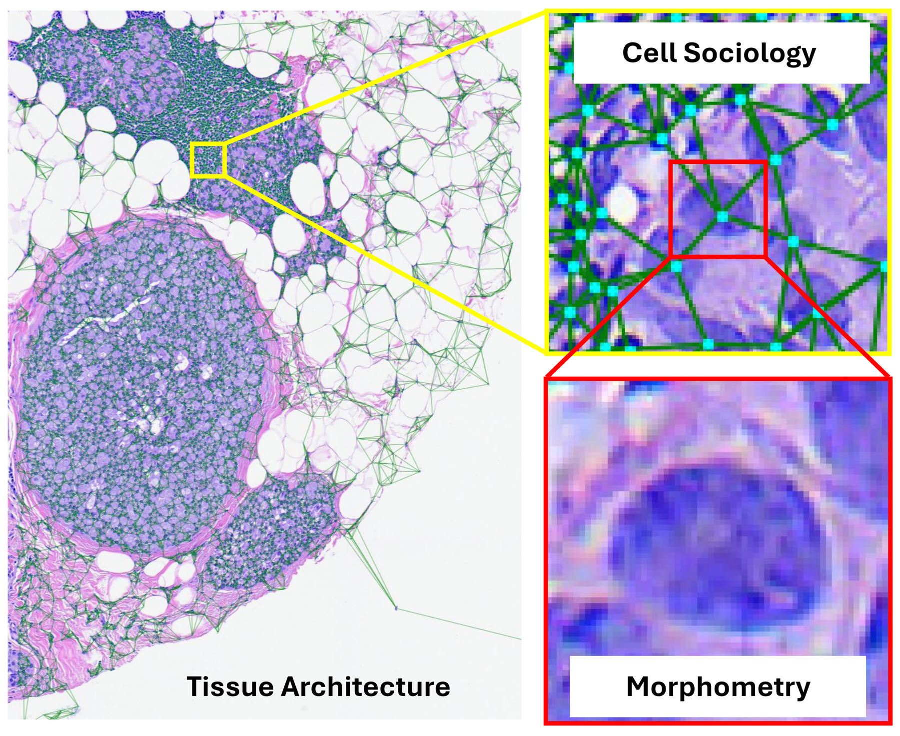
Attention-Enhanced Sequential U-Net for Nuclear Segmentation

Fumi Inaba, Paul Gallagher, Martial Guillaud, Calum MacAulay



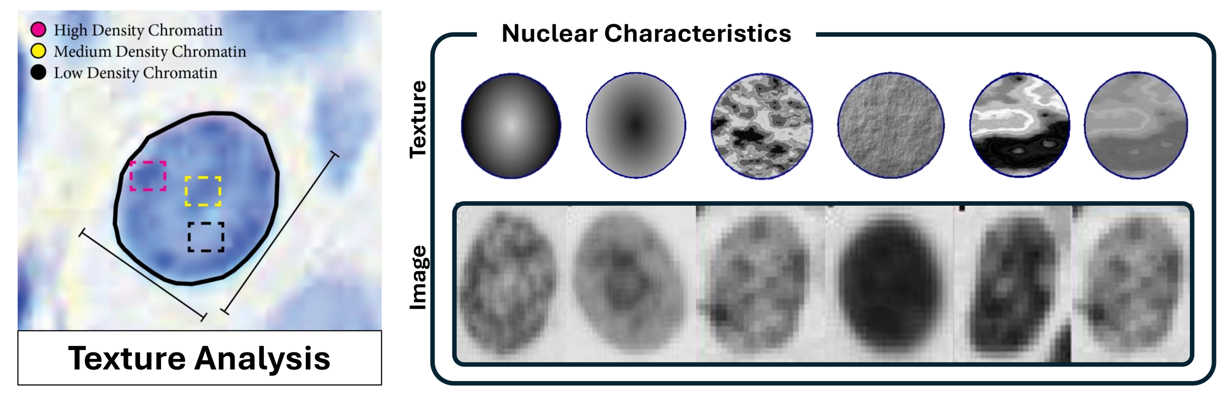


Figure 1) Histology image analysis – tissue architecture, cell sociology, morphometry and texture analysis. Instance segmentation of nuclei are essential processes for these analyses, and segmentation performance may directly affect quality of data for downstream analysis. Deep learning models have shown a lot of promise in segmentation tasks, but require a large volume of data to train, and the large number of parameters can require dedicated computational hardware/cost.

Attention U-Net

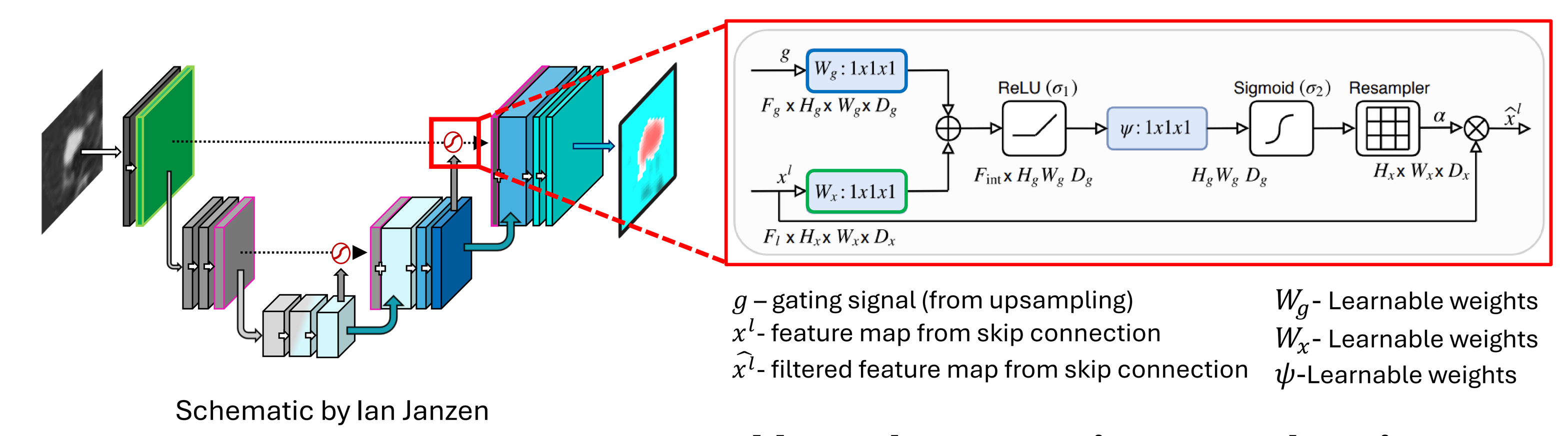


Figure 2) Attention U-net from Oktay et al. Gated attention (right) allows U-Net to focus on specific aspects of the image or features important for segmentation.

How do attention mechanisms:
1. Reduce computational cost?
2. Improve segmentation quality?

Sequential U-Net Pipeline

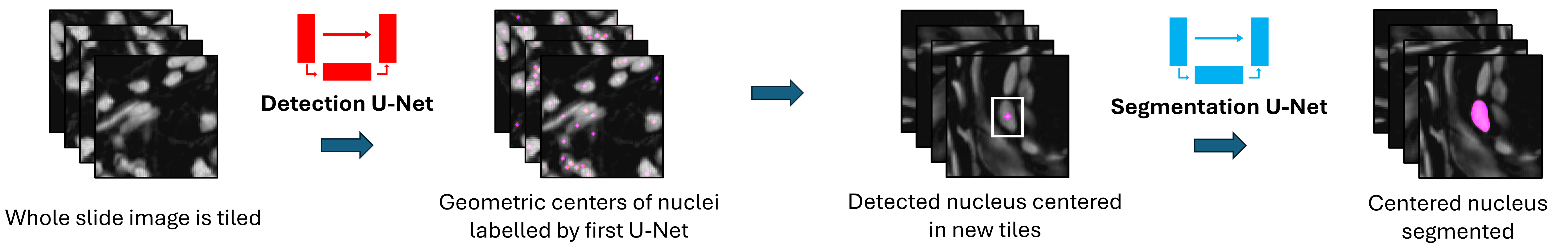
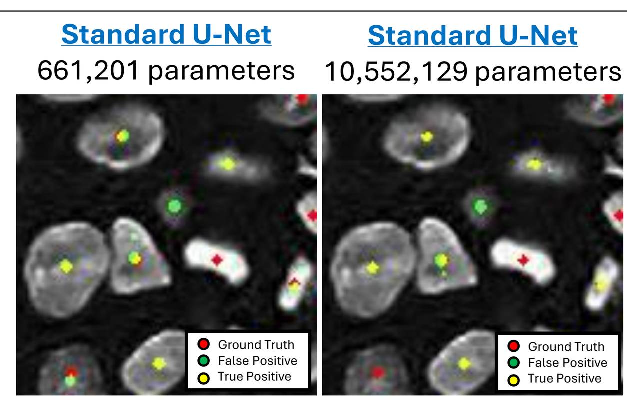


Figure 3) The sequential U-Net model for nuclear instance segmentation, developed by Dr. Calum MacAulay and Paul Gallagher. **Detection U-Net** identifies centers of detected nuclei, where coordinates are extracted to generate new tiles, where detected cells are centered. **Segmentation U-Net** generates a binary mask for the centered nucleus. Thus, two U-Nets are required for this paradigm. We assess the utility of gated attention mechanisms in both U-Net models, and the extent to which the number of parameters can be reduced.

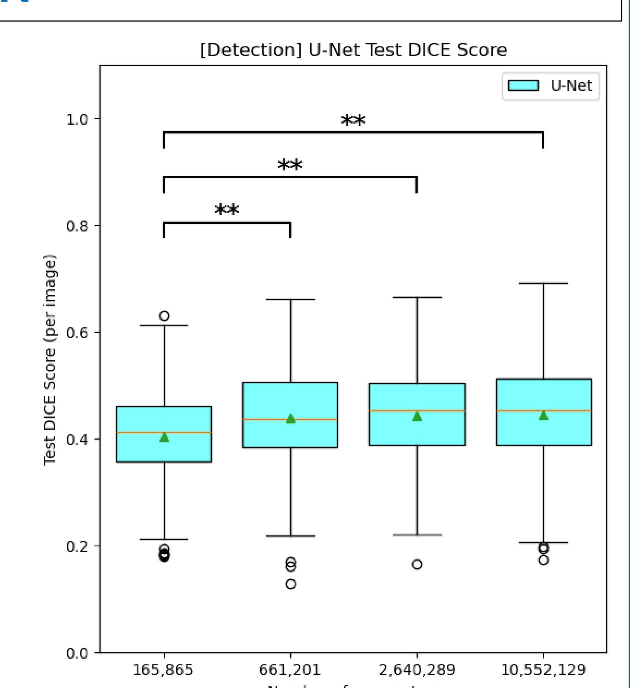
Detection U-Net Results

U-NET PARAMETER REDUCTION

Figure 4) The U-Net with default settings, as originally proposed by Ronneberger et al., has 10.5 million parameters. This test compares U-Nets with different number of feature channels to lower the number of parameters. Verifying with Bonferroni corrected Mann Whitney U-test, the number of parameters can be reduced to 661,201 (16 initial convolution channels), without a significant reduction in the DICE score for the test images. **Right)** DICE score on test images. **Left)** DICE scores and detected cell counts with corresponding test image.

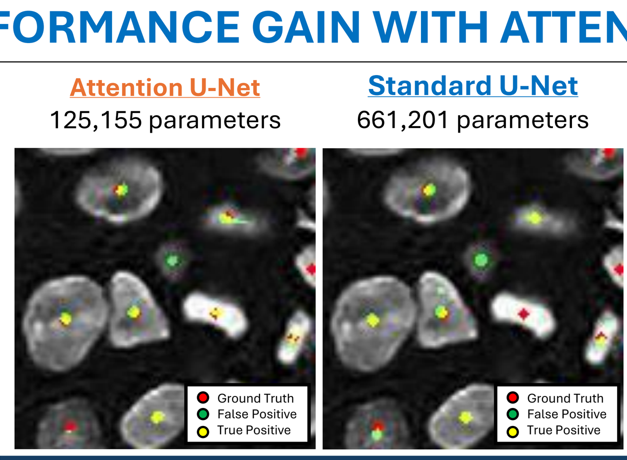


Measure	Number of Parameters	
	661,201	10,552,129
DICE Score	0.397	0.528
True Positives	8	8
False Positives	3	3
False Negatives	3	3

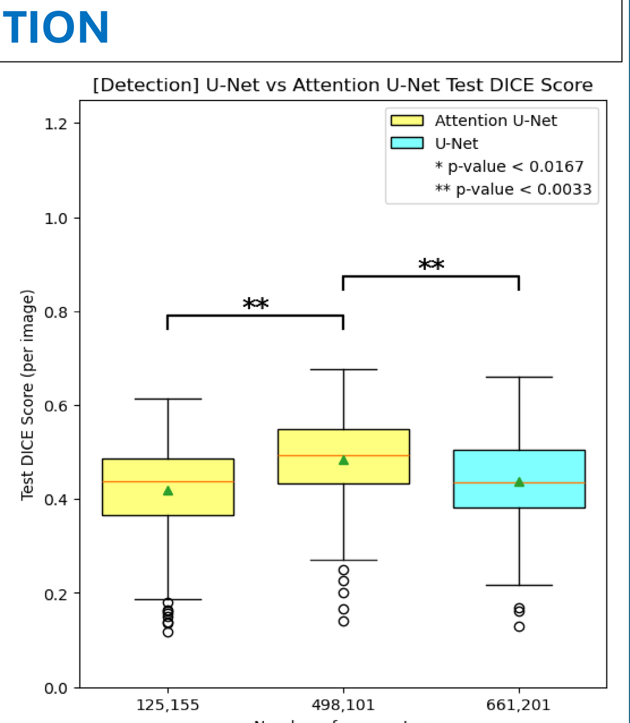


PERFORMANCE GAIN WITH ATTENTION

Figure 5) Comparison of Standard U-Net to Attention U-Net of similar number of parameters. The compared attention U-Net have 498,101 (8 initial convolution channels) and 125,155 (4 initial convolution channels) parameters. The test DICE score does not significantly change between a standard U-Net of 660,201 parameters and attention U-Net of 125,155 parameters. (~80% reduction in parameter count). In terms of the number of detected cells, the attention U-Net with 125,155 performs better than the standard U-Net with 661,201 parameters.

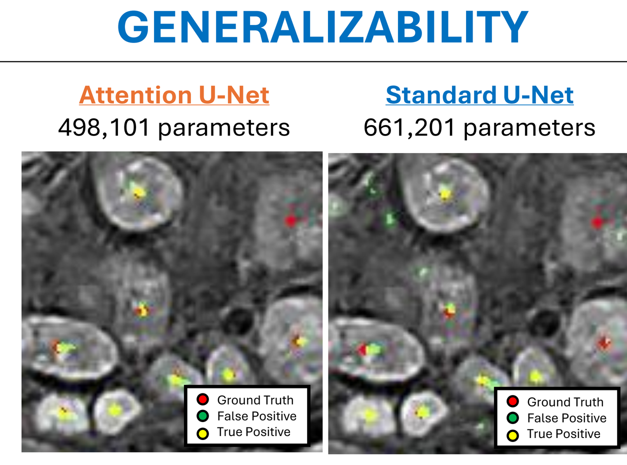


Measure	Number of Parameters	
	125,155	661,201
DICE Score	0.448	0.472
True Positives	9	8
False Positives	1	3
False Negatives	2	3

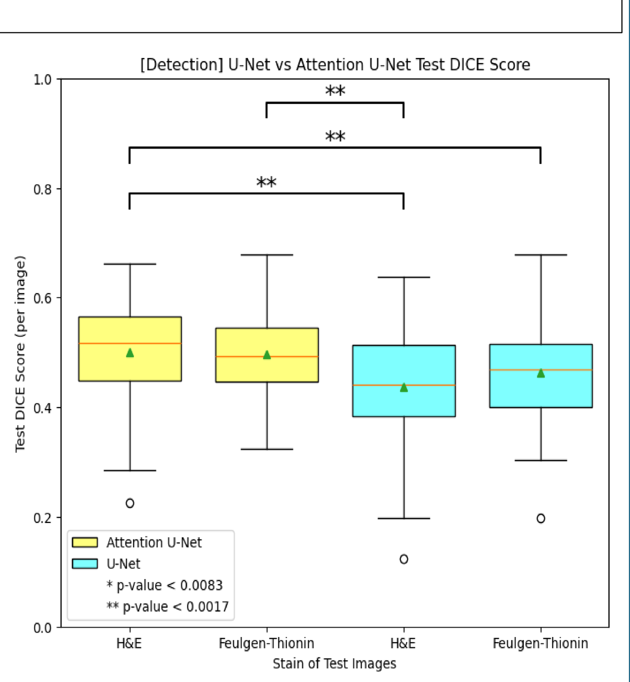


GENERALIZABILITY

Figure 6) The models were trained on Feulgen-thionin images, which only contain nuclear signal. Here, we assess how the U-Net and Attention U-Net, which was trained purely on nuclear-stain images, generalize to hematoxylin and eosin stained (H&E) images. Bonferroni-corrected Mann Whitney U-test shows that there are no significant differences in test DICE score on Feulgen-Thionin stained test images and H&E-stained test images, for both U-Net and Attention U-Net.



Measure	Number of Parameters	
	661,201	10,552,129
DICE Score	0.397	0.528
True Positives	8	8
False Positives	3	3
False Negatives	3	3



INTERPRETABILITY

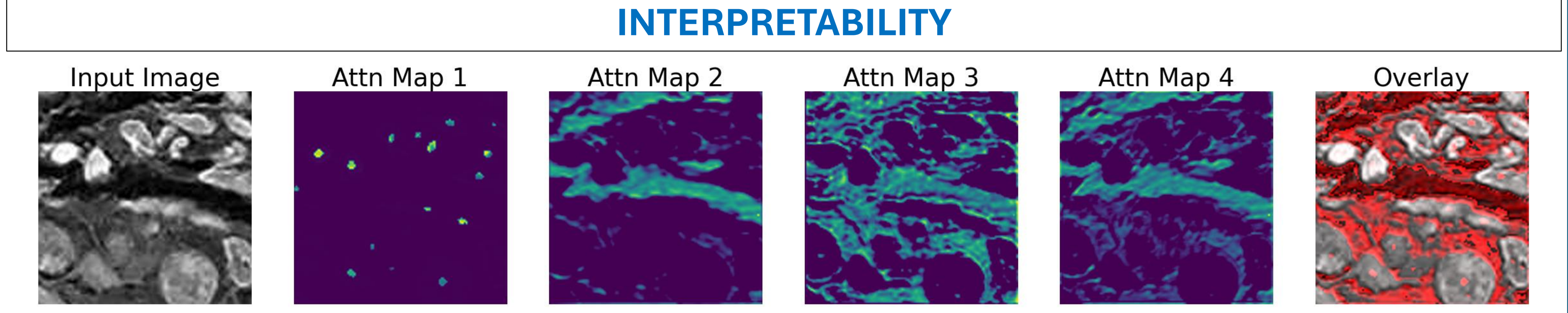
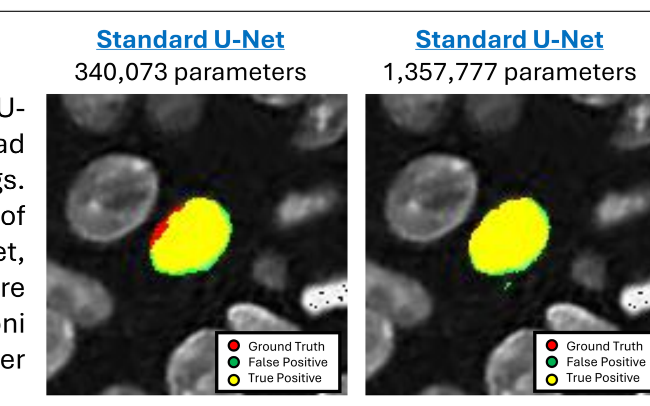


Figure 7) Attention activation maps of an Attention U-Net, with 4 initial convolutional channels. Attention Map 1 shows the model is focusing on regions of the image/features at the center of nuclei in the Input Image. The remaining attention maps identify the background regions of the image. These attention maps provide insight into the specific regions which the model pays attention to make its predictions. (Intensities shown in each attention map are normalized to 0-1 prior to visualization, and are not representative of activation output values)

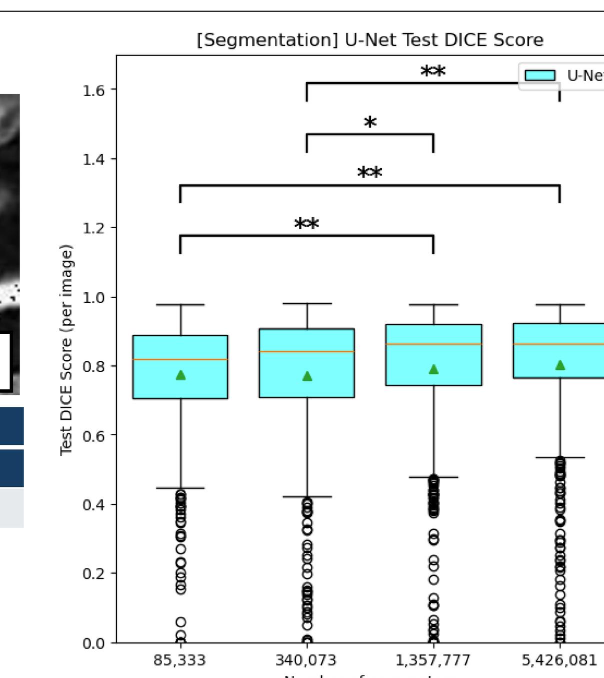
Segmentation U-Net Results

U-NET PARAMETER REDUCTION

Figure 8) The segmentation models use a U-Net architecture with a kernel size of 5 instead of 3 which is the default settings. Consequently, the number of parameters of each U-Net differ from the detection U-Net, even with the same number of feature channels. Verifying again with Bonferroni corrected Mann Whitney U-test, we can lower the number of parameters of the standard U-Net from 5,426,081 (32 initial convolution channels) down to 340,073 (8 initial convolution channels). This marks an approximate ~94% reduction in the number of parameters.

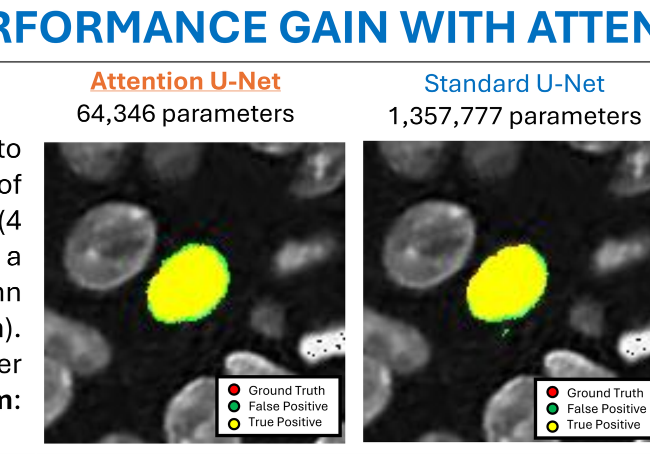


Measure	Number of Parameters	
	340,073	1,357,777
DICE Score	0.853	0.926

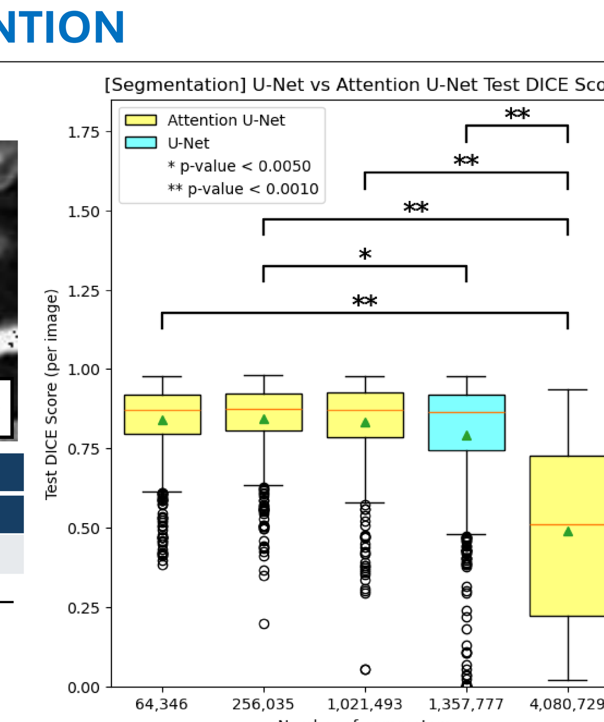


PERFORMANCE GAIN WITH ATTENTION

Figure 9) The standard U-Net compared to the Attention U-Net. The number of parameters can be reduced to 64, 346 (4 initial convolution channels) without a significant reduction in test DICE score (Mann Whitney U-test with Bonferroni correction). This is roughly a 95% reduction in the number of parameters. **Overfitting problem**: Interestingly, the attention U-Net with the largest number of parameters (4,080,729 parameters, 32 initial convolution channels) performs significantly poorer than the rest. This is likely caused by overfitting – suggesting that models with attention may overfit easier than their non-attention counterpart.

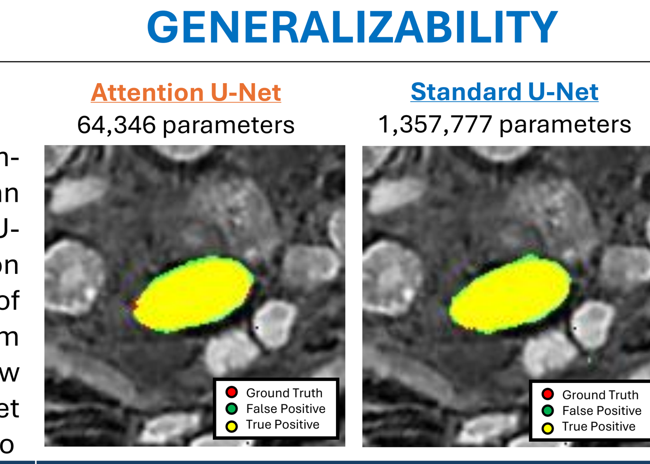


Measure	Number of Parameters	
	64,346	1,357,777
DICE Score	0.944	0.926

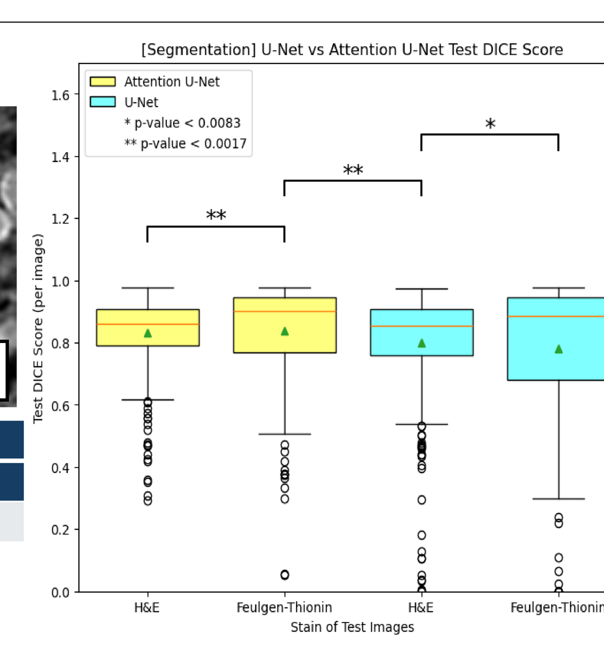


GENERALIZABILITY

Figure 10) Generalizability test of segmentation U-Net to H&E-stained images. Mann Whitney U-test shows both the Attention U-net and standard U-Net perform worse on test H&E images. Upon visual estimation of segmentation examples, there does not seem to be a noticeable difference between how the standard U-Net and Attention U-Net generalize. Although the DICE scores seem to be significantly lower on the H&E images, the model itself still performs extremely well with DICE score over 90%.



Measure	Number of Parameters	
	340,073	1,357,777
DICE Score	0.947	0.960



INTERPRETABILITY

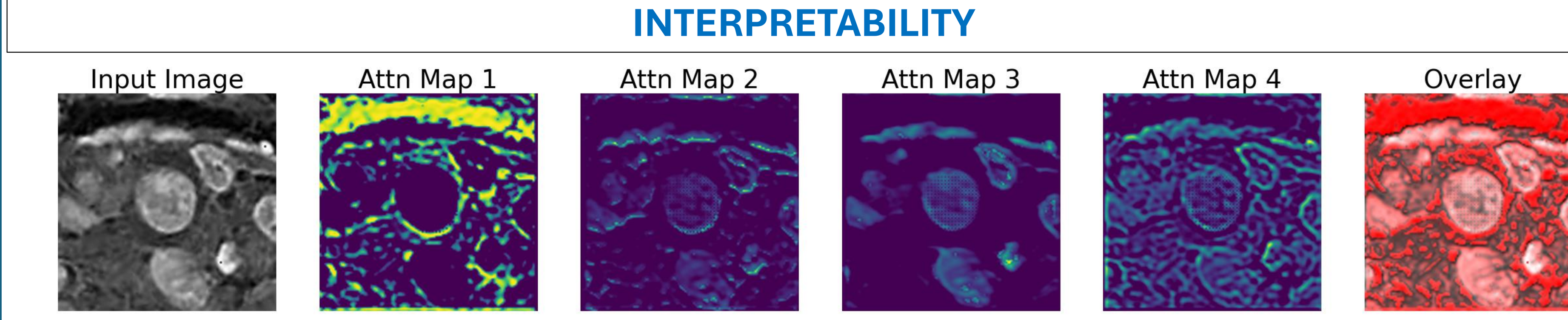


Figure 11) Attention activation maps of an Attention U-Net, with 4 initial convolutional channels. Attention Map 1 focuses on the background where nuclei are absent. The remaining attention maps focus on nuclei present in the image in various ways. For instance, Attention map 3 has removed the background stain signal. Attention map 2 and 4 seem to focus on textural feature elements of the nuclei, and stroma, to varying degrees.

Acknowledgements

References:

1. Macaulay C, Gallagher P. Sequential convolutional neural networks for nuclei segmentation. Published online July 21, 2022. Accessed December 18, 2024. https://patents.google.com/patent/WO2022150909A1/en
2. Oktay O, Schlemper J, Folgoc LL, et al. Attention U-Net: Learning Where to Look for the Pancreas. Published online May 20, 2018. doi:10.48550/arXiv.1804.03999
3. Ronneberger O, Fischer P, Brox T. U-Net: Convolutional Networks for Biomedical Image Segmentation. Published online 2015. doi:10.48550/ARXIV.1505.04597

BC CANCER RESEARCH
Provincial Health Services Authority