H&E-Based Deep Learning Model Predicts Immune Phenotypes in NSCLC

Poster # 168

Josh Och, Bolesław Osinski, Sina Zomorrodian, Rohan Joshi, Nike Beaubier, Martin C. Stumpe

Tempus AI, Inc., Chicago, IL

Disclosures: All authors are employees of and stockholders in Tempus AI, Inc., a for-profit company.

INTRODUCTION

Tumor-infiltrating lymphocytes (TIL) are a biomarker for response to immune checkpoint inhibitor (ICI) therapy. However, visually identifying TILs from hematoxylin and eosin (H&E) stained whole-slide images (WSIs) is expensive and time-consuming. Further, a lack of consensus on TIL scoring criteria can result in high levels of pathologist discordance. Here, we use an imaging-based deep learning model to classify H&E WSIs of non-small cell lung cancer (NSCLC) tumors into Inflamed (TIL-infiltrated) versus Non-Inflamed immune phenotypes.

METHODS

Primary site, resected NSCLC tumors were H&E stained and WSIs were scanned on a Leica Aperio GT450 and enriched for class balance between Inflamed and Non-Inflamed (n=169, Table I).

Slides were assessed as Inflamed or Non-Inflamed by 4 independent pathologists. WSIs were randomly split (50:50) into training and testing sets.

Our model used as input lymphocyte cell counts and tumor region detection from a previously developed deep learning model. If the number of lymphocytes within a tumor region tile was above a certain threshold (tile_threshold) that tile was predicted Inflamed. If the portion of tumor tiles that were predicted Inflamed was above a certain threshold (slide_threshold) for a given slide, the slide was predicted Inflamed. The tile_threshold and slide_threshold were trained on the training set, and evaluation was done on the testing set.

SUMMARY

- Our deep learning model performed as well as a held-out pathologist at separating Inflamed vs. Non-Inflamed samples.
  - Model Mean F1 score: 84.3% (95% CI 81.2% - 87.4%)
  - Held-Out Pathologist Mean F1 score: 86.7% (95% CI 78.5% - 95.0%)
- Inflamed vs. Non-Inflamed status has been shown to correlate with overall survival and progression free survival in response to immune checkpoint inhibitors.[1]

RESULTS

Characteristic	Value	Samples	Percent Inflamed (defined by majority of three pathologists)	Chi-square p-value
Stage	Stage 1	48	77%	0.018
	Stage 2	54	85%	
	Stage 3	50	64%	
	Stage 4	17	53%	
Grade Rollup	Low Grade	105	70%	0.204
	High Grade	64	80%	
Histology	Adenocarcinoma	86	71%	0.681
	Acinar Cell Carcinoma	50	74%	
	Mucinous Adenocarcinoma	33	79%	

Table 1. Cohort characteristics

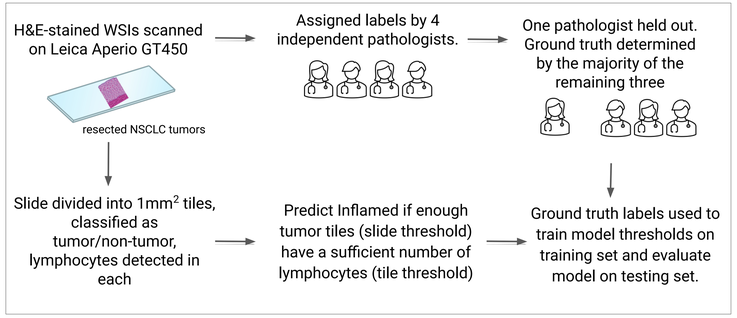


Figure 1. Workflow for generating labels and training and testing model.

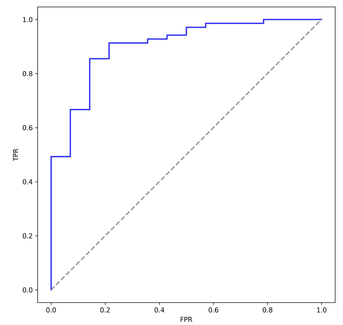
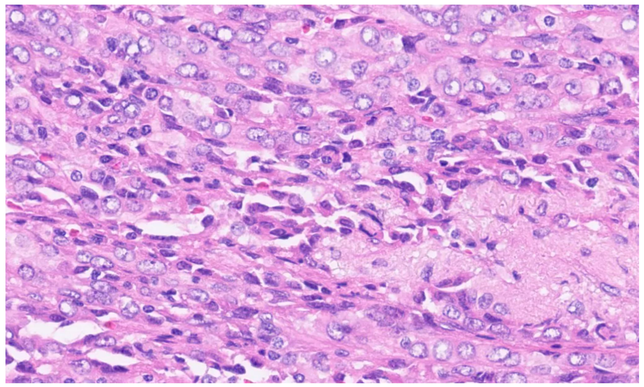


Figure 2. For a given held out pathologist, ROC curve showing the model's predictions against the consensus of three pathologists for a shift in slide_threshold. AUROC = 0.902.



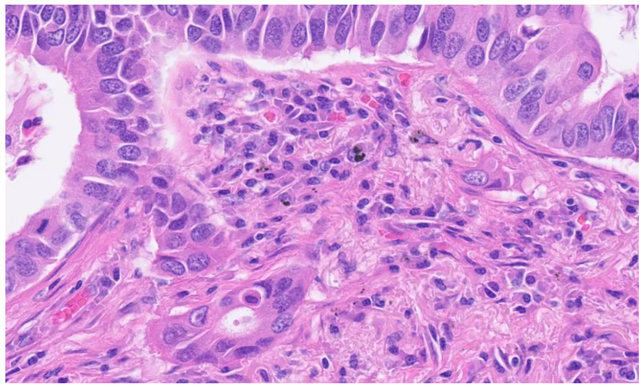


Figure 3. Select examples that highlight room for further model improvement. **Left**: Example of a slide for which the model predicted Inflamed and all four pathologists assessed Non-Inflamed. The model may have mistaken background lung epithelial cells and stromal cells for immune cells. **Right**: Example of a slide for which the model predicted Non-Inflamed and all four pathologists assessed Inflamed. The model may have mistaken immune cells for epithelial cells and stromal cells.

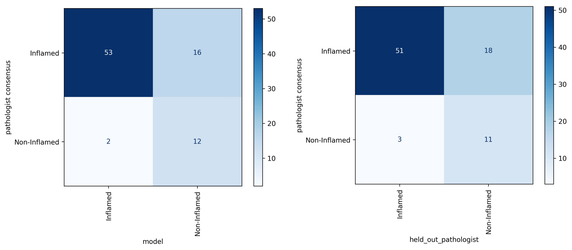


Figure 4. For a given held out pathologist, Left: Model's performance against the consensus of three pathologists. Right: Held out pathologist's performance against the consensus of three other pathologists.

Acknowledgments: We thank Dana DeSantis from the Scientific Communications team for poster development.

1. Park, Sehhoon. "Artificial Intelligence–Powered Spatial Analysis of Tumor-Infiltrating Lymphocytes as Complementary Biomarker for Immune Checkpoint Inhibition in Non–Small-Cell Lung Cancer." National Library of Medicine, 2022,

Correspondence: Josh.Och@tempus.com