A Low-Resource Training Strategy for Cell Segmentation using Patch-Based Attention U-Net

Kai-Lin Chen[1], Yu-Nong Lin[1], Pen-hsiu Grace Chao[2], Kevin T. Chen[1]

Multimodal Medical Imaging Optimization Lab, National Taiwan University, Taipei, Taiwan [1]
Cell Tissue Engineering Lab, National Taiwan University, Taipei, Taiwan [2]

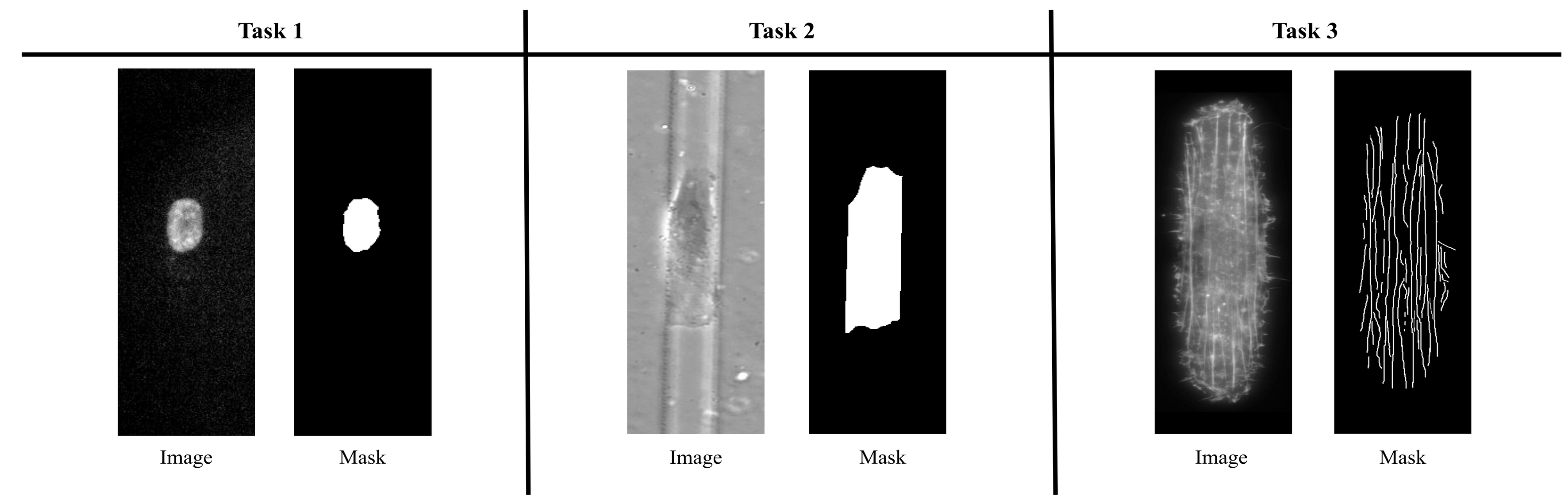
MIDL Salt Lake City 2025

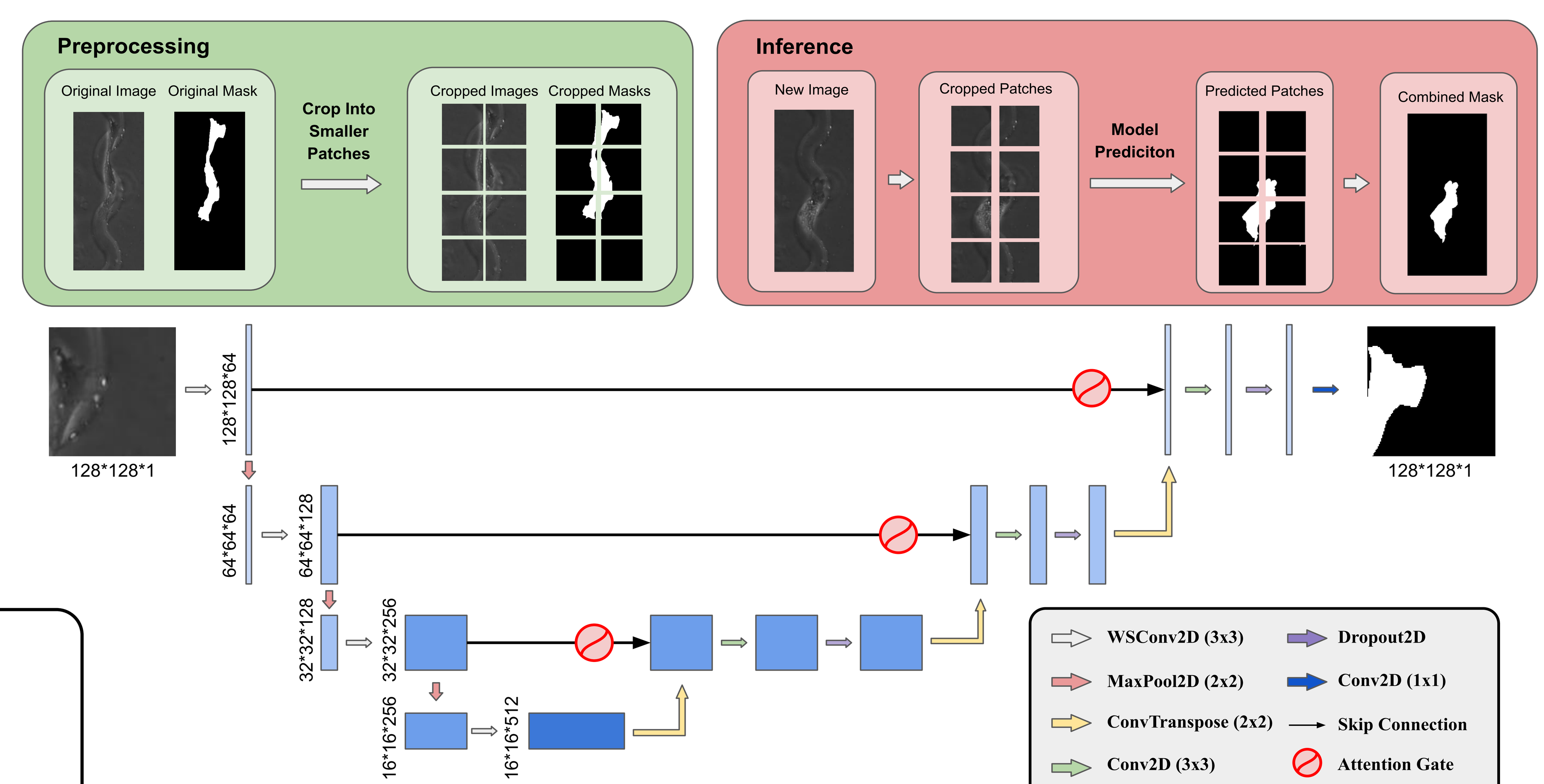
Introduction

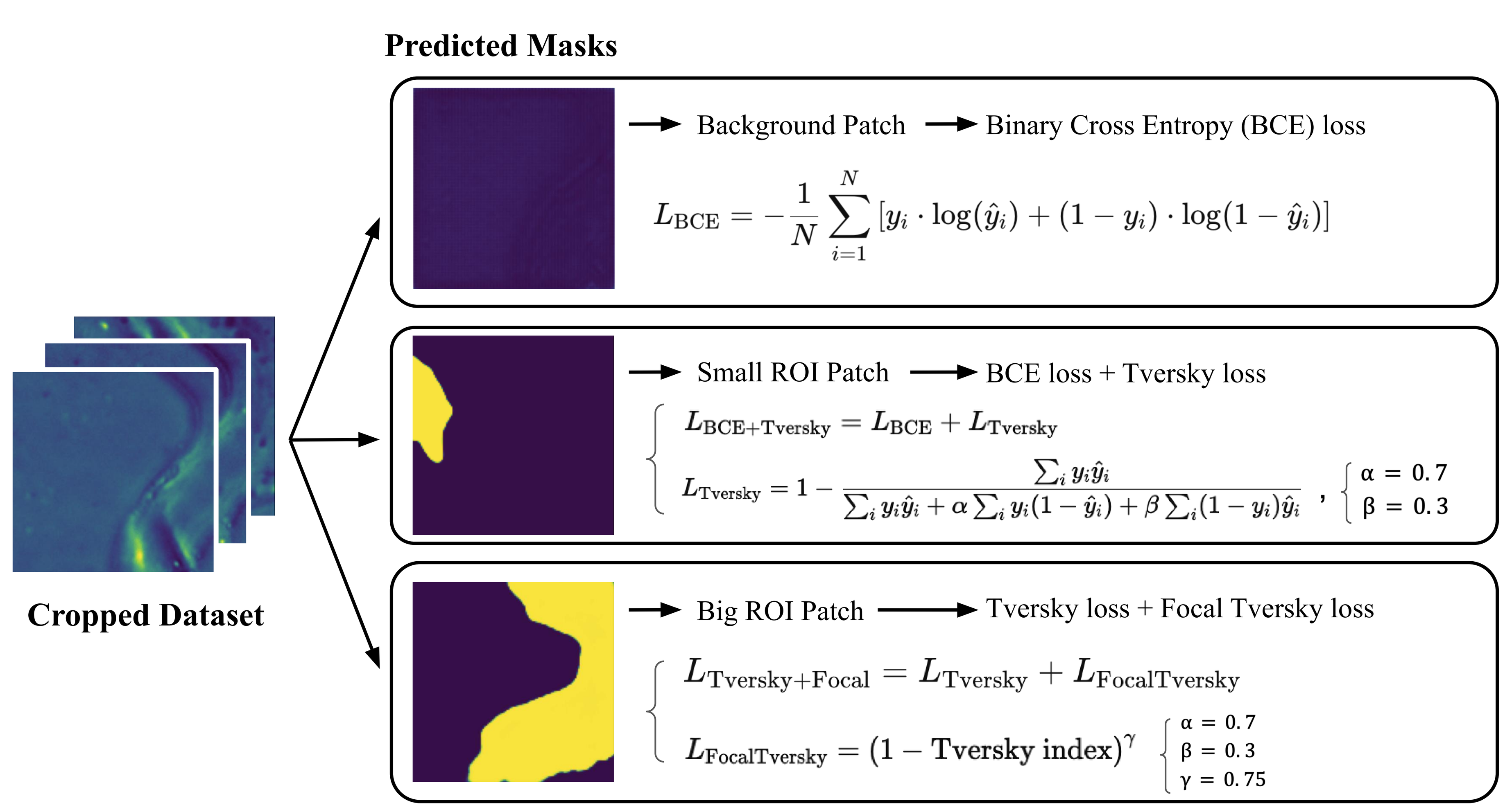
- **Challenge of Image Segmentation Training**: Training deep learning models for image segmentation requires large labeled datasets, which are often difficult and time-consuming to obtain; and without sufficient data, models may suffer from overfitting.
- **Existing Models & Limitations**: Well-known models like Cellpose and Stardist allow fine-tuning but are mainly effective for whole-cell segmentation with regular or round morphology. Furthermore, the finetuned model's performance often depends on dataset size and target morphology.
- **Patch-based Attention U-Net**: We used patch-based cropping to normalize input dimensions with an attention U-Net model architecture to improve generalizability across diverse morphological variations. Validation is done through 5-fold cross-validation across datasets with different regions of interest.

Materials and Methods

- **Datasets:** Three datasets ranging from trivial to challenging: fluorescently labeled cell nuclei (Task 1: 21 training/ 10 test), phase-contrast images of cells (Task 2: 21 training/ 10 test), and fluorescently labeled filaments (Task 3: 10 training/ 4 test)
- **Preprocessing Pipeline:** All images were normalized and the original image and mask were zero-padded such that their edges were multiples of 128 pixels and then cropped into patches of 128 × 128 × 1.
- **Model Architecture:** We employ the gated attention mechanisms in attention U-Nets to focus on relevant regions for segmentation. Additionally, we integrate Weight Standardized Convolutions (WSConv) to stabilize training, and a dynamic loss function to adaptively handle background- and foreground-dominant patches—together improving convergence and segmentation performance. (see below)







- **Dynamic Threshold-based Loss function Selection:** To address class imbalance without discarding any training patches, predicted patches were classified into background, small ROI (less than 6.25% foreground), and large ROI. Background patches used only BCE loss to encourage true negative predictions. Small ROI patches used BCE and Tversky loss to further penalize false positives, large ROI patches combined Tversky loss and Focal Tversky loss to penalize both false positives and false negatives.

Results and Discussion

- Segmentation performance was evaluated using the mean Dice score across three tasks, excluding shape-incompatible cases, with scores weighted by task coverage for fairness.

Model	Task 1	Task 2	Task 3	Average	Task Coverage	Final Score
Stardist	0.8232	0.5904	0.5164	0.6433	1.0000	0.6433
Cellpose	0.9834	0.9036	N/A	0.9435	0.6667	0.6290
Proposed Method	0.9908	0.9327	0.6357	0.8531	1.0000	0.8531

- Our method achieved a final score of 0.8531 (Stardist = 0.6433 and Cellpose = 0.6290) **without the need for large-scale pretraining**, also showing that it **can segment finer details of the regions of interest**.
- Our method's final score and high task coverage, showing its **adaptability to diverse target morphologies** as compared to benchmark models.

Conclusions

- Our proposed method outperformed Cellpose and Stardist in all tasks, achieving a final score of 0.8531 without large-scale pretraining.
- Compared to Stardist and Cellpose, our method showed improvements of 32.60% and 35.62%, respectively, using the same amount of training data.
- The patch-based attention mechanism is proved to be an adaptable solution for segmenting diverse ROI morphologies.

Acknowledgements

This work was financially supported by the Yushan Fellow Program (NTU-112V1015-3, NTU-113V1015-4) and the Higher Education Sprout Program ("Center for Advanced Computing and Imaging in Biomedicine: NTU-114L900703" from the Featured Areas Research Center Program and the National Taiwan University Career Development Project, NTU-114L7872) of the Ministry of Education, National Science and Technology Council (113-2221-E-002-055), and National Health Research Institutes, R.O.C. Taiwan (NHRI-EX114-11205EC).