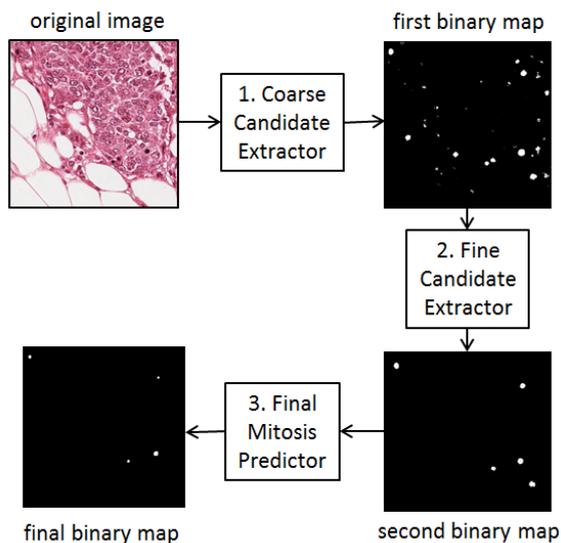


Figure 1: Architecture of our RCNN-based mitosis detection pipeline. Each of the three stages produces a binary map in which the 1-pixels (white) indicate potential centers of mitotic figures at that stage of the process.

Our RCNN-based mitosis detection pipeline has three stages as shown in Figure 1. The original whole slide image feeds into Stage 1: the Coarse Candidate Extractor (CCE). The CCE module extracts pixelwise features of several different scales at each pixel of

the image and classifies each pixel as possibly mitosis (positive) or not (negative), producing a binary map the same size as the original image. The binary map is the input to Stage 2: the Fine Candidate Extractor (FCE). The FCE module uses a Convolutional Neural Network (CNN) that quickly checks a fixed size region centered at each of the positive pixels in the map and decides if that region contains a possible mitosis (positive) or not (negative). It then outputs a second binary map that will, in general, have less positive pixels than the first one. The second binary map is input to Stage 3: the Final Mitosis Predictor (FMP). The FMP also begins with a fixed size region centered at each of the positive pixels in the map and uses the same CNN (with a higher threshold), but it checks multiple pixels and multiple rotations as described in Section 3.3.

3.1 Coarse Candidate Extractor

Distinguishing between mitotic-like pixels and background pixels is not as challenging as mitosis classification and can be performed using color and texture features of small regions about each pixel of the original color image. Our stage-1 feature extraction and classification step can quickly extract mitotic candidate pixels that will be sent to our relatively more expensive stages 2 and 3 CNN classifiers. Previous studies have proven that color, edge, and texture descriptors are very effective in segmenting regions like nuclei (Xing and Yang, 2016). Precision is not the priority in this stage, since we are aiming to quickly eliminate as many irrelevant pixels as possible and include all mitoses in our mitotic candidate regions.

Features: We began with 37 multi-scale features including color features (pixel values in Gaussian-smoothed color images), edge features (Gaussian Gradient Magnitude, Laplacian of Gaussian and Difference of Gaussian) and texture features (Structure Tensor Eigenvalues and Hessian of Gaussian Eigenvalues) of the three color channels (R, G, B) as our initial features. We used the feature selection method of Peng et al. (Peng et al., 2005) to reduce the number of selected features from 37 to 10. Table 1 lists the 10 selected features, their channels, and their scales.

Classification: A random forest classifier is trained to label pixels as mitotic candidates or not, given the above 10 features. The classifier has six trees, and its maximum number of features is four. The threshold for accepting a pixel as a possible mitosis at this stage is θ_{rfc} . See the Experiments section for details.

Table 1: Selected 10 low-cost pixel-wise features

Number	Feature Name	Scale	Channel
1	Smooth Pix Val	3.5px	R
2	Gauss Grad Mag	0.7px	R
3, 4	Gauss Grad Mag	1.0px	R,B
5, 6	Gauss Grad Mag	1.6px	R,G
7	Struc Tens Eig	0.7px	B
8, 9	Struc Tens Eig	3.5px	R,B
10	Struc Tens Eig	10.0px	G

3.2 Fine Candidate Extractor

The Fine Candidate Extractor uses a Convolutional Neural Network (CNN) model that is trained using ground truth training data that were provided by the organizers of the ICPR12 and ICPR14 contests and that we color normalize according to the color transfer method described in Reinhard et al. (Reinhard et al., 2001). Each pixel of each training image is labeled as either mitosis (pixel of mitotic region and within $8\mu\text{m}$ distance to the centroid of the mitosis) and non-mitosis (elsewhere). For the ICPR14 data set, in which mitoses were labeled with probabilities, only those mitotic regions with certainty above 0.6 were included in our training samples. Non-mitosis samples include regions with a certainty < 0.6 and regions outside the officially labeled regions. We applied the two-stage sampling techniques discussed in (Ciresan et al., 2013) to build our final training image sample set. Data sets are discussed in Section 4 under Experiments.

The CNN has five convolutional layers with the architecture shown in Table 2. It takes raw input image patches from the RGB image centered at the pixels identified by the stage-1 classifier and of fixed size 101×101 pixels as its inputs and produces a probability value between 0 and 1, which is later thresholded with a threshold θ_{low} (see Experiments) to produce an output binary image with pixels that are clearly *not* mitosis removed. As shown in Table 2, the convolutional layers are followed by rectified linear units (RELU) to improve model convergence and then max pooling layers. The model is trained with back-propagation implemented in the open-source library Caffe (Jia et al., 2014). The training starts with a learning rate of 0.01, which is reduced by 10 times every 20 iterations until the end of iteration 60. A dropout technique is also applied to each layer in every iteration to prevent inter-dependencies from emerging between nodes and to improve the robustness of the model.

Table 2: Architecture of our 5-layer CNN Classifier Model for Color-normalized patches.

Type	Neurons	Filter Size
Input	$3 \times 101 \times 101$	--
Conv	$16 \times 100 \times 100$	2×2
Relu	$16 \times 100 \times 100$	--
MaxPool	$16 \times 50 \times 50$	2×2
Conv	$16 \times 48 \times 48$	3×3
Relu	$16 \times 48 \times 48$	--
MaxPool	$16 \times 24 \times 24$	2×2
Conv	$16 \times 22 \times 22$	3×3
Relu	$16 \times 22 \times 22$	--
MaxPool	$16 \times 11 \times 11$	2×2
Conv	$16 \times 10 \times 10$	2×2
Relu	$16 \times 10 \times 10$	--
MaxPool	$16 \times 5 \times 5$	2×2
Conv	$16 \times 4 \times 4$	2×2
Relu	$16 \times 4 \times 4$	--
MaxPool	$16 \times 2 \times 2$	2×2
FullyConn	100	--
FullyConn	2	--

3.3 Final Mitosis Predictor

Stage 3, the Final Mitosis Predictor, inputs the binary image produced by Stage 2 and uses the same trained CNN, but in a sequence of scans of multiple different pixels of the input image. The idea is that each pixel that passes Stage 2 should lead to a detailed search of a region around it for evidence of a mitotic figure. A spiral scan path, as shown in Fig 2, carries out this detailed search. Specifically, the pixels that pass Stage 2 become the centers for the first "scan" in Stage 3, which is looking for 101×101 regions that pass this stage by a threshold of θ_{high} . If the first point does not satisfy this threshold, it keeps "scanning", moving by a distance I_c to the next point in the scan path (Fig 2) and making that point the center of a new 101×101 region to be tried. If one of these succeeds at the θ_{high} threshold, the system goes into a finer scanning mode (yellow path in Fig 2) in which the distances between points to try are smaller and the threshold is θ_{final} , the final threshold for calling a region a mitotic figure. If it finds one, it succeeds, else it continues on this yellow path till that path runs out, then returns to the green path it was on. If it totally finishes the green path, it fails. Threshold values are given in the Experiments section.

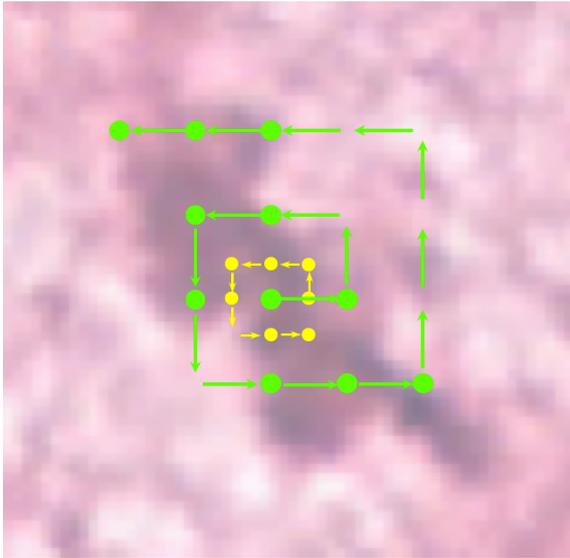


Figure 2: Scan path of final mitosis classifier.

4 EXPERIMENTS

Experiments were performed on a 3.30GHz Intel® Xeon® X5680 CPU with and without a NVIDIA GeForce GTX 1080 graphics processing unit (GPU). In order to compare our light-weight CNN to the CNN of Chen and Hao (Chen et al., 2016), simulated experiments were run on the same machine. Their CaffeNet stage is given the same region proposals that our RCNN gets from our Stage 1, and execution times of the two systems are thus directly compared.

4.1 Datasets

We used the ICPR 2012 and ICPR 2014 contest data sets. We discovered errors in the ground truth of the AMIDA13 data set and so chose to omit it from our experiments.

4.1.1 2012 ICPR Mitosis Dataset

We evaluate our method on the public MITOS dataset (Roullier et al., 2010) including 50 images corresponding to 50 high-power fields (HPF) in 5 different breast cancer slides stained with Hematoxylin & Eosin. A total of 300 mitoses are visible in MITOS. Each field represents a $512 \times 512 \mu m^2$ area, and is acquired using three different setups: two slide scanners and a multispectral microscope. Here we consider images acquired by the Aperio XT scanner, the most widespread and accessible solution among the three. It has a resolution of $0.2456 \mu m$ per pixel, resulting in a 2084×2084 RGB image for each field.

Expert pathologists manually annotated visible mitoses. However, the annotated mitoses are those for which pathologists agreed with each other. So there might be some mitosis-like regions unannotated in these images. Our training and testing splits are the same as those provided in the contest. There are 100 mitoses from 15 images (3 images from each slide) in our training set and 200 mitoses from 35 images (7 images from each slide) in our testing set. Every mitotic pixel is labeled as such. For this dataset, $\theta_{rfc} = 0.5$, $\theta_{low} = 0.2$, $\theta_{high} = 0.4$, and $\theta_{final} = 0.8$. Threshold values were chosen to maximize F-measure on the training set.

4.1.2 2014 ICPR Mitosis Dataset

The 2014 ICPR Mitosis dataset include a total of 1200 training images from 16 different biopsies and 496 testing images from 5 different biopsies of 40X magnification. The images are, however, much smaller than those from ICPR 2012: 1539×1376 pixels. The mitoses are marked by probability values of 1.0, 0.8, 0.6, 0.2 or 0. The labels were made by two different pathologists with the criteria described in (Roux, 2014). Compared with 2012 ICPR Mitosis dataset, the 2014 ICPR Mitosis dataset includes many more variations in tissue appearance, which are affected by different conditions from tissue acquisition process, including staining, lighting conditions, tissue folding and the inclusion of many mitotic-like non mitosis regions. In addition, the training and testing images of the 2014 ICPR Mitosis dataset are taken from different biopsies, whereas the aim of the 2012 ICPR challenge is to train and test mitotic detection algorithms in the same group of images. For this dataset, $\theta_{rfc} = 0.5$, $\theta_{low} = 0.1$, $\theta_{high} = 0.25$, and $\theta_{final} = 0.5$. Threshold values were again chosen to maximize F-measure on the training set.

For the 2014 ICPR Mitosis dataset, we perform tests separately on their “in-group”, which is the data set for which they provide ground truth for comparison and on their “out-group”, which is the data set for which the ground truth is withheld and for which the results must be sent to the contest organizers for evaluation. During in-group experiments, we randomly split images of each slide in the training set into 60% for training and 40% for testing. Since our model works on images with normalized color, we did color alignment with a selected reference image from the 2012 ICPR Mitosis dataset before training, and we used the prediction map from the model pre-trained with the 2012 ICPR Mitosis dataset as the sampling map to build training samples from the 2014 ICPR Mitosis dataset. In our out-group experiment, we took our training patches from the entire official training

dataset. The predictions on the testing set were sent to the organizer of the 2014 ICPR Mitosis dataset contest for scoring.

4.2 Quantitative Evaluation Metrics

According to the criteria of the 2014 ICPR Mitosis Detection Challenge, a detected mitosis would be counted as correct if its centroid is localized within a range of $8\mu m$ of the ground truth region. Multiple separate regions of the same cell (e.g. mitosis after Metaphase in cell division) are counted as a single mitosis. The evaluation metrics are defined as: recall $R = N_{TP}/(N_{TP} + N_{FN})$, precision $P = N_{TP}/(N_{TP} + N_{FP})$ and F_1 measurement $F_1 = 2 * P * R / (P + R)$, where N_{TP} , N_{FN} and N_{FP} are the number of true positives, false negatives and false positives, respectively.

4.3 Results

We report results in both accuracy and time. Our goal in this work was to design a light-weight network that could achieve similar performance to the network of Chen and Hao (Chen et al., 2016) but would execute much faster and thus be able to run in a medical environment with no GPUs available.

4.3.1 Accuracy

We first tested our RCNN model on the test sets for the 2012 ICPR mitosis contest. Accuracy results are reported in Table 3. Our F-measurements are comparable to, but slightly lower than, the F-measure from CaffeNet (Chen et al., 2016) (our F-measure is 0.784, and theirs is 0.788). We then tested our model on the 2014 ICPR mitosis test sets: both the in-group and the out-group. For the in-group test, we achieved an F-measurement of 0.659; Chen and Hao did not report results on the in-group. For the out-group, we achieved an F-measure of 0.427 compared to Chen’s 0.482, which while not as good, is in the same ballpark. As we have discussed above, the 2014 ICPR mitosis dataset includes more challenges, such as more complicated background tissue appearance and more color/pattern variations from mitoses. This accounts for why these F-measurements are lower than those from the 2012 ICPR mitosis datasets. However, by including more training images from more biopsies with different color and texture variation on mitoses, we should be able to achieve higher classification accuracy.

Table 3: Preliminary Results on ICPR 2014 test datasets

		Chen & Hao	Ours
ICPR12	Precision	0.80	0.78
	Recall	0.77	0.79
	F-Measure	0.788	0.784
ICPR14 in group	Precision	not rep.	0.654
	Recall	not rep.	0.663
	F-Measure	not rep.	0.659
ICPR14 out group	Precision	0.46	0.40
	Recall	0.51	0.45
	F-measure	0.482	0.427

4.3.2 Computation Time

Computation time in mitosis detection is always a very important factor in clinical applications (Veta et al., 2015). In the meantime, the accessibility to a high performance computer with a GPU processor has been an important factor in mitosis detection system design. Our RCNN-based mitosis detection pipelines are designed to reduce redundant computation and reduce processing time for scanning each slide, and are compatible with computers with *only CPU*. Our pipelines are mainly implemented in Python and C++. The feature-based step is implemented by ilastik (Sommer et al., 2011), and the deep learning model is implemented in the Caffe library. The details of computational efficiency are listed in Table 4.

For a 4Mpixels high-power field (size 2048×2048) image from the 2012 ICPR Mitosis dataset, the whole pipeline takes a total of 6.93s with GPU support and 12.13s without GPU support per image. All of our experiments running on CPUs are executed on a single CPU thread. In practice, different steps could be executed with different threads running in parallel. The random forest classifier could also be optimized with GPU processing.

4.4 Comparison to Chen and Hao

Our 3-stage system achieves accuracy similar to Chen and Hao’s CaffeNet (Chen et al., 2016) on the ICPR 2012 data set and almost as good on the ICPR 2014 data set, but it is smaller and faster. Our RCNN is a smaller model of 5 convolutional layers with only 16 kernels in each layer, while CaffeNet is also composed of 5 layers, but with 96-384 kernels in each layer. This enables our model to be tested with a regular GPU or only CPU with reasonable processing time, whereas CaffeNet requires large GPU memories and much more computation. Table 5 shows details about the comparison on computational cost between

Table 4: Computational time of our RCNN on the ICPR 2012 and ICPR 2014 mitosis datasets. Time is in seconds per image. Note that ICPR 2014 times are smaller due to the smaller image size.

	Dataset	proc time CPU	Number of regions	Number CNN calls	CNN time	
					GPU	CPU
Coarse Candidate Extractor	ICPR12	3.24s	-	0	0	0
Coarse Candidate Extractor	ICPR14	0.91s	-	0	0	0
Fine Candidate Extractor	ICPR12	1.27s	987	987	0.45s	3.31s
Fine Candidate Extractor	ICPR14	0.41s	266	266	0.14s	0.90s
Final Mitosis Extractor	ICPR12	1.59s	11.8	756	0.38s	2.72s
Final Mitosis Extractor	ICPR14	0.34s	2.4	198	0.10s	0.72s

our RCNN model and CaffeNet, where FLOP indicates the number of floating point operations required for the inference time of each neural network. The time comparisons are for inference time on a 101 x 101 patch of an image.

Table 5: Performance comparison between our RCNN model and Chen and Hao’s CaffeNet. Times are per 101 × 101 pixel regions.

	Chen & Hao	Ours
FLOPs (million)	720.32	8.47
Memory (MB)	3.251	0.949
Parameters (million)	56.87	0.0097
Inf. time on GPU (ms)	1.56	0.46
Inf. time on CPU (s)	0.1381	0.0032

In order to compare them on the same machine, we performed a simulated experiment to compare the RCNN stages of our 3-stage method with the CaffeNet-based method (Chen et al., 2016). Computational performance comparison details can be found in Table 6. In this experiment, we assume that CaffeNet receives the same region proposals generated from stage 1 as our RCNN model does. Thus a 4Mpixels HPF ICPR12 image includes an average of 987 proposed regions. As the prediction on each mitotic-like region is computed from the average probability of three CaffeNet models of different fully connected layers on 10 image variations, the combined fine discrimination model from (Chen et al., 2016) takes an average of 4.62s with GPU support and 408.91s without GPU support in total for its final prediction on each full image. This prediction stage is equivalent to stage 2 and stage 3 from our pipeline, which only takes 0.83s and 6.03s, respectively, in total to produce its final prediction. The simulated experiment on CaffeNet was conducted on the same computer in which our own experiments were run. The computation time reported above is only for neural network inference time after the stage-one region-finding phase. The stage-one processes of the

two systems are quite different, since ours is a simple feature-based classifier and theirs is a neural network of three convolutional layers; they are not comparable.

Notice the striking difference in CPU inference time between our RCNN system and the CaffeNet system. Our pipeline provides a *67.81 times speedup* over theirs. This is critical for medical applications in which a large number of images must be routinely processed.

Table 6: Performance comparison between Stages 2 and 3 of our pipeline and the fine discrimination model of Chen’s pipeline (Chen et al., 2016) for inference on a full image. Numbers shown below are taken as the average on 4Mpixels HPF images from the ICPR12 dataset.

	Chen & Hao	Ours
CNN calls	2961	1743
GPU inference time	4.62s	0.83s
CPU inference time	408.91s	6.03s

5 CONCLUSION

Automatic mitosis detection from breast cancer histology images can help to improve the accuracy and efficiency of breast cancer diagnosis. In this paper, we have proposed a hybrid mitosis detection pipeline, which combine efficient handcrafted-feature-based pixel classifiers and 5-layer neural networks in multi-stage pipelines. Compared to the state-of-the-art methods, our approach reduced the computation amounts and hardware requirements, which makes it more practical in clinical applications. Future work includes optimizing CPU implementation of normalization layer in neural network model and optimizing our random-forest based coarse mitotic candidate extractor with GPU implementations.

ACKNOWLEDGEMENTS

Research reported in this publication was supported by the National Cancer Institute awards R01 CA172343, R01 CA140560 and R01 CA200690. The content is solely the responsibility of the authors and does not necessarily represent the views of the National Cancer Institute or the National Institutes of Health. We thank Ventana Medical Systems, Inc. (Tucson, AZ, USA), a member of the Roche Group, for the use of iScan Coreo Au™ whole slide imaging system, and HD View SL for the source code used to build our digital viewer. For a full description of HD View SL, please see <http://hdviewsl.codeplex.com/>.

REFERENCES

- Baak, J. P. A., Gudlaugsson, E., Skaland, I., Guo, L. H. R., Klos, J., Lende, T., Soiland, H., Janssen, E. A. M., and zur Hausen, A. (2009). Proliferation is the strongest prognosticator in node-negative breast cancer: Significance, error sources, alternatives, and comparison with molecular prognostic markers. *Breast Cancer Res Treat*, 115:241–254.
- Chen, H., Dou, Q., Wang, X., Qin, J., and Heng, P. A. (2016). Mitosis detection in breast cancer histology images via deep cascaded networks. *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence*. AAAI Press, pages 1160–1166.
- Ciresan, D., Alessandro, G., Gambardella, L. M., and Schmidhuber, J. (2013). Mitosis detection in breast cancer histology images with deep neural networks. *Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2013*, pages 411–418.
- Ciresan, D., Meier, U., and Schmidhuber, J. (2012). Multi-column deep neural networks for image classification. In *Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference*, pages 3642–3649.
- Girshick, R., Donahue, J., Darrell, T., and Malik, J. (2016). Region-based convolutional networks for accurate object detection and segmentation. *IEEE transactions on pattern analysis and machine intelligence*, 38(1):142–158.
- Irshad, H., Roux, L., and Racoceanu, D. (2013). Multi-channels statistical and morphological features based mitosis detection in breast cancer histopathology. In *Engineering in Medicine and Biology Society 35th Annual International Conference*. IEEE.
- 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. *Proceedings of the 22nd ACM international conference on Multimedia*.
- Malon, C., Brachtel, E., Cosatto, E., Graf, H. P., Kurata, A., Kuroda, M., Meyer, J. S., Saito, A., Wu, S., and Y, Y. Y. (2012). Mitotic figure recognition: agreement among pathologists and computerized detector. *Anal Cell Pathology*, 35(2):97–100.
- Medri, L., Volpi, A., Nanni, O., Vecchi, A. M., Mangia, A., Schittulli, F., Padovani, F., Giunchi, D. C., Vito, A., Amadori, D., Paradiso, A., and Silvestrini, R. (2003). Prognostic relevance of mitotic activity in patients with node-negative breast cancer. *Modern Pathology*, 16(11):1067–1075.
- Meyer, J. S., Alvarez, C., Milikowski, C., Olson, N., Russo, I., Russo, J., Glass, A., Zehnauer, B. A., Lister, K., and Parwaresch, R. (2005). Breast carcinoma malignancy grading by bloom-richardson system vs proliferation index: Reproducibility of grade and advantages of proliferation index. *Modern Pathology*, 18:1067–1078.
- Peng, H., Long, F., and Ding, C. (2005). Feature selection based on mutual information criteria of max-dependency, max-relevance, and min-redundancy. *IEEE Transactions on pattern analysis and machine intelligence*, 27(8):1226–1238.

- Reinhard, E., Adhikhmin, M., Gooch, B., and Shirley, P. (2001). Color transfer between images. *IEEE Computer graphics and applications*, 21(5):34–41.
- Robbins, P., Pinder, S., de Klerk, N., Dawkins, H., Harvey, J., Sterrett, G., Ellis, I., and Elston, C. (1995). Histological grading of breast carcinomas: A study of interobserver agreement. *Human Pathology*, 26(8):873–879.
- Roullier, V., Lzoray, O., Ta, V. T., and Elmoataz, A. (2010). Mitosis extraction in breast-cancer histopathological whole slide images. In *Advances in Visual Computing*, pages 539–548. Springer Berlin.
- Roux, L. (2014). Mitosis atypia 14 grand challenge. <https://mitos-atypia-14.grand-challenge.org/>.
- Roux, L., Racoceanu, D., Lomnie, N., Kulikova, M., Irshad, H., Klossa, J., Capron, F., Genestie, C., Naour, G. L., and Gurcan, M. N. (2013). Mitosis detection in breast cancer histological images: An ICPR 2012 contest. *Journal of pathology informatics*, 4(8).
- Sertel, O., Catalyurek, U. V., Shimada, H., and Gurcan, M. N. (2009). Computer-aided prognosis of neuroblastoma: Detection of mitosis and karyorrhexis cells in digitized histological images. In *Engineering in Medicine and Biology Society Annual International Conference*, pages 1433–1436. IEEE.
- Simo-Serra, E., Trulls, E., Ferraz, L., Kokkinos, I., Fua, P., and Moreno-Noguer, F. (2005). Discriminative learning of deep convolutional feature point descriptors. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 118–126.
- Sommer, C., Straehle, C., Koethe, U., and Hamprecht, F. A. (2011). Ilastik: Interactive learning and segmentation toolkit. In *Biomedical Imaging: From Nano to Macro. 2011 IEEE International Symposium on*, pages 230–233.
- Veta, M., Diest, V., Willems, P. J., Wang, S. M., Madabhushi, H., Cruz-Roa, A., Gonzalez, A., Larsen, F., Vestergaard, A. B., Dahl, J. S., and Cirean, D. C. (2015). Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Medical image analysis*, 20(1):237–248.
- Wang, H., Cruz-Roa, A., Basavanthally, 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. *SPIE Medical Imaging*, 90410B.
- Wu, B., Kausar, T., Xiao, Q., Wang, M., Wang, W., Fan, B., and Sun, D. (2017). FF-CNN: an efficient deep neural network for mitosis detection in breast cancer histological images. In *Medical Image Understanding and Analysis - 21st Annual Conference, MIUA 2017, Edinburgh, UK, July 11-13, 2017, Proceedings*, pages 249–260.
- Xing, F. and Yang, L. (2016). Robust nucleus/cell detection and segmentation in digital pathology and microscopy images: A comprehensive review. *IEEE Rev. Biomed. Eng.*, 9:234–263.
- Yadav, K. S., Gonuguntla, S., Ealla, K. K., Velidandla, S. R., Reddy, C. R., Prasanna, M. D., and Bommu, S. R. (2012). Assessment of interobserver variability in mitotic figure counting in different histological grades of oral squamous cell carcinoma. *Contemporary Dental Practice*, 13(3):339–344.