

# Artificial Intelligence–Based Model Exploiting Hematoxylin and Eosin Images to Predict Rare Gene Mutations in Patients With Lung Adenocarcinoma

Peiling Yu, MD<sup>1</sup>; Weixing Chen, PhD<sup>2</sup> ; Nan Liu, MD<sup>3</sup>; Yang Yu, MD<sup>1,4</sup>; Hongyu Guo, MD<sup>1</sup>; Yinan Yuan, MD<sup>1</sup>; Weilin Guo, MD<sup>1</sup>; Yini Alatan, MD<sup>1</sup> ; Jinming Zhao, MD<sup>1</sup>; Hongbo Su, MD<sup>1</sup>; Siru Nie, MD<sup>1</sup>; Xiaoyu Cui, PhD<sup>2</sup>; and Yuan Miao, MD<sup>3</sup> 

DOI <https://doi.org/10.1200/JCO.2025.00093>

## ABSTRACT

**PURPOSE** Accurately identifying gene mutations in lung cancer is crucial for treatment, while molecular diagnostic methods are time-consuming and complex. This study aims to develop an advanced deep learning model to address this issue.

**METHODS** In this study, the ResNeXt101 model framework was established to predict the gene mutation status in lung adenocarcinoma. The model was trained and validated using data from two cohorts: cohort 1, comprising 144 patients from the First Affiliated Hospital of China Medical University, and cohort 2, which includes 69 patients from the The Cancer Genome Atlas–Lung Adenocarcinoma public database. The model was trained and validated on the two data sets, respectively, and they served as external test sets for each other to further verify the performance of the model. Additionally, we tested the trained model on a metastatic cancer data set, which included metastases to organs outside the lungs. The performance of the model was evaluated using the AUC, accuracy, precision, recall, and F1 score.

**RESULTS** In cohort 1, the model achieved an AUC ranging from 0.93 to 1. In the external test on cohort 2, it performed well in predicting five of the six genes (AUC = 0.85–1). When tested on the metastatic cancer data set, it successfully predicted mutations of three of the six genes (AUC = 0.72–0.80).

**CONCLUSION** The artificial intelligence model developed in this study has a high accuracy in predicting gene mutations in lung adenocarcinoma, which is conducive to improving the management of patients with lung adenocarcinoma and promoting precision medicine.

## ACCOMPANYING CONTENT

 Appendix

Accepted August 13, 2025

Published September 26, 2025

JCO Clin Cancer Inform

9:e2500093

© 2025 by American Society of  
Clinical Oncology

Creative Commons Attribution  
Non-Commercial No Derivatives  
4.0 License

## INTRODUCTION

Lung cancer is the most prevalent cancer globally and remains the leading cause of cancer-related mortality, with a dismal 5-year survival rate of only 15%. This disease is characterized by significant heterogeneity and is broadly classified into two major types based on tumor histology: small cell lung cancer, which accounts for approximately 15% of cases, and non–small cell lung cancer (NSCLC).<sup>1</sup> Among the subtypes of NSCLC, lung adenocarcinoma (LUAD) is the most common.<sup>2</sup>

Treatment options for LUAD include surgery, radiotherapy, chemotherapy, targeted therapy, immunotherapy, or a combination of these approaches.<sup>3</sup> Targeted therapies, such as those targeting *EGFR*, *ALK*, *RET*, and *ROS1*, have shown significant success.<sup>4–6</sup> The 2022 National Comprehensive Cancer Network (NCCN) guidelines emphasize the importance of identifying genomic alterations including mutations

in *EGFR*, *ALK*, *HER2*, *c-MET*, *ROS1*, *KRAS*, and *RET* for personalized treatment.<sup>7–9</sup> However, gene testing is costly and time-consuming, and biopsy tissue is often insufficient for subsequent molecular testing.

Hematoxylin and eosin (H&E)–stained histology sections reflect rich tumor morphology. Specific driver mutations may endow characteristic histopathologic features (eg, nuclear pleomorphism, growth patterns), which may be undetectable by human eyes. Deep convolutional networks trained on H&E images can learn complex genotype–phenotype associations to predict molecular alterations from routine pathology slides. Artificial intelligence (AI) can extract microscopic features from these images to predict gene mutations.<sup>10</sup> Deep learning–based AI models have been widely applied in tumor pathology.<sup>11–14</sup> Some artificial intelligence models have been developed and used for the histologic subtype classification and prognosis prediction of lung cancer.<sup>15,16</sup> Neural networks have also demonstrated

## CONTEXT

### Key Objective

To develop a deep learning model using hematoxylin and eosin (H&E)–stained slides to predict six rare gene mutations in patients with lung adenocarcinoma.

### Knowledge Generated

The ResNeXt101-based model achieved high accuracy (AUC = 0.93–1 in cohort 1) and could predict mutations in metastatic cancer, outperforming previous models.

### Relevance (P.-M. Putora)

The presented Artificial Intelligence-based model can predict mutations of treatment-relevant genes in lung adenocarcinoma using H&E-stained slides. Such an approach may be the basis to reduce costs and improve workflows.\*

\*Relevance section written by JCO Clinical Cancer Informatics Associate Editor Paul–Martin Putora, MD, PhD, MA.

potential in predicting gene mutations in LUAD. For instance, a study by Coudray et al,<sup>17</sup> demonstrated that deep learning could predict mutations in six genes associated with LUAD, achieving an AUC of 0.73 to 0.85. In 2022, Terada et al<sup>18</sup> developed a model capable of predicting *ALK* mutations in NSCLC, with an AUC of 0.73. Beyond these studies, other investigations have explored AI-based biomarker detection in oncology, including colorectal cancer<sup>19</sup> and breast cancer.<sup>20</sup> Incorporating a broader spectrum of AI applications in pathology strengthens the foundation for deploying deep learning in mutation prediction and highlights areas for further research, such as domain adaptation and cross-tissue generalizability. Building on these advances, we explored the use of neural networks for gene mutation prediction and developed a convolutional neural network–based model capable of predicting six relatively rare gene mutations relevant to patient treatment as outlined in the NCCN guidelines. Our model outperforms previously published works in terms of accuracy. Notably, it maintained high performance when tested on an independent external data set, demonstrating exceptional generalization ability and the potential to benefit a broader patient population through targeted therapies. Additionally, we established a test set consisting of patients with metastatic LUAD. To the best of our knowledge, this is the first time a gene mutation prediction model trained on pathological images from primary LUAD has been applied to predict gene mutations in metastatic cancer, representing a novel exploration in this field.

## METHODS

### Data Preprocessing

In this study, eligible lung adenocarcinoma patient slides from the First Affiliated Hospital of China Medical University and TCGA database were collected to build a database. This study was conducted in accordance with the Declaration of

Helsinki and received ethical approval from the Ethics Committee of the First Hospital of China Medical University (Approval No. AF–SOP–07–1.2–01; institutional review board No. [2024]–1077).

Regions of interest (ROIs) for invasive carcinoma were annotated, excluding necrotic areas, benign stroma, and epithelium. ASAP software marked tumor boundaries, and annotations were exported as XML with X and Y coordinates. These were used to extract ROIs from slide images, ensuring training and testing on annotated tumor regions only.

Slides from both cohort 1 and cohort 2 were stratified by mutation type and then randomly split at the patient level. Specifically, for cohort 1 (n = 144 patients), 70% of patients were allocated to the training set, 10% to the validation set, and 20% to the internal test set; the cohort 2 test set was reserved as an external test set for cohort 1. Conversely, for cohort 2 (n = 69 patients), patients were similarly divided into 70% training, 10% validation, and 20% internal test sets, whereas the cohort 1 test set served as its external test set. At 400× magnification, the slide images were divided into nonoverlapping tiles of 224 × 224 pixels. Regions containing excessive background, no tissue, or nontumor areas were excluded. Data augmentation was applied to the training tiles by rotating them at 90°, 180°, and 270°, as well as flipping them horizontally and vertically. Additionally, normalization was performed using the image mean and standard deviation, following the same settings as ImageNet.

### ResNeXt101 Model Framework

Our framework is designed to predict the mutation status of six genes (*ALK*, *HER2*, *KRAS*, *RET*, *MET*, and *ROS1*) from whole-slide images (WSIs) of patients with LUAD using neural networks. We employed the ResNeXt101 model for

this task. This model uses an end-to-end learning method, eliminating the need for manual extraction of heuristic features. Compared with traditional machine learning methods, it simplifies the classification and recognition tasks, making it more efficient and accurate.

The model training was conducted using two strategies: transfer learning and full training. For transfer learning, the model parameters were initialized with the optimal values from the ImageNet Competition. Full-parameter fine-tuning was performed by adjusting the parameters of the last network layer using backpropagation on the training samples. To evaluate the impact of ROIs on training, we implemented three distinct training schemes to predict gene mutation status: a. Unannotated Binary Classifier: Tiles were assigned labels based on the overall WSI label (positive or negative) without considering ROIs during training. b. Annotated Binary Classifier: Only tiles within the ROIs were used for training. External regions, such as stromal cells, necrotic areas, or mixed regions of tumor and normal cells, were excluded from the training process. WSI labels were applied solely to the tiles within the ROIs. c. Annotated Three-Branch Classifier: Both ROI and non-ROI regions were used to train a multibranch classifier. Tiles within the ROIs were labeled as positive or negative based on the WSI label, whereas tiles outside the ROIs were labeled as “others,” independent of the WSI label. A similar strategy was adopted to train a binary classifier for mutation prediction. The model used cross-entropy as the loss function, and mutation probabilities were calculated for each tile. The training was conducted using stochastic gradient descent optimization with a learning rate of 0.0001 and a weight decay of 0.05. Additionally, the dropout probability for the linear layer was set to 0.5. The training process for each model required approximately 1.2 hours using an NVIDIA RTX A5000 graphics processing unit.

## Visual Interpretation of the Diagnosis

To assess the morphological patterns used by the annotated model, we used the Class Activation Mapping (CAM) technique. Areas of the image that the model focused on are visualized in red, indicating high attention, whereas regions of lesser importance are displayed in blue. This approach provides a visual representation of the model’s decision-making process, offering insights into the features influencing its predictions.

## Statistical Methods

To comprehensively evaluate the model’s accuracy in predicting LUAD pathology, particularly its ability to predict gene mutation status, we employed the receiver operating characteristic (ROC) curve as a primary evaluation tool. Several key metrics were calculated, including the AUC, accuracy, precision, recall, and F1 score.

## RESULTS

### Data Preprocessing and Cohort Partitioning

To predict rare gene mutations in the H&E WSIs of LUAD, we compiled a data set consisting of 213 H&E slides from patients with LUAD from patients with LUAD with known mutation statuses for *ALK*, *HER2*, *KRAS*, *RET*, *MET*, and *ROS1* genes. The data set was sourced from the First Affiliated Hospital of China Medical University and the TCGA database, forming two distinct cohorts.

Cohort 1 consisted of 144 patients selected from a total of 4,064 individuals who underwent gene testing between August 2018 and August 2022 at the First Affiliated Hospital of China Medical University. The patients’ age ranged from 23 to 81 years, with a median age of 60 years. The male-to-female ratio was 59:85, and 25 patients had a history of smoking. There were 70 surgical specimens and 74 biopsy specimens. The specific specimen distribution is shown in Appendix [Table A1](#).

Cohort 2 was derived from The Cancer Genome Atlas–Lung Adenocarcinoma database. After a pathologist review, 69 patients were selected for inclusion in this cohort. The patients’ age ranged from 41 to 85 years, with a median age of 65 years. The male-to-female ratio was 34:35, and 50 patients were smokers. All specimens are surgical specimens.

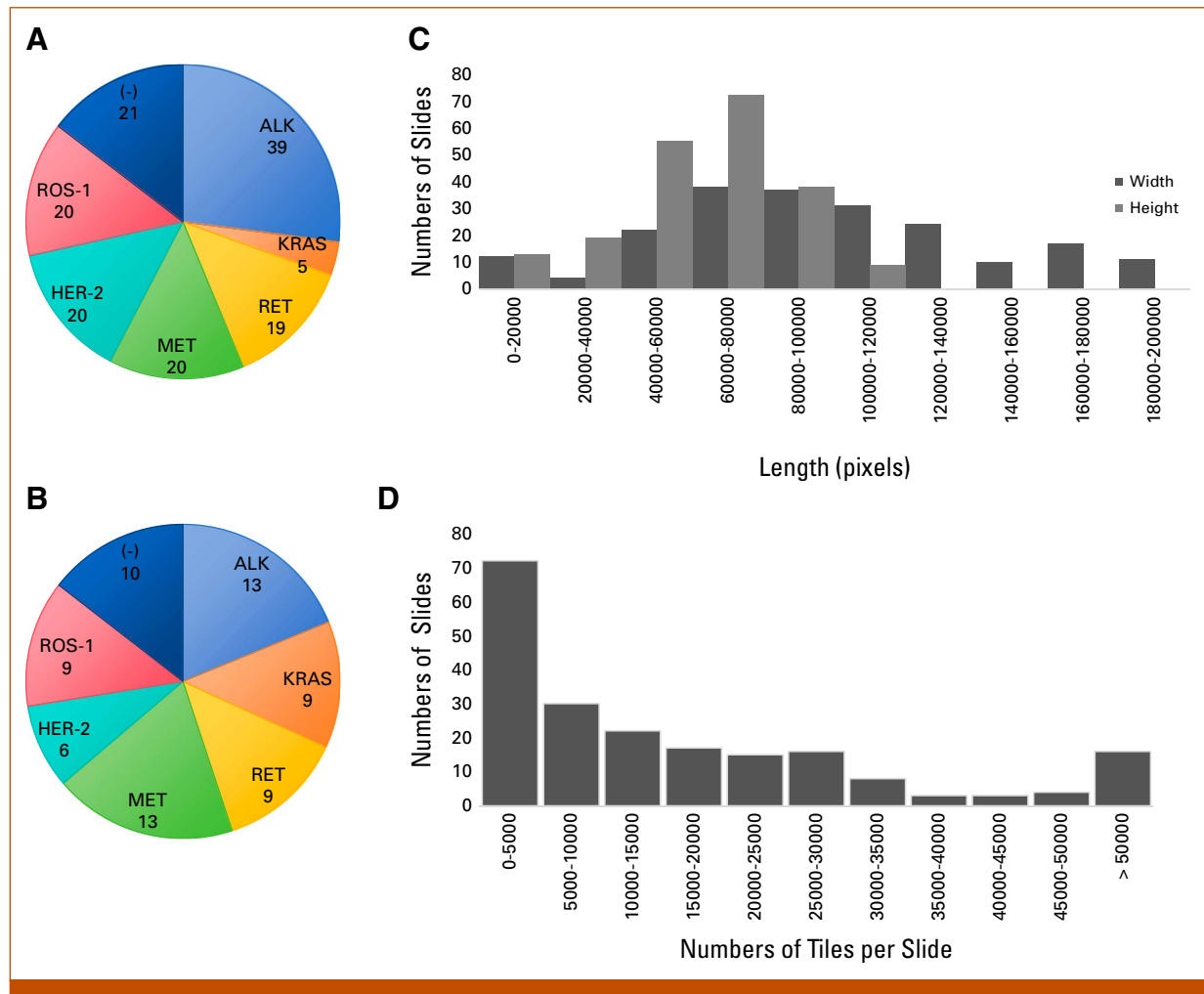
The gene distribution in each cohort is summarized as follows ([Fig 1](#)):

Cohort 1, 39 cases of *ALK* mutation, 20 cases of *HER2* mutation, five cases of *KRAS* mutation, 19 cases of *RET* mutation, 20 cases of *MET* mutation, 20 cases of *ROS1* mutation, and 21 cases with no mutation in any of the above genes. Gene detection was performed using polymerase chain reaction.

Cohort 2, 13 cases of *ALK* mutation, six cases of *HER2* mutation, nine cases of *KRAS* mutation, nine cases of *RET* mutation, 13 cases of *MET* mutation, nine cases of *ROS1* mutation, and 10 cases with no mutation in any of the genes. Gene detection was carried out using next-generation sequencing (NGS).

Experienced pathologists annotated, segmented, and denoised the tumor regions in the H&E-stained slides. The annotated tumor regions were then cropped into  $224 \times 224$ -pixel image tiles, which were randomly divided into training, validation, and test sets ([Table 1](#)).

To assess model generalizability to metastatic disease, we assembled an independent metastatic test set comprising 26 WSIs from 26 patients. Slides were drawn from metastatic lung adenocarcinoma sites and stratified by driver mutation. At the slide level, the cohort includes *ALK*-mutant: nine



**FIG 1.** Data distribution: (A) Proportion of cases in each gene mutation group in cohort 1; (B) proportion of cases in each gene mutation group in cohort 2; (C) size distribution of the image widths (black) and heights (gray); (D) distribution of the number of tiles per slide.

cases; *HER2*-mutant: one case; *KRAS*-mutant: two cases; *RET*-mutant: three cases; *MET*-mutant: two cases; *ROS1*-mutant: six cases; mutation-negative: three cases. For patch-level evaluation, annotated tumor regions were tiled into  $224 \times 224$  px patches, yielding *ALK*: 14,696 patches; *HER2*: 8,101 patches; *KRAS*: 3,637 patches; *RET*: 4,297 patches; *MET*: 3,477 patches; *ROS1*: 18,902 patches; no mutation: 3,348 patches. This metastatic cohort enables

rigorous testing of the primary LUAD-trained model on distant disease sites.

### Model Architecture and Training Performance

The goal of this study was to develop a deep learning model that uses H&E-stained slides as the sole input to predict rare gene mutations in NSCLC. We trained the model using 144

**TABLE 1.** Patch Distribution in the Model

| Dataset and Cohort  | ALK     | HER2    | RET     | MET     | ROS1    | KRAS   | (-)     |
|---------------------|---------|---------|---------|---------|---------|--------|---------|
| Cohort 1            |         |         |         |         |         |        |         |
| Training set, No.   | 238,571 | 148,980 | 228,877 | 214,237 | 166,802 | 69,600 | 124,299 |
| Validation set, No. | 40,750  | 26,715  | 10,332  | 56,886  | 97,089  | 26,494 | 33,463  |
| Testing set, No.    | 133,991 | 95,278  | 93,224  | 143,457 | 128,698 | 70,590 | 21,953  |
| Cohort 2            |         |         |         |         |         |        |         |
| Training set, No.   | 204,810 | 31,403  | 161,088 | 153,307 | 256,629 | 80,851 | 102,396 |
| Validation set, No. | 11,481  | 50,569  | 7,372   | 32,237  | 124,109 | 3,167  | 967     |
| Testing set, No.    | 29,051  | 477     | 32,586  | 56,858  | 47,816  | 63,507 | 40,598  |

WSIs from our institution and 69 WSIs from the TCGA database to train the model. Because the size of a WSI is too large (Fig 1C) to be directly used as input for the neural network, we divided each WSI into  $224 \times 224$  pixel tiles at  $400\times$  magnification for training, validation, and testing. Depending on the original slide size, this results in dozens to thousands of tiles per slide (Fig 1D).

Using the calculation strategy depicted in Figure 2, we developed a deep learning algorithm to classify gene mutations in *ALK*, *HER2*, *KRAS*, *RET*, *MET*, and *ROS1*. The model was trained separately on the two cohorts. In cohort 1, the AUC values for predicting each gene mutation ranged from 0.93 to 1 (Fig 3A). In cohort 2, the AUC values ranged from 0.54 to 0.98 (Fig 3B).

To further assess the generalization ability of our model, we used the two cohorts as independent external test sets for one another. When the model was trained with cohort 1 and tested with cohort 2, it constituted cohort 3. When trained with cohort 2 and tested with cohort 1, it formed cohort 4. Before testing, we observed noticeable differences in the color characteristics of the pathologic images between the two cohorts, particularly among the negative cases. To address this, we preprocessed the color characteristics of the test set images (a comparison of the

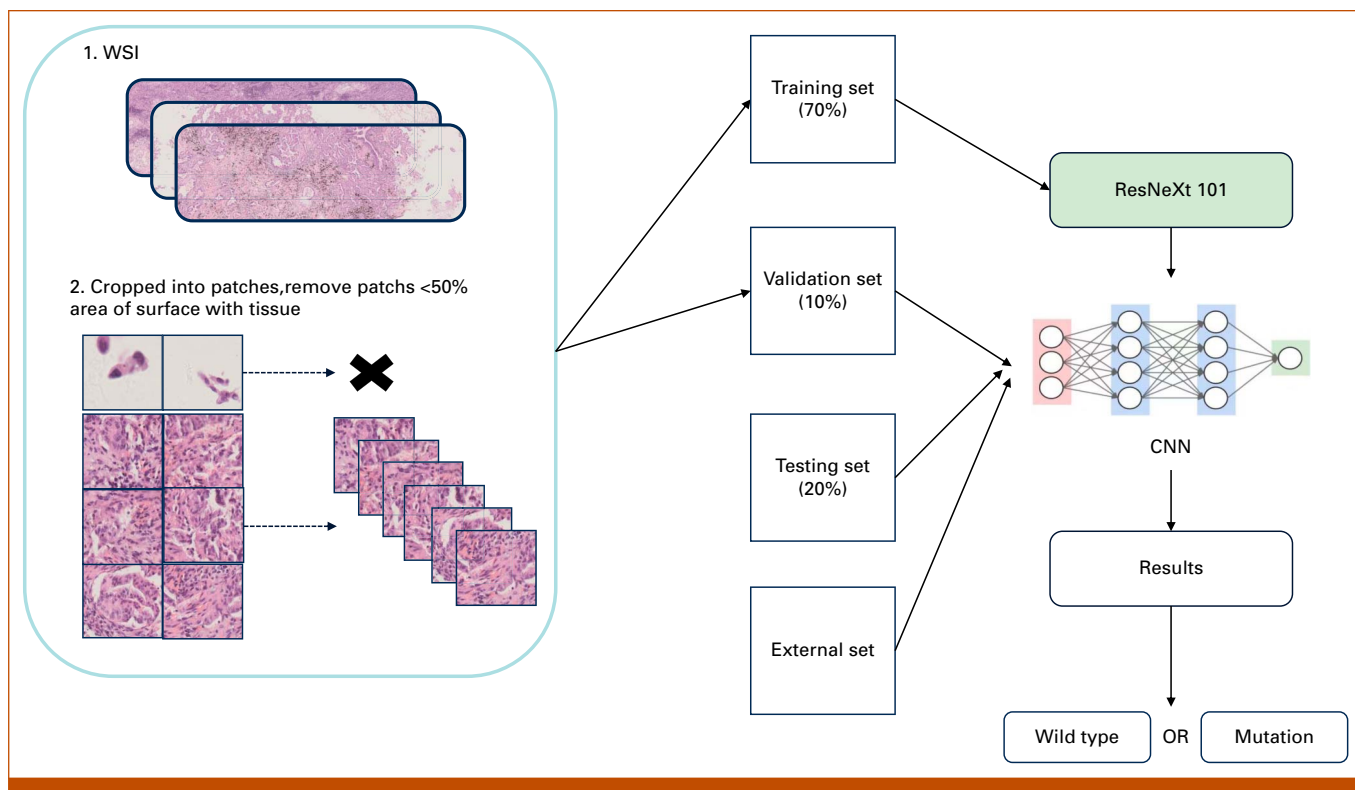
images before and after processing is shown in Appendix Fig A1).

When the model trained in cohort 1 was tested in cohort 2, the model exhibited strong predictive capability for five of the six genes (AUC = 0.85–1; Fig 3C). However, its performance in predicting *HER2* mutations is poor (AUC = 0.44), which indicates that there may be challenges in extracting features related to this specific genetic alteration from H&E-stained sections. This may stem from low *HER2* prevalence in the data set and subtle morphologic features. Future studies could explore higher-resolution imaging, additional staining methods, or multimodal models integrating genetic and histopathologic data.

However, when the model trained in cohort 2 was tested in cohort 1, its performance was poorer, with AUC values ranging from 0.26 to 0.82 (Fig 3D). This discrepancy may be attributed to the smaller size of cohort 2, which limited the number of features learned by the model, thus affecting its generalizability.

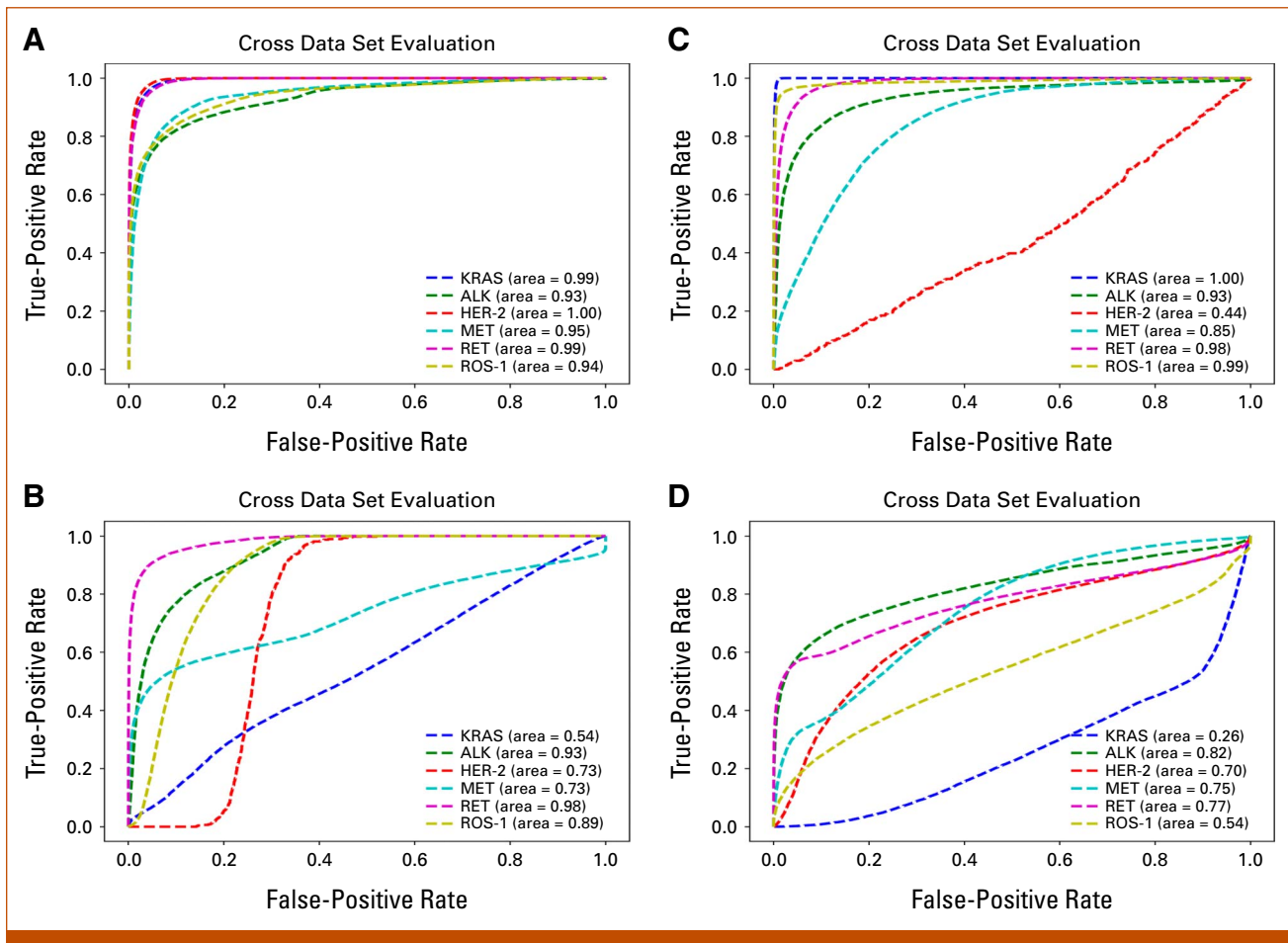
### Interpretability of the Model

To enhance the understanding of the model's predictions and build trust, we used CAM to conduct a visual analysis of



**FIG 2.** Calculation strategy: First, the patient's WSI was cut into tile-like patches, and patches with <50% of the tissue volume were removed in the preprocessing stage. Then, the cases were divided into a 70% training set, 10% validation set, and 20% test set. Finally, the output of the model is wild type or mutant. To verify the generalization of the model, we test the two queues as external test sets for each other. WSI, whole-slide image.





**FIG 3.** Model performance: (A) The ROC curve of each gene mutation predicted by the internal test of the cohort 1 training model; (B) the ROC curve of each gene mutation predicted by the internal test of the cohort 2 training model; (C) the ROC curve in cohort 3 (when the model was trained with cohort 1 and tested with cohort 2); (D) the ROC curve in cohort 4 (cohort 2 training model cohort 1 makes an external test set). ROC, receiver operating characteristic.

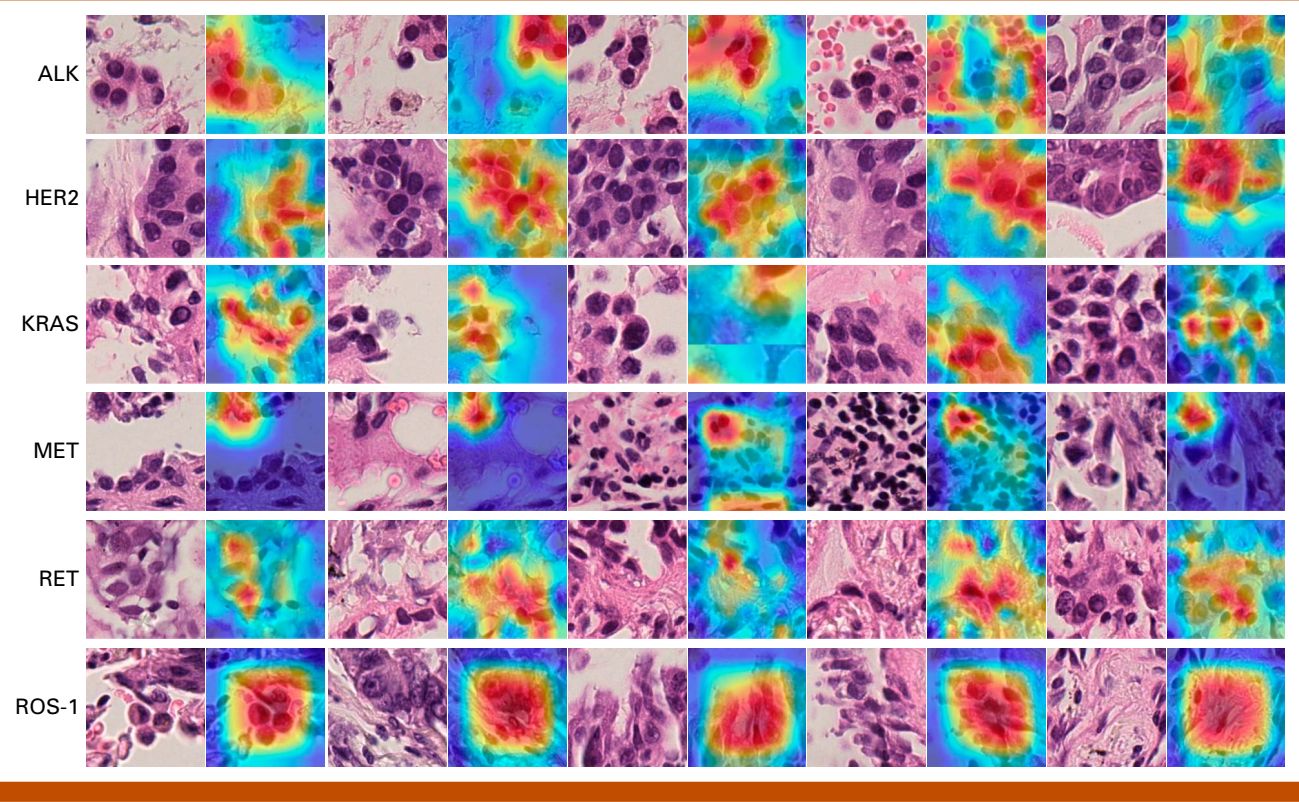
the image tiles that were correctly classified with high confidence.

As shown in Figure 4, by observing the CAM images of different genes, we found that the model mainly focuses on changes in the cellular regions, while paying less attention to the stromal regions. This observation indicates that alterations in the cellular regions are associated with gene expression. After a detailed analysis of the CAM maps of each gene, we discovered the following associations between morphologic abnormalities and gene mutations: in regions with *ALK* mutations, uniform tumor cells with round-to-ovoid nuclei, moderate nuclear size, indistinct cell borders, and a solid growth pattern features consistent with *ALK*-rearranged adenocarcinoma.<sup>21</sup> *HER2*-mutated cases exhibit large pleomorphic nuclei, abundant eosinophilic cytoplasm, multilayered clusters with loss of polarity, and occasional nuclear vacuoles indicating high proliferation.<sup>22</sup> *KRAS* mutation shows obvious nuclear irregularity, pleomorphism, and thickened chromatin, which may reflect the activation of *MAPK/ERK* and *PI3K/AKT* pathways.<sup>23</sup> Regions with *MET*

abnormalities exhibit high-grade nuclear features, as described in previous literature.<sup>23</sup> Regions with *RET* mutations are characterized by angulated nuclei and the presence of mucin in the cytoplasm. *ROS1*-related tumors show eccentric pale nuclei with loose cell-cell contacts, areas prioritized by the model because of *ROS1*'s impact on cytoskeletal and junctional proteins.<sup>24</sup> Although our current analysis has not clearly identified the specific features used by the network to recognize mutations, the high AUC values of the model in predicting gene mutations indicate an underlying correlation between genotypes and phenotypes. This correlation can be detected by deep learning methods.

### Prediction of Gene Mutation Status in Metastatic Cancer

Lung cancer is often diagnosed at the metastatic stage, and molecular testing of metastatic cancer is crucial but challenging because of limited tissue samples.<sup>25</sup> We explored the application of our deep learning model for predicting gene mutations in metastatic LUAD. Our results showed that the model could successfully predict gene mutations for three



**FIG 4.** Visualization of gene mutation prediction results. Grad-CAM is used to display the activation of the last convolution layer of the block-level prediction model. It shows the activation of the last convolutional layer using Grad-CAM for gene mutation prediction evaluation. CAM, Class Activation Mapping.

genes (*HER2*, *KRAS*, and *RET*) with AUC values exceeding 0.7 (Table 2). Although the data set used for this analysis was small, these findings suggest that AI can be an effective tool for predicting gene mutations in metastatic cancer, potentially offering a noninvasive alternative to traditional molecular testing in clinical settings.

DISCUSSION

Targeted therapy is a promising treatment approach for lung cancer. Techniques such as immunohistochemistry, fluorescence in situ hybridization, and NGS are used to identify gene mutations. Although these detection methods have their respective advantages, they have limitations in clinical

practice, such as high cost and time-consuming nature, which restrict the application of targeted therapy.<sup>26,27</sup>

Our model is capable of evaluating the status of key biomarkers through WSIs, significantly shortening the time interval from diagnosis to the initiation of targeted therapy. It can assist pathologists by highlighting samples that may harbor potential mutations, allowing for prioritized examination. Once routine molecular testing is completed, samples with discrepancies between the results of the routine test and the AI model can undergo additional confirmation tests before final reporting to the patient. Thus, AI not only enhances the accuracy of biomarker detection but also aids in adjusting the priority of patient testing, ultimately improving the efficiency and precision of clinical decision making. In addition, our model can also be applied in clinical trials to screen patients with rare molecular subtypes. This provides a more precise group of patients for targeted drug development, accelerating the research and development process of new targeted drugs. We also applied the model to pathologic images of metastatic cancer, which had not been included in the training process, an approach not explored in previous studies. Among the six rare mutant genes tested, *HER2*, *KRAS*, and *RET* exhibited good performance in the metastatic cancer test set, with AUC values ranging from 0.72 to 0.80. This suggests that our model can identify similarities in the pathologic images between

**TABLE 2.** Test Results in Metastatic Cancer

| Gene  | ACC   | SPE   | PRE   | Recall | F1 Score | AUC   |
|-------|-------|-------|-------|--------|----------|-------|
| ALK   | 37.74 | 61.11 | 47.81 | 46.76  | 36.29    | 47.35 |
| HER2  | 40.71 | 94.03 | 60.33 | 56.35  | 39.48    | 78.14 |
| KRAS  | 59.51 | 91.43 | 66.94 | 60.78  | 56.03    | 80.25 |
| RET   | 57.14 | 93.73 | 67.99 | 61.18  | 54.29    | 71.90 |
| MET   | 32.78 | 54.90 | 29.14 | 33.19  | 29.65    | 15.62 |
| ROS-1 | 56.48 | 24.70 | 46.34 | 43.41  | 42.70    | 49.62 |

Abbreviations: ACC, accuracy; PRE, precision; SPE, specificity.

metastatic and primary LUAD for these three gene mutation states. This effort demonstrates that AI has a broader potential for predicting gene mutations in patients with LUAD, extending beyond primary lung specimens to include metastasized tissue from other parts of the body.

We chose ResNeXt101 for gene mutation prediction because of its efficient grouped convolutions, which improve feature extraction and reduce computational complexity. Alternative models like EfficientNet and Vision Transformers were considered but rejected due to higher computational costs or limited validation in pathology-based mutation prediction. Future work may compare architectures to validate the optimal approach.

In conclusion, the LUAD pathologic image recognition model we developed, on the basis of ResNeXt101 architecture,

demonstrated good accuracy and feasibility in predicting *ALK*, *HER2*, *KRAS*, *RET*, *MET*, and *ROS1* at the patch level of tumor cells.

This study has limitations. The small sample size and rare mutation subset may cause class imbalance. Despite using data augmentation and weighted loss functions, larger multi-institutional data sets are needed to enhance generalizability. The black-box characteristic of deep learning limits interpretability because the model's prediction basis remains unclear, risking trust issues. There is no formal comparison of CAM confidence scores between primary and metastatic samples; future work will expand samples, standardize scoring, and test activation differences. Prospective clinical trials are essential to validate the model's real-world applicability, and clinical implementation still has a long way to go.

## AFFILIATIONS

<sup>1</sup>Department of Pathology, the First Affiliated Hospital of China Medical University, Shenyang, China

<sup>2</sup>College of Medicine and Biological Information Engineering, Northeastern University, Shenyang, China

<sup>3</sup>Department of Pathology, the First Affiliated Hospital and College of Basic Medical Sciences of China Medical University, Shenyang, China

<sup>4</sup>Department of Pathology, Cancer Hospital of China Medical University, Liaoning Cancer Hospital and Institute, Shenyang, China

## CORRESPONDING AUTHOR

Yuan Miao, MD; e-mail: cmumiaoyuan@163.com.

## EQUAL CONTRIBUTION

P.Y., W.C., N.L. and Y.Y. contributed equally to this work as co-first authors. Y.M., X.C. and S.N. contributed equally to this work as co-senior authors.

## SUPPORT

Supported by the Central-Guided Funding Project for Local Science and Technology Development in Liaoning Province (No. 2024JH6/100800022 to Y.M.).

## DATA SHARING STATEMENT

To ensure patient confidentiality and compliance with ethical guidelines, all pathological image data were anonymized before analysis, with patient identifiers removed and encrypted where applicable. Additionally, data access is restricted to authorized researchers, and sharing is subject to the institutional review board (IRB) approval. Although public data release is not feasible due to privacy constraints, de-identified metadata, and model training scripts will be made available upon request to promote research transparency and reproducibility. Additionally, the model code, including data preprocessing scripts and hyperparameter configurations, will be made

available in a public repository upon acceptance of this manuscript. This transparency aims to facilitate independent validation and adaptation of our approach for related research applications.

## AUTHOR CONTRIBUTIONS

**Conception and design:** Nan Liu, Yang Yu, Yinan Yuan, Siru Nie, Xiaoyu Cui, Yuan Miao

**Financial support:** Xiaoyu Cui, Yuan Miao

**Administrative support:** Xiaoyu Cui

**Provision of study materials or patients:** Yang Yu, Weilin Guo, Yini Alatan, Xiaoyu Cui

**Collection and assembly of data:** Nan Liu, Yang Yu, Yinan Yuan, Weilin Guo, Yini Alatan, Siru Nie, Xiaoyu Cui, Yuan Miao

**Data analysis and interpretation:** Peiling Yu, Weixing Chen, Nan Liu, Yang Yu, Hongyu Guo, Yinan Yuan, Jinming Zhao, Hongbo Su, Siru Nie, Xiaoyu Cui, Yuan Miao

**Manuscript writing:** All authors

**Final approval of manuscript:** All authors

**Accountable for all aspects of the work:** All authors

## AUTHORS' DISCLOSURES OF POTENTIAL CONFLICTS OF INTEREST

The following represents disclosure information provided by authors of this manuscript. All relationships are considered compensated unless otherwise noted. Relationships are self-held unless noted. I = Immediate Family Member, Inst = My Institution. Relationships may not relate to the subject matter of this manuscript. For more information about ASCO's conflict of interest policy, please refer to [www.asco.org/rwc](http://www.asco.org/rwc) or [ascopubs.org/cci/author-center](http://ascopubs.org/cci/author-center).

Open Payments is a public database containing information reported by companies about payments made to US-licensed physicians ([Open Payments](http://OpenPayments.org)).

No potential conflicts of interest were reported.

## REFERENCES

- Chen Z, Fillmore CM, Hammerman PS, et al: Non-small-cell lung cancers: A heterogeneous set of diseases. *Nat Rev Cancer* 14:535-546, 2014
- Davidson MR, Gazdar AF, Clarke BE: The pivotal role of pathology in the management of lung cancer. *J Thorac Dis* 5:S463-S478, 2013 (suppl 5)
- Herbst RS, Morgensztern D, Boshoff C: The biology and management of non-small cell lung cancer. *Nature* 553:446-454, 2018



4. Kwak EL, Bang Y-J, Camidge DR, et al: Anaplastic lymphoma kinase inhibition in non-small-cell lung cancer. *N Engl J Med* 363:1693-1703, 2010
  5. Bergethson K, Shaw AT, Ou S-HI, et al: ROS1 rearrangements define a unique molecular class of lung cancers. *J Clin Oncol* 30:863-870, 2012
  6. Drilon A, Wang L, Hasanovic A, et al: Response to cabozantinib in patients with RET fusion-positive lung adenocarcinomas. *Cancer Discov* 3:630-635, 2013
  7. Ettinger DS, Wood DE, Aisner DL, et al: Non-small cell lung cancer, version 3.2022, NCCN clinical practice guidelines in oncology. *J Natl Compr Canc Netw* 20:497-530, 2022
  8. Luo W, Wang Z, Zhang T, et al: Immunotherapy in non-small cell lung cancer: Rationale, recent advances and future perspectives. *Precis Clin Med* 4:258-270, 2021
  9. Lindeman NI, Cagle PT, Aisner DL, et al: Updated molecular testing guideline for the selection of lung cancer patients for treatment with targeted tyrosine kinase inhibitors: Guideline from the College of American Pathologists, the International Association for the Study of Lung Cancer, and the Association for Molecular Pathology. *J Thorac Oncol* 13:323-358, 2018
  10. Luo X, Zang X, Yang L, et al: Comprehensive computational pathological image analysis predicts lung cancer prognosis. *J Thorac Oncol* 12:501-509, 2017
  11. Bychkov D, Linder N, Turkki R, et al: Deep learning based tissue analysis predicts outcome in colorectal cancer. *Sci Rep* 8:3395, 2018
  12. Araújo T, Aresta G, Castro E, et al: Classification of breast cancer histology images using convolutional neural networks. *PLoS One* 12:e0177544, 2017
  13. Khameneh FD, Razavi S, Kamasak M: Automated segmentation of cell membranes to evaluate HER2 status in whole slide images using a modified deep learning network. *Comput Biol Med* 110: 164-174, 2019
  14. Couture HD, Williams LA, Geradts J, et al: Image analysis with deep learning to predict breast cancer grade, ER status, histologic subtype, and intrinsic subtype. *NPJ Breast Cancer* 4:30, 2018
  15. Du H, Wang X, Wang K, et al: Identifying invasiveness to aid lung adenocarcinoma diagnosis using deep learning and pathomics. *Sci Rep* 15:4913, 2025
  16. Pan X, AbdulJabbar K, Coelho-Lima J, et al: The artificial intelligence-based model ANORAK improves histopathological grading of lung adenocarcinoma. *Nat Cancer* 5:347-363, 2024
  17. Coudray N, Ocampo PS, Sakellaropoulos T, et al: Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat Med* 24:1559-1567, 2018
  18. Terada Y, Takahashi T, Hayakawa T, et al: Artificial intelligence-powered prediction of ALK gene rearrangement in patients with non-small-cell lung cancer. *JCO Clin Cancer Inform* 10.1200/CCI.22.00070
  19. Niehues JM, Quirke P, West NP, et al: Generalizable biomarker prediction from cancer pathology slides with self-supervised deep learning: A retrospective multi-centric study. *Cell Rep Med* 4: 100980, 2023
  20. Wu S, Yue M, Zhang J, et al: The role of artificial intelligence in accurate interpretation of HER2 immunohistochemical scores 0 and 1+ in breast cancer. *Mod Pathol* 36:100054, 2023
  21. Yoshida A, Tsuta K, Nakamura H, et al: Comprehensive histologic analysis of ALK-rearranged lung carcinomas. *Am J Surg Pathol* 35:1226-1234, 2011
  22. Lee Y, Lee B, Choi Y-L, et al: Clinicopathologic and molecular characteristics of HER2 (ERBB2)-altered non-small cell lung cancer: Implications for precision medicine. *Mod Pathol* 37:100490, 2024
  23. Yatabe Y: Molecular pathology of non-small cell carcinoma. *Histopathology* 84:50-66, 2024
  24. Pan Y, Zhang Y, Li Y, et al: ALK, ROS1 and RET fusions in 1139 lung adenocarcinomas: A comprehensive study of common and fusion pattern-specific clinicopathologic, histologic and cytologic features. *Lung Cancer* 84:121-126, 2014
  25. Forest F, Stachowicz M-L, Casteillo F, et al: EGFR, KRAS, BRAF and HER2 testing in metastatic lung adenocarcinoma: Value of testing on samples with poor specimen adequacy and analysis of discrepancies. *Exp Mol Pathol* 103:306-310, 2017
  26. Ahn M-J: Molecular testing in lung cancer: Still big gap in implementation for real-world use. *J Thorac Oncol* 15:1399-1400, 2020
  27. Smeltzer MP, Wynes MW, Lantuejoul S, et al: The International Association for the Study of Lung Cancer global survey on molecular testing in lung cancer. *J Thorac Oncol* 15:1434-1448, 2020
-

APPENDIX

TABLE A1. The Specimen Types of Each Gene Mutation in Cohort 1

| Specimen Type | ALK | HER2 | RET | MET | ROS1 | KRAS | No-Mutation |
|---------------|-----|------|-----|-----|------|------|-------------|
| SS            | 12  | 15   | 11  | 15  | 9    | 4    | 4           |
| BS            | 27  | 5    | 8   | 5   | 11   | 1    | 17          |

Abbreviations: BS, biopsy specimens; SS, surgical specimens.

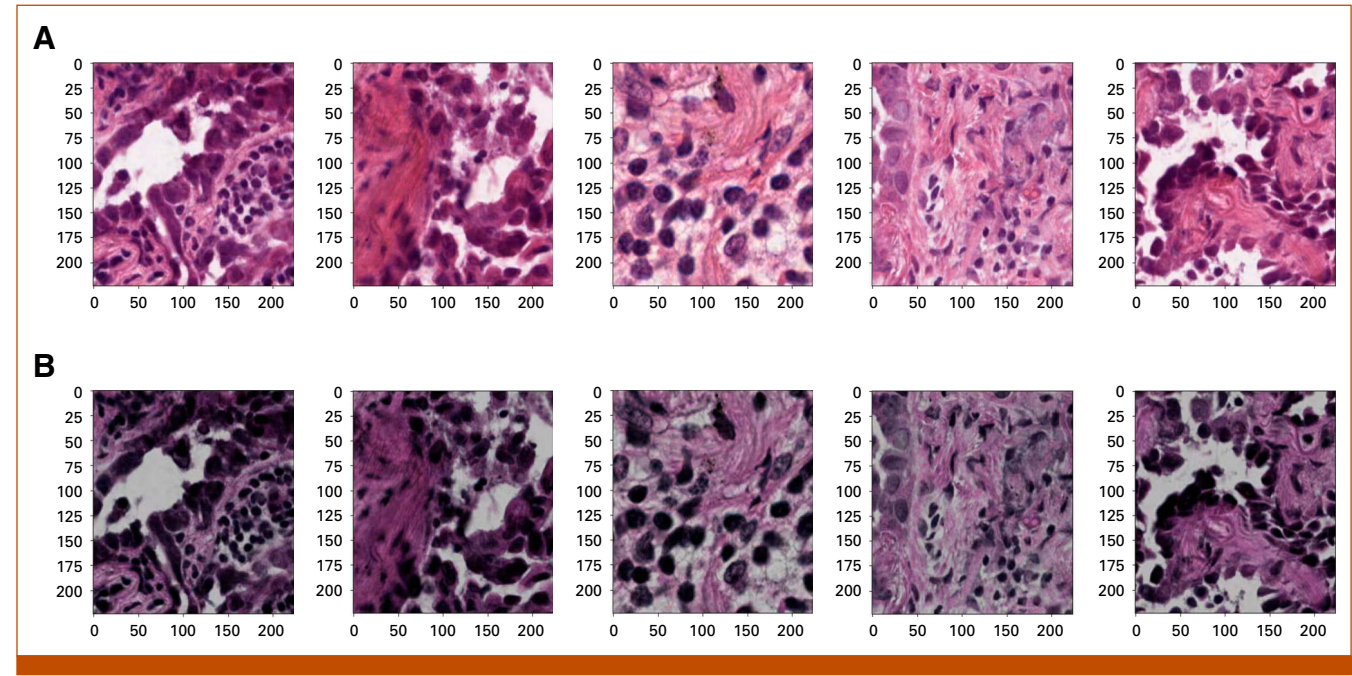


FIG A1. Color preprocessing of H&E-dyed images. (A) The original image; (B) The processed image. H&E, hematoxylin and eosin.