**OCELOT 2023 Report: Enhancing Cell Detection Performance through Multi-Task Learning**

Sunggu Kyung1\*

1Department of Biomedical Engineering, Asan Medical Institute of Convergence Science and Technology, Asan Medical Center, University of Ulsan College of Medicine, Seoul, South Korea

[babbu3682@gmail.com](mailto:babbu3682@gmail.com)

github: https://github.com/babbu3682/MICCAI2023\_OCELOT

# Introduction

The task of cell detection has been a challenging endeavor in the field of pathology. Traditional methods often depend on cell-level labeling, with tissue-level areas considered during the process. Inspired by this practice, we aim to enhance cell detection performance through a multi-task learning approach. This report details our methodology, data processing, hyperparameter tuning, preliminary results in the validation set, and other crucial information needed for reproducing the results.

# Methodology

Inspired by the practice of pathology labelers who consider tissue-level areas when performing cell-level labeling, our approach seeks to perform cell detection while simultaneously conducting tissue-level segmentation. By inducing the model to capture essential parts at the tissue level, we aim to assist the cell detection task. Furthermore, to better grasp the intrinsic features of pathology images, we introduce an image restoration task using an encoder-decoder mechanism.

## Multi-Task Framework

지도, 스크린샷, 도표이(가) 표시된 사진

자동 생성된 설명

Figure 1. Our model works with a total of 4 multi-tasks: Point counting, cell detection task, tissue segmentation task and, reconstruction task. We used MaxViT-Xlarge[1] as backbone network for the encoder and skip connection for the rest of the decoder except for reconstruction.

As shown in Figure 1, our model operates on four different multi-task functions:

* Cell Detection: The primary task where the model learns to detect individual cells within the image.
* Tissue-Level Segmentation: An auxiliary task that aids in understanding the broader tissue context, thereby improving cell detection.
* Image Restoration: A task employing an encoder-decoder structure to restore the pathology image, ensuring the model understands the natural structure and appearance of the images.
* Cell Point Count Matching: A novel task aimed at limiting false positives (FP) by aligning the number of detected cell points with actual counts, further diversifying the multi-task approach.

This combined approach has resulted in a gradual increase in performance; however, some false positives still exist.

## Data Processing

The preprocessing of the images involved a series of transformations tailored for both training and other modes (e.g., validation, testing). Here are the details of the transformations:

* **Training Mode**

For the training mode, the following transformations were applied sequentially:

1. Contrast Limited Adaptive Histogram Equalization (CLAHE): Adjusted the contrast of the images, with a clip limit of 2.0 and a tile grid size of (8,8), ensuring compatibility with 3-channel images.
2. Random Rotations and Flips: Included random 90-degree rotations, horizontal flips, and vertical flips with a probability of 0.5.
3. Blurring and Noise: A random choice of Gaussian blur (with a blur limit between 3 and 7), median blur (with a blur limit of 7), or Gaussian noise (with a variance limit of 50.0) was applied with a 50% probability.
4. Color Adjustments: Color jittering was used to alter brightness, contrast, saturation, and hue, based on this paper, along with a probability of converting to grayscale.
5. Normalization: The images were normalized to a mean of (0.5,0.5,0.5) and a standard deviation of (0.5, 0.5, 0.5), with a maximum pixel value of 255.
6. Tensor Conversion: Finally, the images were converted to tensors, with masks transposed if necessary.

* **Other Modes (e.g., Validation, Testing)**

For other modes, a more simplified transformation pipeline was utilized:

1. Fixed CLAHE: A specific CLAHE function was applied to every image.
2. Normalization: The images were normalized as in the training mode.
3. Tensor Conversion: Images were converted to tensors as in the training mode.

These transformations were crafted with consideration of the medical context, as well as specific requirements of the model and tasks. By employing a combination of spatial and color transformations, the preprocessing step ensures that the model is exposed to a diverse and representative sample of the data.

## Hyperparameter Tuning

Due to lack of time, I couldn't finetuning separately. Based on the mf1 score of the highest score in the self-divided validation set, the highest score epoch network was selected.

# Conclusion

Our multi-task learning approach has shown promising results in enhancing cell detection performance by considering tissue-level segmentation and introducing additional tasks such as image restoration and cell point count matching. Future work may focus on further reducing false positives and exploring additional auxiliary tasks to continue improving performance.

# References

1. Zhengzhong Tu, Hossein Talebi, Han Zhang, Feng Yang, Peyman Milanfar, Alan Bovik, and Yinxiao Li. Maxvit: Multi-axis vision transformer. arXiv preprint arXiv:2204.01697, 2022.