

# SK-Unet Model with Fourier Domain and Weight Perturbation for Mitosis Detection

Sen Yang<sup>1</sup>, Feng Luo<sup>2</sup>, Jun Zhang<sup>1</sup>, and Xiyue Wang<sup>3</sup>

<sup>1</sup>College of Biomedical Engineering, Sichuan University, Chengdu, China

<sup>2</sup>Shenzhen International Graduate School, Tsinghua University, Shenzhen, China

<sup>3</sup>College of Computer Science, Sichuan University, Chengdu 610065, China

**Counting mitotic cells (cells undergoing division) can represent the tumor proliferation speed, which is a key cancer prognosis marker in breast cancer, prostate carcinoma, lymphoma, lung carcinoma, melanoma, etc. However, the various phases of mitosis have different visual manifestations and are easily confused with other normal cells (e.g., apoptotic or necrotic nuclei). In addition, the manifestation of cancerous tissue varies greatly across origins (domain shift problem), which makes visual assessment of mitosis very difficult and then a high inter-pathologist variability. To reduce the domain dependencies, this work proposes a new mitosis cell detection method, where a modified U-Net architecture combined with a Fourier-based data augmentation and a weight perturbation loss function are built for better generalization ability. Our method is developed on the MIDOG 2022 challenge and achieves F1-score of 0.7445 and 0.7526 on the preliminary test set in Task 1 and Task 2, respectively.**

Correspondence: wangxiyue@stu.scu.edu.cn

## Introduction

Mitotic count is the one of the most important morphological feature of tumor proliferation assessment. So pathologists usually search for mitosis in a complete slide with a high-power fields of view (HPF) manually to count. However, a large number of HPF in a single complete slide and the appearance difference of mitotic cells make the task time-consuming and tedious. In addition, it is objective to judge mitotic cell and are prone to reach a consensus on mitotic count among pathologists.

Recent advances in deep learning and digital scans have paved the way and many automatic mitosis detection methods have been proposed[1][2][3]. Although achieving great success, a drop in performance is often observed when the trained model is tested on data from another domain(i.e. different slide scanners and sample preparation from clinical centers). This problem makes it hard for mitosis detection algorithms to be widely used in real diagnosis process.

To solve the problem, we construct a SK-Unet mitosis detection architecture based on a Fourier-based data augmentation and a weight perturbation loss function. Our method is developed using the released data in the MIDOG 2022 challenge. We transform the object detection problem into a semantic segmentation one. For the supervised signal, in Task 1, we convert mitosis detection labels (bounding boxes) to mitotic cell segmentation masks using a simple inner tangent circles inside the boxes. In Task 2, a well-pretrained HoVer-Net is

used to produce the segmentation masks of mitotic nuclei. The used Fourier-based data augmentation is inspired by [4], in which we swap the low-frequency spectrum of source and target images to alleviate the discrepancy between different scanners. The weight perturbation loss is motivated by [5], where we add a small perturbation on the network weight matrix to derive more robust segmentation model. Experimental results show that our method can address the domain shift in mitosis detection. It achieves F1-score of 0.7445 and 0.7526 on the preliminary test set in Task 1 and Task 2 of MIDOG 2022 challenge, respectively.

## Methodology

Regarding mitosis detection as segmentation, the proposed algorithm can be divided into image pre-processing, Fourier-based data augmentation, SK-Unet architecture, and image post-processing. Image pre-processing and image post-processing are processes of converting bounding boxes and masks. Fourier-based data augmentation aims to generate more diverse samples by combining information across domains. As for the network, SK-Unet[6] equipped with weight perturbation loss is modified for the segmentation task. It is noted that, our solution for the MIDOG 2022 is an improvement of our previous solution for the MIDOG 2021 [7]. The main differences are the segmentation mask generation and recycling training strategy (in Task 1), and the weight perturbation loss used in both tasks.

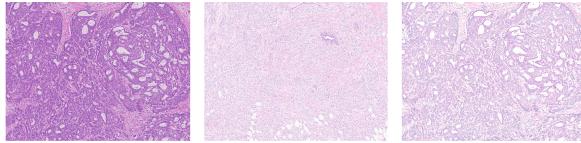
**Image Pre-processing.** Due to the fact that segmentation model is more robust, we convert mitotic detection to segmentation, thus masks of mitotic cells are required.

**Task 1:** Since Task 1 prohibits the use of external data, the segmentation labels of mitotic nuclei are obtained by taking the inner cut circle of the bounding boxes provided by the organizer.

**Task 2:** First, all cells in an image are segmented with pre-trained HoVer-Net [8] which is publicly available <sup>1</sup>. Then we get cells which need to reserve according bounding boxes of the image. In specific, a cell is reserved when the Intersection of Union (IOU) of the cell and any bounding box is over 0.8.

<sup>1</sup>[https://github.com/simongraham/hovernet\\_inference](https://github.com/simongraham/hovernet_inference)

**Fourier-based data augmentation.** In order to solve the problem of domain adaptation, a simple method for unsupervised domain adaptation is adopted, which is swapping the low-frequency spectrum of one with the other [4]. To be specific, there are three steps. First, given an image  $I_s$ , its amplitude and phase components can be calculated using FFT algorithm[9]. Second, the center region of  $I_s$ 's amplitude component is replaced by that of another image  $I_t$ . This means that low-frequency information of the two images is swapped. Third, the modified amplitude component and its unaltered phase component are used to reconstitute an image with similar style of  $I_t$  using inverse FFT (iFFT). The motivation of swapping process is that high-level semantics represented by high-frequency spectrum is the real cue for mitosis while low-level semantics is closer to background information. So combining one high-frequency spectrum with several low-frequency components can generate images with different styles and the same label, which enlarges the amount of training data and enhances the generalization ability of our model. Some generated samples are shown in Fig. 1.



**Fig. 1.** A FDA sample. Images are the source image, reference image and generated image from left to right.

**SK-Unet architecture.** For the mitosis segmentation, SK-Unet is adopted by us. The method proposed a combination of feature maps from different scales in the encoder-decoder network to improve the segmentation results. The SE-ResNeXt50 model is employed as the encoder. In each block of decoder, a SK block is added for multi-scale feature extraction. Considering the class imbalance problem, our segmentation loss ( $L$ ) combines the Dice and Focal loss. Another import strategy is the addition of a weight perturbation loss for the final loss function, that is

$$\mathcal{L} = L(f_w(x), y) + \beta L(f_{w+v}(x), y) \quad (1)$$

where  $x$  and  $y$  denote the input sample and its ground truth, respectively,  $f$  denotes our segmentation backbone,  $w$  represents the network weights,  $v$  means the perturbation signal.

**Post-processing.** The image post-processing process aims to refine the result of cell segmentation and convert it to bounding boxes. Initially, the hole filling technique is applied to attain accurate segmentation masks. Then, connected component analysis for all the obtained masks is performed and each connected component is regarded as a cell. Last, centers of all minimum bounding rectangles for connected components are calculated as our final result.

## Experiment

**Dataset.** Our algorithm is evaluated on the MICCAI 2022 MIDOG challenge (MIDOG2) [10, 11]. The MIDOG2 training subset consists of 354 Whole Slide Images (WSIs) from six cancer types, including breast cancer, lymphoma, lung carcinoma, and mast cell tumor, and neuroendocrine. To validate the model, we randomly select one cancer type in the training samples as the validation set and the remaining samples are used to train the model. In addition, there is a preliminary test set from MIDOG2 to evaluate the prior model. It contains 20 WSIs, which are from four cancer types that are unknown for participants.

**Experiment Setup.** A sliding window scheme with overlap is used to crop each WSI into small patches of size 512x512 pixels. Standard real-time data augmentation methods such as horizontal flipping, vertical flipping, random rescaling, random cropping, and random rotation are performed to make the model invariant to geometric perturbations. Moreover, RandomHSV is also adopted to randomly change the hue, saturation, and value of images in the hue-saturation-value (HSV) color space, making the model robust to color perturbations. The Adam optimizer [12] is used as the optimization method for model training. The initial learning rate is set to 0.0003, and reduced by a factor of 10 at the 30th and the 50th epoch, with a total of 80 training epochs. The min-batch size is set as 24. The parameter  $\beta$  in the final loss function is set to 0.1. All models are implemented using the PyTorch framework [13] and all experiments are performed on a workstation equipped with an Intel(R) Xeon(R) E5-2680 v4 2.40GHz CPU and four 32 GB memory NVIDIA Tesla V100 GPU cards.

In Task 1, our model is fully trained twice, which is called recycling training or iterative refinement. In the first time, the segmentation masks adopts the incircles of the bounding boxes. In the second time, the segmentation mask tasks the intersection of bounding boxes and the predicted segmentation masks in the first training.

**Experiment Results.** We present the ablation results on the validation test set and preliminary test set in Task 1 and Task 2, which are shown in Table 1. It is seen that our recycling training used in Task 1 brings an improvement of around 2% in the validation set, and the weight perturbation used in both tasks improves around 1% performance gains in the two sets. The effectiveness of the Fourier-based data augmentation can be seen in other solution on the MIDOG 2021 challenge.

**Table 1.** Ablation results on validation set and preliminary test set (F1-score)

		Validation set	Preliminary test set
Task 1	SK-Unet+mask (incircle)	0.7144	/
	SK-Unet+mask (incircle)+recycling (iterative refinement)	0.7339	0.7304
	SK-Unet+mask (incircle)+recycling (iterative refinement)+weight perturbation (Ours)	0.7445	0.7414
Task 2	SK-Unet+mask (HoVer-Net)	0.7465	/
	SK-Unet+mask (HoVer-Net)+weight perturbation (Ours)	0.7554	0.7526

## References

- Chao Li, Xinggang Wang, Wenyu Liu, and Longin Jan Latecki. Deepmitosis: Mitosis detection via deep detection, verification and segmentation networks. *Medical image analysis*, 45:121–133, 2018.
- Chao Li, Xinggang Wang, Wenyu Liu, Longin Jan Latecki, Bo Wang, and Junzhou Huang. Weakly supervised mitosis detection in breast histopathology images using concentric loss. *Medical image analysis*, 53:165–178, 2019.
- Meriem Sebai, Xinggang Wang, and Tianjiang Wang. Maskmitosis: a deep learning framework for fully supervised, weakly supervised, and unsupervised mitosis detection in histopathology images. *Medical & Biological Engineering & Computing*, 58:1603–1623, 2020.
- Yanchao Yang and Stefano Soatto. Fda: Fourier domain adaptation for semantic segmentation. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, pages 4085–4095, 2020.
- Dongxian Wu, Shu-Tao Xia, and Yisen Wang. Adversarial weight perturbation helps robust generalization. *Advances in Neural Information Processing Systems*, 33:2958–2969, 2020.
- Xiyue Wang, Sen Yang, Yuqi Fang, Yunpeng Wei, Minghui Wang, Jing Zhang, and Xiao Han. Sk-unet: An improved u-net model with selective kernel for the segmentation of lge cardiac mr images. *IEEE Sensors Journal*, 21(10):11643–11653, 2021.
- Sen Yang, Feng Luo, Jun Zhang, and Xiyue Wang. Sk-unet model with fourier domain for mitosis detection. In *MICCAI*, pages 86–90. Springer, 2021.
- Simon Graham, Quoc Dang Vu, Shan E Ahmed Raza, Ayesha Azam, Yee Wah Tsang, Jin Tae Kwak, and Nasir Rajpoot. Hover-net: Simultaneous segmentation and classification of nuclei in multi-tissue histology images. *Medical Image Analysis*, 58:101563, 2019.
- Matteo Frigo and Steven G Johnson. Fftw: An adaptive software architecture for the fft. In *Proceedings of the 1998 IEEE International Conference on Acoustics, Speech and Signal Processing, ICASSP'98 (Cat. No. 98CH36181)*, volume 3, pages 1381–1384. IEEE, 1998.
- M. Aubreville, C. Bertram, K. Breininger, S. Jabari, N. Stathonikos, and M. Veta. Mitosis domain generalization challenge 2022. *Advances in Neural Information Processing Systems*, 2022. doi: 10.5281/zenodo.6362337.
- Marc Aubreville, Nikolas Stathonikos, Christof A Bertram, Robert Kloppeisch, Natalie ter Hoeve, Francesco Ciompi, Frauke Wilms, Christian Marzahl, Taryn A Donovan, Andreas Maier, et al. Mitosis domain generalization in histopathology images—the midog challenge. *arXiv preprint arXiv:2204.03742*, 2022.
- Diederik P Kingma and Jimmy Ba. Adam: A method for stochastic optimization. *arXiv preprint arXiv:1412.6980*, 2014.
- Adam Paszke, Sam Gross, Soumith Chintala, Gregory Chanan, Edward Yang, Zachary DeVito, Zeming Lin, Alban Desmaison, Luca Antiga, and Adam Lerer. Automatic differentiation in pytorch. 2017.