



Classification of Mouse Lung Metastatic Tumor with Deep Learning

Ha Neul Lee¹, Hong-Deok Seo², Eui-Myoung Kim³, Beom Seok Han⁴ and Jin Seok Kang^{1,*}

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 a 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 training and experience; thus, better diagnostic tools are required (Mazer *et al.*, 2019). For example, when exploring carcinogenicity, a pathologist must examine up to 50,000 slides per test (Boorman *et al.*, 1994). Digital pathology refers to information collected using digitized slides. This increases the accuracy, reproducibility, and standardization of pathology-based trial entry criteria; preclinical and clinical endpoints; diagnostics; chemical assessment; and drug development (Pell *et al.*, 2019). Digitized slides are created by a digital scanner; the process is termed whole-slide imaging (WSI). WSI technology, which is replacing the glass slide examination of a traditional pathologist (Bradley and Jacobsen, 2019), requires co-operation among pathologists, technologists, and executives (Zarella *et al.*, 2019). WSI analysis has been automated in recent decades,

increasing speed, reducing costs (Farahani *et al.*, 2015), and eliminating human error (Aeffner *et al.*, 2019). WSI has been used to quantitate lung fibrosis (Seger *et al.*, 2018) and to segment specific areas in pathological images (Veta *et al.*, 2013).

Various applications incorporating artificial intelligence (AI) seek to assist diagnosis and lesion detection/segmentation (Komura and Ishikawa, 2019). Given the rapid development of computer vision, automated deep learning is now used to classify microscopic images, including medical images (Brent and Boucheron, 2018). Convolutional neural networks (CNNs) have been used to build medical imaging algorithms. The Inception-v3 model features parallel convolutional paths of different sizes and depths that simultaneously process multi-scale data and collect various feature maps to facilitate classification (Szegedy *et al.*, 2016). Another CNN-based system using a pre-learned auto-encoder, or VGGNet or ResNet, has been proposed (Hoefling *et al.*, 2021).

Mouse lung metastasis models are commonly used to assess therapies and track cancer cell numbers in real time, thus to dynamically monitor metastasis (Bos *et al.*, 2010). AI-based

Open Access <https://doi.org/10.4062/biomolther.2021.130>

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/4.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received Aug 2, 2021 Revised Aug 31, 2021 Accepted Sep 13, 2021

Published Online Nov 2, 2021

***Corresponding Author**

E-mail: kang@nsu.ac.kr

Tel: +82-41-580-2721, Fax: +82-41-580-2932

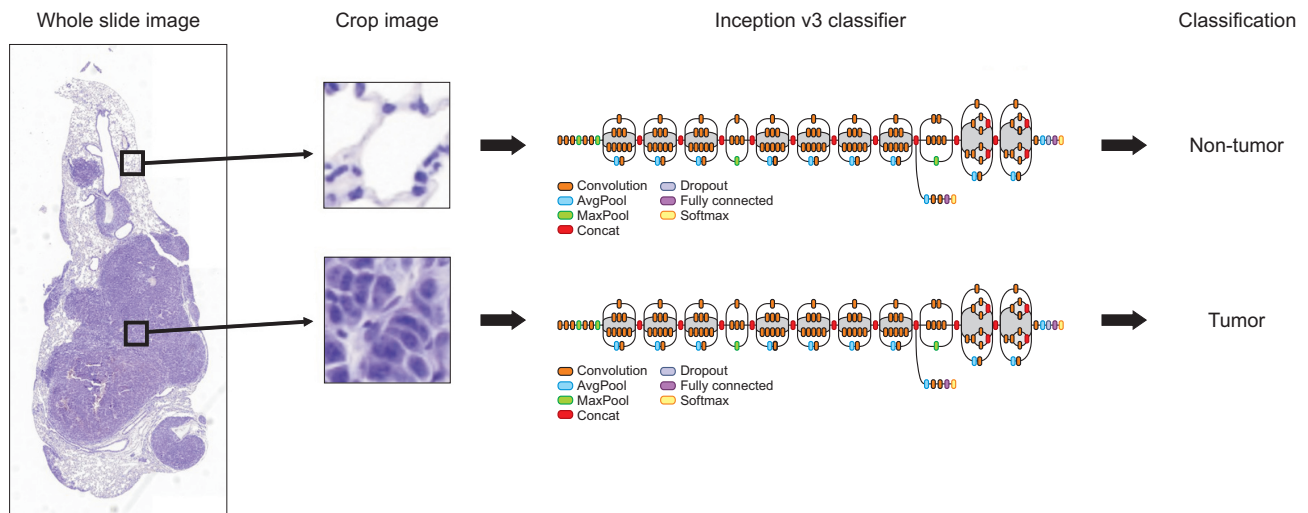


Fig. 1. Overview of the WSI classification of histopathological patterns. We used a sliding window approach to generate small patches, classified each patch using the Inception-v3 neural network, aggregated the patch predictions, and employed a heuristic to identify the predominant and minor patterns. All patch predictions were independent of those of adjacent patches and patch location in the WSI.

Table 1. Distribution of training and test set data Number of training and test data between tumor and non-tumor lesion of lung

Diagnosis	Number of training data	Number of test data
Tumor	13,186	9,504
Non-tumor	7,831	8,712
Sum	21,017	18,216

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 20x 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 Incep-

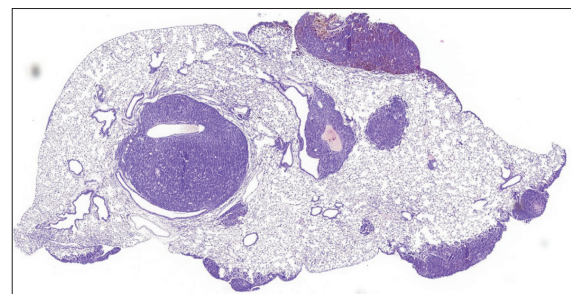


Fig. 2. Representative WSI histopathological figures.

tion-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 151x151 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-

ages are shown in Fig. 3.

Inception-v3 classification of tumor and non-tumor tissues

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)

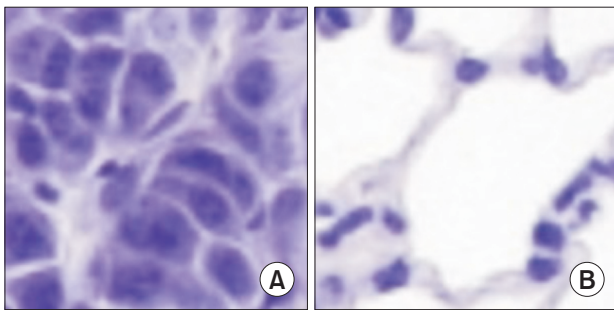


Fig. 3. Representative cropped images of tumorous and normal tissues. (A) tumor tissue; (B) normal tissue.

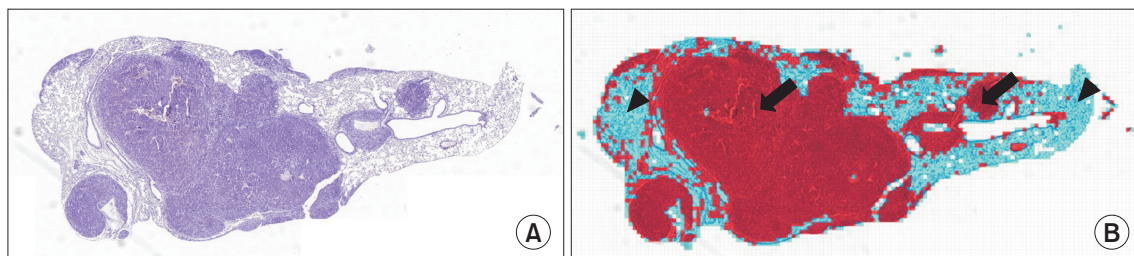


Fig. 4. Visualization of lung tumorous and non-tumorous tissues. (A) Original H&E images. (B) Visualized images. Red: tumorous tissue (arrow); blue: non-tumorous tissue (arrowhead).

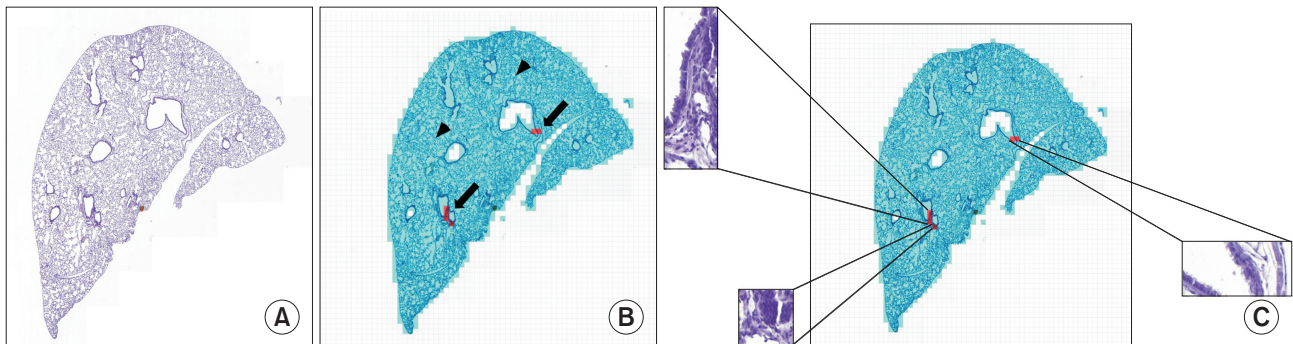


Fig. 5. Visualization of lung non-tumorous tissues. (A) Original H&E images. (B) Visualized images. Red: tumorous tissue (arrow); blue: non-tumorous tissue (arrowhead). (C) Image-matching misclassifications of normal epithelial cells as tumor cells.

(Fig. 5). The accuracy was very high. However, a few normal epithelial cells were misclassified as tumorous (red color), particularly when epithelial or stromal connective tissue cells were present at high densities.

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 (Hoeffling *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).

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 tumorous 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 (Schaer *et al.*, 2019). Human training requires time and effort; user-friendly deep learning algorithms are urgently required. The current limitations of AI-based computer-aided pathology can be overcome by collaboration among AI researchers, software engineers, and medical experts (Nam *et al.*, 2020).

Our deep learning model distinguished mouse metastatic from normal tissues. This promising approach will facilitate the rapid and accurate analysis of tissue characteristics.

CONFLICT OF INTEREST

Authors declare no conflict of interest.

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).

REFERENCES

- Aeffner, F., Zarella, M. D., Buchbinder, N., Bui, M. M., Goodman, M. R., Hartman, D. J., Lujan, G. M., Molani, M. A., Parwani, A. V., Lillard, K., Turner, O. C., Vemuri, V. N. P., Yuil-Valdes, A. G. and Bowman, D. (2019) Introduction to digital image analysis in whole-slide imaging: a white paper from the Digital Pathology Association. *J. Pathol. Inform.* **10**, 9.
- Boorman, G. A., Maronpot, R. R. and Eustis, S. L. (1994) Rodent carcinogenicity bioassay: past, present, and future. *Toxicol. Pathol.* **22**, 105-111.
- Bos, P. D., Nguyen, D. X. and Massague, J. (2010) Modeling metastasis in the mouse. *Curr. Opin. Pharmacol.* **10**, 571-577.
- Bradley, A. and Jacobsen, M. (2019) Toxicologic Pathology Forum*: opinion on considerations for the use of whole slide images in GLP pathology peer review. *Toxicol. Pathol.* **47**, 100-107.
- Brent, R. and Boucheron, L. (2018) Deep learning to predict microscope images. *Nat. Methods* **15**, 868-870.
- Coudray, N., Ocampo, P. S., Sakellaropoulos, T., Narula, N., Snuderl, M., Fenyö, D., Moreira, A. L., Razavian, N. and Tsirigos, A. (2018) Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat. Med.* **24**, 1559-1567.
- Farahani, N., Parwani, A. V. and Pantanowitz, L. (2015) Whole slide imaging in pathology: advantages, limitations, and emerging perspectives. *Pathol. Lab. Med. Int.* **7**, 23-33.
- Gilhodes, J. C., Jule, Y., Kreuz, S., Stierstorfer, B., Stiller, D. and Wolin, L. (2017) Quantification of pulmonary fibrosis in a bleomycin mouse model using automated histological image analysis. *PLoS ONE* **12**, e0170561.
- Hart, I. R. and Fidler, I. J. (1980) Role of organ selectivity in the determination of metastatic patterns of B16 melanoma. *Cancer Res.* **40**, 2281-2287.
- Hoefling, H., Sing, T., Hossain, I., Boisclair, J., Doelemeyer, A., Flandre, T., Piaia, A., Romanet, V., Santarossa, G., Saravanan, C., Sutter, E., Turner, O., Wuersch, K. and Moulin, P. (2021) HistoNet: a deep learning-based model of normal histology. *Toxicol. Pathol.* **49**, 784-797.
- Kannan, S., Morgan, L. A., Liang, B., Cheung, M. G., Lin, C. Q., Mun, D., Nader, R. G., Belghasem, M. E., Henderson, J. M., Francis, J. M., Chitalia, V. C. and Kolachalama, V. B. (2019) Segmentation of glomeruli within trichrome images using deep learning. *Kidney Int. Rep.* **4**, 955-962.
- Komura, D. and Ishikawa, S. (2019) Machine learning approaches for pathologic diagnosis. *Virchows Arch.* **475**, 131-138.
- Kuo, K. H. and Leo, J. M. (2019) Optical versus virtual microscope for medical education: a systematic review. *Anat. Sci. Educ.* **12**, 678-685.
- Mazer, B. L., Homer, R. J. and Rimm, D. L. (2019) False-positive pathology: improving reproducibility with the next generation of pathologists. *Lab. Invest.* **99**, 1260-1265.
- Nam, S., Chong, Y., Jung, C. K., Kwak, T. Y., Lee, J. Y., Park, J., Rho, M. J. and Go, H. (2020) Introduction to digital pathology and computer-aided pathology. *J. Pathol. Transl. Med.* **54**, 125-134.
- Pell, R., Oien, K., Robinson, M., Pitman, H., Rajpoot, N., Rittscher, J., Snead, D. and Verrill, C. (2019) The use of digital pathology and image analysis in clinical trials. *J. Pathol. Clin. Res.* **5**, 81-90.
- Schaer, R., Otolara, S., Jimenez-Del-Toro, O., Atzori, M. and Muller, H. (2019) Deep learning-based retrieval system for gigapixel histopathology cases and the open access literature. *J. Pathol. Inform.* **10**, 19.

- 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**, e70221.
- Zarella, M. D., Bowman, D., Aeffner, F., Farahani, N., Xthona, A., Absar, S. F., Parwani, A., Bui, M. and Hartman, D. J. (2019) A practical guide to whole slide imaging: a white paper from the Digital Pathology Association. *Arch. Pathol. Lab. Med.* **143**, 222-234.
- Zarella, M. D., Yeoh, C., Breen, D. E. and Garcia, F. U. (2017) An alternative reference space for H&E color normalization. *PLoS ONE* **12**, e0174489.