Journal of Medical Imaging

MedicalImaging.SPIEDigitalLibrary.org

Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features

Haibo Wang
Angel Cruz-Roa
Ajay Basavanhally
Hannah Gilmore
Natalie Shih
Mike Feldman
John Tomaszewski
Fabio Gonzalez
Anant Madabhushi

SPIE.

Mitosis detection in breast cancer pathology images by combining handcrafted and convolutional neural network features

Haibo Wang,[a,*] Angel Cruz-Roa,[b] Ajay Basavanhally,[a] Hannah Gilmore,[a] Natalie Shih,[c] Mike Feldman,[c] John Tomaszewski,[d] Fabio Gonzalez,[b] and Anant Madabhushi[a]
[a]Case Western Reserve University, Center for Computational Imaging and Personalized Diagnostics, 2071 Martin Luther King Jr. Drive, Cleveland, Ohio 44106, United States
[b]Universidad Nacional de Colombia, Aulas de Ingenieria, MINDLab, 114 Edif. 453, Bogota, Colombia
[c]Hospital of the University of Pennsylvania, 3400 Spruce Street, Philadelphia, Pennsylvania 19104, United States
[d]University at Buffalo, School of Medicine and Biomedical Sciences, Buffalo, New York 14214, United States

Abstract. Breast cancer (BCa) grading plays an important role in predicting disease aggressiveness and patient outcome. A key component of BCa grade is the 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 (HPFs) 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). Although 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 supervised feature generation methods, there is an appeal in 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. 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 the performance by leveraging the disconnected feature sets. Evaluation on the public ICPR12 mitosis dataset that has 226 mitoses annotated on 35 HPFs (400× magnification) by several pathologists and 15 testing HPFs yielded an F-measure of 0.7345. Our approach is accurate, fast, and requires fewer computing resources compared to existent methods, making this feasible for clinical use.
 [DOI: 10.1117/1.JMI.1.3.034003]

Keywords: mitosis; breast cancer; convolutional neural networks; cascaded ensemble; handcrafted feature; digital pathology.

Paper 14061PRR received May 16, 2014; revised manuscript received Sep. 14, 2014; accepted for publication Sep. 16, 2014; published online Oct. 10, 2014.

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 (HPFs, 400× magnification). However, the manual identification of mitotic nuclei often suffers from poor inter-interpreter 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.[4]

The detection of mitotic nuclei in H&E stained histopathology is a difficult task (see Fig. 1). First, mitosis is a complex biological process during which the cell nucleus undergoes various morphological transformations. This leads to highly variable sizes and shapes 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., nonmitotic nuclei). In this paper, we present a new automatic mitosis detection approach to address the aforementioned challenges and

*Address all correspondence to: Haibo Wang, E-mail: hbwang1427@gmail.com

0091-3286/2014/$25.00 © 2014 SPIE

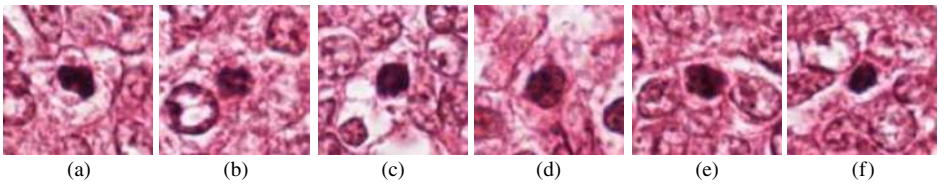


Fig. 1 An illustration of the visual similarity between true mitotic processes and confounding false positives. (a)–(c) True mitoses. (d)–(f) Confounding nonmitotic figures.

outperform the majority of state-of-the-art approaches in mitosis detection.

The organization of the rest of this paper is as follows. In Sec. 2, we describe motivations of the proposal. In Sec. 3, we describe details of the new methodology. In Sec. 4, we present experimental results. Finally, in Sec. 5, we present our concluding remarks.

2 Motivation and Previous Work

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[5,6] to identify the best automated mitosis detection algorithm, a variety of approaches competed against each other. These approaches 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.[7–10]

Although 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),[11,12] CNN are multilayer neural networks that learns a bank of convolutional filters at each layer.[13,14] In contrast to handcrafted features, CNN is fully data-driven, therefore, it is more accurate in representing training samples and is able to find feature patterns that handcrafted features fail to describe. However, CNN is computationally demanding and sensitive to the scalability of the training data. The winner[14] 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 (several weeks) for both training and testing.

Other methods achieved an F-measure of up to 0.71, based primarily on combining various handcrafted features. Although handcrafted feature approaches are faster, drawbacks include (1) the fact that the identification of salient features is 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 training deep CNN models. 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 Nippon Electric Company (NEC) team,[13] 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 combinations 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 Fig. 2. The first step is to

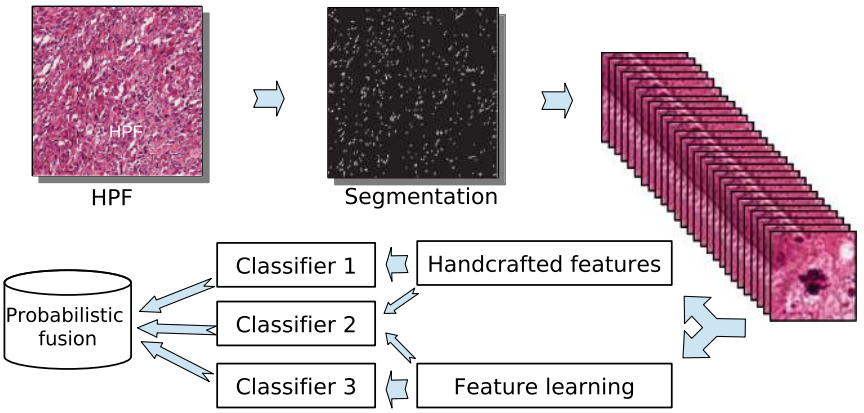


Fig. 2 Workflow of our methodology. Blue-ratio thresholding[15] is first applied to segment mitosis candidates. On each segmented blob, handcrafted features are extracted and classified via a random forests classifier. Meanwhile, on each segmented 80×80 patch, convolutional neural networks (CNN)[11] 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.

segment likely mitosis regions. This initial phase serves as a triage to remove obviously nonmitotic 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 handcrafted 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 toward the feature set with the larger number of attributes. Our approach is less prone to this issue. Second, CNN works on a 80×80 pixels patch size while handcrafted features are extracted from clusters of segmented nuclei (normally $\leq 30 \times 30$ pixels). This way we capture attributes of not only mitotic nuclei, but also of the local context. Local context around candidate mitoses is an important factor for pathologists in correctly identifying mitoses. In summary, key 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.

3 Methodology

3.1 *Candidate Segmentation*

We segment likely mitosis candidates by first converting RGB images into blue-ratio images.[15] By assigning a higher value to a pixel with a high blue intensity relative to its red and green components, blue-ratio is proven capable of highlighting nuclei regions.[15] Laplacian of Gaussian[16] 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. One segmentation example is shown in Fig. 3. We can see that most dark-blue spots are retained as potential mitotic figures.

3.2 *Detection with Convolutional Neural Networks*

3.2.1 *CNN architecture*

First, each HPF is converted from the RGB space to the YUV space and normalized to a mean of 0 and variance of 1. The CNN architecture employs three layers (Fig. 4): two consecutive convolutional and pooling layers and a final fully connected layer. The convolution layer applies a two-dimensional convolution of the input feature maps and a convolution kernel. The pooling layer applies an 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 nonmitosis) activated by a logistic regression model.

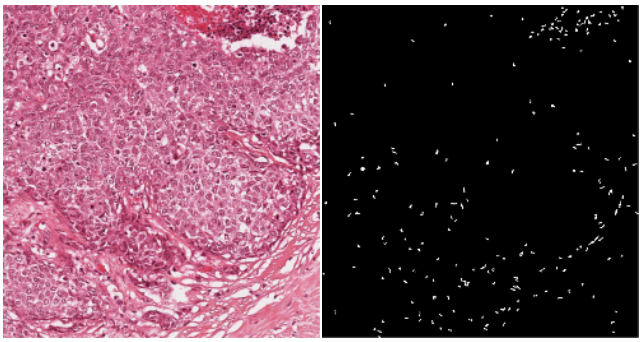


Fig. 3 Example of blue-ratio segmentation. (a) is the original high power field (HPF) slice while (b) is the segmentation mask. Note that a majority of the objects identified via this approach in (b) are indeed mitotic figures.

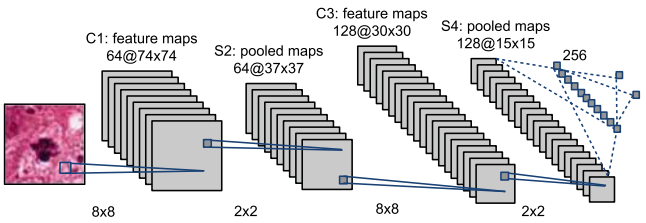


Fig. 4 Architecture of the CNN model. The CNN architecture comprises three layers: two consecutive convolutional-pooling layers and a fully connected classification layer. The two convolutional-pooling layers use the same fixed 8×8 convolutional kernel and 2×2 pooling kernel, but have 64 and 128 neurons, respectively. The last layer has 256 neurons, which are all connected to the final two neurons for mitosis/nonmitosis classification.

The three-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.

3.2.2 *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[17] 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.

3.2.3 *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.

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

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 seven color channels
Texture	26×7	Concurrence features: mean and standard deviation of 13 Haralick[19] gray-level concurrence features grabbed at four orientations. Run-length features:[20] mean and standard deviation of gray-level run-length matrices at four orientations

3.3 *Detection with Handcrafted Features*

3.3.1 *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 the binary mask of the mitosis candidate, which is generated by blue-ratio thresholding[15] and local nonmaximum 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 Ref. 7. The intensity features capture statistical attributes of mitosis intensity and the texture features capture textural attributes of the 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).[18] The best features are retained in PCA by keeping 98% of the total component variations.

3.3.2 *Class balancing and classifier*

We correct for the classification bias that occurs due to the relatively small number of mitotic nuclei compared to nonmitotic nuclei. To train a balanced classifier, we (1) reduce nonmitotic nuclei by replacing overlapping nonmitotic nuclei with their clustered center; (2) oversample mitotic cells by applying the synthetic minority oversampling technique,[21] 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.

3.4 *Cascaded Ensemble*

The cascaded ensemble consists of two stages (shown in Fig. 5). First, we perform classification with CNN and handcrafted features individually. During training, we denote via L_d and L_h the classification labels associated with using CNN and handcrafted features, 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, as nonmitosis.

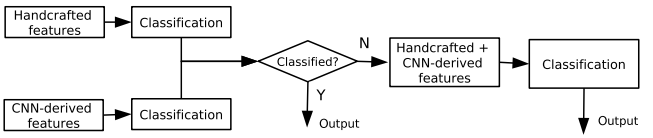


Fig. 5 Workflow of the cascaded ensemble, which comprises two stages. First, we perform classification with CNN-learned and handcrafted features individually, and if the two classification scores are consistent, a binary decision (mitosis/nonmitosis) will be made directly. Second, for those instances whose individual classification scores are highly inconsistent, we classify them again by combining their CNN and handcrafted features.

4 Experimental Results

4.1 *ICPR Dataset*

The dataset includes 50 images corresponding to 50 HPF in five different biopsy slides stained with H&E (illustrated in Fig. 6). Each field represents a 512×512 μm^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 μm/pixel, resulting in a 2084×2084 pixels RGB image for each field. A total of 326 mitotic nuclei are manually annotated by an 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) are used for evaluation.

4.2 *Performance Measures*

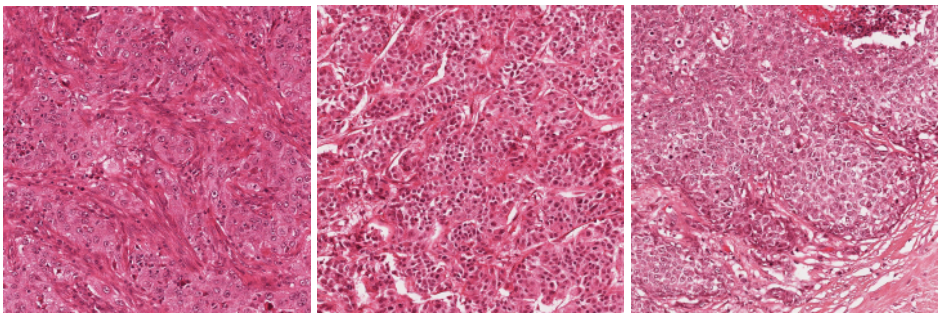
Evaluation is performed according to the ICPR 2012 contest criteria, where true positives (TP) are defined as detected mitoses whose coordinates are closer than 5 μm (20.4 pixel) to the ground truth centroid. Nuclei that do not meet this criteria are defined as false positive (FP) and false negative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

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

We compare the proposed approach (HC + CNN) with approaches using handcrafted features only (HC), using CNN only (CNN), as well as the reported approaches in Ref. 5.

4.3 *Results*

The mitosis detection results on the ICPR12 dataset are shown in Table 2. The HC + CNN approach yields a higher F-measure (0.7345) than all other methods except that of Istituto Dalle Molle di Studi sull'Intelligenza Artificiale (IDSIA) (0.7821).



(a) (b) (c)

Fig. 6 Hematoxylin and eosin-stained HPF examples from the ICPR dataset. The HPFs are acquired by a Aperio XT scanner with a resolution of 0.2456 μm per pixel. Each HPF has a size of 2084×2084 pixels, representing a 512×512 μm^2 area annotated by pathologists.

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

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[14]	70	9	30	0.89	0.70	0.7821
	IPAL[7]	74	32	26	0.70	0.74	0.7184
	SUTECH	72	31	28	0.70	0.72	0.7094
	NEC[13]	59	20	41	0.75	0.59	0.6592

Note: The bold value highlights the result of the proposed HC + CNN approach.

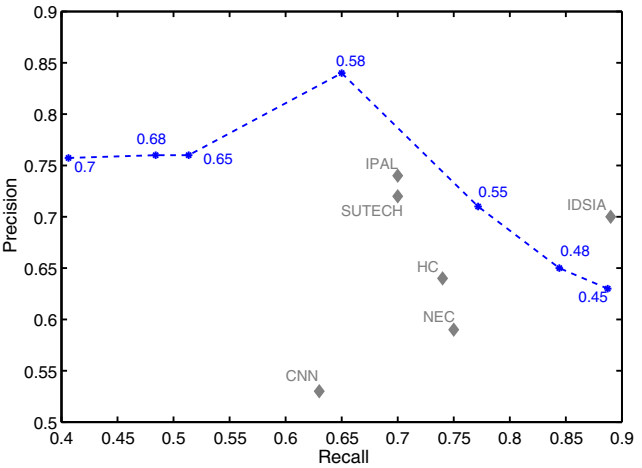


Fig. 7 Precision–recall curve of the proposed HC + CNN method. The performance of the other methods is also plotted for comparison. The curve is generated by varying the classification threshold between 0.45 and 0.7. (The threshold for each point is marked along the curve.) The fact that the performance of the other methods (except IDSIA) lie in the interior of the areas spanned by the curve suggests that the performance of HC + CNN is resilient to the precise choice of the classification threshold.

The FN rate associated with HC + CNN is relatively high compared to other methods. As Table 3 illustrates, this is partially due to the fact that the blue-ratio segmentation has an FN error of seven mitoses. 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 nonmitotic nuclei.

The most critical parameter of the HC + CNN classifier is the classification threshold that is used to decide mitosis/nonmitosis. Based off our empirical results, the optimal threshold was identified to be ≈0.6. In general, a larger threshold will lead to fewer TPs, FPs and more FNs, and vice versa. In order to evaluate the influence of this threshold parameter, we generate the precision-recall curves by varying the threshold from 0.45 to 0.7. Figure 7 shows that the performances of the other methods (except IDSIA) lie in the interior of the areas spanned by the curve. This fact suggests that the performance of HC + CNN

Table 3 Performances of the blue-ratio segmentation module and the detection module. The blue-ratio segmentation finds 2484 mitosis candidates, among which 93 are true mitoses while the other 2391 are nonmitoses. Seven true mitoses are lost in this step. The detection module identifies 65 true mitoses and 12 false mitoses from these 2484 candidates. Twenty-eight mitoses are misclassified as nonmitotic figures in this module.

Segmentation module			Detection module			Final		
TP	FP	FN	TP	FP	FN	TP	FP	FN
93	2391	7	65	12	28	65	12	35

Table 4 The influence of the number of RF trees.

Number of trees	10	20	30	50	100	200
F-measure	0.57	0.57	0.67	0.7345	0.65	0.66

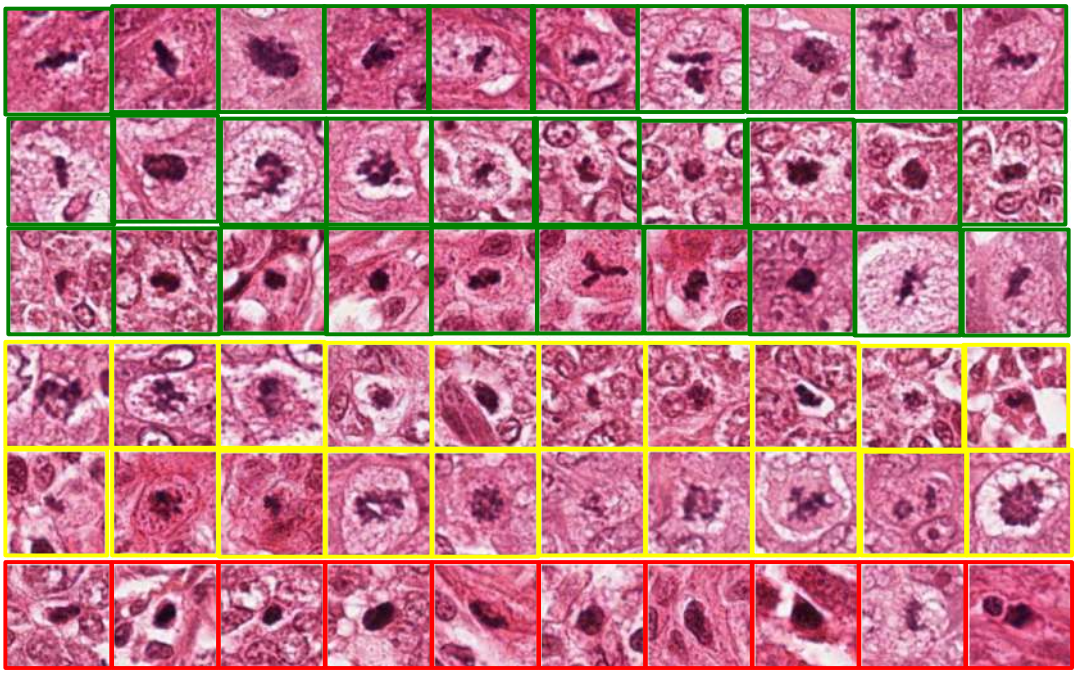


Fig. 8 Mitoses identified by HC + CNN as TP (green rectangles), FN (yellow rectangles), and FP (red rectangles) on the ICPR12 dataset. The TP examples have distinctive intensity, shape, and texture while the FN examples are less distinctive in intensity and shape. The FP examples are visually more alike to mitotic figures than the FNs.

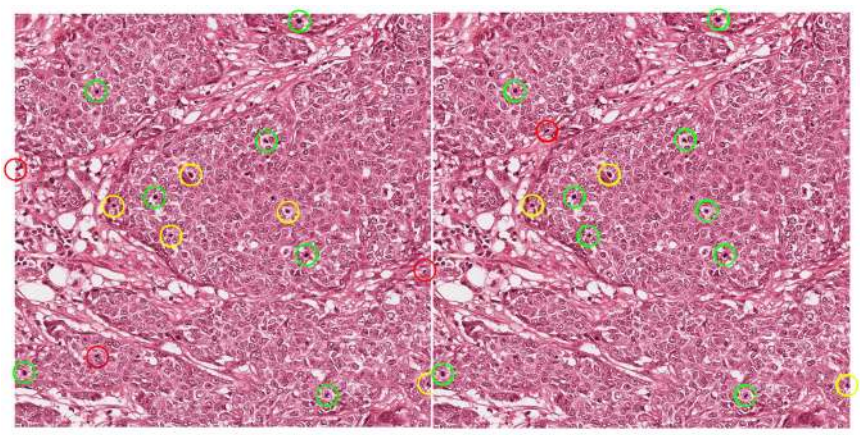


Fig. 9 Mitoses identified by CNN and HC + CNN as TP (green circles), FN (yellow circles), and FP (red circles) on a HPF of ICPR12 dataset. (a) Only using CNN leads to 7 TPs, 5 FNs and 3 FPs. (b) Using HC and CNN leads to 9 TPs, 3 FNs and 1 FP, which clearly outperforms the use of CNN alone.

is resilient to the precise choice of the classification threshold. Table 4 shows the influence of the number of random forests trees on mitosis detection. We can clearly see that fewer trees will most likely lead to worse classification, while more trees may cause overfitting.

Figure 8 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 prespecified HPF resolution. Figure 9 shows a mitosis detection example using CNN and HC + CNN, respectively, revealing the improvement obtained by integrating handcrafted features and CNN in HC + CNN. Figure 10 shows two mitotic detection results of HC + CNN, while also revealing some FN examples. Both the segmentation and detection steps contribute to the loss of these mitotic figures.

The two 11-layers neural networks used by IDSIA[14] requires roughly 30 epochs, which takes 2 days for training with GPU optimization. Our three-layer CNN needs <10 epochs, and requires only 11.4 h using nine epochs without GPU optimization. Including the time needed to extract handcrafted features (6.5 h in pure MATLAB implementation), the training stage for HC + CNN was completed in <18 h. At the detection stage, the

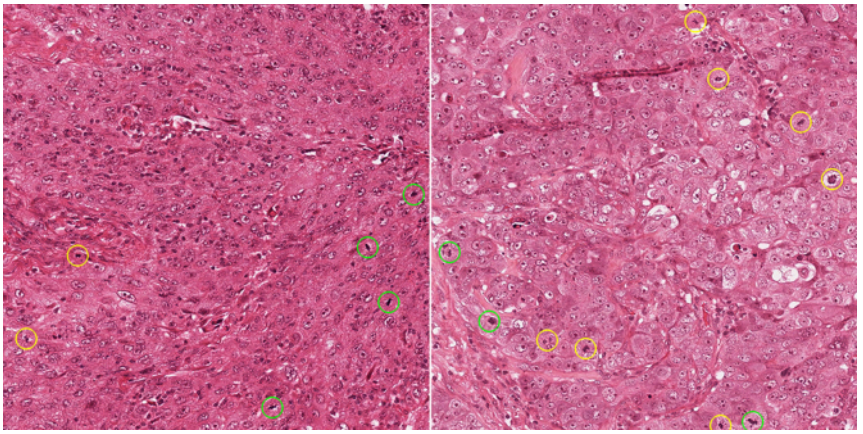


Fig. 10 Mitoses identified by HC + CNN as TP (green circles) and FN (yellow circles) on two HPFs of the ICPR12 dataset. Mitoses on the left HPF have distinctive intensities and shapes, and confounding nuclei are few. Therefore, most mitoses can be correctly detected on this HPF. Comparatively, intensity 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.

MATLAB implementation of HC + CNN takes about 1.5 min to process each H&E image, which is roughly 5× faster than the winner of the ICPR challenge.[14]

5 Concluding Remarks

Mitosis detection is one of the three key factors in BCa 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 an F-measure of 0.7345, which would have secured the second rank in the ICPR contest, and is higher than the NEC approach that combines CNN and handcrafted features at the 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.

Experimental results shows that it is still necessary to improve the accuracy of the presented approach. Future work will use a GPU to implement a multilayer (>3) CNN model.

Acknowledgments

Research reported in this publication was supported by the National Cancer Institute of the National Institutes of Health under Award Nos. R01CA136535-01, R01CA140772-01, NIH 1R21CA179327-01A1, and R21CA167811-01; the National Institute of Diabetes and Digestive and Kidney Diseases under Award No. R01DK098503-02, the DOD Prostate Cancer Synergistic Idea Development Award (PC120857), the DOD CDMRP Lung Cancer Research Idea Development Award New Investigator (LC130463); the QED award from the University City Science Center and Rutgers University, the Ohio Third Frontier Technology development grant, and the CTSC Coulter Annual Pilot grant. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health. Angel Cruz-Roa is supported by the doctoral Grant from Administrative Department of Science, Technology and Innovation of Colombia (Colciencias) (No. 528/2011).

References

1. H. J. Bloom and W. W. Richardson, "Histological grading and prognosis in breast cancer: a study of 1409 cases of which 359 have been followed for 15 years," *Br. J. Cancer* **11**(3), 359–377 (1957).
2. C. W. Elston and I. O. Ellis, "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 (1991).
3. C. Genestie et al., "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), 571–576 (1998).
4. M. N. Gurcan et al., "Histopathological image analysis: a review," *IEEE Rev. Biomed. Eng.* **2**, 147–171 (2009).
5. L. Roux et al., "Mitosis detection in breast cancer histological images: an ICPR 2012 contest," *J. Pathol. Inf.* **4**(1), 8–14 (2013).
6. http://ipal.cnrs.fr/ICPR2012/.
7. H. Irshad, "Automated mitosis detection in histopathology using morphological and multi-channel statistics features," *J. Pathol. Inf.* **4**(1), 10–15 (2013).
8. C. Sommer et al., "Learning-based mitotic cell detection in histopathological images," in *Int. Conf. on Pattern Recognition (ICPR)*, pp. 2306–2309, IEEE, Tsukuba, Japan (2012).
9. H. Irshad et al., "Automated mitosis detection using texture, SIFT features and HMAX biologically inspired approach," *J. Pathol. Inf.* **4**(2), 12–18 (2013).
10. C.-H. Huang and H.-K. Lee, "Automated mitosis detection based on eXclusive Independent Component Analysis," in *Int. Conf. on Pattern Recognition (ICPR)*, pp. 1856–1859, IEEE, Tsukuba, Japan (2012).
11. Y. LeCun et al., "Gradient-based learning applied to document recognition," *Proc. IEEE* **86**(11), 2278–2324 (1998).
12. A. A. Cruz-Roa et al., "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)*, Vol. 8150, pp. 403–410, Springer, Nagoya, Japan (2013).
13. C. Malon and E. Cosatto, "Classification of mitotic figures with convolutional neural networks and seeded blob features," *J. Pathol. Inf.* **4**(1), 9–13 (2013).
14. D. C. Ciresan et al., "Mitosis detection in breast cancer histology images with deep neural networks," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, Springer, Nagoya, Japan (2013).
15. H. Chang, L. Loss, and B. Parvin, "Nuclear segmentation in H&E sections via multi-reference graph cut (MRGC)," in *Proc. of the Sixth IEEE Int. Conf. on Symposium on Biomedical Imaging (ISBI'2012)*, IEEE, New York (2012).
16. R. M. Haralick and L. G. Shapiro, *Computer and Robot Vision*, Vol. **1**, pp. 346–351, Addison-Wesley Publishing Company (1992).

17. L. Bottou and O. Bousquet, "The tradeoffs of large scale learning," in *Conference on Neural Information Processing Systems*, Curran Associates, Inc., New York (2007).
18. K. Pearson, "On lines and planes of closest fit to systems of points in space," *Philos. Mag.* **2**, 559–572 (1901).
19. R. Haralick, K. Shanmugam, and I. Dinstein, "Textural features for image classification," *IEEE Trans. Syst. Man Cybern.* **3**(6), 610–621 (1973).
20. M. M. Galloway, "Texture analysis using gray level run lengths," *Comput. Graphics Image Process.* **4**(2), 172–179 (1975).
21. N. V. Chawla et al., "SMOTE: synthetic minority over-sampling technique," *J. Artif. Intell. Res.* **16**(1), 321–357 (2002).

Haibo Wang is a senior research associate at Case Western Reserve University since 2012. He received his PhD degree in computer vision at the Institute of Automation, Chinese Academy of Sciences, in 2011, and his joint PhD degree at the University of Science and Technology of Lille in 2010. He has authored over 20 publications in TCSVT, TMI, IJIM, and MICCAI. His research interests include computer vision, machine learning, healthcare data processing, and virtual reality.

Angel Cruz-Roa is currently a PhD student in systems and computing engineering in MindLab Group, National University of Colombia, Bogota (Colombia). He received his BS degree in systems engineering from University of the Llanos, Villavicencio (Colombia), in 2005, and his MS degree in biomedical engineering from National University of Colombia, Bogota (Colombia), in 2011. His research interests include image analysis and understanding, visual mining, machine learning, and digital pathology.

Ajay Basavanhally is a research associate at Department of Biomedical Engineering, Case Western Reserve University since 2014. He received his PhD degree in biomedical engineering from Rutgers University in 2013. He has published more than 20 papers in peer-reviewed journals and conferences. His research interests are in developing histology image processing techniques for breast cancer diagnosis.

Hannah Gilmore is a breast pathologist and director of surgical pathology at University Hospitals Case Medical Center and an assistant professor of pathology at Case Western Reserve University School of Medicine. She received her MD from the Case Western Reserve University in 2003. Her academic interests include both clinical and translation breast cancer research with a particular interest in the integration of pathologic, radiologic, and genomic data to study the biology of breast cancer.

Natalie Shih is a pathologist at the Department of Surgical Pathology, Hospital of the University of Pennsylvania, USA. Her main area of interests has been in the field of digital imaging, with a particular focus on the integration of pathology and computer assisted diagnostic algorithms for prostate and breast cancer diagnosis.

Mike Feldman is an associate professor of Pathology and Laboratory Medicine at the Hospital of the University of Pennsylvania. He received his MD from the University of Medicine and Dentistry of New Jersey, in 1992. His main area of interest has been in the field of digital imaging, with a particular focus on exploring interactions between pathology/radiology, development and utilization of computer assisted diagnostic algorithms for machine vision in prostate and breast cancer.

John Tomaszewski received his MD from the University of Pennsylvania School of Medicine in 1977. At University of Penn, he rose through the ranks to become professor and Interim Chair of Pathology and Laboratory Medicine. In 2011, he was appointed as professor and chair of the Department of Pathology and Anatomical Sciences at the State University of New York at Buffalo. His research interests are in the computational modeling of human structure across scale.

Fabio Gonzalez is an associate professor at the Department of Computer Systems and Industrial Engineering, National University of Colombia. He earned his MS degree in mathematics from the National University of Colombia in 1998, and his PhD degree in computer science from the University of Memphis, USA, in 2003. His research interests are in the field of machine learning and information retrieval, with a particular focus on the representation, indexing and automatic analysis of multi-modal data.

Anant Madabhushi is a professor at Case Western Reserve University. He received his PhD from the University of Pennsylvania in 2004. He has authored over 200 peer-reviewed publications, and issued 7 patents. He is an associate editor for *IEEE TBME*, *IEEE TBME Letters*, *BMC Cancer*, and *Medical Physics*. His research interests include medical image analysis, computer-aided diagnosis, and computer vision. He is a Wallace H. Coulter fellow and a senior IEEE member.