

# Detecting Mitosis against Domain Shift using a Fused Detector and Deep Ensemble Classification Model for MIDOG Challenge

Jingtang Liang<sup>1</sup>, Cheng Wang<sup>1</sup>, Yujie Cheng<sup>2</sup>, Zheng Wang<sup>2</sup>, Fang Wang<sup>2</sup>,  
Liyu Huang<sup>1</sup>, Zhibin Yu<sup>2</sup>, and Yubo Wang<sup>1</sup>

<sup>1</sup> School of Life Science and Technology, Xidian University, Xi'an, Shannxi, China

<sup>2</sup> College of Electrical Engineering, Ocean University of China, Qingdao, Shandong, China

huangly@mail.xidian.edu.cn, yuzhibin@ouc.edu.cn, ybwang@xidian.edu.cn

**Abstract.** Mitotic figure count is an important marker of tumor proliferation and has been shown to be associated with patients' prognosis. Deep learning based mitotic figure detection methods have been utilized to automatically locate the cell in mitosis using hematoxylin & eosin (H&E) stained images. However, the model performance deteriorates due to the large variation of color tone and intensity in H&E images. In this work, we proposed a two-stage mitotic figure detection framework by fusing a detector and a deep ensemble classification model. To alleviate the impact of color variation in H&E images, we utilize both stain normalization and data augmentation, aiding the model to learn color-irrelevant features. The proposed model obtains a F1 score of 0.7550 on the preliminary testing set released by the MIDOG challenge.

**Keywords:** Mitosis · Domain shift · data augmentation · deep ensemble model

## 1 Introduction

Tumor proliferation obtained from hematoxylin & eosin stained (H&E) histopathological images provides valuable information regarding the patient prognosis and treatment planning, especially in breast cancers [12]. Mitotic activity of tumor cells observed in high-power field view is an epiphenomenon of the cell proliferation, is therefore selected to quantify the tumor proliferation and has been shown to be associated with the patients' prognosis [8].

In pathology labs, experienced pathologists count the cell under active mitosis phase in 10 consecutive high-power field views to calculate the mitotic index. This practice is time-consuming and prone to human error, inducing high inter-observer disagreement [11]. Recently, owing to its superior performance in computer vision tasks, deep learning based methods have been extensively studied for automatically identifying tumor cells in mitosis using whole-slide images [4, 6, 13].

The large variability observed in H&E stained pathological images still impeded the application of automatic mitosis detection in clinical settings. The variability in H&E images may arise due to the difference in the scanner to obtain WSI, or the procedure employed by different labs for tissue preparation and staining [2, 5]. In fact, it was recently shown that the WSI itself contains enough information to differentiate where it has been collected [5]. Hence, deep mitosis detection algorithms that can generalize well across different scanners or from different institutions are desired.

Despite the source of variation in the H&E images, it mostly manifests as large variation in color tone among different H&E images. Hence, earlier attempts primarily focused on unifying the color space by utilizing color normalization techniques [7]. Among them, the Macenko stain normalization has been employed in the preprocessing pipeline when training deep learning models [3]. Furthermore, the model can also acquire ability to generalize to out-of-distribution data by using data augmentation or using specially designed structure, such as domain-adversarial training [14].

In this work, we presented our approach to the MIDOG challenge [1]. Inspired by earlier works [6, 4, 10], we constructed a two stages mitosis detection model by using the detectorRS [9] as the base model to coarsely identify the mitosis figure in the images. The results of detector model is later refined by a deep ensemble classification model to illuminate false positives and improve the overall performance. To address the domain shift problem, we employed both stain normalization and data augmentation focusing on inducing color variation. Our results suggested two-stages model equipped with both stain normalization and data augmentation can be an potential solution to address the domain shift in detecting mitosis figures in H&E images.

## 2 Data-set

The data-set was provided by the MIDOG challenge [1]. In brief, all images were obtained from human breast cancer tissue samples after routine Hematoxylin & Eosin staining. The Training set consists of 200 H&E images obtained from four different scanners, including Hamamatsu XR nanozoomer 2.0, Hamamatsu S360 (0.5 NA), Aperio ScanScope CS2, and Leica GT450. Each scanner provided 50 H&E images. Mitotic figures were annotated for the first three scanners. In total, annotation of 1721 and mitotic figures and 2714 non-mitotic figures (hard negative cases) were provided. To train our model, we randomly selected 5 images from each scanner with annotation as the validation set. The rest of training images were used to optimize the model.

The preliminary test set released by the the MIDOG challenge consisted of 20 WSIs from four scanners, in which two scanners were part of the training set and the remaining two scanners are unknown.

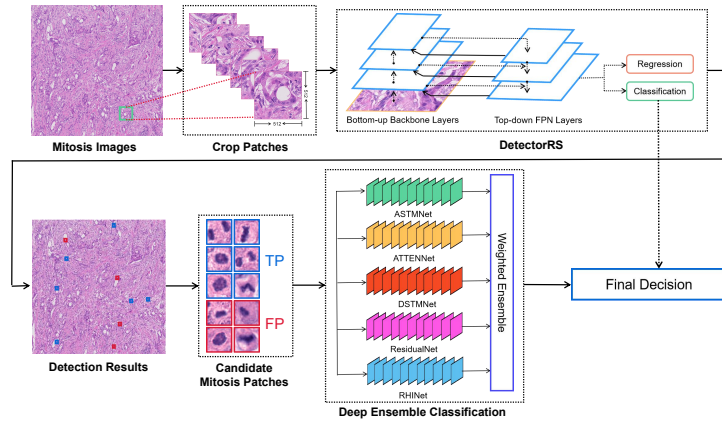


Fig. 1: Pipeline of the proposed mitotic figure detection framework

### 3 Proposed Model

The proposed model is shown in Fig.2. Our whole H&E image processing pipeline consisted of five steps. Firstly, we cropped the original training images into patches of the size  $512 \times 512$  pixels, centered at the ground truth mitotic figures and hard negative cases. For each annotated cases, we randomly shifted the center of each patch within the range of  $\pm 205$  pixels. Then, a detectorRS model [9] was trained to identify the location of mitotic figures using a bounding box with a size of  $50 \times 50$  pixels. In the training phase, all training images were normalized with respect to the first images of the first scanner (*001.tiff*) by using Macenko stain normalization. Then, we augmented the training patches by using random rotation, elastic deformation, scaling, Gaussian blur and a brightness and contrast enhancement. The detector was trained by using SGD with a learning rate of 0.02 for 12 epochs. Once the detector was trained, we employed the trained model to the whole training images to identify all suspected mitotic figures. To be noted, the model was trained using patches containing annotations of ground truth mitotic and hard negative cases, whereas the trained model scanned through the training images can produce many previously un-annotated false mitotic figures. This observation also motivated us to employed a second stage classifier to refine the results produced by the detector.

The overall structure of deep ensemble model consists of five convolution networks, adopted from [10]. The input to the deep ensemble model was the suspected mitotic figures found by the trained detector on the training images with a classification threshold of 0.3. The positive cases for training the classification model consisted of all samples with ground truth mitotic figures. The negative cases were the false positive cases identified by the detector on the whole training images and the hard negative cases. Training samples for the classification model were construed by shifting the center of obtained patches from detector within

the range of  $[-5, 5]$  pixels. To balance the training samples between positive and negative cases, we adjusted the number of times of applying offset to balance the number of cases in positive and negative classes. Finally, the samples were resized to  $120 \times 120$  and fed to the deep ensemble model. To overcome the domain shift caused by different scanner, we heavily utilized online augmentation methods that can induce color variation to increase the diversity of the training samples. The augmentation employed were horizontal and vertical flipping, random clipping and color jitter augmentation with luminance, contrast, hue and saturation disturbance intensity. Each individual model was optimized using AdamW with a learning rate of  $2 \times 10^{-4}$  and was trained for 100 epochs. The optimal weights for each individual convolution network was selected based on their performance on the validation set. The output of the ensemble model was the weighted sum of soft-max score produced by each convolution networks. The final decision of the proposed two-stage mitotic figure detection was obtained by combing the classification score obtained from both detector and deep ensemble model as,

$$S_{final} = \alpha * S_{DE} + (1 - \alpha) * S_{Dect} \quad (1)$$

where  $\alpha \in [0, 1]$  is the weights to balance the decision made by the detector and the deep ensemble model and optimized on the validation set,  $S_{final}$  is the final score to produce the final decision,  $S_{DE}$  and  $S_{Dect}$  are the classification score for the deep ensemble modular and detection modular, respectively.

## 4 Results

We first tested the performance of detection modular on the validation set. The results of F1 score, precision and recall were given in Table.1. It can be observed that detector alone was able to retrieve almost 80% of mitosis figures. In the meantime, it also produced many false positives resulted in a inferior precision score and a significantly degraded F1-score. The ability to refine the results obtained from the detector by the ensemble classification model weighted by different  $\alpha$  was shown in Fig.2. It can be observed that by varying the value of  $\alpha \in [0, 1]$ , the optimal performance on F1 score was found when  $\alpha = 0.9$ . The obtained model obtained a F1 score of 0.7550 on the preliminary testing set.

|           | Detector Only | Proposed model on validation set | Proposed model on preliminary testing set |
|-----------|---------------|----------------------------------|---|
| F1-Score  | 0.5909        | 0.7128                           | 0.7550                                    |
| Precision | 0.4719        | 0.7270                           | 0.7238                                    |
| Recall    | 0.7904        | 0.6993                           | 0.7892                                    |

Table 1: Model performance on validation set and preliminary testing set

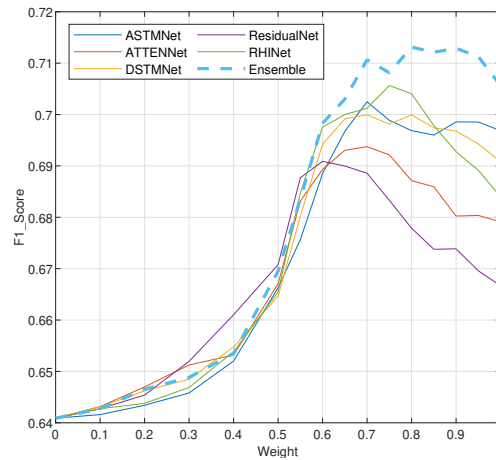


Fig. 2: Performance of individual and ensemble classification model on the validation set

## 5 Conclusion

In conclusion, we presented a fused detector and deep ensemble classification model with image preprocessed by stain normalization and heavy data augmentation to address the domain shift problem for mitosis figure detection. Experiment results showed that the fused model performs reasonably well on the preliminary testing set released by the MIDOG challenge.

## References

1. Aubreville, M., Bertram, C., Veta, M., Klopfleisch, R., Stathonikos, N., Breininger, K., ter Hoeve, N., Ciompi, F., Maier, A.: Mitosis DOrain Generalization Challenge. 24th International Conference on Medical Image Computing and Computer Assisted Intervention (MICCAI 2021). pp. 1–15 (2021). <https://doi.org/10.5281/zenodo.4573978>
2. Aubreville, M., Bertram, C., Veta, M., Klopfleisch, R., Stathonikos, N., Breininger, K., ter Hoeve, N., Ciompi, F., Maier, A.: Quantifying the Scanner-Induced Domain Gap in Mitosis Detection pp. 2017–2020 (2021), <http://arxiv.org/abs/2103.16515>
3. Azevedo Tosta, T.A., de Faria, P.R., Neves, L.A., do Nascimento, M.Z.: Computational normalization of H&E-stained histological images: Progress, challenges and future potential. *Artificial Intelligence in Medicine* **95**(October 2018), 118–132 (2019). <https://doi.org/10.1016/j.artmed.2018.10.004>
4. Chen, H., Dou, Q., Wang, X., Qin, J., Heng, P.A.: Mitosis detection in breast cancer histology images via deep cascaded networks. 30th AAAI Conference on Artificial Intelligence, AAAI 2016 pp. 1160–1166 (2016)
5. Howard, F.M., Dolezal, J., Kochanny, S., Schulte, J., Chen, H., Heij, L., Huo, D., Nanda, R., Olopade, O.I., Kather, J.N., Cipriani, N., Grossman, R.L., Pearson, A.T.: The impact of site-specific digital histol-

- ogy signatures on deep learning model accuracy and bias. *Nature Communications* **12**(1), 1–13 (2021). <https://doi.org/10.1038/s41467-021-24698-1>, <http://dx.doi.org/10.1038/s41467-021-24698-1>
6. Li, C., Wang, X., Liu, W., Latecki, L.J.: DeepMitosis: Mitosis detection via deep detection, verification and segmentation networks. *Medical Image Analysis* **45**, 121–133 (2018). <https://doi.org/10.1016/j.media.2017.12.002>, <https://doi.org/10.1016/j.media.2017.12.002>
  7. Macenko, M., Niethammer, M., Marron, J.S., Borland, D., Woosley, J.T., Guan, X., Schmitt, C., Thomas, N.E.: A method for normalizing histology slides for quantitative analysis. In: 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro. pp. 1107–1110. IEEE (2009)
  8. Medri, L., Volpi, A., Nanni, O., Vecchi, A.M., Mangia, A., Schittulli, F., Padovani, F., Giunchi, D.C., Vito, A., Amadori, D.: Prognostic relevance of mitotic activity in patients with node-negative breast cancer. *Modern pathology* **16**(11), 1067–1075 (2003)
  9. Qiao, S., Chen, L.C., Yuille, A.: DetectoRS: Detecting Objects with Recursive Feature Pyramid and Switchable Atrous Convolution (2020), <http://arxiv.org/abs/2006.02334>
  10. Sohail, A., Khan, A., Nisar, H., Tabassum, S., Zameer, A.: Mitotic nuclei analysis in breast cancer histopathology images using deep ensemble classifier. *Medical Image Analysis* **72**, 102121 (2021). <https://doi.org/10.1016/j.media.2021.102121>, <https://doi.org/10.1016/j.media.2021.102121>
  11. Theissig, F., Kunze, K.D., Haroske, G., Meyer, W.: Histological Grading of Breast Cancer: Interobserver, Reproducibility and Prognostic Significance. *Pathology Research and Practice* **186**(6), 732–736 (1990). [https://doi.org/10.1016/S0344-0338\(11\)80263-3](https://doi.org/10.1016/S0344-0338(11)80263-3), [http://dx.doi.org/10.1016/S0344-0338\(11\)80263-3](http://dx.doi.org/10.1016/S0344-0338(11)80263-3)
  12. Van Diest, P.J., Van Der Wall, E., Baak, J.P.A.: Prognostic value of proliferation in invasive breast cancer: A review. *Journal of Clinical Pathology* **57**(7), 675–681 (2004). <https://doi.org/10.1136/jcp.2003.010777>
  13. Veta, M., van Diest, P.J., Willems, S.M., Wang, H., Madabhushi, A., Cruz-Roa, A., Gonzalez, F., Larsen, A.B.L., Vestergaard, J.S., Dahl, A.B., Cireşan, D.C., Schmidhuber, J., Giusti, A., Gambardella, L.M., Tek, F.B., Walter, T., Wang, C.W., Kondo, S., Matuszewski, B.J., Precioso, F., Snell, V., Kittler, J., de Campos, T.E., Khan, A.M., Rajpoot, N.M., Arkoumani, E., Lacle, M.M., Viergever, M.A., Pluim, J.P.W.: Assessment of algorithms for mitosis detection in breast cancer histopathology images. *Medical Image Analysis* **20**(1), 237–248 (2015). <https://doi.org/10.1016/j.media.2014.11.010>
  14. Zhou, K., Liu, Z., Qiao, Y., Xiang, T., Loy, C.C.: Domain Generalization in Vision: A Survey pp. 1–18 (2021), <http://arxiv.org/abs/2103.02503>