



## Automatic breast carcinoma detection in histopathological micrographs based on Single Shot Multibox Detector



Mio Yamaguchi <sup>a</sup>, Tomoaki Sasaki <sup>b,c</sup>, Kodai Uemura <sup>b</sup>, Yuichiro Tajima <sup>b</sup>, Sho Kato <sup>c,d</sup>, Kiyoshi Takagi <sup>a</sup>, Yuto Yamazaki <sup>e</sup>, Ryoko Saito-Koyama <sup>f</sup>, Chihiro Inoue <sup>e</sup>, Kurara Kawaguchi <sup>c</sup>, Tomoya Soma <sup>c,g</sup>, Toshio Miyata <sup>c</sup>, Takashi Suzuki <sup>a,e,h,\*</sup>

<sup>a</sup> Department of Pathology and Histotechnology, Graduate School of Medicine, Tohoku University, Sendai, Miyagi 980-8575, Japan

<sup>b</sup> NEC Solution Innovators, Ltd., Koto-ku, Toyko 136-8627, Japan

<sup>c</sup> Department of Molecular Medicine and Therapy, Graduate School of Medicine, Tohoku University, Sendai, Miyagi 980-8575, Japan

<sup>d</sup> Renascence Inc., Chuo-ku, Tokyo 103-0023, Japan

<sup>e</sup> Department of Anatomic Pathology, Graduate School of Medicine, Tohoku University, Sendai, Miyagi 980-8575, Japan

<sup>f</sup> Department of Pathology, Sendai Medical Center, Sendai, Miyagi 983-8520, Japan

<sup>g</sup> NEC Corporation, Minato-ku, Tokyo 108-8001, Japan

<sup>h</sup> Department of Pathology, Tohoku University Hospital, Sendai, Miyagi 980-8574, Japan

### ARTICLE INFO

#### Keywords:

Breast cancer

Pathology

Artificial intelligence

Deep learning

### ABSTRACT

**Background:** A diagnosis with histological classification by pathologists is very important for appropriate treatments to improve the prognosis of patients with breast cancer. However, the number of pathologists is limited, and assisting the pathological diagnosis by artificial intelligence becomes very important. Here, we presented an automatic breast lesions detection model using microscopic histopathological images based on a Single Shot Multibox Detector (SSD) for the first time and evaluated its significance in assisting the diagnosis.

**Methods:** We built the data set and trained the SSD model with 1361 microscopic images and evaluated using 315 images. Pathologists and medical students diagnosed the images with or without the assistance of the model to investigate the significance of our model in assisting the diagnosis.

**Results:** The model achieved 88.3% and 90.5% diagnostic accuracies in 3-class (benign, non-invasive carcinoma, or invasive carcinoma) or 2-class (benign or malignant) classification tasks, respectively, and the mean intersection over union was 0.59. Medical students achieved a remarkably higher diagnostic accuracy score (average 84.7%) with the assistance of the model compared to those without assistance (average 67.4%). Some people diagnosed images in a short time using the assistance of the model (shortened by average 6.4 min) while others required a longer time (extended by 7.2 min).

**Conclusion:** We presented the automatic breast lesions detection method at high speed using histopathological micrographs. The present system may conveniently support the histological diagnosis by pathologists in laboratories.

### Introduction

Breast cancer is one of the most common malignancies in females worldwide.<sup>1</sup> Benign breast disorders include ductal hyperplasia, intraductal papilloma, adenosis, and fibroadenoma.<sup>2</sup> Conversely, invasive carcinoma accounts for >80% of all breast carcinoma diagnoses, and non-invasive carcinoma account for 10%–20%.<sup>3,4</sup> Histological examination of specimens (Hematoxylin and eosin (H&E)-stained tissues) is conventionally used under light microscopy in pathological diagnosis.<sup>5</sup> Early detection and accurate diagnosis of breast carcinoma with histological classification by pathologists are very important for appropriate treatments to improve the

prognosis of patients. However, the number of pathologists is limited and they have too many tasks and are under the stress in many laboratories.<sup>6–9</sup>

In addition, pathological diagnosis tends to depend on the subjectivity of their experience. Therefore, assisting the pathological diagnosis by objectively double-checking with artificial intelligence (AI) becomes very important.

Several studies about automatic classification for breast pathological images, especially using artificial neural network (ANN) approach, have been previously reported.<sup>10–13</sup> K. Kiambe proposed a 2-stage model for 4-class classification (normal, benign, non-invasive carcinoma, and invasive carcinoma) of breast histopathological images, and the model achieved 99.84% accuracy.<sup>14</sup> Yun Jiang et al. designed a convolutional neural

\* Corresponding author at: Department of Pathology and Histotechnology, Tohoku University Graduate School of Medicine, 2-1 Seiryomachi, Aoba-ku, Sendai, Miyagi-ken 980-8575, Japan. E-mail address: [t-suzuki@patholo2.med.tohoku.ac.jp](mailto:t-suzuki@patholo2.med.tohoku.ac.jp) (T. Suzuki).

network (CNN) with a small SE-ResNet module to classify BreakHis dataset into 8 subtypes and revealed that the model achieved a 90.66% and 93.81% accuracy.<sup>15</sup>

However, various problems remained to overcome to practically introduce AI into the clinical workflow for pathological diagnosis. Firstly, most tissues include various types of components although many showed the results of classification for whole images and/or heat maps, indicating potential areas where carcinoma cells may be present.<sup>16–23</sup> Especially, a single result of classification for a whole image by AI leads many users to feel the AI as a black box and be hesitant to use AI.<sup>24</sup> Thus, providing more information about the processes that lead to the results by showing the areas of each component may be important to make the users feel less inhibited and for practical use as assistance for doctors.

Furthermore, the impact of AI assistance on clinical utility has not been fully investigated while the performance to classify histopathological images has been mainly focused on. Pathological diagnosis is always regarded as the final diagnosis to decide on patient treatments. Moreover, pathological diagnosis is complicated and pathologists diagnose based on their breadth of knowledge and diagnostic experience as stated above. Therefore, completely replacing the role of pathologists with AI may be currently difficult, and confirming the significance of AI as assistance for pathologists and understanding these disadvantages, as well as benefits, is important.

Finally, introducing AI to clinical practice remains a big hurdle regarding cost and trouble. The technology has been developed to digitize an entire glass slide (Whole slide imaging; WSI),<sup>25</sup> and various AIs to classify WSI have been reported.<sup>19,26–28</sup> WSI gives various advantages, such as the automated WSI scanner that automatically scans.<sup>29,30</sup> However, introducing WSI scanners and digital pathology system and managing huge amounts of digital data need a lot of money. In addition, scanning speed is widely different by machines. For example, while some scanners with high specifications scan 1 slide in 0.5–1.5 min, some WSI scanners take 7–9 min for scanning 1 slide.<sup>31</sup> Although WSI technology is becoming widespread in clinical practice, actually, using WSI is sometimes difficult in some hospitals.<sup>32</sup> Thus, proposing the significance of the AI model, which is easy to introduce and simpler to use using micrographs, could be important.

Single Shot Multibox Detector (SSD) is an object detection method that provides detection at high speed in real-time.<sup>33</sup> The SSD network is relatively simple and can be trained and integrated into systems with comparative ease. The significance of SSD in pathology has not been investigated although it may assist the pathological diagnosis. Therefore, the present study demonstrated an automatic breast lesion detection model using microscopic histopathological images based on SSD for the first time and investigated the effects of the present model in assisting the diagnosis (Fig. 1).

## Materials and methods

### Clinical cases

The H&E-stained<sup>34</sup> glass slides of histopathological specimens of human breast diseases were obtained from Tohoku University Hospital, Sendai, Japan. All specimens were obtained from patients who had undergone surgical treatment or biopsy at Tohoku University Hospital and had been fixed with 10% formalin neutral buffer solution and embedded in paraffin wax.<sup>35</sup> Experiments and analyses were performed following the Helsinki declaration, and the research protocol of this study was approved by the Ethics Committee at the Tohoku University Graduate School of Medicine (approval no. 2021-1-1046).

### Datasets and annotations

Table 1 showed the dataset summary. The micrographs were taken in PNG formation from glass slides by microscope (Olympus BX53, Olympus Inc., Tokyo, Japan) with Olympus DP26 digital camera (Olympus) and the software cellSens Dimension (Olympus) by an expert pathologist, especially for breast disorders (annotation pathologist). The size of each image is 2448 × 1920 pixels (72 dpi) at 40× magnification. The images were annotated by a pathologist using the LabelImg v1.8.1 tool (<https://github.com/tzutalin/labelImg>), drawing bounding boxes around the breast epithelium tissues that conformed with one of these 3 labels (annotation label): invasive carcinoma, non-invasive carcinoma, and benign, including benign lesions and normal breast epithelium tissues in the present study. Representative examples of annotation for various images were shown in Fig. 2A–C.

Then, a single-label diagnosis of the whole image (image label) was provided for each image. Many images include various annotation labels, and an image label was determined following the prioritization, such as invasive carcinoma, non-invasive carcinoma, and benign. Images, including the area of invasive carcinoma, were labeled with “invasive carcinoma” regardless of the presence of non-invasive carcinoma and benign lesions. Images, including components of non-invasive carcinoma without invasive carcinoma cells, were labeled with “non-invasive carcinoma” regardless of the presence of a benign lesion. Images, which did not contain carcinoma cells, were labeled as “benign” in this study. These determination methods followed the same rules as for clinical diagnosis.

### SSD network structure and training

The SSD is an object detection method that accurately detects at high speed in real-time.<sup>33</sup> SSD is based on a forward propagation CNN network, and, the network produces scores for the presence of each object category in each bounding box and performs adjustments to the boxes to match

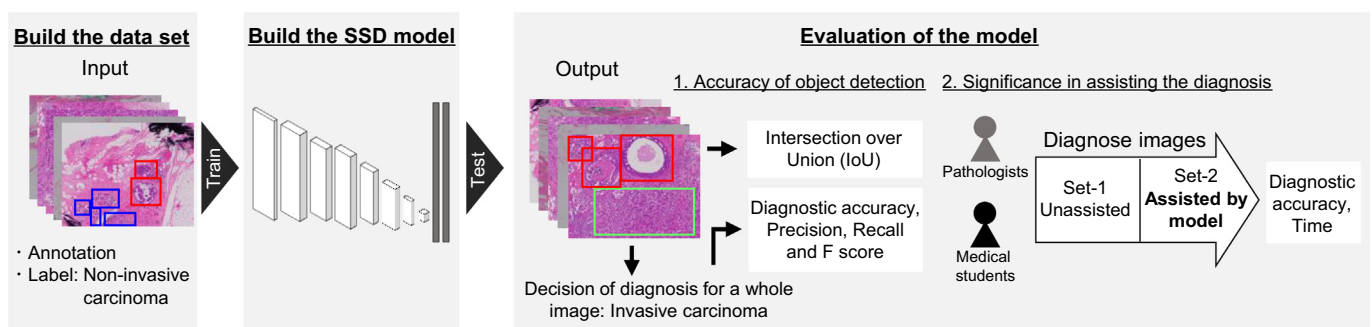
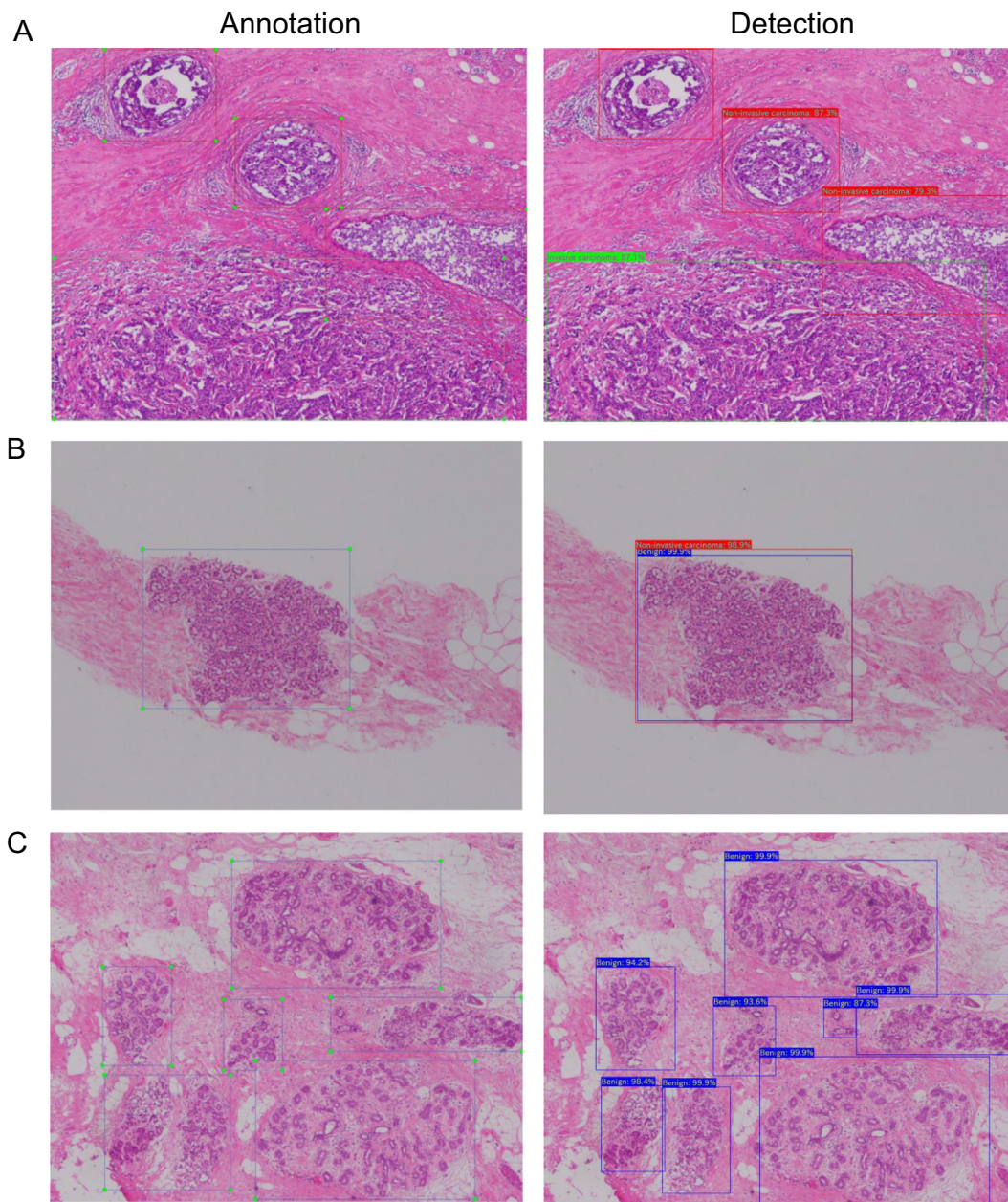


Fig. 1. Summary of the present study. We built the data set. The micrographs were taken from glass slides by a pathologist and annotation and image label (benign, non-invasive carcinoma, and invasive carcinoma) were provided for each image. The Single Shot Multibox Detector (SSD) model was trained using 1361 images and evaluated using 315 images. The model performance was evaluated by the intersection over union (IoU) and diagnostic accuracy using detection of the model. To investigate the significance of our model in assisting the diagnosis, 3 pathologists and 5 medical students diagnosed images with or without assistance of the model.

**Table 1**  
The dataset summary.

	Number of pictures				Number of annotations			
	Benign	Non-invasive carcinoma	Invasive carcinoma	Total	Benign	Non-invasive carcinoma	Invasive carcinoma	Total
Training	608	337	416	1361	3516	1485	1294	6295
Test	164	40	111	315	853	154	247	1254
Examination set-1	14	15	14	43	-	-	-	-
Examination set-2	14	15	14	43	-	-	-	-



**Fig. 2.** Example images of annotation by pathologists and detection by the trained SSD model. Examples of the annotation (left) and the model detection (right). The blue, red, and green boxes indicated benign, non-invasive carcinoma, and invasive carcinoma, respectively. **A:** The image of accurately detection by the model. **B:** The image that some boxes with different labels were detected in same region. **C:** The image that the model accurately detected the benign area by bounding boxes with different shape against the annotation.

the object shape for prediction. Furthermore, the network combines predictions from multiple feature maps with various resolutions, thereby processing the objects at diversified sizes. SSD is simple because it sums up all

computation in a single network, leading to detection at high speed. Thus, SSD can train and integrate into systems with comparative ease without using machines of too high specifications. We performed 300 000

iterations of the train using 1361 images in the training set and conducted random horizontal flipping for images during training for data augmentation. The test set used 315 images to evaluate the model. An example image of accurate detection by the trained model was shown in Fig. 2A.

### Evaluation measures

This study evaluated the performance of the model in diagnosing each image. The diagnosis of each image was determined following the same rules, as well as giving an image label in building the data set, using the detection results of the trained model. Images were classified as benign, non-invasive carcinoma, or invasive carcinoma in the 3-class task and classified as benign or malignant, including non-invasive carcinoma and invasive carcinoma, in the 2-class task. Images that the model did not detect the lesions were considered as “Benign.” Detection with higher confidence score was used for diagnosis and that with lower confidence score was not used, if some boxes with different labels were detected in the same region (Fig. 2B). Accuracy, recall, precision, and F1-score were used following the formula below to evaluate the diagnosis of images.<sup>36</sup>

$$\text{Accuracy (Diagnostic accuracy)}(\%) = \frac{\text{Accurately classified images}}{\text{Total images}} \times 100$$

$$\text{Recall}(\%) = \frac{\text{Accurately classified target images}}{\text{Total target images}} \times 100$$

$$\text{Precision}(\%) = \frac{\text{Accurately classified target images}}{\text{Total predicted target images}} \times 100$$

$$\text{F1-score} = 2 * (\text{Recall} * \text{Precision}) / (\text{Recall} + \text{Precision})$$

Moreover, the intersection over union (IoU) score was used to evaluate the performance of breast lesion detection.<sup>37</sup> The IoU score was calculated using the below formula.

$$\text{IoU score} = \frac{\text{Area of overlap}}{\text{Area of Union}}$$

### Impact of the SSD model on the diagnosis by pathologists and medical students

We compared the diagnostic accuracy between the trained SSD model, pathologists, and medical school students, and investigated the effects of this model for assisting the diagnosis. A total of 3 experienced pathologists and 5 medical school students who conducted the pathological studies using breast carcinoma tissues participated in the experiments. They were not involved in building a data set. They diagnosed the Examination set-1 and set-2, both with 43 images. The accuracies of the trained model were the same score between Examination set-1 and set-2. First, they diagnosed images in Examination set-1 without the assistance of the trained model. Then, they classified the images in Examination set-2 with the assistance of model detection. The trained model was implemented as a simple tool, which showed the result of object detection and confidence score by inputting an image, and the threshold of detection confidence score could be freely changed in the tool. They were given information about the performance of the trained model, and the selection of the threshold of detection confidence score for display was left to them. All experiments were conducted without time constraints and the time taken for diagnosis of all images in each set was recorded. The diagnosis accuracy and needed time were analyzed between the trained model, pathologists, and medical students, as well as between the model-assisted and unassisted patterns.

### Software, tool, and statistical analysis

All the SSD model experiments were performed using a PC with the following specifications: Intel(R) Core (TM) i5-10300H processor with 16 GB RAM and NVIDIA(R) GeForce RTX (TM) 2060 GPU. The models were

mounted and trained by TensorFlow, and a tool to simply use this trained model was implemented. IoU score was calculated in python using a shapely package. A student t-test was used to examine differences in the diagnostic accuracies of pathologists and medical school students.

## Results

### Accuracy of diagnosis for images using the SSD model

The SSD model was trained using 1361 images in the training set and evaluated using 315 images in the test set. We investigated the diagnostic accuracy for images using detection by the trained model in the different thresholds of available confidence scores of detected objects. Thresholds of confidence score were changed from 0.1 to 0.9 at intervals of 0.1. The present model achieved a diagnostic accuracy of 85.4% using detection confidence score thresholds of 0.3 and 0.4, as shown in Table 2. The calculated F1-score with recall and precision score showed 89.6% for images labeled with benign (the confidence score threshold of 0.4), 71.0% for non-invasive carcinoma (the confidence score threshold of 0.5), and 88.8% for invasive carcinoma (the confidence score threshold of 0.1). The model achieved a higher diagnostic accuracy of 88.3% using the threshold of detection confidence score when each component showed the highest F1-score (benign for 0.4, non-invasive carcinoma for 0.5, and invasive carcinoma for 0.1 of confidence score thresholds). Furthermore, the trained model showed 90.5% accuracy under the same conditions in a 2-class task to diagnose as benign or malignant (non-invasive carcinoma or invasive carcinoma). Conversely, the model showed a relatively low precision score (68.0%) for the classification of images of non-invasive carcinoma. Additionally, the model sometimes could not detect lesions in the images that showed remarkably weak H&E staining in this study (Supplementary figure).

### Performance of object detection by the SSD model

We evaluated the trained model by IoU score, which is a better detection evaluation metric.<sup>37</sup> The mean IoU of benign detection was 0.52, non-invasive carcinoma was 0.44, and invasive carcinoma was 0.62 when the threshold of detection confidence score for benign was 0.4, non-invasive carcinoma was 0.5, and invasive carcinoma was 0.1, as shown in Table 3. The 3-class average IoU was 0.59.

The mean average precision (mAP)<sup>38</sup> is also useful for evaluating the target localization and detection model, but we did not use it to evaluate this model. Breast lesions could not be divided and counted simply due to the intricacies of these forms and there is a wide range of variations of drawing bounding boxes around them. For example, the model accurately detected the benign area by bounding boxes with different shapes against the annotation boxes as shown in Fig. 2C. Hence, we considered that mAP was not suitable to accurately evaluate this model.

### Comparison of the SSD model with humans and its effects on diagnosis

We evaluated the trained SSD model as compared with 3 pathologists and 5 medical students using 43 images in Examination set-1. They diagnosed these images without time constraints, and diagnostic accuracy was evaluated. All results were shown in Table 4 and Fig. 3. The model showed higher diagnostic accuracy than medical students while pathologists got remarkably higher accuracy than the model and medical students in both 3-class (Fig. 3A) and 2-class (Fig. 3B) tasks.

We then investigated the significance of the model in assisting the diagnosis by humans by evaluating its effects on diagnostic accuracy and the time taken for diagnosis. They conducted the diagnosis for images in Examination set-2, which contains different 43 images from Examination set-1 but keeps the same composition, with the assistance of the model detection. Medical students showed remarkably higher diagnostic accuracy scores (3-class task: average 84.7%, 2-class task: average 88.4%) with the model assistance compared to that without model assistance (3-class task: average

**Table 2**  
Model performance for diagnosis of images in different threshold of confidence score.

			Threshold of detection confidence score.											
			0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	Benign; 0.4 Non-invasive carcinoma; 0.5 Invasive carcinoma; 0.1		
3-class task	Diagnostic accuracy (%)		81.6	83.8	85.4	85.4	84.8	82.9	82.5	80.0	77.8	<b>88.3</b>		
	Benign	Recall										90.9		
		Precision										90.9		
	F1-score	Non-invasive carcinoma	F1-score										90.9	
			Recall										85.0	
		Invasive carcinoma	Precision										68.0	
F1-score												75.6		
2-class task	Diagnostic accuracy (%)		<b>88.8</b>	88.3	88.2	85.4	83.7	81.3	79.8	74.4	68.2	89.6		
	Recall (%)	Recall										85.6		
		Precision										94.1		
	F1-score	F1-score										89.6		
		Recall (%)										90.5		
	Precision (%)	Precision										90.1		
F1-score											90.1			

**Table 3**  
Model performance for the breast lesions detection.

Intersection over Union (IoU) <sup>a</sup>			
Benign	Non-invasive carcinoma	Invasive carcinoma	3-class average
0.52 (0.32)	0.44 (0.39)	0.62 (0.34)	<b>0.59 (0.27)</b>

<sup>a</sup> Threshold of detection confidence score; Benign 0.4, Non-invasive carcinoma 0.5, Invasive carcinoma 0.1. Data were presented as mean (STD).

67.4%, 2-class task: average 79.1%) in both 3-class (Figure 3C) and 2-class (Figure 3D) tasks. Notably, no significant difference was found between the accuracy of pathologists and medical students with the assistance of the model in the 2-class task. Furthermore, some medical students achieved the same or higher accuracy scores than pathologists when using the assistance of the model for diagnosis (Fig. 3E, F). Pathologists achieved a diagnostic accuracy score of more than the 90% average with or without the assistance of the model. Two participants diagnosed images in a short time using the assistance of the model (shorten by average 6.4 min) while six participants required a longer time (extended by average 7.2 min) when we investigated the time taken for diagnosis of all images in each set (Fig. 3G). The average times taken for diagnosis were extended by 5.0 min and 3.1 min in pathologists and medical students, respectively.

**Discussion**

The present study firstly demonstrated the SSD model as an object detection method to detect the breast lesions in micrographs of human breast tissues to our best knowledge and evaluated the significance of the model in assisting the diagnosis. Many currently employed AI models that classify

**Table 4**  
Impact of the SSD model on the diagnosis by pathologists and medical students.

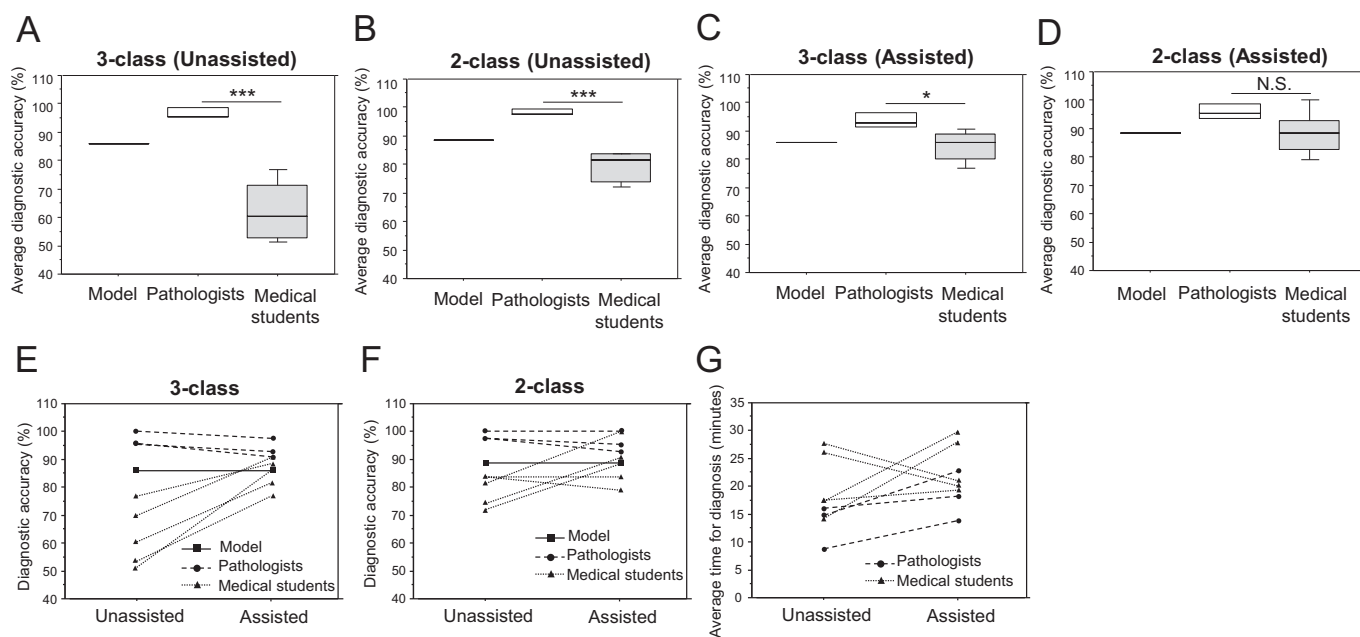
		Examination set-1 (Unassisted)			Examination set-2 (Assisted)		
		Model only <sup>a</sup>	Pathologists (n = 3)	Medical students (n = 5)	Model only <sup>a</sup>	Pathologists (n = 3)	Medical students (n = 5)
3-class task	Diagnostic accuracy (%)	86.0	97.7	67.4	86.0	93.8	84.7
	Diagnostic accuracy (%)	88.4	98.4	79.1	88.4	96.1	88.4
	Precision (%)	92.9	98.9	78.0	100	96.7	89.6
2-class task	Recall (%)	89.7	98.9	96.6	82.8	97.7	93.8
	F1-score	91.2	98.9	86.2	90.6	97.1	91.6
Time taken for diagnosis (min)		-	12.9	20.2	-	17.9	23.3
Time differences between unassisted or assisted tasks (min)		-	-	-	-	+5.0	+3.1

<sup>a</sup> Threshold of detection confidence score; Benign 0.4, Non-invasive carcinoma 0.5, Invasive carcinoma 0.1. Data of pathologists and medical students were presented as average scores.

breast pathological images were developed using an algorithm, such as ANN, including CNN,<sup>39-42</sup> ResNet,<sup>43,44</sup> AlexNet,<sup>14,45</sup> and Inception-V3,<sup>46,47</sup> however, the present model provides more detailed information not only about the classification of images but also location and type of lesion in images, and it could be useful for assisting the pathologists. The 3-class and 2-class classification tasks obtained 88.3% and 90.5% diagnostic accuracies for images, respectively. The diagnosis for images using the model achieved higher accuracy scores than medical students in experiments for model evaluation compared with pathologists and medical students, which would be helpful for pathological diagnosis.

Conversely, pathologists achieved higher accuracies than the model. The mean IoU is 0.59 and there is room for improvement in the performance of the model. There may be a limit to completely annotate and accurately detect lesions by object detection methods using bounding boxes because of the intricacies of these forms. Using other methods, such as semantic segmentation,<sup>48</sup> with pixel-level annotated images may be efficient if the speed or simplicity of machine specs for implementation is not so important for users. For example, Priego-Torres et al. presented a new framework for carcinoma cell segmentation in breast histopathological images and the estimated segmentation accuracy was 95.62%.<sup>49</sup> Additionally, new released object detection methods after SSD also may be useful in the aspect of accuracy.<sup>50</sup>

We then investigated the significance of the present model in assisting the diagnosis. Steiner et al evaluated the potential impact of assistance for pathologists by automatic detection of breast carcinoma metastasis in lymph nodes and revealed some benefits of the assistance by AI in pathology.<sup>51</sup> Very recently, Mantrala et al investigated the concordance rate in breast carcinoma grading as determined by the Nottingham Grading System between AI and pathologists using WSI.<sup>27</sup> However, the impact of assistance by AI in diagnosing various breast lesions on clinical utility has



**Fig. 3.** Performance of diagnosis by the SSD model, pathologists, and medical students with or without assistance of the model. **A–D:** The average diagnostic accuracy of the model, pathologists ( $n = 3$ ), and medical students ( $n = 5$ ) without assistance (**A, B**) or with assistance (**C, D**) of the model in the 3-class (**A, C**) or 2-class (**B, D**) classification tasks. A student t-test was used to examine differences in the diagnostic accuracies and the data were presented as the mean  $\pm$  S.D. \* $P < 0.05$ , \*\*\* $P < 0.001$  vs model, respectively. N.S.; not significant. **E, F:** The change of each accuracy score with or without assistance of the model in 3-class (**E**) and 2-class (**F**) classification tasks. **G:** The change of the time taken for diagnosis with or without assistance of the model.

not been fully investigated. The present study revealed that medical students upregulated accuracy scores when using the assistance of the model and some students achieved the same or higher accuracy scores than pathologists, although the SSD model and medical students showed lower diagnostic accuracies than pathologists. Notably, some medical students showed the best performance when using the assistance of the present model compared with the model only or without the assistance of the model. These results indicated that the present model may be helpful for medical students or pathologists who are not so skilled in breast disorders to accurately diagnose using micrographs, and using AI as assistance has the potential to support pathological diagnosis in laboratories. Additionally, the present model uses micrographs of breast tissues although WSI scanning still takes a lot of time,<sup>32</sup> and there is a wide range of variation of methods to use. For example, pathologists can take micrographs and quickly confirm the detection of the model only if they hesitate to judge certain components. A microscope is used in routine diagnosis in clinical practice, and the present model may be comparatively easy to introduce and simpler to use at high speed using micrographs. Conversely, the average accuracy of pathologists achieved  $>90\%$  and could not be changed by the assistance of the model. Further amelioration of the performance of the model can be helpful for pathologists, with a broader range of experience levels, including virtuosic pathologists. Additionally, these experiments were conducted without time constrain and AI may be more effective in stressful situations, such as when many images are needed to be diagnosed under time constrain in clinical practice.

Some medical students achieved high diagnostic accuracy in a short time using the model when pathologists and medical students diagnosed images using the assistance of the model, and this result indicated that assistance by the present model may have the potential to improve both accuracy and efficiency of pathological diagnosis. On the other hands, diagnosing images with assistance took a long time on average than without assistance. The present study used the model by inputting one image at a time. More simple operation methods to use the model could further improve the efficiency. Additionally, further investigation about appropriate conditions in using AI tools will be needed to improve the efficiency of diagnosis for more people. Too much information from AI can lead to

confusion among pathologists and increase the workload to process information, thus an appropriate amount of information should be supplied for them. For example, by setting the suitable threshold of confidence score. The relationship between accuracy and efficiency of diagnosis using AI as assistance, when the amount of information from AI is changed, should be explored in the future.

One of the limitations of the present study is the created dataset with specimens obtained from a single institution and it may not represent the heterogeneity of specimens in various facilities. The present model could not detect lesions in images taken by specimens whose hematoxylin staining is weak compared with other specimens. The model should be trained and evaluated using other images taken from specimens in other facilities in the future.

In conclusion, we presented the automatic breast lesions detection method using histopathological micrographs based on SSD, which is an object detection algorithm, for the first time. The model can detect breast lesions in micrographs at high speed with a low cost for the introduction. The model showed 88.3% diagnostic accuracies for images in 3-class classification tasks and medical students improved their diagnosis performance using the assistance of this model. Therefore, the present system conveniently supports the histological diagnosis by pathologists in laboratories.

#### Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

#### Disclosure

The authors report no conflicts of interest in this work.

#### Acknowledgments

We would like to thank the volunteer medical school students (Freeha Khalid, Shiori Fujisawa, Reina Taguchi, Takumi Fukasawa and Kento Iida)

who have been participated in our study and Enago ([www.enago.jp](http://www.enago.jp)) for the English language review.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jpi.2022.100147>.

## References

- Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J Clin* 2018;68:394–424. <https://doi.org/10.3322/caac.21492>.
- Gaymard B, Sahin AA. Benign breast diseases: classification, diagnosis, and management. *Oncologist* 2006;11:435–449. <https://doi.org/10.1634/theoncologist.11.5-435>.
- Shackney SE, Silverman JF. Molecular evolutionary patterns in breast cancer. *Adv Anat Pathol* 2005;205:248–254. <https://doi.org/10.1097/00125480-200309000-00003>.
- Simpson PT, Reis-Filho JS, Gale T, Lakhani SR. Molecular evolution of breast cancer. *J Pathol* 2005;205:248–254. <https://doi.org/10.1002/path.1691>.
- Veta M, Plum J, van Diest PJ, Viergever MA. Breast cancer histopathology image analysis: a review. *IEEE Trans Biomed Eng* 2014;61:1400–1411. <https://doi.org/10.1109/TBME.2014.2303852>.
- Murakami M, Matoba K, Terazawa K. Raising future forensic pathologists (first report): results of a preliminary questionnaire survey. *Hokkaido Igaku Zasshi* 2012;87:153–157.
- Mudenda V, Malyangu E, Sayed S, Fleming K. Addressing the shortage of pathologists in Africa: creation of a MMed programme in pathology in Zambia. *Afr J Lab Med* 2020;9:974. <https://doi.org/10.4102/ajlm.v9i1.974>.
- Märkl B, Füzesi L, Huss R, Bauer S, Schaller T. Number of pathologists in Germany: comparison with European countries, USA, and Canada. *Virchows Arch* 2021;478:335–341. <https://doi.org/10.1007/s00428-020-02894-6>.
- Naritoku WY, Furlong MA, Knollman-Ritschel B, Kaul KL. Enhancing the pipeline of pathologists in the United States. *Acad Pathol* 2021;8. <https://doi.org/10.1177/23742895211041725>.
- Spanhol FA, Oliveira LS, Petitjean C, Heutte L. A dataset for breast cancer histopathological image classification. *IEEE Trans Biomed Eng* 2016;63:1455–1462. <https://doi.org/10.1109/TBME.2015.2496264>.
- Kausar T, Wang M, Idrees M, Lu Y. HWDCCNN: multi-class recognition in breast histopathology with Haar wavelet decomposed image based convolution neural network. *Biocybernet Biomed Eng* 2019;39:967–982. <https://doi.org/10.1016/j.bbe.2019.09.003>.
- Zhou X, Li C, Rahaman MM, et al. A comprehensive review for breast histopathology image analysis using classical and deep neural networks. *IEEE Access* 2020;8:90931–90956. <https://doi.org/10.1109/ACCESS.2020.2993788>.
- Krithiga R, Geetha P. Breast cancer detection, segmentation and classification on histopathology images analysis: a systematic review. *Arch Computat Method Eng* 2021;28:2607–2619. <https://doi.org/10.1007/s11831-020-09470-w>.
- Kiambe K. Breast histopathological image feature extraction with convolutional neural networks for classification. *ICSES Trans Image Process Pattern Recognit* 2018;4:4–12.
- Jiang Y, Chen L, Zhang H, Xiao X. Breast cancer histopathological image classification using convolutional neural networks with small SE-ResNet module. *PLoS One* 2019;14:e0214587. <https://doi.org/10.1371/journal.pone.0214587>.
- Spanhol FA, Oliveira LS, Petitjean C, Heutte L. Breast cancer histopathological image classification using convolutional neural networks. *Proc Int Joint Conf Neural Netw* 2016;2560–2567.
- Nahid AA, Mehrabi MA, Kong Y. Histopathological breast cancer image classification by deep neural network techniques guided by local clustering. *Biomed Res Int* 2018;2018:2362108. <https://doi.org/10.1155/2018/2362108>.
- Yan R, Ren F, Wang Z, et al. Breast cancer histopathological image classification using a hybrid deep neural network. *Methods* 2020;173:52–60. <https://doi.org/10.1016/j.ymeth.2019.06.014>.
- Iizuka O, Kanavati F, Kato K, Rambeau M, Arihiro K, Tsuneki M. Deep learning models for histopathological classification of gastric and colonic epithelial tumours. *Sci Rep* 2020;10:1504. <https://doi.org/10.1038/s41598-020-58467-9>.
- Li X, Shen X, Zhou Y, Wang X, Li TQ. Classification of breast cancer histopathological images using interleaved DenseNet with SENet (IDSNet). *PLoS One* 2020;15:e0232127. <https://doi.org/10.1371/journal.pone.0232127>.
- Liu W, Juhas M, Zhang Y. Fine-grained breast cancer classification with bilinear convolutional neural networks (BCNNs). *Front Genet* 2020;11:547327. <https://doi.org/10.3389/fgene.2020.547327>.
- Cruz-Roa A, Basavanthally A, González F, et al. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. *Proc Medical Imaging* 2014; Digital Pathol 2014;9041:904103. <https://doi.org/10.1117/12.2043872>.
- Romero FP, Tang A, Kadoury S. Multi-level batch normalization in deep networks for invasive ductal carcinoma cell discrimination in histopathology images. *Proc. 2019 IEEE 16th International Symposium on Biomedical Imaging*; 2019. p. 1092–1095. <https://doi.org/10.1109/ISBI.2019.8759410>.
- Nensa F, Demircioglu A, Rischpler C. Artificial intelligence in nuclear medicine. *J Nucl Med* 2019;60:29S–37S. <https://doi.org/10.2967/jnumed.118.220590>.
- Mukhopadhyay S, Feldman MD, Abels E, et al. Whole slide imaging versus microscopy for primary diagnosis in surgical pathology: a multicenter blinded randomized noninferiority study of 1992 cases (Pivotal Study). *Am J Surg Pathol* 2018;42:39–52. <https://doi.org/10.1097/PAS.0000000000000948>.
- Ehteshami Bejnordi B, Veta M, Johannes van Diest P, et al. Diagnostic assessment of deep learning algorithms for detection of lymph node metastases in women with breast cancer. *JAMA* 2017;318:2199–2210. <https://doi.org/10.1001/jama.2017.14585>.
- Mantrala S, Ginter PS, Mitkar A, et al. Concordance in breast cancer grading by artificial intelligence on whole slide images compares with a multi-institutional cohort of breast pathologists. *in press. Arch Pathol Lab Med* 2022. <https://doi.org/10.5858/arpa.2021-0299-OA>.
- Abdollahi M, Salehi M, Shokatian I, Reiazi R. Artificial intelligence in automatic classification of invasive ductal carcinoma breast cancer in digital pathology images. *Med J Islam Repub Iran* 2020;34:140. <https://doi.org/10.34171/mjiri.34.140>.
- Pantanowitz L, Valenstein PN, Evans AJ, et al. Review of the current state of whole slide imaging in pathology. *J Pathol Inform* 2011;2:36. <https://doi.org/10.4103/2153-3539.83746>.
- Pantanowitz L, Sinard JH, Henricks WH, et al. Validating whole slide imaging for diagnostic purposes in pathology: guideline from the College of American Pathologists Pathology and Laboratory Quality Center. *Arch Pathol Lab Med* 2013;137:1710–1722. <https://doi.org/10.5858/arpa.2013-0093-CP>.
- Atallah NM, Toss MS, Verrill C, Salto-Tellez M, Snead D, Rakha EA. Potential quality pitfalls of digitalized whole slide image of breast pathology in routine practice. *Mod Pathol* 2022;35:903–910. <https://doi.org/10.1038/s41379-021-01000-8>.
- Kumar N, Gupta R, Gupta S. Whole slide imaging (WSI) in pathology: current perspectives and future directions. *J Digit Imaging* 2020;33:1034–1040. <https://doi.org/10.1007/s10278-020-00351-z>.
- Liu W, Angelou D, Erhan D, et al. SSD: Single shot multibox detector. *in Proc. European Conference on Computer Vision*; 2016. p. 21–37.
- Feldman AT, Wolfe D. Tissue processing and hematoxylin and eosin staining. *Methods Mol Biol* 2014;1180:31–43. [https://doi.org/10.1007/978-1-4939-1050-2\\_3](https://doi.org/10.1007/978-1-4939-1050-2_3).
- Yamaguchi M, Takagi K, Sato M, et al. Androgens enhance the ability of intratumoral macrophages to promote breast cancer progression. *Oncol Rep* 2021;46:188. <https://doi.org/10.1038/s12078-021-01319-1>.
- Yang Y, Liu X. A re-examination of text categorization methods. *Proc. 22nd ACM International Conference on Research and Development in Information Retrieval*; 1999. p. 42–49.
- Rezatofghi H, Tsoi N, Gwak J, Sadeghian A, Reid I, Savarese S. Generalized intersection over union: A metric and a loss for bounding box regression. *Proc. IEEE Computer Vision and Pattern Recognition*; 2019. p. 658–666.
- Padilla R, Netto SL, Da Silva EAB. A survey on performance metrics for object-detection algorithms. *Proc. 2020 International Conference on Systems, Signals and Image Processing (IWSSIP)*; 2020. p. 237–242. <https://doi.org/10.1109/IWSSIP48289.2020.9145130>.
- Du B, Qi Q, Zheng H, Huang Y, Ding X. Breast cancer histopathological image classification via deep active learning and confidence boosting. *Proc. International Conference on Artificial Neural Networks*; 2018. p. 109–116.
- Nazeri K, Aminpour A, Ebrahimi M. Two-stage convolutional neural network for breast cancer histology image classification. *Proc. International Conference Image Analysis and Recognition*; 2018. p. 717–726.
- Roy K, Banik D, Bhattacharjee D, Nasipuri M. Patch-based system for classification of breast histology images using deep learning. *Comput Med Imaging Graph* 2019;71:90–103. <https://doi.org/10.1016/j.compmedimag.2018.11.003>.
- Saxena S, Shukla S, Gyanchandani M. Pre-trained convolutional neural networks as feature extractors for diagnosis of breast cancer using histopathology. *Int J Imag Syst Technol* 2020;1-15. <https://doi.org/10.1002/ima.22399>.
- Gandomkar Z, Brennan P, Mello-Thoms C. A framework for distinguishing benign from malignant breast histopathological images using deep residual networks. *Proc. 14th International Workshop on Breast Imaging (IWBI 2018)*, 10718. ; 2018. p. 204–210. <https://doi.org/10.1117/12.2318320>.
- Bhuiyan M, Shamsujjoha M, Ripon SH, Proma FH, Khan F. Transfer learning and supervised classifier based prediction model for breast cancer. *Proc. Big Data Analytics for Intelligent Healthcare Management*; 2019. p. 59–86. <https://doi.org/10.1016/B978-0-12-818146-1.00004-0>.
- Ahmad H, Ghuffar S, Khurshid K. Classification of breast cancer histology images using transfer learning. *Proc. 2019 16th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*; 2019. p. 328–332. <https://doi.org/10.1109/IBCAST.2019.8667221>.
- Vang YS, Chen Z, Xie X. Deep learning framework for multi-class breast cancer histology image classification. *Proc In International Conference Image Analysis and Recognition*; 2018. p. 914–922.
- Rakhlin A, Shvets A, Igloukov V, Kalinin A. Deep convolutional neural networks for breast cancer histology image analysis. *Proc. International Conference Image Analysis and Recognition*; 2018. p. 737–744.
- Noh H, Hong S, Han B. Learning deconvolution network for semantic segmentation. *Proc. Proceedings of the IEEE International Conference on Computer Vision*; 2015. p. 1520–1528.
- Priego-Torres BM, Sanchez-Morillo D, Fernandez-Granero MA, Garcia-Rojo M. Automatic segmentation of whole-slide H&E-stained breast histopathology images using a deep convolutional neural network architecture. *Expert Syst Appl* 2020;151, 113387. <https://doi.org/10.1016/j.eswa.2020.113387>.
- SSA Zaidi, MS Ansari, A Aslam, N Kanwal, M Asghar, B Lee. A Survey of Modern Deep Learning based Object Detection Model. [10.48550/arXiv.2104.11892](https://arxiv.org/abs/2104.11892)
- Leiner DF, MacDonald R, Liu Y, et al. Impact of deep learning assistance on the histopathologic review of lymph nodes for metastatic breast cancer. *Am J Surg Pathol* 2018;42:1636–1646. <https://doi.org/10.1097/PAS.0000000000001151>.