MuTILs: a multiresolution approach for computational TILs assessment using clinical guidelines

Mohamed Amgad[1]; **Lee A.D. Cooper**[1]

[1] *Department of Pathology, Northwestern University, Chicago, IL*

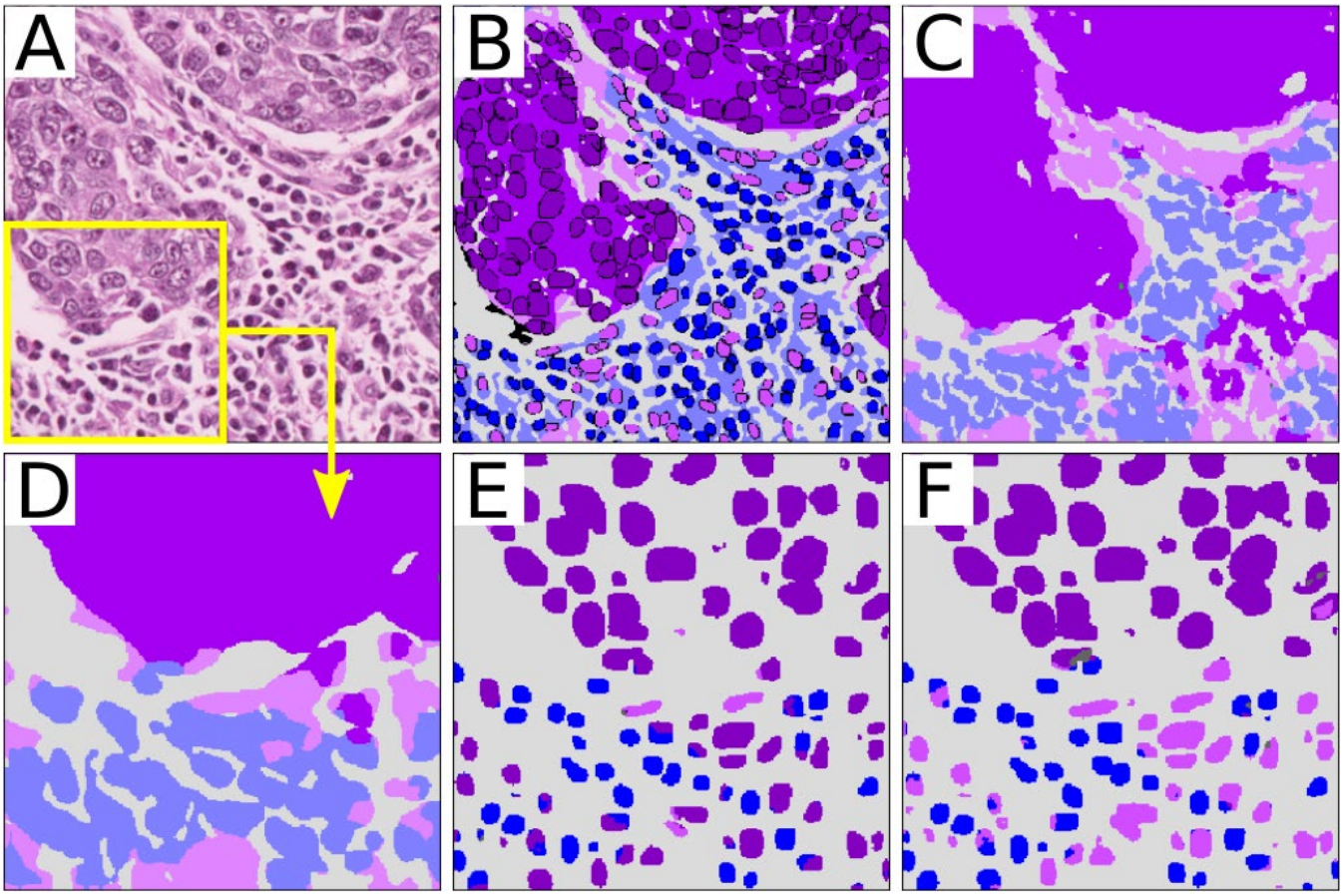
Background: Tumor-infiltrating lymphocytes (TILs) are an important diagnostic and predictive biomarker in solid tumors. Manual TILs assessment suffers from inter- and intra-rater variability, motivating the development of computational assessment tools. Most existing algorithms diverge from clinical recommendations, which include focused assessment within intra-tumoral stroma, and de-emphasizing TILs in distant stroma or hotspots. Previous work either relied on cellular classification alone, or naive combination of independent region and nuclear models without enforcing biological compatibility.

Technology: MuTILs relies on jointly-trained U-Net deep learning models for segmentation of tissue regions at 10x and cell nuclei at 20x magnification. We rely on average pooling of the 10x tumor and stromal predictions to prioritize high-power fields for assessment. The 10x model is used to improve the 20x predictions by concatenating intermediate feature map representations, consistent with the HookNet architecture. Additionally, upsampled region predictions are used to impose biological constraints to improve training efficiency and predictive accuracy.

Methods: We analyzed scans of FFPE H&E stained slides from 144 breast cancer patients from the TCGA dataset. Ground truth annotations were obtained from our public crowdsourcing datasets of tissue regions and nuclei (NuCLS). Segmented classes included tumor, stroma, TILs, normal acini, necrosis and others. 5-fold internal-external cross-validation was used to measure generalization performance.

Results: Figure 1 shows a sample hold-out set region prediction (C, D), as well as corresponding nuclear predictions before (E) and after imposition of the biological compatibility constraint (F). For reference, panel B shows the true region and nuclei classes. Tumor, stroma, and TILs are respectively colored violet, purple and blue. The biological constraint prevents misclassifications of large fibroblasts and activated lymphocytes as tumor cells. Pearson's correlation between computational TILs scoring using the ground truth and model predictions is high, both when global scoring is done using the 10x predictions ($R=0.761$, $p<0.001$) as well as when focused scoring is done using 20x predictions ($R=0.825$, $p<0.001$).

Conclusions: MuTILs enables highly accurate computational assessment of TILs in breast cancer in a fashion that is consistent with clinical scoring recommendations, paving the way for improved prognostication and therapeutic targeting.


A
B
C
D
E
F