Classification of Mouse Lung Metastatic Tumor with Deep Learning

Ha Neul Lee¹, Hong-Deok Seo², Eui-Myoung Kim³, Beom Seok Han⁴ and Jin Seok Kang¹,*

¹Department of Biomedical, Laboratory Science, Namseoul University, Cheonan 31020,
²Department of Industrial Promotion, Spatial Information Industry Promotion Agency, Seongnam 13487,
³Department of Spatial Information Engineering, Namseoul University, Cheonan 31020,
⁴Department of Pharmaceutical Engineering, Hoseo University, Asan 31499, Republic of Korea

Abstract
Traditionally, pathologists microscopically examine tissue sections to detect pathological lesions; the many slides that must be evaluated impose severe work burdens. Also, diagnostic accuracy varies by pathologist training and experience; better diagnostic tools are required. Given the rapid development of computer vision, automated deep learning is now used to classify microscopic images, including medical images. Here, we used an Inception-v3 deep learning model to detect mouse lung metastatic tumors via whole slide imaging (WSI); we cropped the images to 151 by 151 pixels. The images were divided into training (53.8%) and test (46.2%) sets (21,017 and 18,016 images, respectively). When images from lung tissue containing tumor tissues were evaluated, the model accuracy was 98.76%. When images from normal lung tissue were evaluated, the model accuracy ("no tumor") was 99.87%. Thus, the deep learning model distinguished metastatic lesions from normal lung tissue. Our approach will allow the rapid and accurate analysis of various tissues.

Key Words: Mouse, Lung tumor, Digital pathology, Classification, Deep learning

INTRODUCTION
Pathologists with the requisite long-term training and pre-clinical and/or clinical experience are overwhelmed by large numbers of pathology slides (Kuo and Leo, 2019). Also, diagnostic accuracy varies by pathologist training and experience; better diagnostic tools are required. Given the rapid development of computer vision, automated deep learning is now used to classify microscopic images, including medical images. Here, we used an Inception-v3 deep learning model to detect mouse lung metastatic tumors via whole slide imaging (WSI); we cropped the images to 151 by 151 pixels. The images were divided into training (53.8%) and test (46.2%) sets (21,017 and 18,016 images, respectively). When images from lung tissue containing tumor tissues were evaluated, the model accuracy was 98.76%. When images from normal lung tissue were evaluated, the model accuracy ("no tumor") was 99.87%. Thus, the deep learning model distinguished metastatic lesions from normal lung tissue. Our approach will allow the rapid and accurate analysis of various tissues.
recognition of metastatic tumors will reduce the slide reading time and increase diagnostic accuracy; such systems will find many preclinical applications. Here, we used the Inception-v3 deep learning model to characterize mouse lung metastatic tumors; we evaluated the accuracy of the approach.

MATERIALS AND METHODS

Data collection

We retrieved 20 hematoxylin and eosin (H&E)-stained slides (10 containing lung metastatic tumors and 10 normal lung tissue) from experiment using C57BL/6 female mice that was treated with mouse melanoma cells (B16-F10 cell) intravenously. The samples were retrieved from the study (NSU-19-05), which was approved by the animal experiment committee of Namseoul University (Cheonan, Korea) based on the Animal Protection Act.

All were scanned using a Panoramic Whole-slide Scanner (3D Histotech Co. Ltd., Budapest, Hungary) at 20× magnification in the Department of Biomedical Laboratory Science of Namseoul University. The staining intensity, contrast, and thresholding were not adjusted.

Deep learning

We used Inception-v3 model for training and testing, and evaluated classification accuracies. An overview of the approach is shown in Fig. 1. H&E-stained sections were scanned, converted, cropped, and used for supervised training of Inception-v3. Using this model, cropped images were classified as non-tumor or tumor. For assessment of Inception-v3 performance, classification accuracies were calculated.

Slide annotation

We cropped the images of 20 lung tissue samples to 151×151 pixels. For computational learning of images, we divided the images two parts as training and test set. The 39,233 images were randomly divided into training (53.8%) and test (46.2%) sets (21,017 and 18,016 images, respectively) (Table 1).

Visualization of predicted patches

We visualized metastatic tumors in WSIs by adding colored dots to patches predicted to be tumorous. This allowed pathologists to understand the classification method.

RESULTS

Collection of training and evaluation data

A representative WSI histopathological profile is shown in Fig. 2. Metastatic tumor cells (several lesions) are evident. We cropped the images as described above; representative im-
Inception-v3 evaluated square patches and predicted the metastatic probabilities. The accuracies were 98.76% for images with tumor tissues and 99.87% for images with normal tissues.

Visualization of tumor and normal tissues

The data for tumorous tissues are shown in Fig. 4. Inception-v3 colored tumor cells as red color and non-tumor cells as blue color. The accuracy was very high; however, a few normal bronchiolar epithelial cells were misclassified as tumor cells. Almost all normal cells were so identified (blue color).

DISCUSSION

Inception-v3 reliably (and automatically) identified H&E-stained tumorous lesions on slides. The identification accuracy was 98.76% in tumor tissues and that of normal tissues was 99.87%.

In images containing tumor tissues, a few normal epithelial cells were misclassified as tumorous when the cell concentrations were dense. Similar results were obtained when normal lung tissue was evaluated. Inception-v3 was as good as an experienced pathologist; even such professionals find it difficult to distinguish normal cells from tumorous cells when the cells are in dense arrays.

To reduce misclassification, tumor-similar tissues should be excluded prior to deep learning. Inception-v3 and ResNet-50 outperformed VGG-16 in classification of tissue images, showing Inception-v3 identified the tissue from query images, with an accuracy up to 83.4% and misclassification of histologically related tissues is more common at higher magnifications (Hoefling et al., 2021). For example, prior to analysis, images containing large bronchi and vessels were manually excluded before fibrosis quantification (Gilhodes et al., 2017), and cartilage, periosteal, and skeletal muscle tissues were excluded before measurement of bone marrow cellularity (Smith et al., 2021).

Fig. 3. Representative cropped images of tumorous and normal tissues. (A) tumor tissue; (B) normal tissue.

Fig. 4. Visualization of lung tumorous and non-tumorous tissues. (A) Original H&E images. (B) Visualized images. Red: tumorous tissue (arrow); blue: non-tumorous tissue (arrowhead).

Fig. 5. Visualization of lung non-tumorous tissues. (A) Original H&E images. (B) Visualized images. Red: tumorous tissue (arrow); blue: non-tumorous tissue (arrowhead). (C) Image-matching misclassifications of normal epithelial cells as tumor cells.
Inception-v3 was effective when the image sizes were 299×299, 151×151, and 79×79 pixels (Szegedy et al., 2016). We found that the accuracy was better when the image size was 151×151 than 299×299 pixels (data not shown). This may be because we augmented the images to generate many training samples, and/or because of variations among the pathological images evaluated. The great variabilities in biological tissues per se, and the tissue preparation protocols, often trigger significant imaging changes that hamper computational effectiveness (Zarella et al., 2017).

For training and testing of deep learning models, it is common to use 70% of input data for training and to test the model using the remaining 30% that are not used for training. We used tumor-containing lung tissue images for training. We employed both normal lung and tumor-containing lung tissue slides for testing. Inception-v3 afforded very accurate classifications.

Mouse tail vein injection of B16 melanoma cells is followed by cell lodgment in lung capillaries, allowing assessment of lung extravasation and colonization (Hart and Fidler, 1980). It was possible to accurately distinguish metastatic tumor cells from normal lung cells on entire tissue slides by exploiting their morphological differences.

Inception-v3 uses a CNN for patch classification and a whole-slide-inference mechanism to determine predominant and minor cellular subtypes; this expedites tumor classification by automatically detecting tumors or non-tumorous patterns. However, we found it difficult to measure tumor cell numbers. Image segmentation is required. Quantification of metastatic, lung tumor cell numbers is extremely important when evaluating therapeutic efficacy and toxicity, but it can be challenging. An appropriate algorithm must be chosen and staining must be uniform. A CNN precisely segmented glomeruli in digitized trichrome-stained kidney sections from patients with chronic kidney disease (Kannan et al., 2019).

We studied mouse lung metastases; our data may not reflect the lung tumor histology of other animals. AI can integrate genomic/transcriptomic with epidemiological data, aiding non-clinical researchers. Inception-v3 not only diagnosed lung tumors but also predicted mutations in specific genes (Coudray et al., 2018). The genetic patterns of laboratory animal lung tumors require further analysis. In the interim, WSI analysis quantitatively and reproducibly measures small-scale histological features. However, the technique is time-consuming and not user-interactive. Pathologists seek to improve diagnostic quality and save time using AI-based applications and not user-interactive. Pathologists seek to improve diagnostic quality and save time using AI-based applications.

REFERENCES

ACKNOWLEDGMENTS

We would like to thank Ms. Nahyeon Gu and Kanghee Ryu for their technical assistance (Namseoul University). This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2019R1F1A1058721).

CONFLICT OF INTEREST

Authors declare no conflict of interest.
Seger, S., Stritt, M., Vezzali, E., Nayler, O., Hess, P., Groenen, P. M. A. and Stalder, A. K. (2018) A fully automated image analysis method to quantify lung fibrosis in the bleomycin-induced rat model. *PLoS ONE* **13**, e0193057.

Smith, M. A., Westerling-Bui, T., Wilcox, A. and Schwartz, J. (2021) Screening for bone marrow cellularity changes in cynomolgus macaques in toxicology safety studies using artificial intelligence models. *Toxicol. Pathol.* **49**, 905-911.

Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. and Wojna, Z. (2016) Rethinking the inception architecture for computer vision. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition, pp. 2818-2826.

Veta, M., van Diest, P. J., Kornegoor, R., Huisman, A., Viergever, M. A. and Pluim, J. P. (2013) Automatic nuclei segmentation in H&E stained breast cancer histopathology images. *PLoS ONE* **8**, e0174489.