



OPEN

Deep UV-excited fluorescence microscopy installed with CycleGAN-assisted image translation enhances precise detection of lymph node metastasis towards rapid intraoperative diagnosis

Junya Sato ^{1,4,5,7}, Tatsuya Matsumoto ^{2,7}, Ryuta Nakao ^{2,7}, Hideo Tanaka ², Hajime Nagahara ^{1,3}, Hirohiko Niioka ^{1,3}✉ & Tetsuro Takamatsu ^{2,6}✉

Rapid and precise intraoperative diagnosing systems are required for improving surgical outcomes and patient prognosis. Because of the poor quality and time-intensive process of the prevalent frozen section procedure, various intraoperative diagnostic imaging systems have been explored. Microscopy with ultraviolet surface excitation (MUSE) is an inexpensive, maintenance-free, and rapid imaging technique that yields images like thin-sectioned samples without sectioning. However, pathologists find it nearly impossible to assign diagnostic labels to MUSE images of unfixed specimens; thus, AI for intraoperative diagnosis cannot be trained in a supervised learning manner. In this study, we propose a deep-learning pipeline model for lymph node metastasis detection, in which CycleGAN translate MUSE images of unfixed lymph nodes to formalin-fixed paraffin-embedded (FFPE) sample, and diagnostic prediction is performed using deep convolutional neural network trained on FFPE sample images. Our pipeline yielded an average accuracy of 84.6% when using each of the three deep convolutional neural networks, which is a 18.3% increase over the classification-only model without CycleGAN. The modality translation to FFPE sample images using CycleGAN can be applied to various intraoperative diagnostic imaging systems and eliminate the difficulty for pathologists in labeling new modality images in clinical sites. We anticipate our pipeline to be a starting point for accurate rapid intraoperative diagnostic systems for new imaging modalities, leading to healthcare quality improvement.

Abbreviations

DUV Deep-ultraviolet
FFPE Formalin-fixed paraffin-embedded
DCNN Deep convolutional neural network

¹Graduate School of Information Science and Technology, Osaka University, 1-5, Yamadaoka, Suita, Osaka 565-0871, Japan. ²Department of Pathology and Cell Regulation, Kyoto Prefectural University of Medicine, 465 Kajicho, Kawaramachi-Hirokoji, Kamigyo-ku, Kyoto 602-8566, Japan. ³Institute for Dataability Science, Osaka University, 2-8 Yamadaoka, Suita 565-0871, Japan. ⁴Department of Radiology, Osaka University Graduate School of Medicine, 2-2, Yamadaoka, Suita, Osaka 565-0871, Japan. ⁵Department of Artificial Intelligence Diagnostic Radiology, Osaka University Graduate School of Medicine, 2-2, Yamadaoka, Suita, Osaka 565-0871, Japan. ⁶Department of Medical Photonics, Kyoto Prefectural University of Medicine, 465 Kajicho, Kawaramachi-Hirokoji, Kamigyo-ku, Kyoto 602-8566, Japan. ⁷These authors contributed equally: Junya Sato, Tatsuya Matsumoto and Ryuta Nakao. ✉email: niioka@ist.osaka-u.ac.jp; ttakam@koto.kpu-m.ac.jp

MUSE	Microscopy with ultraviolet surface excitation
TPEF	Two-photon excited autofluorescence
SHG	Second harmonic generation
THG	Third harmonic generation
SRS	Stimulated Raman scattering

The number of patients with cancer is increasing with global population growth and ageing, and a total of 19.3 million patients were newly diagnosed in 2020¹. The gold standard treatment for radical cure of cancer is surgery, and detailed surgical procedures are determined according to the tumour size and metastasis to surrounding lymph nodes.

Rapid intraoperative diagnosis plays an important role in detecting the presence of sentinel lymph node metastases for some malignancies (e.g., gastric and breast cancers). It is a direct decision-making aid for surgeons that affects surgical progress and patient prognosis. Although pathological diagnosis based on tissue sections and staining is the gold standard for cancer diagnosis, these procedures take several days to complete². For rapid intraoperative diagnosis of lymph node metastases, a frozen section procedure is widely used; however, this process is associated with some problems, such as sample degradation due to the formation of ice crystals and the requirement of experience in handling the samples. Recently, the OSNA™ (One-Step Nucleic Acid Amplification) method has been developed, enabling the determination of metastasis based on the amplification and detection of cytokeratin 19 mRNA contained in the solubilised lymph nodes. This method has been mainly applied to the sentinel lymph node of breast cancer, and highly accurate results have been obtained without pathological procedures³. However, these two methods, namely frozen section and OSNA, generally require approximately 20 min for the entire detection process. Another approach is the highly sensitive and specific 5-aminolevulinic acid (5-ALA)-induced rapid fluorescence imaging method^{4,5}. We have previously developed a device that automatically detects lymph node metastasis by quantitatively evaluating the 5-ALA-induced protoporphyrin IX (PpIX) fluorescence intensity while eliminating tissue autofluorescence. However, the activity of the ATP-binding cassette transporter G2 is very high in breast cancer cells, making 5-ALA unsuitable for detecting breast cancer lymph node metastasis⁶.

Microscopy with ultraviolet surface excitation (MUSE) is a less expensive and maintenance-free imaging system with the potential for rapid intraoperative diagnosis⁷. As deep-ultraviolet (DUV) light in the 250–300 nm wavelength range is easily absorbed and scattered by biological samples, resulting in excitation only near the sample surface, MUSE produces images like thin-sliced sample sections without sectioning. Observing cross-sections of the excised tissue is adequate for the pathologist to distinguish clearly between the nucleoplasm, nucleolus, and cytoplasm. Our group has reported a staining method for MUSE using Terbium ion and Hoechst/DAPI, which provided clear contrast images and required only 5 min of staining time⁷.

In our previous study, formalin-fixed paraffin-embedded (FFPE) sections of lymph nodes from gastric cancer patients stored in a hospital were stained and observed with MUSE to create a fluorescent image dataset⁸. These images were utilised for training a deep convolutional neural network (DCNN), achieving high diagnostic accuracy comparable to that of a pathologist using HE images. The advantage of MUSE is that metastatic lesions can be detected in a short time using unfixed specimen images instead of FFPE images. The contrast of an unfixed specimen image differs from that of an FFPE image, such as that at the cell boundary. Thus, if the unfixed specimen image of the test data can be translated to an FFPE-like image, a high detection rate can be achieved. FFPE images allow us to train image classification models, such as DCNN, because supervised labels can be assigned to cancerous and other regions by referring to the serially sectioned HE images. As labelling is not possible for MUSE images of unfixed specimens, image classification models are not trainable in a supervised manner.

In this study, we propose a rapid diagnostic pipeline model that combines a CycleGAN⁹ to translate MUSE images of unfixed lymph nodes to FFPE sample images with a DCNN trained on FFPE sample images for diagnostic prediction. CycleGAN is a successful image translation method that uses unpaired image data. Some CycleGAN-based methods have been reported to correct the colour of HE images at different facilities^{10–12}, translate the images to a different staining method^{13–16}, and improve accuracy by augmentation¹⁷. First, we trained an image translation model that translates between unfixed and FFPE MUSE images using CycleGAN and a diagnostic prediction DCNN model using FFPE sample images. Next, the CycleGAN and DCNN models were coupled to form a pipeline model, and we compared the diagnostic prediction accuracy based on MUSE images of unfixed lymph nodes with and without CycleGAN.

Methods

Clinical specimens

All clinical experiments were conducted with the approval of the Ethics Committees of the Kyoto Prefectural University of Medicine (approval no.: ERB-C-1038) as well as in accordance with the guidelines from the committees and regional laws related to clinical research. The lymph nodes used in this study were obtained from patients during gastric cancer surgery at the University Hospital, Kyoto Prefectural University of Medicine. Informed consent was obtained from all participants. Patients were diagnosed as metastasis-positive or metastasis-negative based on postoperative pathological examination. None of the patients had received preoperative radiotherapy or chemotherapy. The number of prepared FFPE MUSE images for metastasis-positive and metastasis-negative samples was 38 and 18, respectively. These MUSE images were identical to those used in our previous study⁸. They were chosen such that they included all the common histological types of gastric carcinoma. For developing the CycleGAN translation model, the FFPE lymph nodes mentioned above and 19 metastasis-positive and 9 metastasis-negative unfixed lymph nodes obtained from 28 patients were used. For the MUSE images of unfixed lymph nodes, we had 45 metastasis-positive and 46 metastasis-negative images.

Staining protocol and fluorescence microscopy

Complete details of the sample preparation and imaging methods have been described in a previous study by Kumamoto et al.⁷. Briefly, surgically obtained lymph nodes were sliced and immersed in 95% or 99.5% ethanol, rinsed with HEPES buffer (10 mM HEPES, pH 7 adjusted with NaOH), immersed in 100% D₂O HEPES buffer solution containing TbCl₃ (TBH03XB, Kojundo Chemical Laboratory, Saitama, Japan) and Hoechst 33342 (Dojindo Molecular Technologies), and then rinsed with 100% D₂O HEPES buffer. Stained specimens were then set on an inverted microscope (IX71, Olympus) equipped with an objective lens (UPLFLN 10×, Olympus). The specimen's surface was illuminated with a DUV light (30 mW and 5 mm in power and diameter, respectively) emitting from an LED (M285L5, Thorlabs). Fluorescence emitted from the specimen was collimated with the objective lens and imaged using a CMOS camera (UI-3180CP-C-HQ Rev. 2, OnSemi). An optical filter (FF01-464/547-25, Semrock) was set between the objective and imaging lenses to attenuate DUV light. A schematic representation of the image acquisition protocol is shown in Fig. S1.

FFPE sections were obtained after standard formalin fixation and paraffin embedding, followed by thin sectioning at 4 μm. Before the MUSE imaging, the FFPE sections were deparaffinised and stained using the same method as that used for unfixed lymph nodes.

Training of CycleGAN

CycleGAN is a model with two generators and two discriminators that learns to translate between the domains of the two image datasets. The generator outputs an image in the other domain from the input image data domain, and the discriminator predicts which domain the output image belongs to. Two pairs of generator-discriminator perform domain translations and learn bi-directional translations. By training these pair models simultaneously, CycleGAN can output images that are restyled from one to the other. Each of the MUSE images of unfixed lymph nodes was divided into small patches composed of 256 × 256 pixels. The training dataset consisted of 256 × 256 pixels size images cropped from all the MUSE images of unfixed lymph nodes data used in this study. In this process, the 256 × 256 pixels size cropping window was scanned vertically and horizontally in 10 pixels steps. Thus, 1,364,540 patch images were prepared. The MUSE image dataset of the FFPE samples was prepared using the patch-image dataset from our previous study⁸. The original patch-image size was 278 × 278. Each patch was then resized into 256 × 256 pixels, and pixel values were scaled from -1 to 1 by dividing by 127.5 and subtracting 1 for normalization in the CycleGAN model.

The loss of CycleGAN consists of adversarial loss l_{adv} , cycle consistency loss l_{cycle} , and identity loss $l_{identity}$. First, adversarial loss, which was introduced by Goodfellow et al.¹⁸, was set up so that the generator G and discriminator D were simultaneously trained by trying to defeat each other. Given the images in domain A (X_A^{real}), those in domain B (X_B^{real}), the generative model that transforms domain A into domain B (G_{AB}), and the model that discriminates whether X_B is true or false (D_B), adversarial loss in domain B l_{adv}^B was expressed as:

$$l_{adv}^B = E_{X_B} [\log D_B(X_B^{real})^2] + E_{X_A} [(\log D_B(G_{AB}(X_A^{real})) - 1)^2],$$

where E_X is the expected value over all real data instances. l_{adv}^A was also calculated in the same way using G_{BA} and D_A . The total adversarial loss was the sum of l_{adv}^A and l_{adv}^B . Next, cycle consistency loss ensures the reversibility of the translation and makes it closer to a one-to-one conversion. The loss l_{cycle} was expressed as follows:

$$l_{cycle} = E_{X_A} [\|G_{BA}(G_{AB}(X_A^{real})) - X_A^{real}\|_1] + E_{X_B} [\|G_{AB}(G_{BA}(X_B^{real})) - X_B^{real}\|_1],$$

where $\|\cdot\|_1$ is L1-norm. The first term refers to cycle loss in domain A and the second term refers to cycle loss in domain B. Third, the identity loss regularises a colour change and may ensure common elements within a domain as proposed by Taigman et al.¹⁹. The loss $l_{identity}$ was expressed as:

$$l_{identity} = E_{X_A} [\|G_{BA}(X_A^{real}) - X_A^{real}\|_1] + E_{X_B} [\|G_{AB}(X_B^{real}) - X_B^{real}\|_1].$$

Using the above three types of loss, the total loss was

$$l_{total} = l_{adv} + \alpha l_{cycle} + \beta l_{identity},$$

where α and β are relative ratios; in this study, we set $\alpha = 10$ and $\beta = 1$.

Adam was used as an optimiser, and the learning rate was set at 0.0002 based on the original study⁹. The model was trained on 48 epochs.

Training of DCNN models

Three image classification DCNN models were adopted, including Inception v3²⁰, DenseNet121²¹, and EfficientNetB4²². These models have been previously applied in medical image classification, achieving high accuracy of 98.4%, 99.1%²³, and 94.9%²⁴, respectively, in the colorectal cancer prediction task. We downloaded the weights pre-trained on ImageNet²⁵ and performed transfer learning. In these models, global average pooling is preceded before the output layer composed of 2 nodes. In the transfer learning of these three models, only the weights of the last fully connected layer are trained first, and then the weights of the whole model are trained. As an optimiser, stochastic gradient descent was used for all models. All MUSE images of FFPE lymph nodes were divided into 27,704 small patches of 256 × 256 pixels in size without them overlapping each other. Of these, 10,502 metastasis-positive and 9931 metastasis-negative patches were used for training, 1988 metastasis-positive and 1455 metastasis-negative patches were used for validation, and 1834 metastasis-positive and

1994 metastasis-negative patches were used for testing. During data partitioning, the patch images were treated together as a group for each patient and not as split. Each image patch was normalized to a range between 0 and 1 by dividing each pixel value by the maximum pixel value of the image before being input into the model. In all DCNN models, training was performed with the training dataset, and the weights were saved when the accuracy for the validation dataset was the highest. Saved weights were loaded during the evaluation of the performance of each DCNN model using the test dataset.

Sliding window mapping

The mapping data was constructed using a sliding window process. The window size was 256×256 pixels square with 10 pixels steps, inputting the patch-image cropped with the window size into the CycleGAN and DCNN pipeline model and sequentially obtaining the classified results. The window scanned the entire image, and the mapping data were output. To prevent the disappearance of the metastasis-positive region present at the peripheral regions of the image during the later majority processing steps, mirroring of 128 pixels was performed at the top, bottom, left, and right of the original pathological image before performing the mapping process. The size of the mapping data finally obtained by the above means was 259×204 pixels from the original MUSE image (2592×2048) obtained with a $10\times$ objective lens.

Majority voting process

The majority voting process was applied to the image after mapping. Of note, mapping by the deep-learning pipeline is a separate process from the majority voting process. If the mapping data were created with a sliding window of 10 pixels steps, a sufficiently small metastasis-positive area in the fluorescence pathology image was converted to a size of about 26 pixels square because the window size was 256×256 . Based on this size, a 26×26 pixels kernel was defined for majority voting, and a new positive/negative pixel was output when the ratio of the positive/negative pixels included in the kernel exceeded the threshold α (a detailed schematic image is shown in Fig. S2). The α was set to 70%, 80%, 90%, and 95%, respectively.

Computer and software used for calculation

For training all models, a custom-made GPU server with a CPU (Xeon Gold 5218, 2.30 GHz, Intel) and a GPU (Quadro RTX 8000, 48 GB, NVIDIA) was used. For the mapping process, another GPU (A100, 80 GB, NVIDIA) was used. The installed OS was Ubuntu 16.04 LTS. We used TensorFlow version 2.2.0 and Keras version 2.4.3 to build the CycleGAN and DCNN pipeline model.

Results

MUSE images of unfixed lymph nodes from gastric cancer

We obtained MUSE images of unfixed lymph nodes from patients with gastric cancer using a previously reported protocol (described in “Methods”, Fig. S1). The MUSE images of unfixed specimens were obtained in the clinical setting approximately 5 min after the start of biopsy specimen processing without thin sectioning (Fig. S1). Figures 1 and 2 show representative MUSE images of metastasis-negative and metastasis-positive tissues obtained from unfixed lymph nodes, along with MUSE and HE images of FFPE specimens prepared from the same lymph nodes, respectively. The red squares in the images have been enlarged and displayed below each image. In the MUSE images from unfixed lymph nodes, gland-formed cancer tissues and normal lymphoid tissues were clearly visible, as in those in the HE images of FFPE specimens. Cancer tissues were identified by experienced pathologists in the HE images of FFPE specimens shown in Fig. 2C,F. Similarly, cancer glands were also identified in the metastasis-positive MUSE images of unfixed specimens (Fig. 2A,D). When comparing the MUSE images derived from unfixed and FFPE lymph nodes, the former images had higher fluorescence background originating from outside the focal plane due to the thicker specimen, resulting in a different contrast for these images. Additionally, the background intensity was not uniform in different image areas. Such background irregularities are thought to be caused by a complex combination of various factors, including cell density, the intensity of fluorescence of individual cells, and light scattering by cellular tissues. Cellular tissue structures also appeared to have shrunken in FFPE images due to the effects of fixation and dehydration employed during sample preparation. These differences in the appearance of the images make it difficult to diagnose the unfixed specimens using DCNN models trained on FFPE images. The issue of degraded performance of classification models due to differences in the appearance of cellular histology is also known to occur with inter-institutional variability in HE staining¹¹.

Image translation using CycleGAN

We adopted CycleGAN to translate between MUSE images of unfixed and FFPE lymph nodes, which allowed us to diagnose unfixed lymph nodes with the classification DCNN model trained using only FFPE images. CycleGAN trains on bi-directional image modality translations simultaneously. Figure 3A shows the MUSE images translated from the unfixed patch images to FFPE-like patch images using the trained CycleGAN. Original unfixed lymph node images had blurred edges of the nucleolus, cytoplasm, and other cell organelles, as well as a low contrast (Fig. 3B,D). Unwanted non-uniform background fluorescence was also observed. On the contrary, the images translated through CycleGAN had a clear contrast and corrected background compared to the original images (Fig. 3C,E). During image translation using CycleGAN, the images of the sizes shown in Fig. 3 (2592×2048) were too large to be translated all at once; hence, the images were cut into patches, translated sequentially, and then stitched together. The quality of the translation was high enough, such that the joints between the patched images were difficult to distinguish. For the training of CycleGAN, the images in Fig. 3B,D were divided into 256×256 patch-image sizes (see Details of CycleGAN training in “Methods” section).

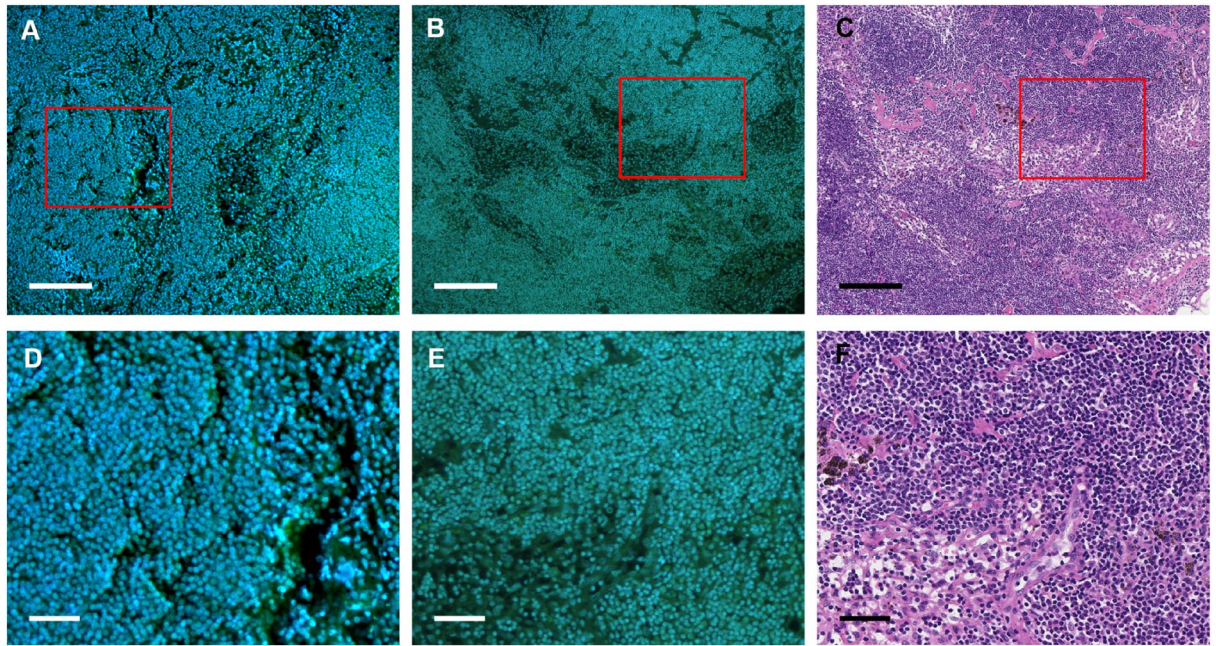


Figure 1. Representative MUSE and HE Images of metastasis-negative lymph nodes. (A,D) MUSE images of unfixed lymph nodes and (B,E) FFPE lymph nodes; (C,F) the corresponding HE images. All images were acquired using the lymph nodes obtained from patients with gastric cancer and were diagnosed as metastasis-negative. Zoomed-in versions of the red squared insets from each image (A–C) are presented below the source images (D–F). Scale bars: 200 μm (A–C) and 100 μm (D–F). *FFPE* formalin-fixed paraffin-embedded, *HE* haematoxylin and eosin staining, *MUSE* microscopy with ultraviolet surface excitation.

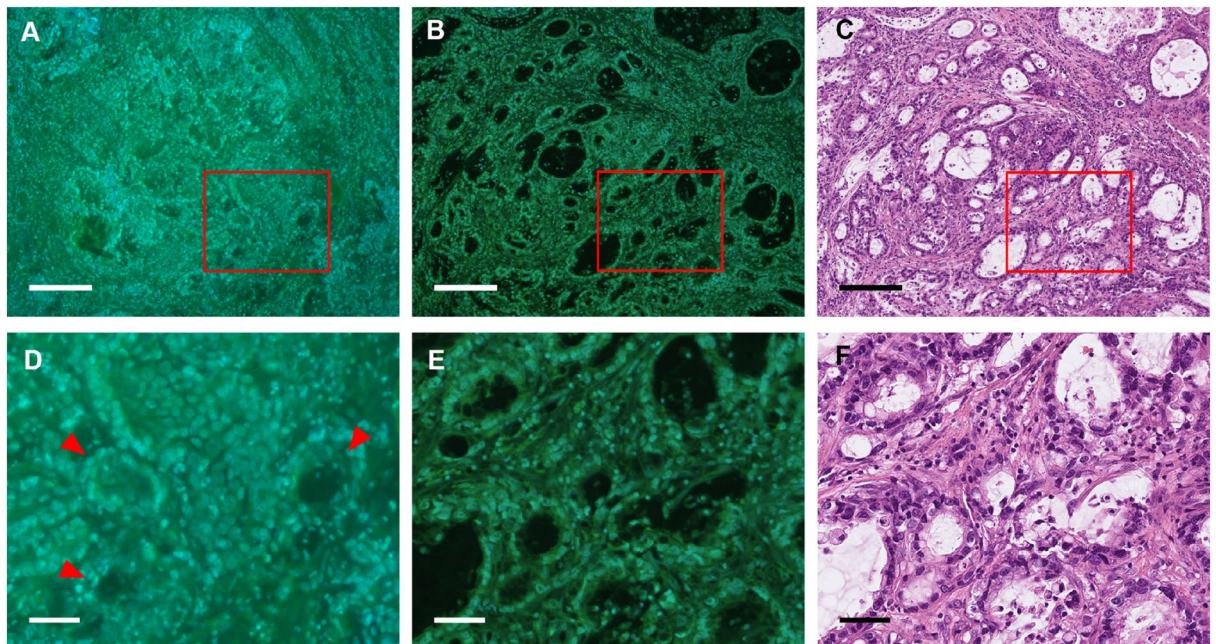


Figure 2. Representative MUSE and HE Images of metastasis-positive lymph nodes. (A,D) MUSE images of the unfixed lymph nodes and (B,E) FFPE lymph nodes; (C,F) the corresponding HE images. All images were acquired using the lymph nodes obtained from patients with gastric cancer and were diagnosed as metastasis-positive. Zoomed-in versions of the red squared insets from each image (A–C) are presented below the source images (D–F). Scale bars: 200 μm (A–C) and 100 μm (D–F). *FFPE* formalin-fixed paraffin-embedded, *HE* haematoxylin and eosin staining, *MUSE* microscopy with ultraviolet surface excitation.

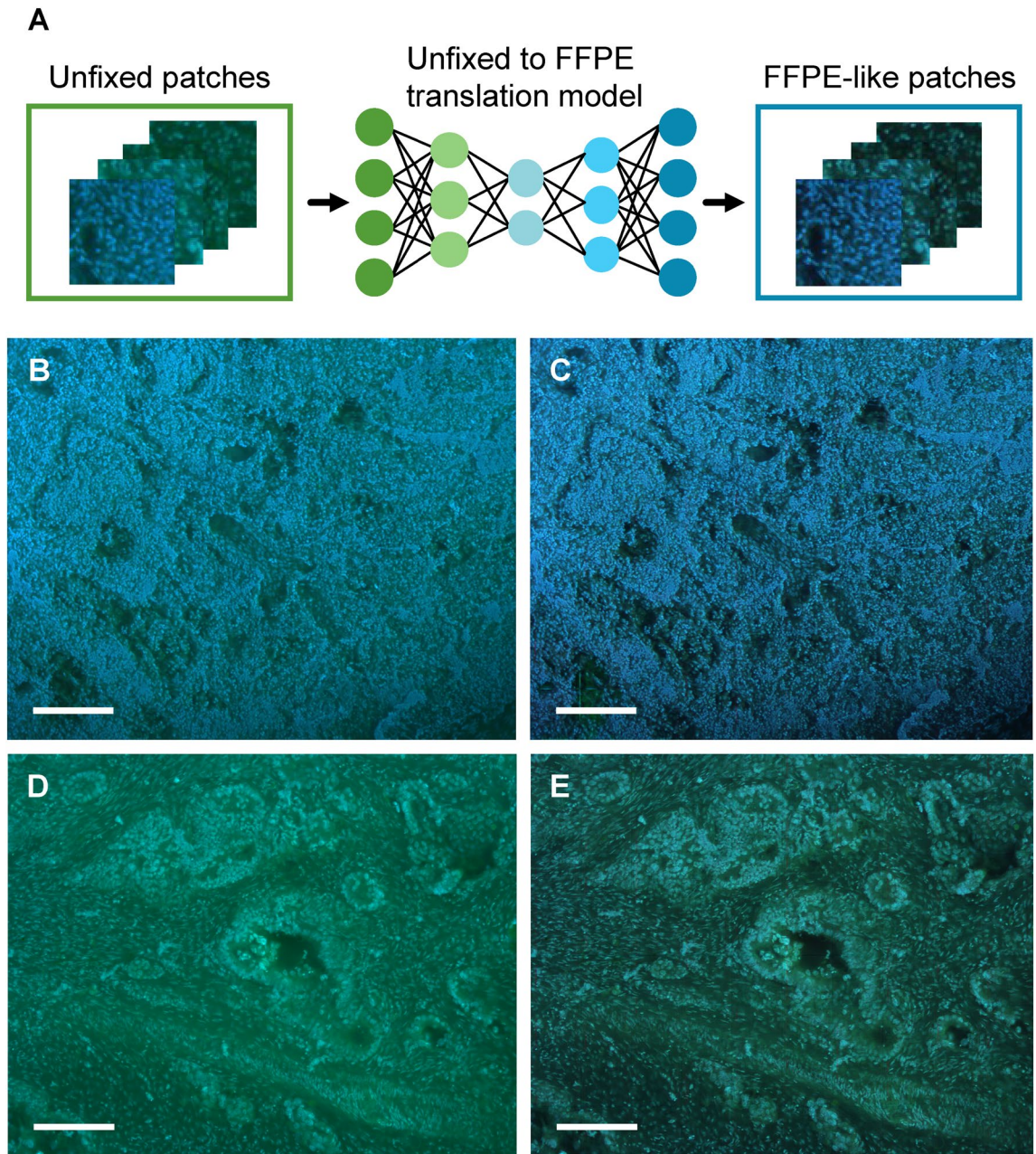


Figure 3. CycleGAN-assisted image translation. (A) Patch images extracted from the whole MUSE image patches are translated into FFPE-like via CycleGAN. (B–E) MUSE images before and after CycleGAN translation. (B) Unfixed metastasis-negative and (D) metastasis-positive images were translated into FFPE-like images (C,E). Scale bars: 200 μm . FFPE formalin-fixed paraffin-embedded, MUSE microscopy with ultraviolet surface excitation.

Training of DCNN models for cancer classification using FFPE MUSE images

We trained three deep learning-based diagnostic prediction models (InceptionV3, EfficientNetB4, and DenseNet-121). The MUSE images of FFPE lymph nodes used for training, which were obtained from our previous study⁸, were labelled by the pathologists using HE images of serial sections. All labelled images used to train the DCNN models were of 256×256 patch size. To account for variations in accuracy due to the initial values of the weights, we trained each model 10 times; the models that yielded the best accuracy when using the test dataset were adopted into the pipeline for subsequent predictions. The accuracies of the three finally selected models based on the test dataset were 98.45%, 97.28%, and 97.78%, respectively, with InceptionV3 showing the best accuracy. The receiver operating characteristic curve (AUC) scores were 0.9985, 0.9982, and 0.9969, respectively, and the F1 scores were 98.48%, 97.24%, and 98.15%, respectively, as shown in Fig. 4B,C. All three models could classify lymph node metastatic cancers with very high accuracy. The threshold was set to a value that maximised the F1 score obtained using the validation dataset.

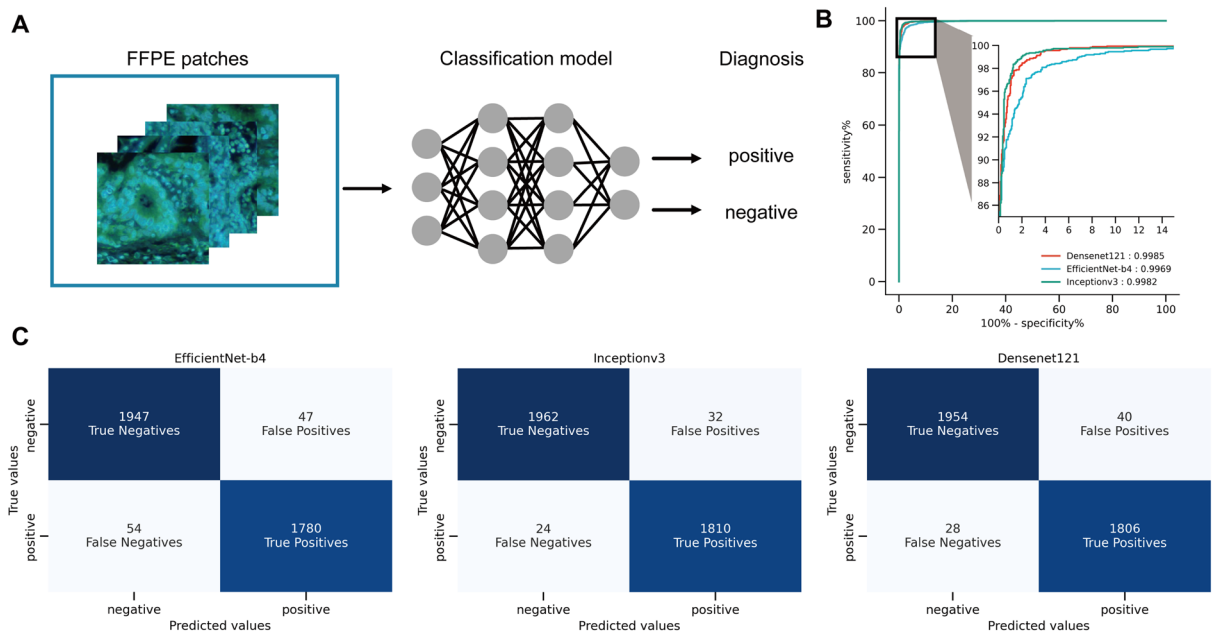


Figure 4. Performance of DCNN Classification Models on MUSE Image patches. **(A)** Conceptual diagram depicting DCNN model for image classification. The MUSE image patches of FFPE lymph nodes are the input, and metastasis-positive or metastasis-negative is obtained as the output for each image input. **(B)** The receiver operating characteristic curves of the three trained DCNN models (InceptionV3, EfficientNetB4, and DenseNet-121). The black square area has been enlarged and displayed as an inset. The areas under the curves are shown in the lower right corner of the image. **(C)** The confusion matrix of the three DCNN models on the test dataset. *DCNN* deep convolutional neural network, *MUSE* microscopy with ultraviolet surface excitation, *FFPE* formalin-fixed paraffin-embedded.

Large-scale image classification and metastasis area mapping through the CycleGAN-DCNN pipeline

As shown in Figs. 3 and 4, the image translation quality of CycleGAN was high, and the DCNN models showed high diagnostic accuracy. CycleGAN and DCNN models were then connected to create a pipeline that can predict samples being metastasis-positive or metastasis-negative based on patch images of unfixed lymph nodes. Mapping was performed using sliding window processing, and the pipeline was used to visualise metastasis-positive areas on large-scale images (Fig. 5A). We employed InceptionV3, which showed the highest prediction accuracy among the three assessed DCNNs. During the mapping process, cancerous areas were densely filled with metastasis-positive predictions, while the other areas showed a binary noise-like pattern due to sparse false positive predictions. As DCNN prediction accuracy is not 100%, such false positive predictions are thought to be inevitable. The noise pattern was removed using 95% majority voting processing for the mapping results obtained using the sliding window processing (see “Methods” section, Fig. S2). Examples of the results of each image processing are shown in Fig. 5B–E. While metastasis-positive areas could not be identified in the image before translation, the CycleGAN-assisted preprocessing in the pipeline enabled the detection of the metastasis-positive regions. The number of immunologically activated histiocytes increases in the metastatic regional lymph nodes, and they sometimes form sarcoidosis-like granulomas. These histiocytes were easily misidentified as metastasis-positive because of their large cytoplasm and low nuclear density compared to lymphocytes⁸. Large blood vessels were also misrecognised as metastasis-positive because they were similar to cancerous glandular tissue (Fig. S2B). These false positives were corrected by the 95% majority voting processing (Fig. S2D). Before majority voting processing, normal tissues surrounding cancerous tissues were widely misidentified as metastasis-positive (Fig. 5D); however, majority voting processing made it possible to recognise only the cancerous tissues (Fig. 5E).

The performance of diagnostic predictions was evaluated for the MUSE images of unfixed lymph nodes observed using a 10× objective lens. For each image and mapping results, the following rules were used for the diagnosis: even if 1 pixel of the metastasis-negative image was predicted to be positive, it was a wrong prediction, and even if 1 pixel of the metastasis-positive image was predicted to be positive, it was a correct prediction. The diagnostic prediction results for the CycleGAN-translated images, along with those obtained without the image translation, are shown in Table 1. A total of 91 images were diagnosed; 45 images from 19 cases were classified as metastasis-positive images, and 46 images from 9 cases as metastasis-negative images. The average predictive accuracy of the three DCNN models was 84.6% and 66.3% for images with and without CycleGAN translation, respectively. Additionally, all models showed a considerable increase in diagnostic accuracy when the translation with CycleGAN was introduced. AUC and F1 scores are also provided in the Supplementary Fig. S4 and Supplementary Table 1. Furthermore, see Supplementary Table 2 for results obtained when the threshold (%) for the majority voting processing was varied. The highest accuracies were achieved at all the threshold values when

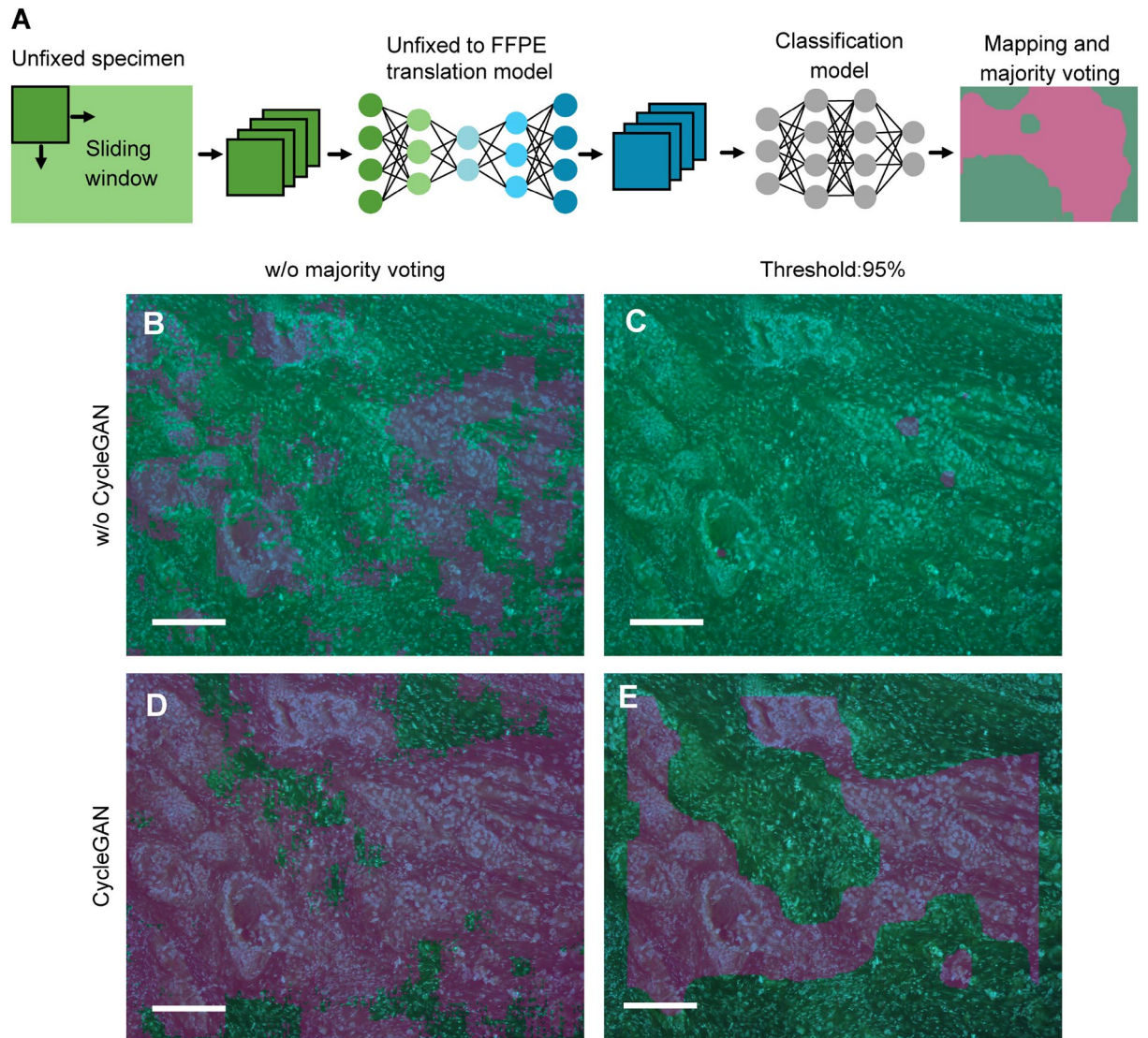


Figure 5. Large-scale metastasis area mapping using CycleGAN-DCNN pipeline. (A) Divided patches using a sliding window approach are used as the input into the trained CycleGAN, and patch-level classification is performed through the DCNN model sequentially. The majority voting processing is used as postprocessing. (B–E) A metastasis-positive image and its mapping results. The top (B,C) and bottom (D,E) images are original and translated images by CycleGAN, respectively. The CycleGAN-assisted preprocessing changed the contrast of the MUSE images and improved the mapping results. (B,D) Show images without majority voting processing, and (C,E) show those with majority voting processing with a 95% threshold. Scale bars: 200 μm .

α (%)			Metastasis (-)	Metastasis (+)	Average (%)
95	Without CycleGAN	InceptionV3	38/46	23/45	67.0
		DenseNet-121	17/46	40/45	62.6
		EfficientNetB4	19/46	44/45	69.2
		Average (%)	53.6	79.3	66.3
	With CycleGAN	InceptionV3	39/46	40/45	86.8
		DenseNet-121	33/46	41/45	81.3
		EfficientNetB4	42/46	36/45	85.7
		Average (%)	82.6	86.7	84.6

Table 1. The result of large-scale image diagnostic accuracies for each of the three classification DCNN models in original unfixed specimens and upon translation by CycleGAN. The threshold of majority voting processing was set to 95%.

employing CycleGAN for preprocessing. The computational time of the three phases—CycleGAN translation, DCNN mapping, and majority voting—is shown in Supplementary Table 3.

We visualised which area of the image before and after CycleGAN processing contributed to the prediction of the DCNN model (InceptionV3) using Grad-CAM²⁶ (Fig. S3). The red areas contributed the most to the prediction, indicating that the image after CycleGAN processing could predict the diagnosis based on cancer characteristics, while those before CycleGAN processing did not recognise cancerous regions.

Discussion

The pipeline involving the image classification model (DCNN) and image translation model (CycleGAN) proposed in this study improved the average diagnostic accuracy from 66.3 to 84.6% for MUSE images acquired using a 10× objective lens. Further improvement in accuracy can be achieved by increasing the number of patients and MUSE image data. Enhancing the quality of the image data also contributes to improved accuracy. The staining and imaging processes are performed by human operators, which can lead to variations in the data set. Mechanising and automating these processes would reduce these variations. Achieving 95% accuracy is a benchmark for operations in clinical practice. Although our study showed improved accuracy through image-translations, it's still not sufficient for clinical use. To improve accuracy, we plan to expand our dataset and comprehensively learn various types of cancer. We hope that this will lead to more stable and accurate translations, ultimately achieving the clinical baseline accuracy.

MUSE does not require thin-sliced sections compared to conventional pathological diagnostic techniques, including HE staining protocol, and the staining process is simple and rapid. Regarding the time required for MUSE image acquisition and mapping using the pipeline, we expect that improvements in staining protocols, imaging systems, mapping algorithms, and computational speed will enable us to reduce the required time. Although the microscope used in this paper allows observation of only an area of approximately $1.2 \times 1.0 \text{ mm}^2$ at a time, there is room for more extensive imaging with improved optics in the future. It is also possible to reduce the time required for staining and pipeline model calculations.

The combination of MUSE and DCNN offers the following advantages. First, the combined method can reduce the burden on pathologists. As the incidence of cancer increases, the number of diagnoses per pathologist would also increase²⁷. Highly accurate rapid intraoperative diagnosis systems combining rapid imaging technology and automatic AI-assisted diagnosis will increase their importance in pathology. Second, variations in diagnostic performance among pathologists can also be minimised. Third, our method can provide expert-level diagnosis where pathology resources are scarce.

A sufficient number of images is required for the training of classification models, such as DCNN; however, it is difficult to collect a large number of MUSE images of unfixed specimens and to cover all pathological findings at a single facility. Notably, the method proposed in our study does not require a large number of images of unfixed specimens for training the DCNNs. Our method only needs a relatively small number of images of unfixed specimens for CycleGAN training and a large amount of image data can be generated from previous FFPE samples stored at the facility for classification model training. By using FFPE samples to prepare large amounts of image datasets, including those for rare diseases, it would become possible to create image classification models that can predict diagnosis with high accuracy for a variety of diseases. Without the CycleGAN image translation process, we will need to devote a significant effort to creating clinical image datasets of rare diseases from unfixed specimens for accurate diagnostic prediction.

Our results show that CycleGAN-based image translation improved diagnostic performance even for different specimens prepared using different processes. CycleGAN has the potential to handle not only the unfixed to FFPE MUSE image translation but also other intraoperative diagnostic imaging modalities, such as two-photon excited autofluorescence (TPEF), second harmonic generation (SHG), third harmonic generation (THG), and stimulated Raman scattering (SRS) microscopy, and would reduce the difficulties associated with clinical image data acquisition. These nonlinear optical microscopies do not require thin sectioning and staining with conventional methods. Additionally, deep learning-based image classification models enable the diagnosis of images obtained by a new imaging method that general pathologists cannot diagnose^{28–31}. Among these new imaging modalities, MUSE imaging offers a cost-effective, accessible alternative to complex nonlinear optical methods. It uses a low-cost LED light source and is as easy to use as standard fluorescence microscopy, making it practical and suitable for clinical applications.

A chief limitation of our study is that we only evaluated lymph node metastasis in gastric cancer. The robustness of the model should be assessed through external validation. To address this, the dataset will be expanded to include other cancer samples obtained from several facilities in the future. A secondary limitation is that the edge of the images is not mapped after the majority voting process. This problem will be solved in the future by replacing the image classification task using DCNN and the majority voting process with segmentation models. As the entire dataset was used to train CycleGAN in this study, further research is warranted to validate the generalisation of its performance for clinical settings.

Conclusion

MUSE microscopy was applied for pathological diagnosis of lymph node metastasis in gastric cancer. The modality translation using CycleGAN improved the pathological diagnosis of non-thin-sliced surface images using DCNN model trained with FFPE thin-sliced images. Our method enables highly accurate diagnosis even with a small number of images of unfixed samples when combined with existing images of FFPE samples and can be potentially applied to rapid intraoperative diagnostic images acquired using various microscopes.

Data availability

The datasets generated and analysed during the current study are not publicly available due to privacy concerns of the patients, but are available from the corresponding authors on reasonable request.

Received: 28 September 2023; Accepted: 24 November 2023

Published online: 04 December 2023

References

- Sung, H. *et al.* Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA Cancer J. Clin.* **71**, 209–249 (2021).
- Novis, D. A., Zarbo, R. J. & Saladino, A. J. Interinstitutional comparison of surgical biopsy diagnosis turnaround time: A College of American Pathologists Q-Probes study of 5384 surgical biopsies in 157 small hospitals. *Arch. Pathol. Lab. Med.* **122**, 951–956 (1998).
- Tamaki, Y. *et al.* Molecular detection of lymph node metastases in breast cancer patients: Results of a multicenter trial using the one-step nucleic acid amplification assay. *Clin. Cancer Res.* **15**, 2879–2884 (2009).
- Koizumi, N. *et al.* Highly sensitive fluorescence detection of metastatic lymph nodes of gastric cancer with photo-oxidation of protoporphyrin IX. *Eur. J. Surg. Oncol.* **42**, 1236–1246 (2016).
- Matsumoto, T. *et al.* 5-ALA-assisted automated detection of lymph node metastasis in gastric cancer patients. *Gastric Cancer* **23**, 725–733 (2020).
- Morita, M. *et al.* Fluorescence-based discrimination of breast cancer cells by direct exposure to 5-aminolevulinic acid. *Cancer Med.* **8**, 5524–5533 (2019).
- Kumamoto, Y., Matsumoto, T., Tanaka, H. & Takamatsu, T. Terbium ion as RNA tag for slide-free pathology with deep-ultraviolet excitation fluorescence. *Sci. Rep.* **9**, 10745 (2019).
- Matsumoto, T. *et al.* Deep-UV excitation fluorescence microscopy for detection of lymph node metastasis using deep neural network. *Sci. Rep.* **9**, 16912 (2019).
- Zhu, J.-Y., Park, T., Isola, P. & Efros, A. A. Unpaired image-to-image translation using cycle-consistent adversarial networks. <http://arXiv.org/1703.10593> (2017).
- de Bel, T., Bokhorst, J.-M., van der Laak, J. & Litjens, G. Residual cycleGAN for robust domain transformation of histopathological tissue slides. *Med. Image Anal.* **70**, 102004 (2021).
- Swiderska-Chadaj, Z. *et al.* Impact of rescanning and normalization on convolutional neural network performance in multi-center, whole-slide classification of prostate cancer. *Sci. Rep.* **10**, 14398 (2020).
- Runz, M. *et al.* Normalization of HE-stained histological images using cycle consistent generative adversarial networks. *Diagn. Pathol.* **16**, 71 (2021).
- Xu, Z. *et al.* Effective immunohistochemistry pathology microscopy image generation using CycleGAN. *Front. Mol. Biosci.* **7**, 571180 (2020).
- Murali, L. K., Lutnick, B., Ginley, B., Tomaszewski, J. E. & Sarder, P. Generative modeling for renal microanatomy. *Proc. SPIE Int. Soc. Opt. Eng.* **11320**, 99 (2020).
- de Haan, K. *et al.* Deep learning-based transformation of H&E stained tissues into special stains. *Nat. Commun.* **12**, 4884 (2021).
- Abraham, T., Shaw, A., O'Connor, D., Todd, A. & Levenson, R. Slide-free MUSE microscopy to H&E histology modality conversion via unpaired image-to-image translation GAN models. <http://arXiv.org/2008.08579> (2020).
- He, J. *et al.* CycleGAN with an improved loss function for cell detection using partly labeled images. *IEEE J. Biomed. Health Inform.* **24**, 2473–2480 (2020).
- Goodfellow, I. J. *et al.* Generative adversarial networks. <http://arXiv.org/1406.2661> (2014).
- Taigman, Y., Polyak, A. & Wolf, L. Unsupervised cross-domain image generation. <http://arXiv.org/1611.02200> (2016).
- Szegedy, C., Vanhoucke, V., Ioffe, S., Shlens, J. & Wojna, Z. Rethinking the inception architecture for computer vision. In *2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR): IEEE* 2818–2826 (2016).
- Huang, G., Liu, Z., van der Maaten, L. & Weinberger, K. Q. Densely connected convolutional networks. <http://arXiv.org/1608.06993> (2016).
- Tan, M. & Le, Q. V. EfficientNet: Rethinking model scaling for convolutional neural networks. <http://arXiv.org/1905.11946> (2019).
- Kallipolitis, A., Revelos, K. & Maglogiannis, I. Ensembling EfficientNets for the classification and interpretation of histopathology images. *Algorithms* **14**, 278 (2021).
- Riasatian, A. *et al.* Fine-tuning and training of densenet for histopathology image representation using TCGA diagnostic slides. *Med. Image Anal.* **70**, 102032 (2021).
- Deng, J. *et al.* ImageNet: A large-scale hierarchical image database. In *2009 IEEE Conference on Computer Vision and Pattern Recognition* 248–255 (2009).
- Selvaraju, R. R. *et al.* Grad-CAM: Visual explanations from deep networks via gradient-based localization. <http://arXiv.org/1610.02391> (2016).
- Metter, D. M., Colgan, T. J., Leung, S. T., Timmons, C. F. & Park, J. Y. Trends in the US and Canadian pathologist workforces from 2007 to 2017. *JAMA Netw. Open* **2**, e194337 (2019).
- You, S. *et al.* Real-time intraoperative diagnosis by deep neural network driven multiphoton virtual histology. *NPI Precis. Oncol.* **3**, 33 (2019).
- Matsui, T. *et al.* Nonlinear optics with near-infrared excitation enable real-time quantitative diagnosis of human cervical cancers. *Cancer Res.* **80**, 3745–3754 (2020).
- Hollon, T. C. *et al.* Near real-time intraoperative brain tumor diagnosis using stimulated Raman histology and deep neural networks. *Nat. Med.* **26**, 52–58 (2020).
- Zhang, L. *et al.* Rapid histology of laryngeal squamous cell carcinoma with deep-learning based stimulated Raman scattering microscopy. *Theranostics* **9**, 2541–2554 (2019).

Author contributions

J.S., T.M., R.N., H.N. and T.T. conceived the project; T.M. and R.N. performed the experiments of fluorescence imaging; J.S., H.N. and H.N. performed deep-learning analysis; T.M. and R.N. contributed to human lymph node collection and preparation; R.N. contributed to histopathology review of human lymph nodes; J.S., T.M., R.N., H.N., H.T. and T.T. discussed the results; J.S., R.N., H.N. and T.T. wrote the manuscript; All the authors approved the final version.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-48319-7>.

Correspondence and requests for materials should be addressed to H.N. or T.T.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023