



上海市胸科醫院
上海交通大學醫學院附屬胸科醫院
SHANGHAI CHEST HOSPITAL

P1.141: An Interpretable Deep-Learning Pipeline for Risk Stratification in Stage I LUAD Using Histologic Distribution on H&E Slides

Mengqi Jiang, Liqiang Qian, Yuanyuan Fu, Yujia Shen, Wen Feng, Yuchen Han, Xuwei Cai, Xiaolong Fu

Shanghai Chest Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai, China

Presenter: jmq19980703@163.com

Correspondence: xlFu1964@hotmail.com



IASLC 2026 World
Conference on Lung Cancer

SEPTEMBER 12 - 15, 2026
SEOUL, REPUBLIC OF KOREA

INTRODUCTION

Introduction:

Despite overall favorable prognosis, disease recurrence remains a significant challenge for patient with stage I lung adenocarcinoma (LUAD). Current risk stratification methods based on clinicopathological factors lacks sufficient precision to guide adjuvant therapy decisions. Recent advances in computational pathology have introduced large-scale foundation models pretrained on massive histopathology datasets. These pathology foundation models have strong generalization capabilities but are often not satisfying in specific clinical scenarios, achieving C-index of only 0.58-0.68 in LUAD prognosis. Moreover, their black-box nature and substantial computational demands limit practical deployment in routine clinical workflows.

This study aims to establish a multi-step (non-end-to-end) and interpretable deep-learning pipeline that leverages the spatial distribution of histological components from whole-slide images (WSIs) to precisely identify patients with stage I LUAD at high risk of postoperative recurrence and those who may benefit from adjuvant therapy.

METHODS

Design:

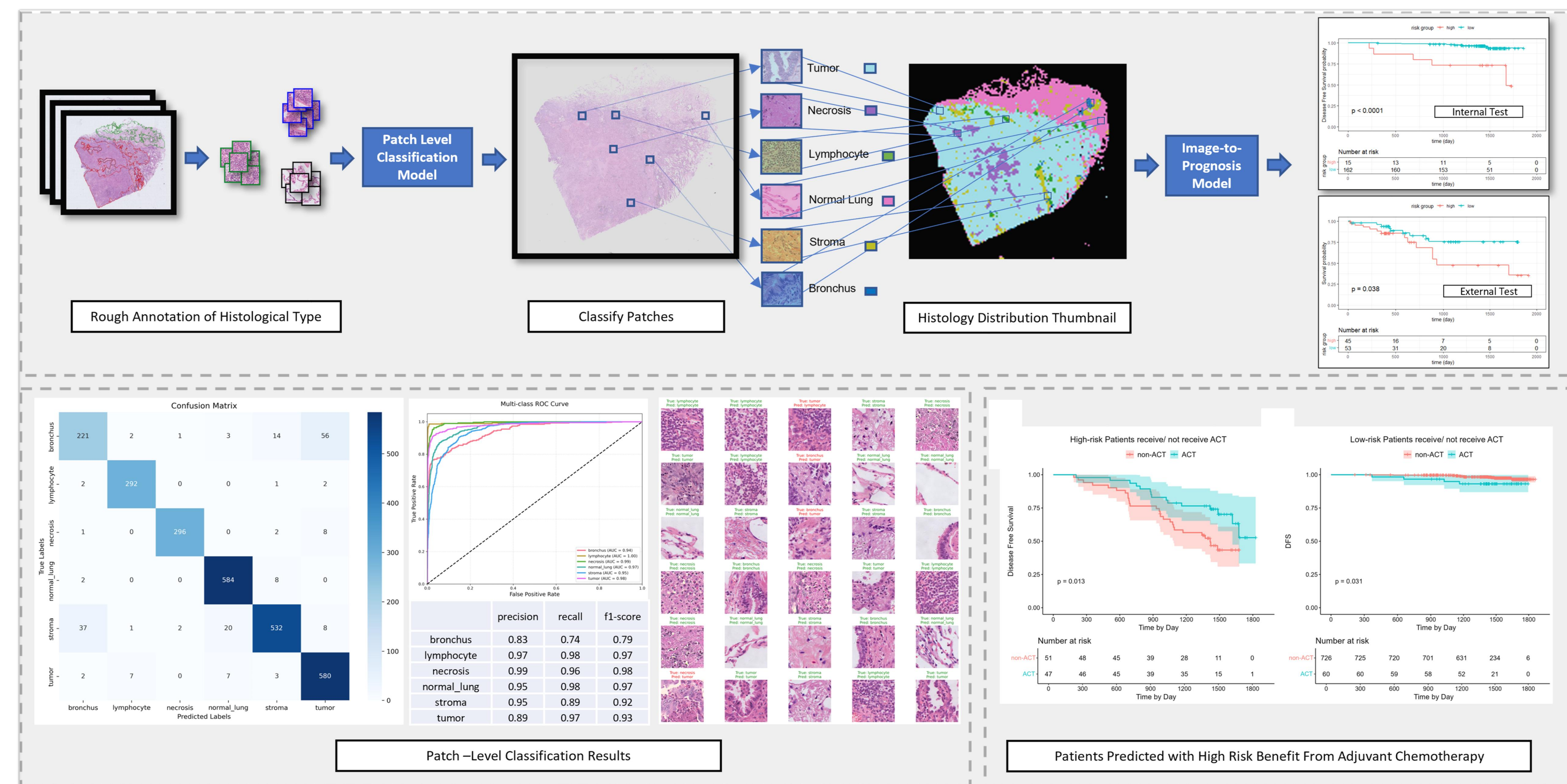
This study retrospectively enrolled a consecutive cohort of stage I LUAD patients who underwent radical resection at Shanghai Chest Hospital between August 1 and December 31, 2018. .

WSIs and clinical data from these patients were collected and histology annotations were roughly created on 50 WSIs using QuPath. A ResNet50-based classifier was first trained to classify image patches into six histological types (normal lung, tumor, stroma, lymphocytes, bronchus, necrosis). This classifier was then deployed to generate a comprehensive, patient-specific spatial distribution map that visualizes the compositional architecture of each WSI.

Subsequently, 256*256 pixel thumbnails of histology map were created to fit in a Vision Transformer (ViT) model, which encode these thumbnails into a global feature representation for risk prediction. Cohort from our institution was randomly split into training set (80%) and internal test set (20%). The CPTAC-LUAD dataset was used for external validation.

RESULTS

- A total of 884 patients were included in this research. The 5-year disease-free survival (DFS) rate was 89% for the cohort.
- The histological patch classifier demonstrated good performance, with an overall accuracy of 0.933 on patch level from external test dataset.
- The prognostic pipeline achieved a C-index of 0.840 for DFS on internal test set and a C-index of 0.659 for predicting overall survival on CPTAC-LUAD dataset.
- Patients was divided into high-risk and low-risk group by this prognostic model. Threshold was defined as the 90th percentile risk-score of internal cohort. The CPTAC cohort was divided into two strata, with significant different prognosis ($p = 0.038$, $HR = 0.415$, 95%CI: 0.176-0.978).
- In our internal dataset, model predicted high-risk patients who received adjuvant chemotherapy had superior DFS compared with those who did not ($p = 0.013$, $HR=0.455$, 95%CI: 0.240-0.862).



DISCUSSION AND CONCLUSION

We developed and externally validated an interpretable, deep-learning pipeline by visualizing spatial histological distribution on WSIs to predict postoperative recurrence risk in stage I LUAD, demonstrating that critical prognostic information is embedded within the spatial architecture of tumor microenvironment. This model could serve as a novel decision-support tool to provide data-driven support for personalizing adjuvant therapy decisions.

Although performance attenuated on external validation, probably due to broader stage in CPTAC and the greater heterogeneity of overall survival endpoint, the model performance remained comparable to large-scale foundation models, which suggests that task-specific models might be better suited for clinical settings.