Cascaded Ensemble of Convolutional Neural Networks and Handcrafted Features for Mitosis Detection

Haibo Wang *[a], Angel Cruz-Roa*[2], Ajay Basavanhally[1], Hannah Gilmore[1], Natalie Shih[3], Mike Feldman[3], John Tomaszewski[4], Fabio Gonzalez[2], and Anant Madabhushi[1]

[1]Case Western Reserve University, USA
[2]Universidad Nacional de Colombia, Colombia
[3]University of Pennsylvania, USA
[4]University at Buffalo School of Medicine and Biomedical Sciences, USA

ABSTRACT

Breast cancer (BCa) grading plays an important role in predicting disease aggressiveness and patient outcome. A key component of BCa grade is mitotic count, which involves quantifying the number of cells in the process of dividing (i.e. undergoing mitosis) at a specific point in time. Currently mitosis counting is done manually by a pathologist looking at multiple high power fields on a glass slide under a microscope, an extremely laborious and time consuming process. The development of computerized systems for automated detection of mitotic nuclei, while highly desirable, is confounded by the highly variable shape and appearance of mitoses. Existing methods use either handcrafted features that capture certain morphological, statistical or textural attributes of mitoses or features learned with convolutional neural networks (CNN). While handcrafted features are inspired by the domain and the particular application, the data-driven CNN models tend to be domain agnostic and attempt to learn additional feature bases that cannot be represented through any of the handcrafted features. On the other hand, CNN is computationally more complex and needs a large number of labeled training instances. Since handcrafted features attempt to model domain pertinent attributes and CNN approaches are largely unsupervised feature generation methods, there is an appeal to attempting to combine these two distinct classes of feature generation strategies to create an integrated set of attributes that can potentially outperform either class of feature extraction strategies individually. In this paper, we present a cascaded approach for mitosis detection that intelligently combines a CNN model and handcrafted features (morphology, color and texture features). By employing a light CNN model, the proposed approach is far less demanding computationally, and the cascaded strategy of combining handcrafted features and CNN-derived features enables the possibility of maximizing performance by leveraging the disconnected feature sets. Evaluation on the public ICPR12 mitosis dataset that has 226 mitoses annotated on 35 High Power Fields (HPF, x400 magnification) by several pathologists and 15 testing HPFs yielded an F-measure of 0.7345. Apart from this being the second best performance ever recorded for this MITOS dataset, our approach is faster and requires fewer computing resources compared to extant methods, making this feasible for clinical use.

1. INTRODUCTION

Bloom Richardson grading,[1] the most commonly used system for histopathologic diagnosis of invasive breast cancers (BCa),[2] comprises three main components: tubule formation, nuclear pleomorphism, and mitotic count. Mitotic count, which refers to the number of dividing cells (i.e. mitoses) visible in hematoxylin and eosin (H & E) stained histopathology, is widely acknowledged as a good predictor of tumor aggressiveness.[3] In clinical practice, pathologists define mitotic count as the number of mitotic nuclei identified visually in a fixed number of high power fields (400x magnification). However, the manual identification of mitotic nuclei often suffers from poor inter-rater agreement due to the highly variable texture and morphology between mitoses. Additionally this is a very laborious and time consuming process involving the pathologist manually looking at and counting mitoses from multiple high power view fields on a glass slide under a microscope. Computerized detection of mitotic nuclei will lead to increased accuracy and consistency while simultaneously reducing the time and cost needed for BCa diagnosis.?

The detection of mitotic nuclei in H & E stained histopathology is a difficult task for several reasons.[3] First, mitosis is a complex biological process during which the cell nucleus undergoes various morphological transformations. This

* indicates equal contributions

Medical Imaging 2014: Digital Pathology, edited by Metin N. Gurcan, Anant Madabhushi,
Proc. of SPIE Vol. 9041, 90410B · © 2014 SPIE · CCC code: 1605-7422/14/$18
doi: 10.1117/12.2043902

leads to highly variable size and shape across mitotic nuclei within the same image. Another issue is rare event detection, which complicates classification tasks where one class (i.e. mitotic nuclei) is considerably less prevalent than the other class (i.e. non-mitotic nuclei). A final difficulty lies in the fact that pathologists routinely employ very high magnification (up to 400x) in conjunction with fine z-axis control to identify mitosis. By contrast, most digitized histopathology images generated by whole-slide scanners are limited to a magnification of 40x and typically do not provide additional z-stack images.

Recently, the development of computerized systems for automated mitosis detection has become an active area of research with the goal of developing decision support systems to be able to relieve the workload of the pathologist. In a contest held in conjunction with the ICPR 2012 conference[3] to identify the best automated mistosis detection algorithm, a variety of approaches competed against each other in the contest, which can be categorized as handcrafted feature based or feature learning based. The commonly used handcrafted features include various morphological, shape, statistical and textural features that attempt to model the appearance of the domain and in particular the appearance of the mitoses within the digitized images.[4–7] While domain inspired approaches (hand crafted) are useful in that they allow for explicit modeling of the kinds of features that pathologists look for when identifying mitoses, there is another category of feature generation inspired by convolutional neural networks (CNN),[8,9] CNN are multi-layer neural networks that learns a bank of convolutional filters at each layer.[10–13] In contrast to handcrafted features, CNN is fully data-driven, therefore being more accurate in representing training samples and able to find feature patterns that handcrafted features fail to describe. However, CNN is computationally demanding and sensitive to the scalability of training data. The winner[13] of the ICPR contest used two 11-layers to achieve an F-measure of 0.78. However, this approach is not feasible for clinical use since each layer of the CNN model comprised hundreds of neurons and required a large amount of time for both training and testing. Other methods achieved an F-measure of upto 0.71, based primarily on combining various handcrafted features. While hand-crafted feature approaches are faster, drawbacks include (1) the fact that the identification of salient features are highly dependent on the evaluation dataset used and (2) the lack of a principled approach for combining disparate features. Hence, it stands to reason that a combination of CNN and handcrafted features will allow us to exploit the high accuracy of CNN while also reducing the computational burden (in terms of time) of handcrafted features. By employing a light CNN model, the proposed approach is far less demanding computationally, and the cascaded strategy of combining handcrafted features and CNN-derived features enables the possibility of maximizing performance by leveraging the disconnected feature sets. Previous work in this approach includes the NEC team,[12] where an attempt was made to stack the CNN-learned features and handcrafted features yielded an F-measure of 0.659, suggesting that more intelligent combination of CNN and handcraft features are required.

In this paper, we present a cascaded approach to combining CNN and handcrafted features for mitosis detection. The workflow of the new approach is depicted in Figure 1. The first step is to segment likely mitosis regions. This initial phase serves as a triage to remove obviously non-mitotic regions. For each candidate region, both CNN-learned and handcrafted features were extracted independently. Independently trained classifiers were constructed using the handcrafted and CNN-learned features alone. For the regions on which the two individual classifiers highly disagree, they are further classified by a third classifier that was trained based on the stacking of handcrafed and CNN-learned features. The final prediction score is a weighted average of the outputs of all the classifiers.

Our approach differs from the NEC system in two key aspects. First, we perform classification via CNN and handcrafted features separately, only using their combination to deal with confounders. Simply stacking handcrafted and CNN features will bias the classifier towards the feature set with the larger number of attributes. Our approach is less prone to this issue. Secondly, CNN works on a 80×80 patch size while handcrafted features are extracted from clusters of segmented nuclei (normally $\leqslant 30 \times 30$). This way we capture attributes of not only mitotic nuclei, but also its local context. Local context around candidate mitoses is an important factor for pathologists in correctly identifying mitoses. In summary, key novel contributions of this work include:

- A cascaded approach for combination of CNN and handcrafted features,
- Learning multiple attributes that characterize mitosis via the combination of CNN and handcrafted features,
- Achieving a high level of mitosis detection while minimizing the computing resources required.

The organization of the rest of this paper is as follows. In Section 2 we describe details of the new methodology. In Section 3 we present experimental results. Finally, in Section 4 we present our concluding remarks.

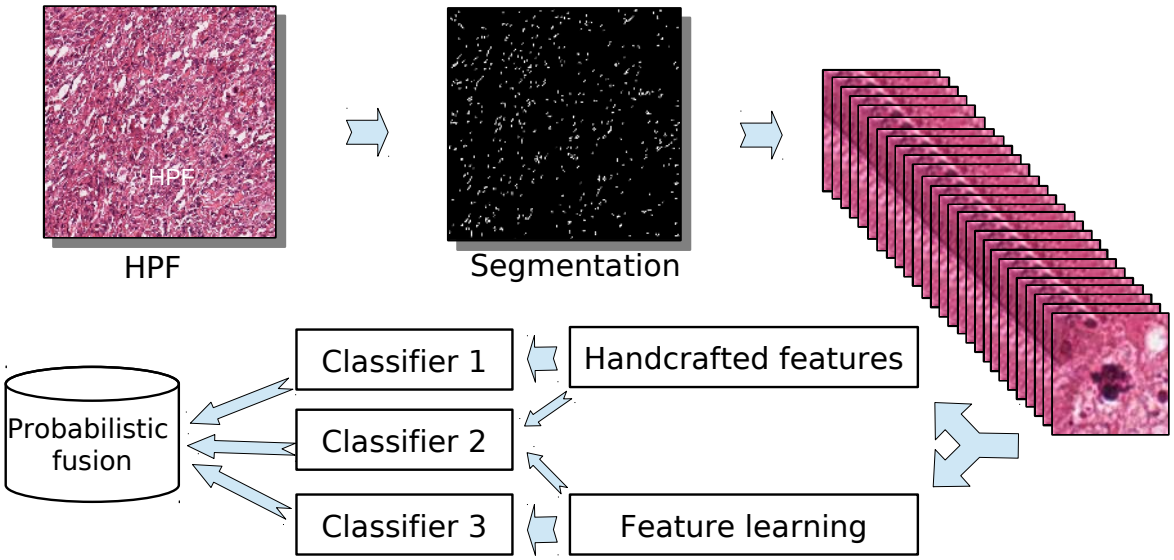

Figure 1: Workflow of our methodology. Blue-ratio thresholding[14] is first applied to segment mitosis candidates. On each segmented blob, handcrafted features are extracted and classified via a Random Forests classifier. Meanwhile, on each segmented 80×80 patch, convolutional neural networks (CNN)[8] are trained with a fully connected regression model as part of the classification layer. For those candidates that are difficult to classify (ambiguous result from the CNN), we train a second-stage Random Forests classifier on the basis of combining CNN-derived and handcrafted features. Final decision is obtained via a consensus of the predictions of the three classifiers.

2. METHODOLOGY

2.1 Candidate Segmentation

We segment likely mitosis candidates by first converting RGB images into blue-ratio images,[14] in which a pixel with a high blue intensity relative to its red and green components is given a higher value. Laplacian of Gaussian (LoG)[15] responses are then computed to discriminate the nuclei region from the background, followed by integrating globally fixed thresholding and local dynamic thresholding to identify candidate nuclei.

2.2 Detection with Convolutional Neural Networks

2.2.1 CNN architecture

First, each HPF is converted from the RGB space to the YUV space and normalized to a mean of zero and variance of one. The CNN architecture employs 3 layers: two consecutive convolutional and pooling layers and a final fully-connected layer. The convolution layer applies a 2D convolution of the input feature maps and a convolution kernel. The pooling layer applies a L2 pooling function over a spatial window without overlapping (pooling kernel) per each output feature map. Learning invariant features will be allowed through the L2 pooling. The output of the pooling layer is subsequently fed to a fully-connected layer, which produces a feature vector. The outputs of the fully-connected layer are two neurons (mitosis and non-mitosis) activated by a logistic regression model. The 3-layer CNN architecture comprises 64, 128, and 256 neurons, respectively. For each layer, a fixed 8×8 convolutional kernel and 2×2 pooling kernel were used.

2.2.2 Training stage

To deal with class-imbalance and achieve rotational invariance, candidate image patches containing mitotic nuclei were duplicated with artificial rotations and mirroring. The whole CNN model was trained using Stochastic Gradient Descent[16] to minimize the loss function: $L(x) = -log\left[\frac{e^{x_i}}{\sum_j e^{x_j}}\right]$, where x_i corresponds to outputs of a fully-connected layer multiplied by logistic model parameters. Thus the outputs of CNN are the log likelihoods of class membership.

2.2.3 Testing stage

An exponential function is applied to the log likelihoods of each candidate nucleus belonging to the positive (mitosis) class in order to calculate the probability that it is mitotic. In our experiments, a candidate nucleus is classified as mitosis if the probability is larger than an empirically-determined threshold of 0.58.

2.3 Detection with handcrafted features

2.3.1 Features and Their Selection

The handcrafted features can be categorized into three groups: morphology, intensity and texture (Table 1). The morphological features are extracted from binary mask of mitosis candidate, which is generated by blue-ratio thresholding[14] and local non-maximum suppression. The morphological features represent various attributes of mitosis shape. Intensity and textural features are extracted from seven distinctive channels of squared candidate patches (Blue-ratio, Red, Blue, Green, L in LAB and V, L in LUV) according to.[4] The intensity features capture statistical attributes of mitosis intensity and the texture features capture textural attributes of mitosis region. The total length of handcrafted features is $15 + 8 \times 7 + 26 \times 7 = 253$. We then perform dimensionality reduction with principal component analysis (PCA).[17] The best features are retained in PCA by keeping 98% of the total component variations.

2.3.2 Class Balancing and Classifier

We correct for the classification bias that occurs due to the relatively small number of mitotic nuclei compared to non-mitotic nuclei. To train a balanced classifier, we (1) reduce non-mitotic nuclei by replacing overlapping non-mitotic nuclei with their clustered center; (2) oversample mitotic cells by applying the Synthetic Minority Oversampling Technique (SMOTE),[18] and (3) use an empirically-selected threshold 0.58. For classification, a Random Forest classifier with 50 trees is used. Using more trees tends to cause overfitting while using less trees leads to low classification accuracy.

Category	Length	Features
Morphology	15	Area, eccentricity, equiv diameter, euler number, extent, perimeter, solidity, major axis length, minor axis length, area overlap ratio, average radial ratio, compactness, hausdorff dimension, smoothness and standard distance ratio.
Intensity	8×7	Mean, median, variance, maximum/minimum ratio, range, interquartile range, kurtosis and skewness of patch intensities at 7 color channels.
Texture	26×7	Concurrence features: mean and standard deviation of 13 Haralick gray-level concurrence features grabbed at four orientations; Run-Length features: mean and standard deviation of gray-level run-length matrices at four orientations;

Table 1: Brief description of handcrafted features used for mitosis detection.

2.4 Cascaded Ensemble

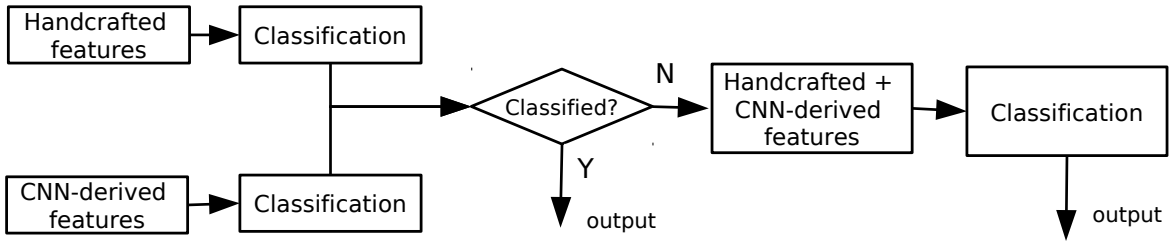

Figure 2: The workflow of cascaded ensemble.

The cascaded ensemble consists of two stages (shown in Fig. 2). First, we perform classification with CNN and handcrafted features individually. During training, let us assume that predicted labels are L_d and L_h, respectively. For instances with $L_d \neq L$ or $L_h \neq L$, where L is the ground truth label, we combine their CNN and handcrafted features to

train a second-stage classifier $\hbar$. During testing, given the output probabilities P_d and P_h of CNN and handcrafted feature classifiers, respectively, we calculate their combined probabilities $P = w_d P_d + w_h P_h$, where w_d and w_h are weighting factors. In the second stage, for instances with $P \in [\lambda_l, \lambda_u]$ (λ_l and λ_u are certain lower and upper bounds, respectively), we let $\hbar$ classify them again. The instance having a final probability p larger than a certain threshold is categorized as mitosis, otherwise, non-mitosis.

3. EXPERIMENTAL RESULTS

3.1 Dataset

3.1.1 ICPR12 dataset

The dataset includes 50 images corresponding to 50 high-power fields (HPF) in 5 different biopsy slides stained with hematoxylin and eosin. 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 widely-used Aperio XT scanner. The Aperio scanner has a resolution of $0.2456 \mu m$ per pixel, resulting in a 2084×2084 RGB image for each field. A total of 326 mitotic nuclei are manually annotated by expert pathologist. The centroids of these mitoses are used as ground truth. According to the test, the first 35 HPF images (226 mitosis) are used for training, while the remaining 15 HPF images (100 mitosis) for evaluation.

3.1.2 AMIDA13 Dataset

The AMIDA13 dataset[2] was released for the MICCAI'13 Grand Challenge on Mitosis Assessment. Slices of 23 breast cancer cases are digitalized by a Aperio ScanScope XT scanner with a $40\times$ magnification. They are then split into individual high power fields (HPF) of a size of 2000×2000 pixels. A total of 1157 mitosis are annotated by two pathologists on all the HPFs. For the challenge, the 23 cases were split onto 2 groups: 12 for training and 11 for testing. A detection is considered a true positive if its distance to a ground truth location is less than $7.5 \mu m$ (30 pixels). All detections that are not within $7.5 \mu m$ of a ground truth location are counted as false positives. And correspondingly, all ground truth locations that do not have a detection within $7.5 \mu m$ are counted as false negatives.

3.2 Performance Measures

Evaluation is performed according to the ICPR 2012 contest criteria, where true positives (TP) are defined as detected mitoses whose coordinates are closer than $5 \mu m$ (20.4 px) to the ground truth centroid. Nuclei that do not meet this criteria are defined as false positive (FP) and false negative (FN) errors. We compute the following performance measures:

$$Recall = \frac{TP}{TP + FN},\ Precision = \frac{TP}{TP + FP},\ F - measure = \frac{2 \times Precision \times Recall}{Precision + Recall}. \quad (1)$$

We compare the proposed approach (**HC+CNN**) with approaches of using handcrafted features only (**HC**), using CNN only (**CNN**), as well as the reported approaches in.[3]

3.3 Results on ICPR12 Dataset

The mitosis detection results on ICPR12 dataset are shown in Table 4reftab:icprrank. The HC+CNN approach yields a higher F-measure (0.7345) than all other methods except that of IDSIA (0.7821). The FN of HC+CNN is relatively high partially because 7 mitoses were not detected during blue-ratio segmentation. In addition, HC+CNN outperforms NEC (F-measure=0.6592), the only other approach to combine CNN and handcrafted features. Note that CNN based approaches (HC+CNN, IDSIA and NEC) tend to produce fewer FP errors, reflecting the capacity of CNN to accurately recognize non-mitotic nuclei.

Figure 3 shows some detected mitosis examples. As one can see, the FNs tend to be poorly colored and textured while the FPs have similar color and shape attributes compared to the TPs. Although the textural patterns between FPs and TPs are different, this difference is not well appreciated at this pre-specified HPF resolution. Figure 5 show two mitotic detection results of HC+CNN, which also revealing some FN examples. Both the segmentation and detection steps

[2] http://amida13.isi.uu.nl/

contribute to the loss of these mitotic figures. Figure 4 shows a mitosis detection example using CNN and HC+CNN, respectively, revealing the improvement obtained by integrating handcrafted features and CNN in HC+CNN.

The two 11-layers neural networks used by IDSIA[13] requires roughly 30 epochs, which takes two days for training with GPU optimization. Our 3-layer CNN needs less than 10 epochs, and requires only 11.4 hours using 9 epochs without GPU optimization. Including the time needed to extract handcrafted features (6.5 hours in pure MATLAB implementation), the training stage for HC+CNN was completed in less than 18 hours.

Dataset	Method	TP	FP	FN	Precision	Recall	F-measure
Scanner Aperio	**HC+CNN**	65	12	35	0.84	0.65	**0.7345**
	HC	64	22	36	0.74	0.64	0.6864
	CNN	53	32	47	0.63	0.53	0.5730
	IDSIA[13]	70	9	30	0.89	0.70	0.7821
	IPAL[4]	74	32	26	0.70	0.74	0.7184
	SUTECH	72	31	28	0.70	0.72	0.7094
	NEC[12]	59	20	41	0.75	0.59	0.6592

Table 2: Evaluation results for mitosis detection using HC+CNN and comparative methods on the ICPR12 dataset.

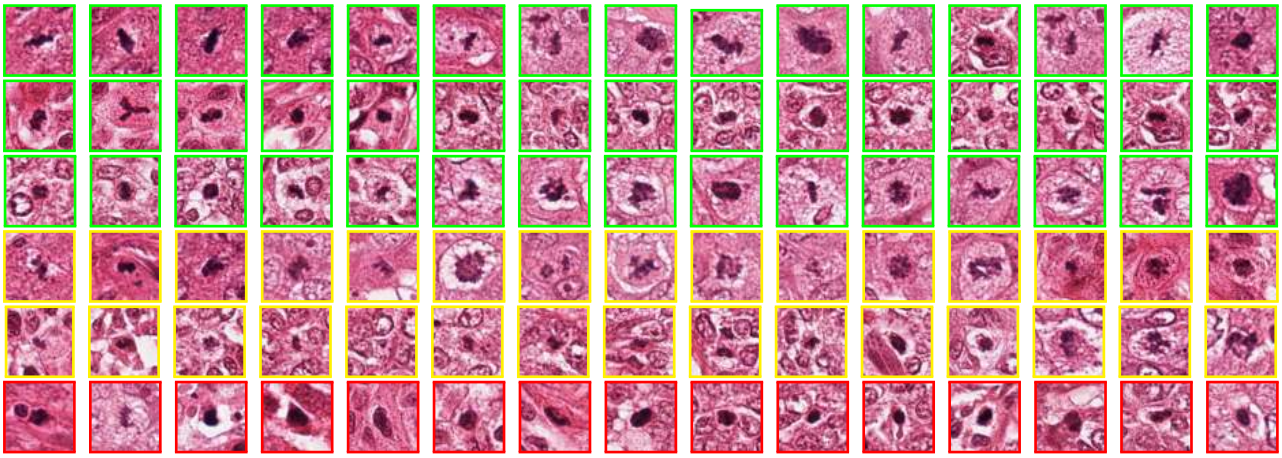

Figure 3: Mitoses identified by HC+CNN as TP (green circles), FN (yellow circles), and FP (red circles) on the ICPR12 dataset. The TP examples have distinctive intensity, shape and texture while the FN examples are less distinctive in intensity and shape. The FP examples are visually more alike to mitotic figures than the FNs.

3.4 Results on AMIDA13 Dataset

On the AMIDA13 dataset, the F-measure of our approach (CCIPD/MINDLAB) is 0.319, which ranks 6 among 14 submissions (shown in Figure 6). The 23 study cases, especially case #3 and #6, have many dark spots that are not mitotic figures. As a result, on these two cases there are many false positives that are clearly apoptotic nuclei, lymphocytes or compressed nuclei. The IDSIA team won this challenge with a F-measure of 0.611, using the same aforementioned CNN models as on the ICPR12 dataset. Note however that there is hardly any difference between the teams that ranked 3-6, in essence all of these teams tying for third place.

Figure 7 shows detection results on two HPF slices. The left HPF has extremely rich dark spots that are not mitotic nuclei but look very similar to mitosis. The existence of these confounder instances tends to increase the false positive hit rate. On the right HPF, non-mitotic nuclei are significantly less but mitotic figures tend to be difficult to identify. Moreover, color differences between the two HPFs increases the difficulty of detecting mitoses on this dataset.

The training time for our approach is about 4 days, which though long is significantly less compared to the training burden of the IDSIA approach. Extracting handcrafted features and training of the CNN model are done in parallel to save time.

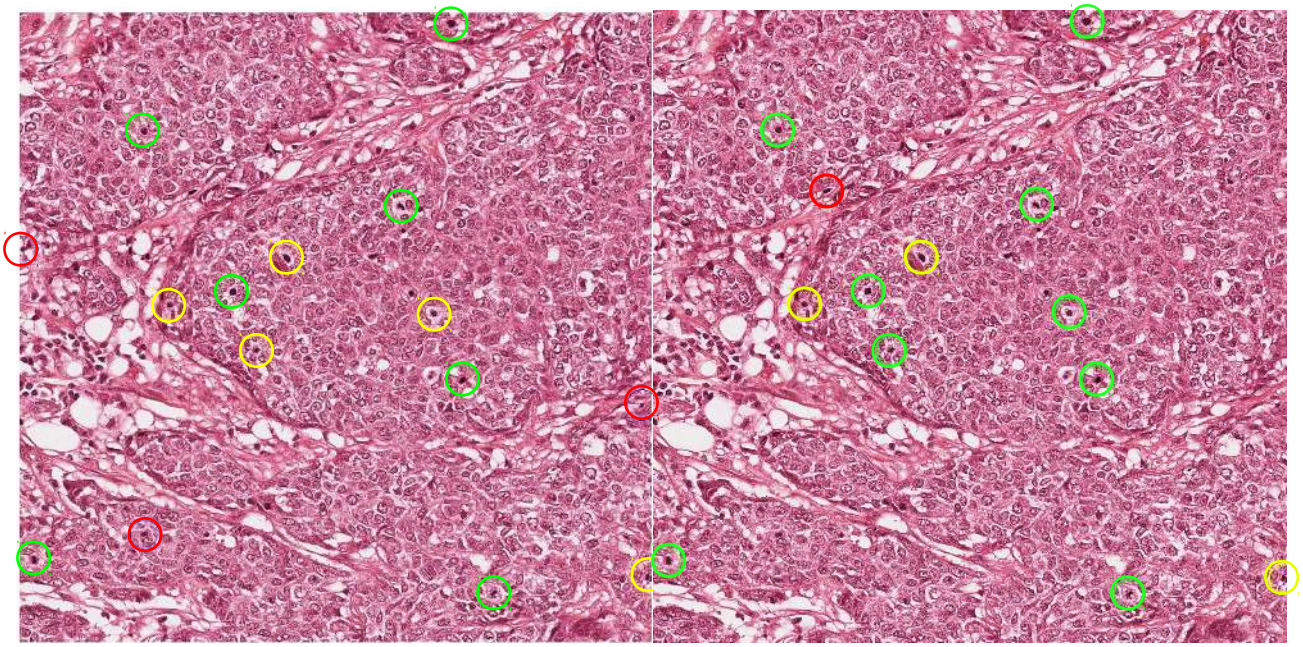

Figure 4: Mitoses identified by CNN and HC+CNN as TP (green circles), FN (yellow circles), and FP (red circles) on a histology slice of ICPR12 dataset. On the left side, only using CNN leads to 7 TPs, 5 FNs and 3 FPs. On the right side, using HC and CNN leads to 9 TPs, 3 FNs and 1 FP, which clearly outperforms the use of CNN alone.

4. CONCLUDING REMARKS

Mitosis detection is one of the three key factors in breast cancer grading. Existing approaches attempt to detect mitosis using either stacked handcrafted features or CNN-learned features. However, the problem of low detection accuracy arises when only handcrafted features are used while CNN-based approaches suffer from the issue of high computational complexity. To tackle these problems, we presented a new approach that combines handcrafted features and a light CNN in a cascaded way. Our approach yields a F-measure of 0.7345, which would have secured the second rank in the ICPR contest, and higher than the NEC approach that combines CNN and handcrafted features at feature level. Compared to the leading methodology (two 11-layer CNN models) at the ICPR contest (F-measure = 0.78), our approach is faster, requiring far less computing resources.

Experiments on the AMIDA13 dataset shows that it is still necessary to improve the accuracy of the presented approach. Future work will use GPU to implement a multi-layer (more than 2) CNN model.

5. ACKNOWLEDGEMENT

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under award numbers R01CA136535-01, R01CA140772-01, and R21CA167811-01; the National Institute of Diabetes and Digestive and Kidney Diseases under award number R01DK098503-02, the DOD Prostate Cancer Synergistic Idea Development Award (PC120857); the QED award from the University City Science Center and Rutgers University, the Ohio Third Frontier Technology development Grant. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

REFERENCES

[1] Genestie, C. e. a., "Comparison of the prognostic value of scarff-bloom-richardson and nottingham histological grades in a series of 825 cases of breast cancer: major importance of the mitotic count as a component of both grading systems," *Anticancer Res* **18**(1B), 571–6 (1998).

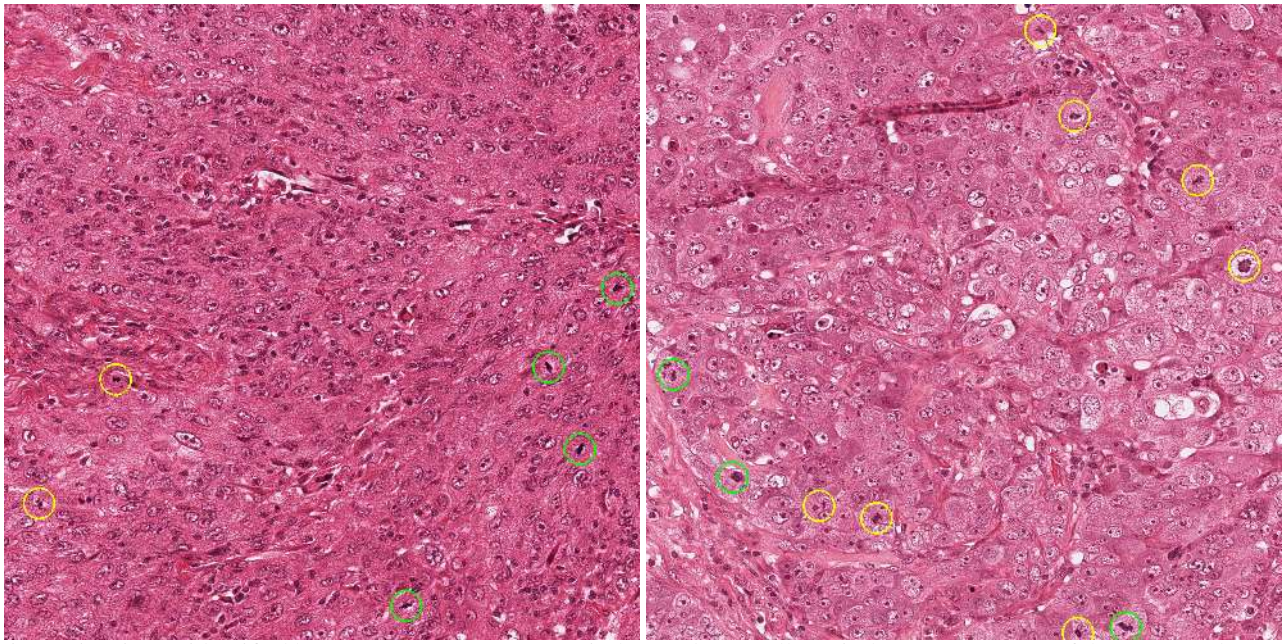

Figure 5: Mitoses identified by HC+CNN as TP (green circles) and FN (yellow circles) on two HPFs of the ICPR12 dataset. Mitoses on the left HPF have distinctive intensities and shapes, and confounding nuclei are few. Therefore, most mitoses can be correctly detected on this HPF. Comparatively, intensity of most mitotic nuclei on the right HPF is not distinctive enough for HC+CNN to identify, as a result, leading to a high FN.

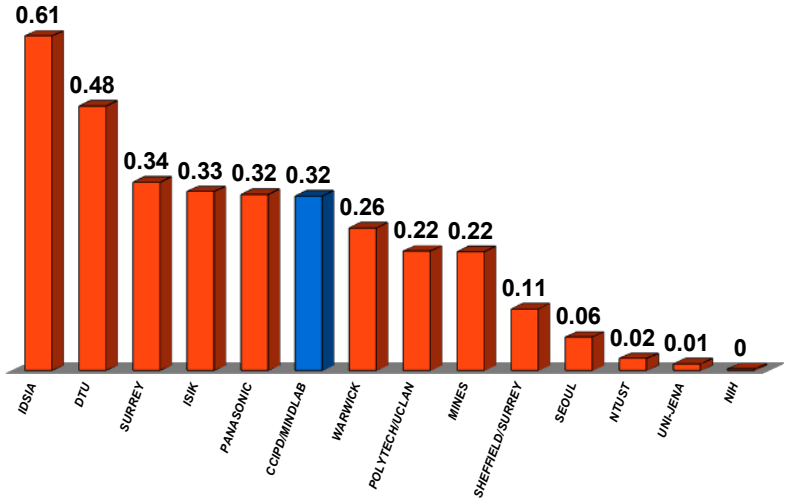

Figure 6: Ranking of MICCAI'13 Grand Challenge on Mitosis according to the overall F-measure. Although our team (CCIPD/MINDLAB) ranks 6 among the 14 participating teams, the F-measure of our approach (blue-bin, F=0.32) is very close to that of the 3rd place (F=0.34). Note that the winner is IDSIA, which also ranked 1 for the ICPR12 dataset.

[2] Basavanhally, A., Ganesan, S., Feldman, M., Shih, N., Mies, C., Tomaszewski, J., and Madabhushi, A., "Multi-field-of-view framework for distinguishing tumor grade in er x002b; breast cancer from entire histopathology slides," *IEEE Transactions on Biomedical Engineering* **60**, 2089–2099 (Aug 2013).

[3] Genestie, C., Racoceanu, D., Capron, F., Naour, G., Irshad, H., Klossa, J., Roux, L., Kulikova, M., Gurcan, M., and Lomnie, N., "Mitosis detection in breast cancer histological images An ICPR 2012 contest," *Journal of Pathology Informatics* **4**(1), 8 (2013).

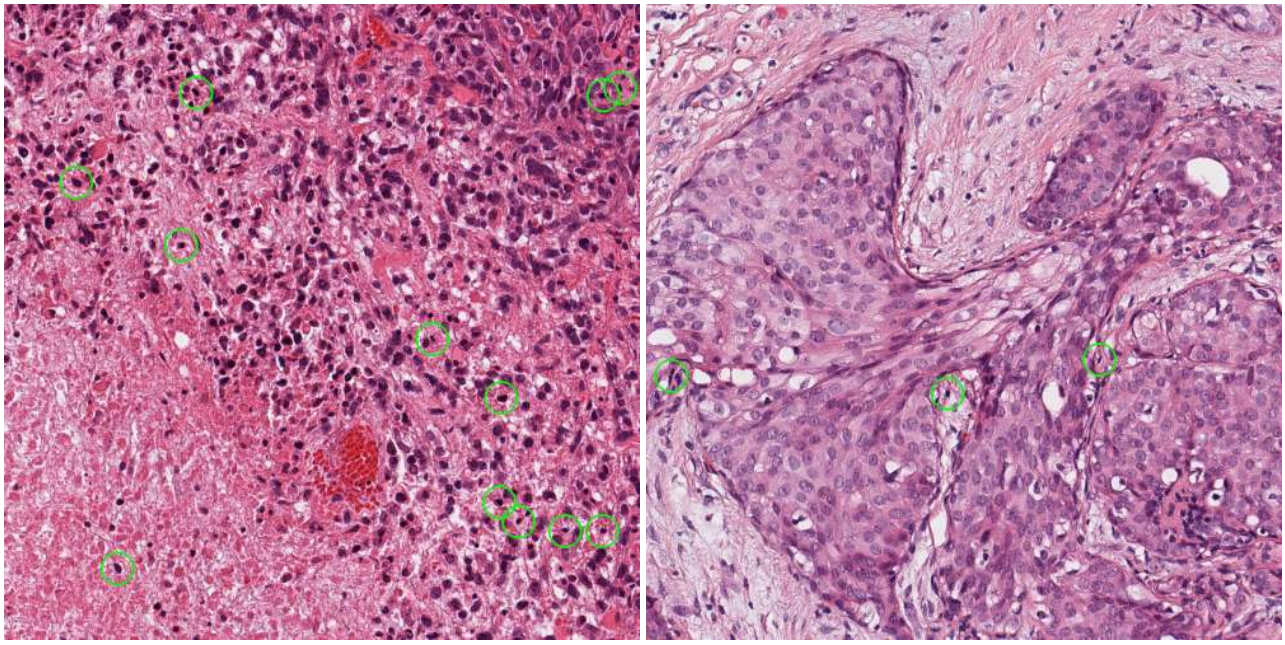

Figure 7: Mitoses (green circles) identified by HC+CNN on two HPFs of the AMIDA13 dataset. The left HPF is rather challenging as it has rich dark spots that are non-mitotic nuclei. The existence of these confusing nuclei makes mitosis detection extremely difficult. On the right HPF, non-mitotic nuclei are significantly less but mitotic figures are difficult to identify. Color difference between the two HPFs further increases the toughness of mitosis detection. Note that we cannot provide TP, FP and FN as groundtruth of the challenge are not publicly available.

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

[5] Sommer, C., Fiaschi, L., Hamprecht, F., and Gerlich, D., "Learning-based mitotic cell detection in histopathological images," in [*International Conference on Pattern Recognition (ICPR)*], 2306–2309 (Nov 2012).

[6] Racoceanu, D., Capron, F., Naour, G., Irshad, H., Hwee, L., Roux, L., and Jalali, S., "Automated mitosis detection using texture, SIFT features and HMAX biologically inspired approach," *Journal of Pathology Informatics* **4**(2), 12 (2013).

[7] Huang, C.-H. and Lee, H.-K., "Automated mitosis detection based on exclusive independent component analysis," in [*International Conference on Pattern Recognition (ICPR)*], 1856–1859 (Nov 2012).

[8] LeCun, Y., Bottou, L., Bengio, Y., and Haffner, P., "Gradient-based learning applied to document recognition," *Proceedings of the IEEE* **86**(11), 2278–2324 (1998).

[9] Cruz-Roa, A., Arevalo Ovalle, J., Madabhushi, A., and Gonzlez Osorio, F., "A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection," in [*Medical Image Computing and Computer-Assisted Intervention MICCAI 2013*], *Lecture Notes in Computer Science* **8150**, 403–410 (2013).

[10] JS, M., C, A., and et al., "Breast carcinoma malignancy grading by bloom-richardson system vs proliferation index: reproducibility of grade and advantages of proliferation index," *Modern Pathology* **18**(8), 1067–1078 (2005).

[11] C, M., E, B., and et al., "Mitotic figure recognition: Agreement among pathologists and computerized detector," *Anal Cell Pathology* , 97–100 (2012).

[12] Malon, C. and Cosatto, E., "Classification of mitotic figures with convolutional neural networks and seeded blob features," *Journal of Pathology Informatics* **4**(1), 9 (2013).

[13] Ciresan, D. C., Giusti, A., Gambardella, L. M., and Schmidhuber, J., "Mitosis detection in breast cancer histology images with deep neural networks," *MICCAI* (2013).

[14] 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).
[15] Haralick, R. and Shapiro, L., "Computer and robot vision," *Addison-Wesley Publishing Company* **1**, 346–351 (1992).
[16] Bottou, L. and Bousquet, O., "The tradeoffs of large scale learning," in [*NIPS*], (2007).
[17] Pearson, K., "On lines and planes of closest fit to systems of points in space," *Philosophical Magazine* **2**, 559–572 (1901).
[18] Chawla, N. V., Bowyer, K. W., Hall, L. O., and Kegelmeyer, W. P., "Smote: synthetic minority over-sampling technique," *Journal of Artificial Intelligence Research* **16**(1), 321–357 (2002).