Automated Non-Tumor Segmentation to Improve Tumor Detection and Analysis using Modified U-NET Networks
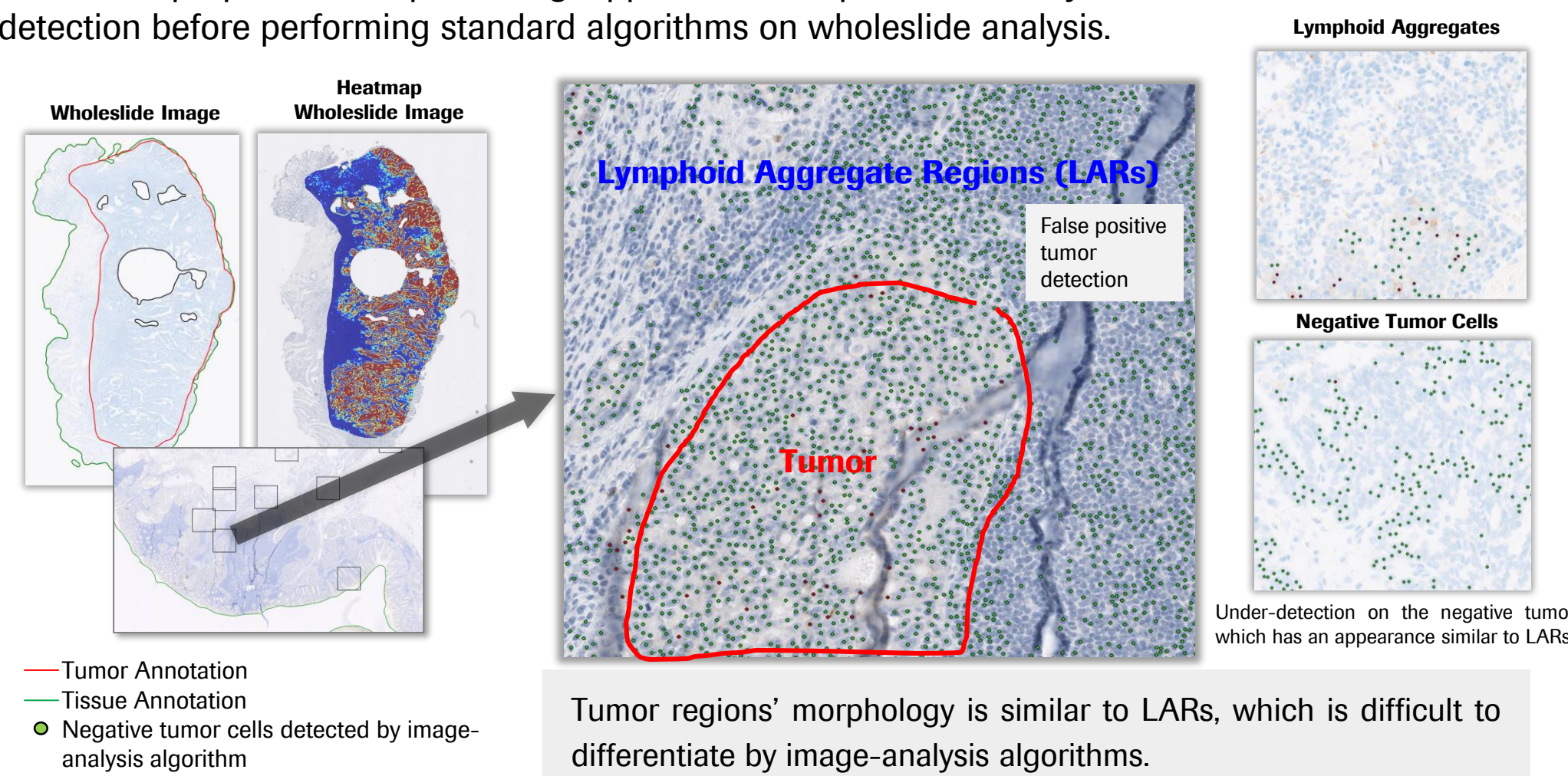
Auranuch Lorsakul[1,*] and Margaret Zhao[1,*]

[1]Imaging and Algorithms, Digital Pathology, Roche Diagnostics Solutions, Santa Clara, CA, USA. E-mail: auranuch.lorsakul@roche.com
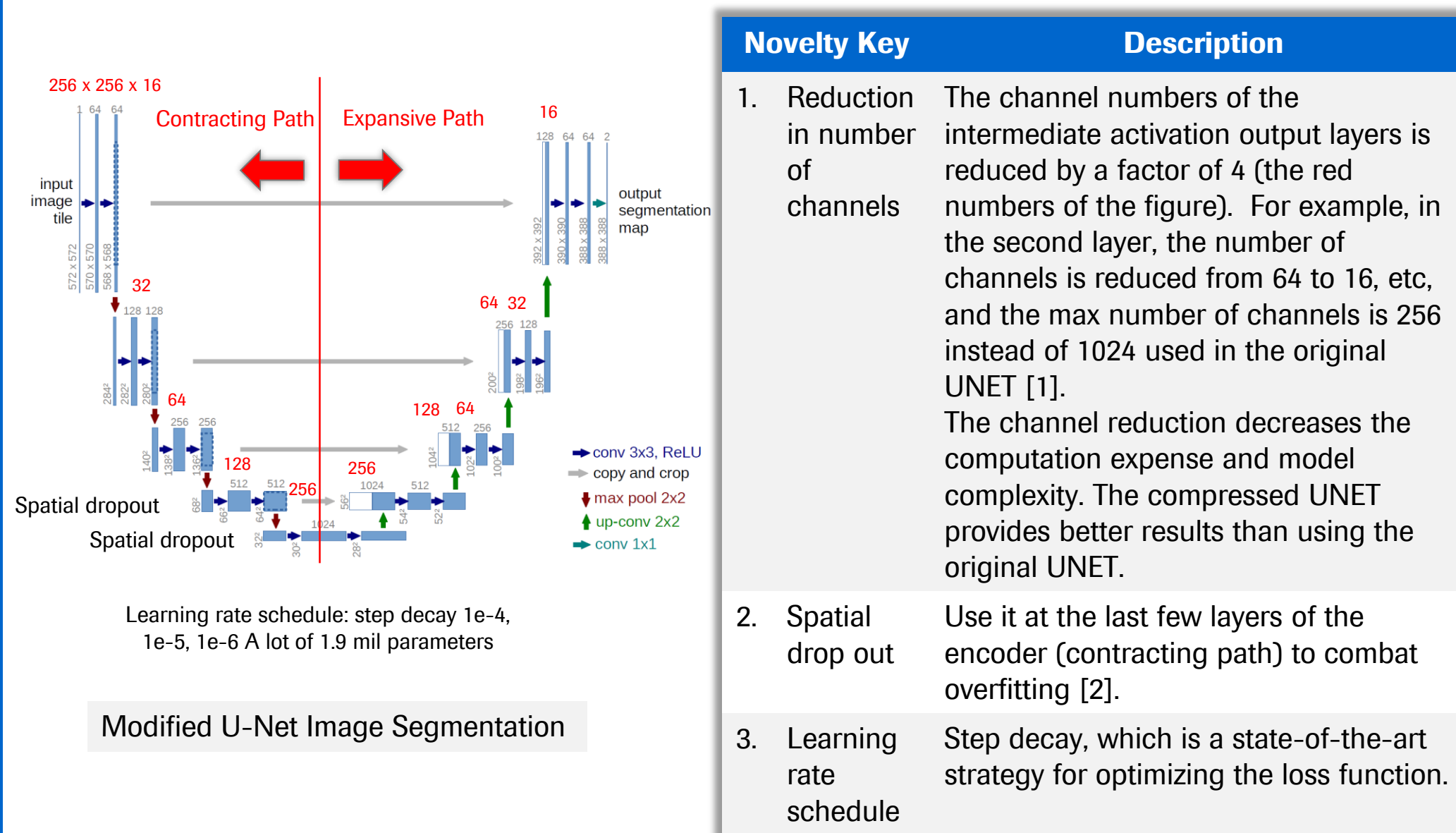
* Equally contributed

1 – Background

In digital pathology, whole-slide analysis with positive and negative tumor cells needs pathologists to initially provide tumor annotations that exclude non-target regions, such as normal tissue. It is difficult to exclude "Lymphoid Aggregate Regions (LARs)," which are clusters of immune cells and their morphology is frequently similar to group of negative tumor cells. As a result, image-analysis algorithms may provide false detection results for these LARs.

Goal: We propose a deep-learning approach to improve accuracy and reduce false non-tumor detection before performing standard algorithms on wholeslide analysis.
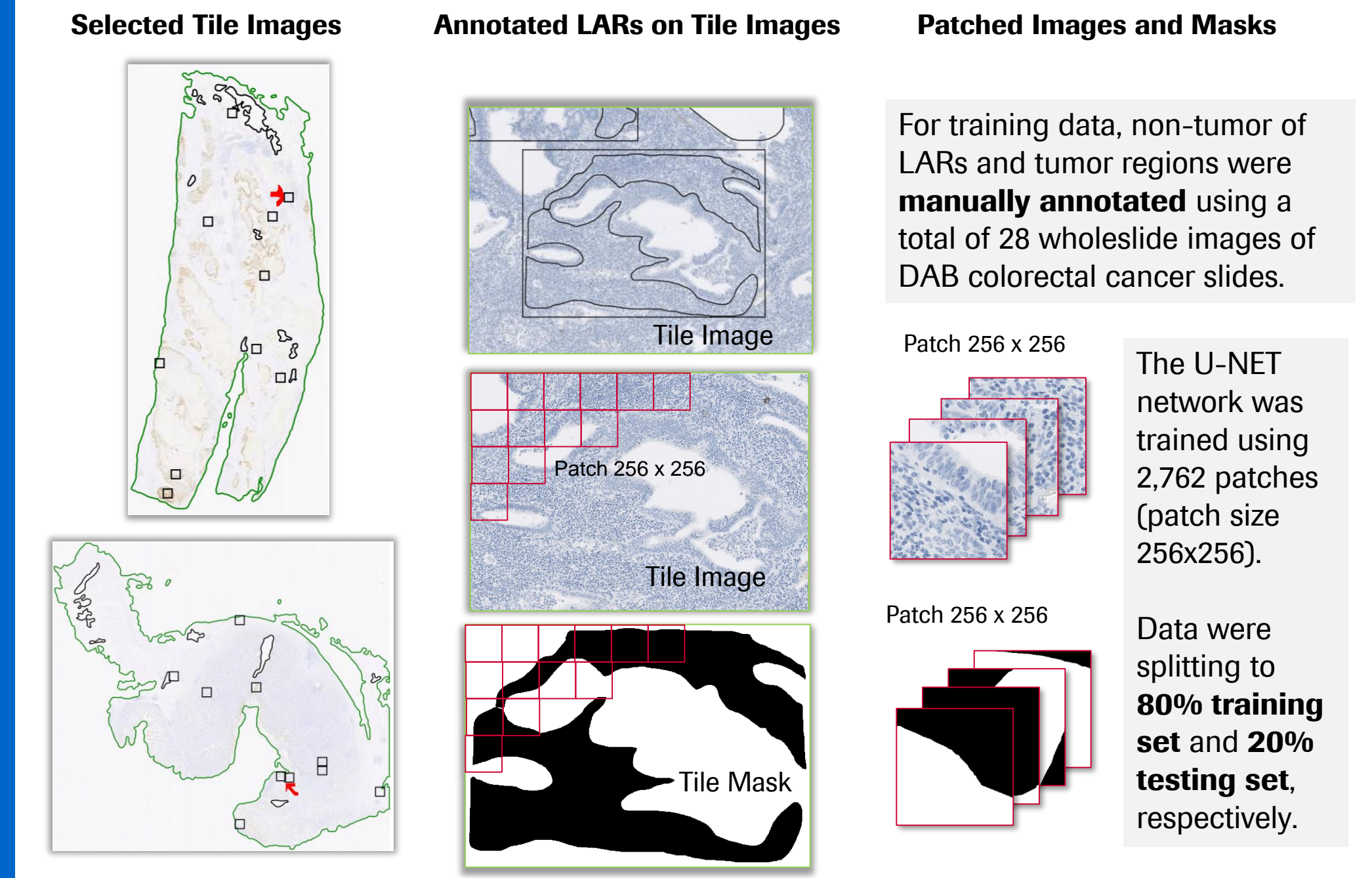

Wholeslide Image | Heatmap Wholeslide Image

Lymphoid Aggregate Regions (LARs)

False positive tumor detection

Tumor

Lymphoid Aggregates

Negative Tumor Cells

Under-detection on the negative tumor which has an appearance similar to LARs

— Tumor Annotation
— Tissue Annotation
○ Negative tumor cells detected by image-analysis algorithm

Tumor regions' morphology is similar to LARs, which is difficult to differentiate by image-analysis algorithms.

2.1 – Method: Modified U-NET, Tile-based Encoder Decoder

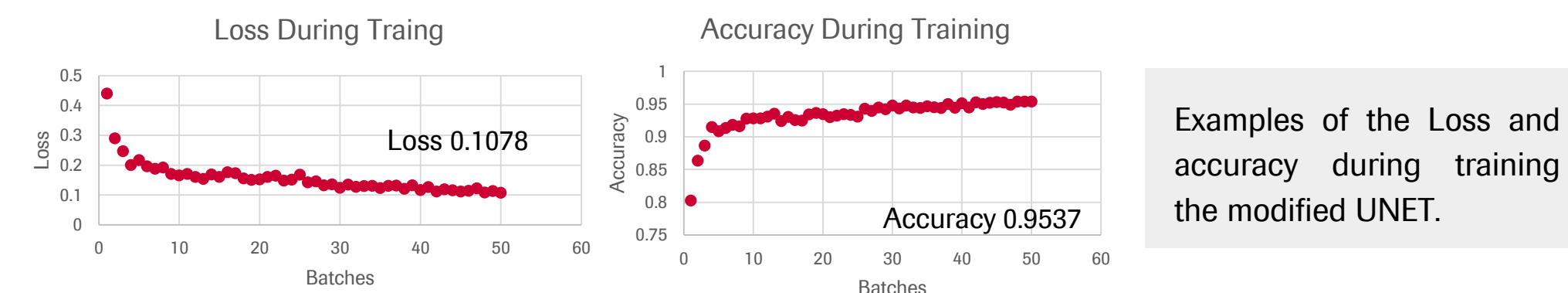We proposed a modified U-NET that automatically detected and masked out LARs.

Contracting Path | Expansive Path

input image tile / output segmentation map

— conv 3x3, ReLU
— copy and crop
↓ max pool 2x2
↑ up-conv 2x2
→ conv 1x1

Spatial dropout

Learning rate schedule: step decay 1e-4, 1e-5, 1e-6 A lot of 1.9 mil parameters

Modified U-Net Image Segmentation

Novelty Key	Description
1. Reduction in number of channels	The channel numbers of the intermediate activation output layers is reduced by a factor of 4 (the red numbers of the figure). For example, in the second layer, the number of channels is reduced from 64 to 16, etc, and the max number of channels is 256 instead of 1024 used in the original UNET [1]. The channel reduction decreases the computation expense and model complexity. The compressed UNET provides better results than using the original UNET.
2. Spatial drop out	Use it at the last few layers of the encoder (contracting path) to combat overfitting [2].
3. Learning rate schedule	Step decay, which is a state-of-the-art strategy for optimizing the loss function.

2.2 – Method: Ground Truth Collections for Training the Networks

Selected Tile Images | **Annotated LARs on Tile Images** | **Patched Images and Masks**

Tile Image
Patch 256 x 256
Tile Image
Tile Mask

Patch 256 x 256
Patch 256 x 256

For training data, non-tumor of LARs and tumor regions were **manually annotated** using a total of 28 wholeslide images of DAB colorectal cancer slides.

The U-NET network was trained using 2,762 patches (patch size 256x256).

Data were splitting to **80% training set** and **20% testing set**, respectively.

2.3 – Method: U-NET Model Training

To optimize the network parameters, we used a binary cross-entropy loss function, 100 epochs, batch size of 1, and learning rate of 1x10e-5 with Adam optimizer. Two levels of image resolution, 20X and 10X, were used to optimize the network parameters.
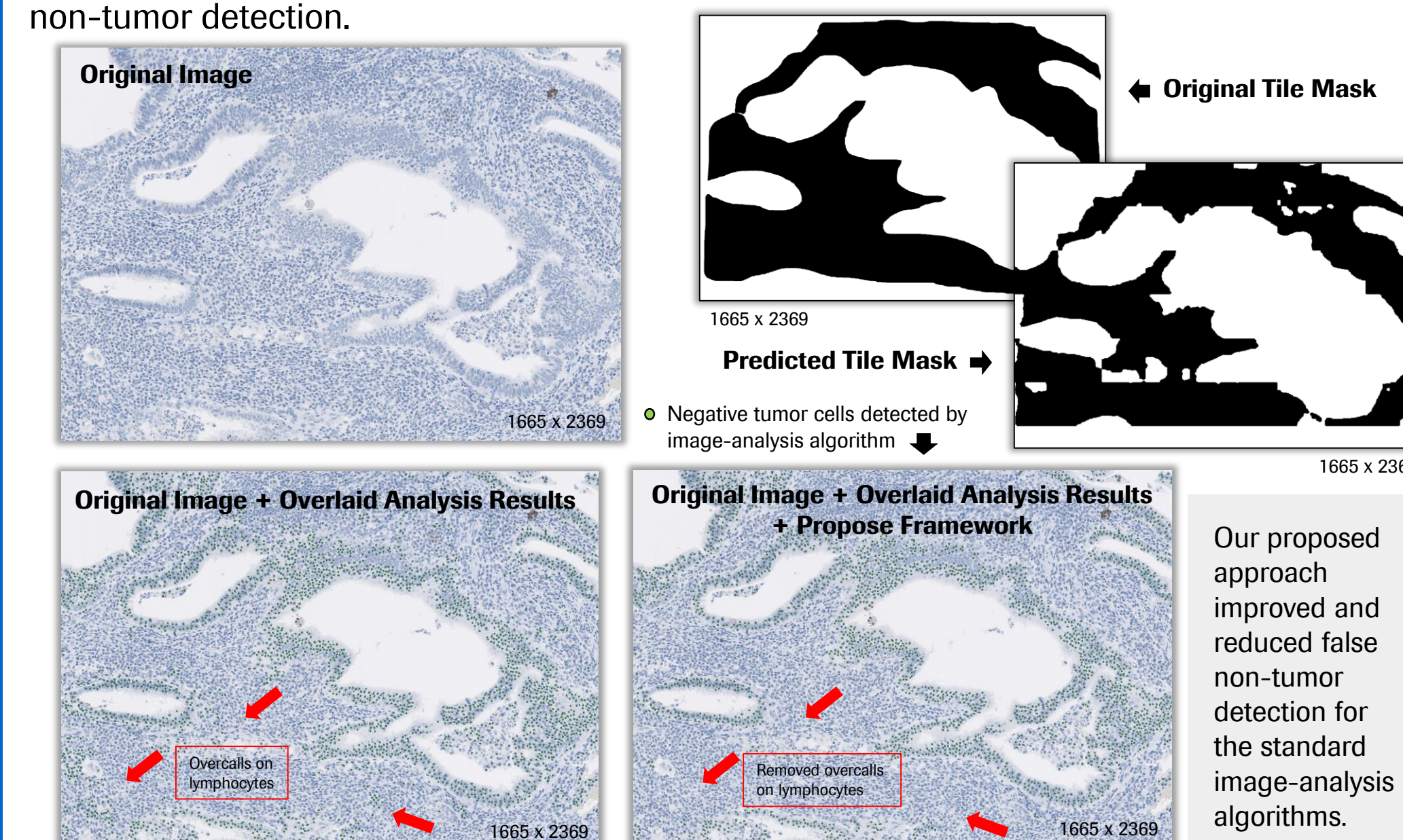

Loss During Traing / Loss 0.1078

Accuracy During Training / Accuracy 0.9537

Examples of the Loss and accuracy during training the modified UNET.

3.1 - Results

The testing results achieved average intersection-over-union (IoU) scores of 0.97 across the tested resolution levels, where the 20X image resolution provided better results. Therefore, our method improves the classification results by reducing false positive detection of LARs.

Original Image | Ground truth | Prediction

256 x 256 | 256 x 256 | 256 x 256

IoU = 0.78

IoU = 0.74

IoU = 1.00
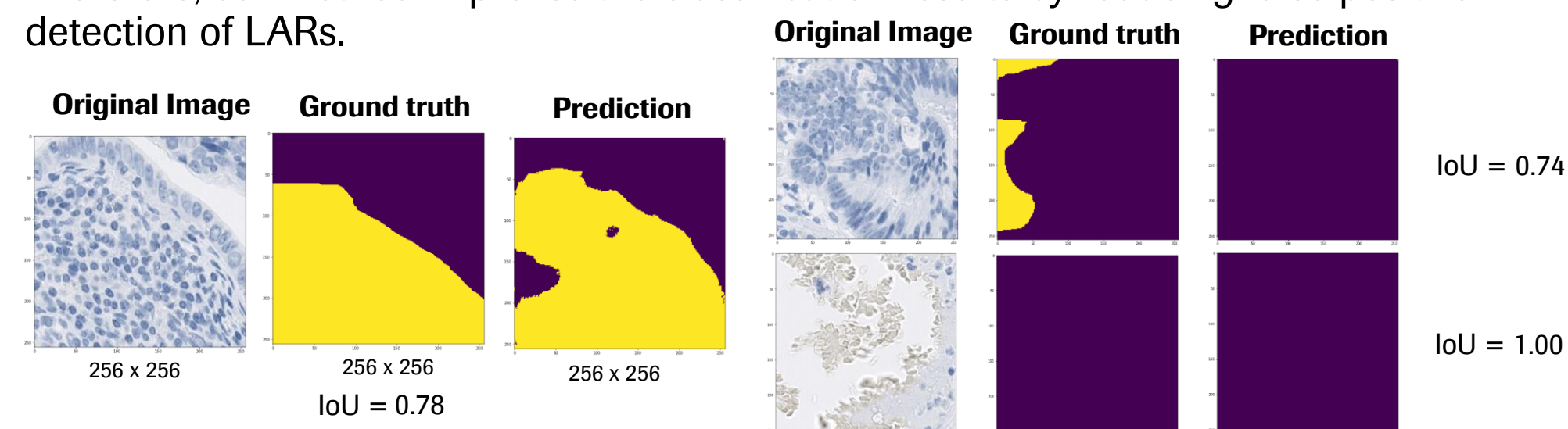
3.2 – Results

The 256 x 256 patch images were stitched to become a tile mask image. This mask was used to correct the analysis results that were over-detected in LARs, which reduced false non-tumor detection.

Original Image / 1665 x 2369

← Original Tile Mask / 1665 x 2369

Predicted Tile Mask ➡ / 1665 x 2369

○ Negative tumor cells detected by image-analysis algorithm

Original Image + Overlaid Analysis Results / 1665 x 2369

Overcalls on lymphocytes

Original Image + Overlaid Analysis Results + Propose Framework / 1665 x 2369

Removed overcalls on lymphocytes

Our proposed approach improved and reduced false non-tumor detection for the standard image-analysis algorithms.

4 – Conclusions

Our proposed method locates and identifies LARs to improve tumor classification tasks. This approach is not limited to segmenting LARs in tissue, but it easily be adapted to other non-tumor areas such as necrosis, scanner artifacts, and tissue folds. For the future work, we can integrate this proposed framework to wholeslide image analysis.

5 – References

[1] Ronneberger O., Fischer P., Brox T. (2015) U-Net: Convolutional Networks for Biomedical Image Segmentation. In: Navab N., Hornegger J., Wells W., Frangi A. (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015. MICCAI 2015. Lecture Notes in Computer Science, vol 9351. Springer, Cham. https://doi.org/10.1007/978-3-319-24574-4_28

[2] https://github.com/keras-team/keras/blob/master/keras/layers/core.py#L178

6 – Acknowledgements