Next Generation H&E Cell Modeling: Leveraging Multiplex Imaging to Create Large-Scale H&E Cell Annotations For Deep Learning

#3518

Becky Arbiv[1,2], Yuval Shachaf[1], Tal Dankovich[1], Ron Elran[1], Shai Bookstein[1], Avi Laniado[1], Tomer Dicker[1], Sun Dagan[1], Amit Bart[1], Oscar Puig[1], Kenneth Bloom[1], Ettai Markovits[1,2]

[1]Nucleai, Tel Aviv, Israel, [2]Corresponding authors: becky@nucleai.ai, ettai.markovits@nucleai.ai

Introduction

- Deep learning-based algorithms are powerful tools for modeling tissue architecture and understanding spatial biology but require a large number of accurate cell annotations for model training.
- Hematoxylin and eosin (H&E) stained slides offer valuable morphological information to pathologists but lack detailed cell type information.
- As a result, pathologist annotations for training these models are limited in both quality and quantity.
- Multiplex immunofluorescence imaging (mIF) provides a more comprehensive and accurate method for classifying cells.
- This study leverages mIF-based cell type predictions to create a large-scale dataset of H&E cell type annotations.

Results

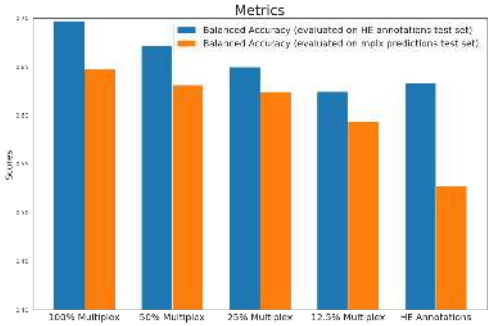

Figure 3: Balanced accuracy on H&E and mIF-based test annotations by amount of training annotations. Increasing the quantity of mIF based annotations improves model performance on both the H&E and mIF-based validation set.

- With equivalent annotation quantities, comparative analyses showed similar accuracy between the model trained on mIF-based annotations and one trained on H&E-based annotations on H&E test annotations
- The mIF-based model outperformed H&E-based model on the mIF-based test set.
- With an expanded training dataset, incorporating more mIF-based annotations, model performance significantly improved, reaching an average AUC of 0.91 on H&E test annotations and 0.90 on mIF predictions.

Conclusion

- This study underscores the value of mIF model predictions in advancing deep learning-based cell typing models in histopathology.
- A substantial dataset of accurate mIF-based annotations was successfully generated, significantly enhancing the accuracy and robustness of H&E cell typing models.
- These results highlight the emergence of mIF-based annotations as the next generation of H&E analysis, enabling pathologists to extract deeper insights from H&E-stained slides.
- Future work can leverage this method to identify challenging and rare cell types, capitalizing on the enhanced accuracy of mIF-based annotations.

Methods

- Employed a deep learning-based mIF analysis pipeline[1] to predict 12 cell types in a CODEX dataset which consisted of 140 tissue cores from 35 colorectal cancer patients, stained with 56 protein markers, and matched with H&E slides[2].
- Generated over 60,000 mIF-based annotations on corresponding H&E-stained slides using the mIF model predictions.
- Trained multiple H&E deep learning classifiers with an increasing amount of mIF-based annotations and compared performance to a classifier trained solely on pathologist-based H&E annotations.
- Evaluation was conducted on 11 test tissue cores using manual H&E-based annotations and mIF cell type predictions.

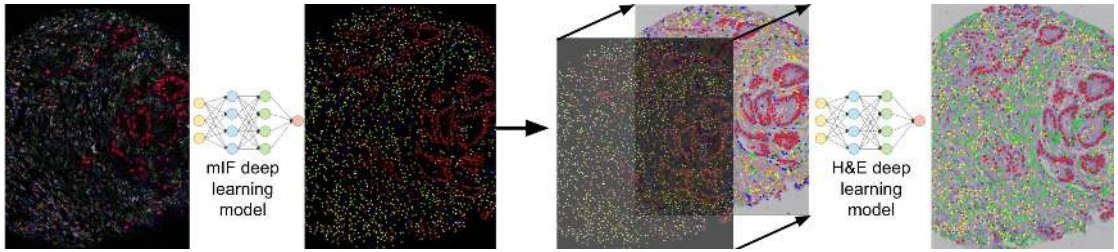

Figure 1: Workflow diagram: mIF slide is analyzed using Nucleai's deep learning analysis pipeline to generate cell type predictions. These predictions are then transferred to H&E slides and used as annotations to train deep learning models for H&E slides.

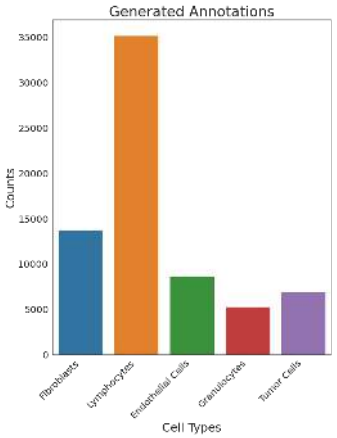

Figure 2: Number of generated annotations by cell class. Over 60,000 annotations were generated across all classes.

References

1. Markovits, E. et al. A novel deep learning pipeline for cell typing and phenotypic marker quantification in multiplex imaging. bioRxiv 2022.11.09.515776 (2022) doi:10.1101/2022.11.09.515776.
2. C. M. Schürch et al., "Coordinated Cellular Neighborhoods Orchestrate Antitumoral Immunity at the Colorectal Cancer Invasive Front," Cell, vol. 182, no. 5, pp. 1341-1359.e19, Sep. 2020, doi: 10.1016/j.cell.2020.07.005.