

Deep-learning-based morphometric profiling of circulating tumor cells predicts EGFR mutation status in non-small cell lung cancer

Y. Feng¹, X. Wu², L. Shen¹, B. Cheng¹, X. Huang³, C. Han⁴, H. Cen³, S. Liu³, J. Xia⁵, J. He¹, W. Liang¹

1. the First Affiliated Hospital of Guangzhou Medical University, Guangzhou/CN ,2. Peking University Shenzhen Hospital, Shenzhen/CN ,3. Guangdong Polytechnic Normal University, Guangzhou/CN ,4. Anfang Biotech Co., Ltd., Guangzhou Life Science Center, Guangzhou/CN ,5. Guangzhou Improve-medical Laboratory,Guangzhou/CN

INTRODUCTION

In non-small cell lung cancer (NSCLC), obtaining accurate genetic information is often hindered by the need for invasive tissue sampling. Here we report an ensemble computational pipeline that predicts EGFR mutation status using morphological features of circulating tumor cells (CTCs).

OBJECTIVE

A total of 1,728 CTC images from 580 patients were analyzed, combining 31 manually defined morphometric traits with deep-learning representations (ResNet-18) within an: AutoGluon-based stacking framework. Our goal was to determine whether CTC morphology can distinguish EGFR-mutant from wild-type NSCLC and whether these signals remain specific when multiple oncogenic drivers are pooled.

STUDY DESIGN

1 CTC capture

Automated microfluidic platform

2 Image QC

1,728 high-quality CTC images

3 Feature extraction

31 morphometric traits + ResNet-18 embedding

4 Ensemble modeling

AutoGluon-based stacking framework

5 Output

EGFR mutation prediction and aggregated driver analysis

RESULTS

The hybrid morphometric deep-learning model distinguished EGFR-mutant from wild-type CTCs with ROC AUC = 0.8986 and accuracy = 85.2%.

Reduced cytoplasmic texture contrast was the most influential indicator (Cohen's d = -0.22).

FIGURE 1

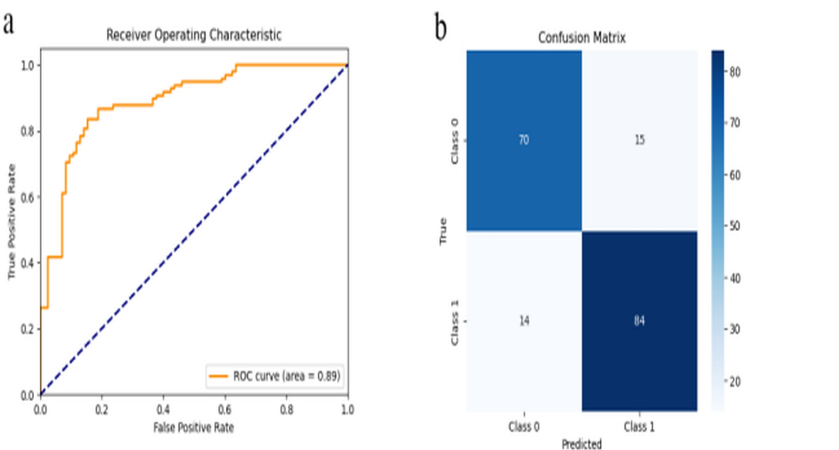


Figure 1: EGFR mutation prediction. The model achieved ROC AUC=0.90 and accuracy=85.2% in distinguishing EGFR-mutant from wild-type CTCs.

FIGURE 2

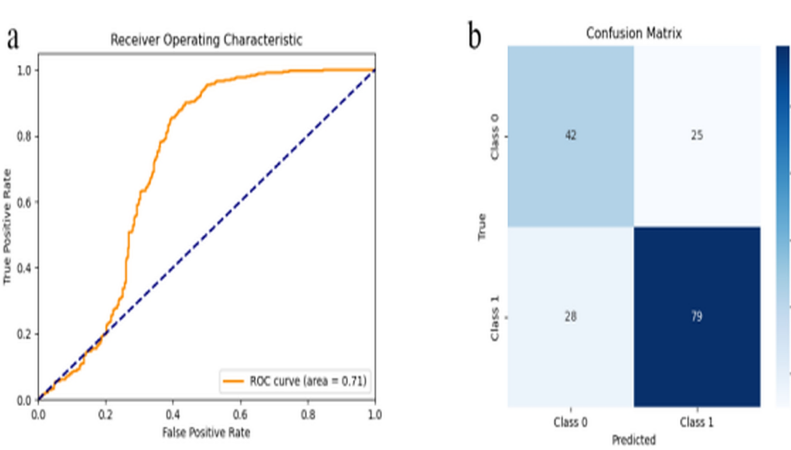


Figure 2: Aggregated driver negative mutation prediction (including KRAS, ALK, and ROS1). Performance declined, with accuracy=73.8% and ROC AUC=0.71.

FIGURE 3

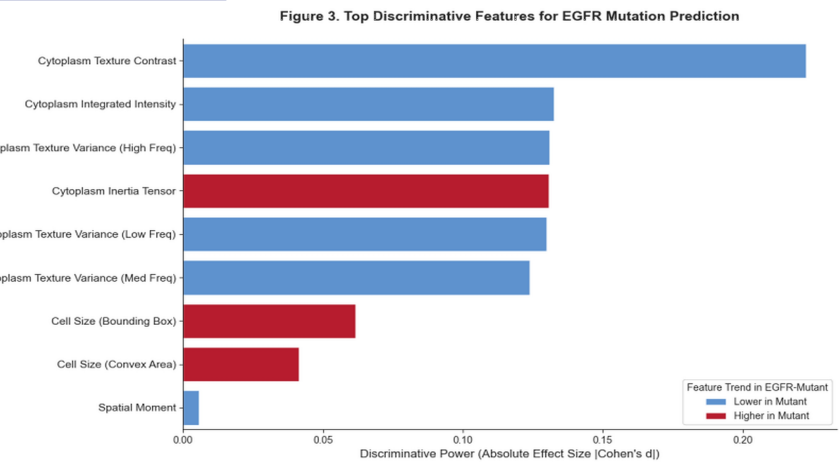


Figure 3: Top discriminative features for EGFR mutation:prediction.:Reduced cytoplasmic texture contrast was the strongest indicator for EGFR-mutant CTCs.

FIGURE 4

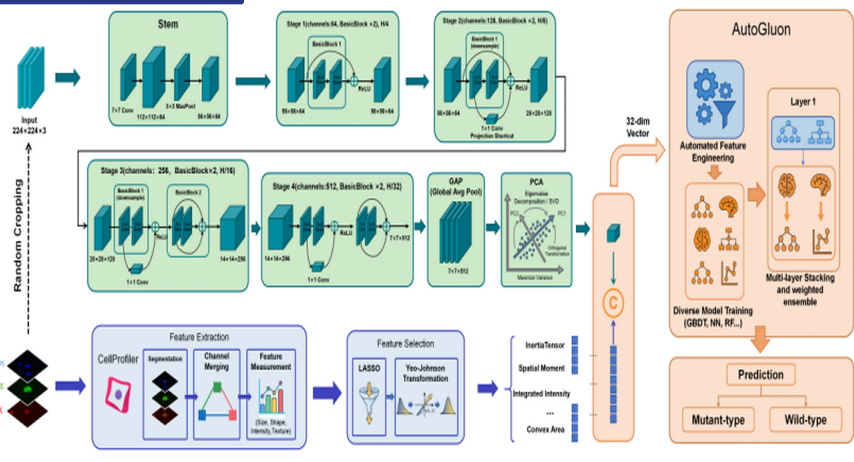


Figure 4: Computational pathology workflow integrating CellProfiler-derived morphometrics, ResNet-18 representations, and an AutoGluon stacking ensemble.

Model Performance Summary

Task	Accuracy	Precision	Recall	F1-score	ROC AUC
EGFR mutation prediction (n=943 images)*	0.8519	0.8482	0.8618	0.8542	0.8986
Aggregated driver prediction (n=869 images)**	0.7342	0.7630	0.7376	0.7500	0.7100

Note:*Model performance evaluated on the full image set. **Subset analysis for heterogeneous oncogenic drivers. ROC AUC, area under the receiver operating characteristic curve.

DISCUSSION & CONCLUSION

CTC morphology appears to encode mutation-specific biological states rather than a generic malignant phenotype. EGFR-mutant CTCs showed reduced cytoplasmic texture contrast, consistent with smoother intracellular organization. This image-based pipeline may serve as a rapid, non-invasive adjunct to conventional genotyping when tissue is scarce or time-sensitive decisions are required.

ACTIONABLE TAKE-AWAY

1. A non-invasive “liquid biopsy of form” may help triage patients for urgent molecular testing.
2. Combining handcrafted traits with ResNet-18 features improves sensitivity to subtle sub-visual patterns.
3. Mutation-specific modeling is more effective than pooling heterogeneous oncogenic drivers.

FUTURE DIRECTIONS

1. Multi-center external validation with standardized CTC imaging.
2. Integration with plasma cfDNA and imaging/metabolic parameters.
3. Prospective evaluation in clinical decision-support workflows.

FUTURE DIRECTIONS

Key references: Sun et al., CA Cancer J Clin 2021; Paweletz, J Cancer2020; Mack et al. N Engl JMed 2017; Imielinski et al., Nat Rev Clin Oncol2012; Hein-Palinsli & Paweletz, Cancer Discov 2016; Keller & Pantel, Nat Rev Cancer 2019; Baird et al, Clin Chem 2012; Esteva et al., Nature2017. Acknowledgements: The figures and tables are original.No conflicts of interest declared.