Journal of Pathology Informatics 14 (2023) 100175

Contents lists available at ScienceDirect

Journal of Pathology Informatics

journal homepage: www.elsevier.com/locate/jpi

Tumor-identification method for predicting recurrence of early-stage lung adenocarcinoma using digital pathology images by machine learning

Hideharu Hattori [a,b,*], Shingo Sakashita [c], Masahiro Tsuboi [d], Genichiro Ishii [e], Toshiyuki Tanaka [a]

[a] Graduate School of Science and Technology, Keio University, 3-14-1 Hiyoshi, Kohoku-ku, Yokohama-shi, Kanagawa 223-8522, Japan
[b] Research & Development Group, Hitachi, Ltd., 1-280 Higashi-koigakubo, Kokubunji-shi, Tokyo 185-8601, Japan
[c] Division of Pathology, National Cancer Center, Exploratory Oncology Research & Clinical Trial Center, 6-5-1 Kashiwanoha, Kashiwa-shi, Chiba 277-8577, Japan
[d] Department of Thoracic Surgery, National Cancer Center Hospital East, 6-5-1 Kashiwanoha, Kashiwa-shi, Chiba 277-8577, Japan
[e] Department of Pathology and Clinical Laboratories, National Cancer Center Hospital East, 6-5-1 Kashiwanoha, Kashiwa-shi, Chiba 277-8577, Japan

ARTICLE INFO

Keywords:
Pathological image
Convolutional neural network
Tumor tissue
Lung adenocarcinoma
H&E staining

ABSTRACT

Lung cancer is one of the cancers with the highest morbidity and mortality in the world. Recurrence often occurs even after complete resection of early-stage lung cancer, and prediction of recurrence after resection is clinically important. However, the pathological characteristics of the recurrence of pathological stage IB lung adenocarcinoma (LAIB) have not yet been elucidated. Therefore, the problem is what type of histological image of lung adenocarcinoma recurs, and it is important to examine the histological image of recurrence. We attempted to predict recurrence of early lung adenocarcinoma after resection on the basis of digital pathological images of hematoxylin and eosin-stained specimens and machine learning applying a convolutional neural network. We constructed a model that extracts the features of two-color spaces and a switching model that automatically switches between our extraction model and one that extracts the features of one-color space for each image. We then developed a tumor-identification method for predicting the presence or absence of LAIB recurrence using these models. We conducted an experiment involving 55 patients with LAIB who underwent surgical resection to evaluate the proposed method. The proposed method determined LAIB recurrence with an accuracy of 84.8%. The use of digital pathology and machine learning can be used for highly accurate prediction of LAIB recurrence after surgical resection. The proposed method has the potential for objective postoperative follow-up observation.

Introduction

Lung cancer is one of the cancers with the highest morbidity and mortality in the world. Recurrence often occurs even after complete resection of early-stage lung cancer, which has become a major clinical problem. Therefore, the problem is what type of histological image of lung adenocarcinoma recurs, and it is important to examine the histological image of recurrence. In previous reports, pleural and vascular invasion have been reported to be prognostic factors. However, they require elastic fiber staining, and it is difficult to predict prognosis with hematoxylin and eosin (H&E) staining alone.[1] It is necessary to confirm the presence or absence of recurrence by follow-up. Among the lung adenocarcinomas, the postoperative recurrence rate of pathological stage IB lung adenocarcinoma (LAIB) is 37.8%.[2] In the tumor-node-metastasis (TNM) classification,[3] LAIB indicates lung adenocarcinoma with a maximum diameter of more than 3 cm and less than 4 cm in the area showing invasive growth without metastasis to the regional lymph nodes or distant metastasis. If this can be predicted, it will lead to an appropriate treatment policy for the patient depending on the presence or absence of recurrence, improving the quality of life of the patient.

Several methods for identifying lung-cancer tumors have been investigated. To predict the recurrence of disease in early-stage non-small cell lung cancer (NSCLC), a method is being investigated to create classifiers using nuclear orientation, texture, shape, and tumor structure.[4] Because this method is intended for tissue microarrays (TMAs) that collect only morphological images of only a portion of the tumor, the challenge with this method is to capture the heterogeneity of tumor presence or absence in NSCLS images for entire slides, such as a whole slide image (WSI), and to assess whether recurrence is predictable. Image features related to cells and cell nuclei are also extracted, and higher-level features are selected using a support vector machine and random-forest classifiers. By using these features, a method for distinguishing long- and short-term survivors of pathological stage I adenocarcinoma is being investigated.[1] This method uses TMAs consisting of only the most representative cases and does not allow for the evaluation of acinar- or papillary-subtype adenocarcinoma.

Abbreviations: LAIB, pathological stage IB lung adenocarcinoma.

* Corresponding author.
E-mail address: hideharu.hattori@keio.jp (H. Hattori).

http://dx.doi.org/10.1016/j.jpi.2022.100175
Received 10 July 2022; Received in revised form 28 November 2022; Accepted 15 December 2022
Available online 23 December 2022

Histological predominant subtypes are known to be lepidic, acinar, papillary, and solid. Among them, the acinar subtype forms mainly glandular structures with a central luminal space surrounded by tumor cells. Tumor cells of the papillary subtype characteristically grow in a papillary fashion. A method of identifying the tumor region of lung cancer using a pre-trained Inception v3[5] model and discovering tumor shape and boundary features that predict survival outcomes is also being investigated.[6] Because this method finds the features that predict the survival outcome only from the RGB features of the H&E-stained image, it cannot find the features for predicting the survival outcome of lung-cancer that cannot be determined from only the RGB features of the H&E-stained image.

We propose a tumor-identification method having the following features.

- The proposed method is capable of extracting features related to the presence or absence of LAIB recurrence.
- By switching between the two models for each image, the proposed method can accurately identify the presence or absence of LAIB recurrence in several predominant subtypes.
- The proposed method determined LAIB recurrence with an accuracy of 84.8% for 55 patients with LAIB who underwent surgical resection.

Materials and methods

Sample information

We used WSIs of specimens from 55 patients (33 with non-recurrent LAIB and 22 with recurrent LAIB). We used 22 of these specimens for training and validation and 33 for evaluation. We also used specimens with a recurrence period of LAIB of about 5 months to about 5 years and 2 months. Of the 55 cases of the histological predominant subtypes, lepidic, acinar, papillary, and solid subtypes were 7, 15, 19, and 14 cases, respectively. From the WSIs of the 55 specimens, we obtained 19 986 partial images for training, 4000 for verification, and 7106 for evaluation. The WSIs were taken using the virtual slide scanner NanoZoomer C9600-02 manufactured by Hamamatsu Photonics. We used images (0.23 μm/pixel) at 40x magnification of the WSIs. Hematoxylin (hematoxylin 3G, Sakura Finetech Japan) and eosin (Eosin, Sakura Finetech Japan) were used for staining of tissue sections. The data for training, verification, and evaluation were the data of the tissue sections stained in one facility. We created training/verification/evaluation data according to the procedure of "Experimental-data creation".

Luminance-image creation

Due to the characteristics of H&E staining, the R image tends to have the highest value and the G image the lowest value among the R, G, and B images. Also, the luminance (L) of an HLS color space image (where, H is hue and S is saturation) is defined as the middle between the largest and smallest values of R, G, and B. The glandular structures or luminal structures are composed of white areas with high luminance on the H&E-stained image. Since the white areas of the glandular structures or luminal structures have large R and G values, and the lung cancer cells forming the glandular structures or luminal structures have large R values and small G values, the L image will emphasize the lung cancer cells forming glandular structures or luminal structures, and the L image will help capture the features of the areas within the glandular structures or luminal structures. When an L*L*L* image is created by repeating three luminance values (L*) of CIE Lab (L*a*b* color space, also referred to as Lab), Diaz[7] reported that L* L* L* images have better identification accuracy than RGB images for a certain model constructed using convolutional neural networks (CNNs).[8] Therefore, capturing features in an image from an image created by repeating a single channel has the potential to improve the model's identification accuracy. Therefore, LLL images were created to capture features in the glandular structures or luminal structures. In other words, an LLL image consists of three channels of L.

Experimental-data creation

We extract features other than RGB from H&E-stained images to improve the accuracy of identifying the presence or absence of recurrence. The method of creating image data for this purpose is described as follows. To extract a partial image containing a large number of cell nuclei, the threshold for dividing the histogram of each partial image in a WSI into three types of brightness (cell nucleus, cytoplasm, and background) is obtained, and the partial image containing the cell nucleus is extracted. If cell nuclei are compared with the cytoplasm or background, they tend to be distributed in dark areas. Therefore, as shown in Fig. 1, WSIs of a pathological image of LAIB are obtained from the tissue section prepared as a continuous section created in the process of normal pathological diagnosis using a virtual slide scanner manufactured by Hamamatsu Photonics. The histogram of a WSI is divided into dark part P1 and bright part P2 using Otsu's binarization[9] for the R component for each WSI. Otsu binarization is applied for P1, and the threshold R_{th} that divides P1 into dark part P11 and bright part P12 is obtained. The number of pixels $R_{\mathrm{th}}N$ exceeding R_{th} is then obtained for each partial H&E-stained image (800 × 800 pixels). Next, to create a partial image including the cell nucleus, if the pixel ratio R_{ratio} obtained by Eq. (1) satisfies the threshold RD_{th} (e.g., 0.7 or less) in a partial H&E-stained image (800 × 800 pixels), RGB partial images are automatically extracted from the WSI. The color space of the H&E-stained image is then converted from RGB to HLS, and the L components of HLS are copied to three channels to create an LLL image. Here, R_{all} in Eq. (1) represents the total number of pixels in the partial image.

$$R_{\mathrm{ratio}} = R_{\mathrm{th}}N / R_{\mathrm{all}} \tag{1}$$

By setting RD_{th} to 0.7 or less, it is possible to suppress the extraction of a partial image, in which most of the image has a white background, and extract a partial image in which most of the image contains a part of the lung adenocarcinoma cells containing the cell nucleus. These partial images are used for training/verification/evaluation data.

The recurrence status of LAIB patients was investigated for about six years, and each image was labeled as having no LAIB recurrence or LAIB recurrence.

Tumor-recurrence-prediction models

It is difficult to capture the features of LAIB recurrence, such as in the glandular structures or luminal structures, from the RGB features alone. The glandular structures or luminal structures are composed of the tissue

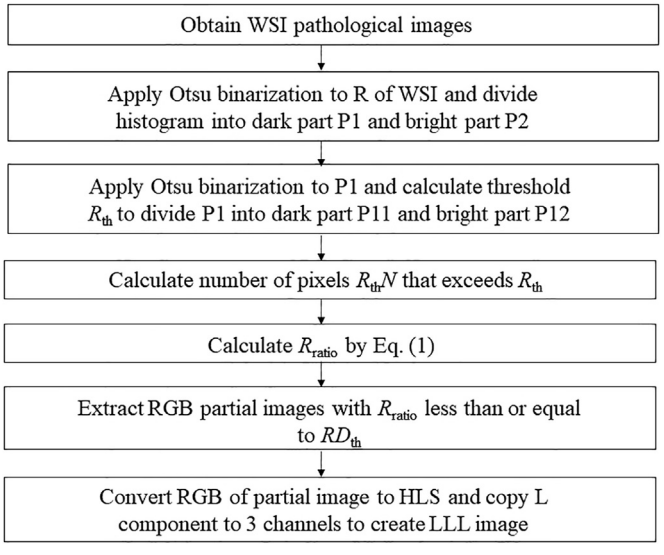

Fig. 1. Extracting partial image from WSI.

containing a white region with high brightness on the H&E-stained image. Hence, to extract the features of the white region with high L, we focused on the L information. We used the features of an L image to capture the features related to LAIB recurrence in the glandular structures or luminal structures and combined the RGB and L features of LAIB. We then constructed the discriminator AB (DAB) model consisting of feature extractors A and B using CNNs. Feature extractor A automatically extracts the RGB features of H&E-stained images using CNNs. Feature extractor B automatically extracts the features of the LLL image using the L component of the HLS of H&E-stained images using CNNs.

Depending on the histological subtypes, i.e., acinar, lepidic, papillary, and solid, lung adenocarcinoma may or may not contain glandular structures or luminal structures. The solid subtype has invasive solid tumor nests without glandular space. Therefore, we constructed the switching discriminator (SD) model that switches between the DAB and discriminator A (DA) models for each image for accurately predicting the presence or absence of LAIB recurrence. The DA model consists of feature extractor A. The proposed tumor-identification method identifies the presence or absence of LAIB recurrence in the pathological images obtained from a virtual slide scanner using the DAB and SD models.

We first explain the DAB model then the DA model. Finally, we explain the SD model and the convolution, pooling, and fully connected (FC) layers of each model.

DAB model

Fig. 2 shows the architecture of the DAB model. As mentioned above, to accurately predict the presence or absence of LAIB recurrence, by using not only the features of the RGB image from the H&E-stained image but also the features other than the features of the RGB image, it is necessary to identify the presence or absence of LAIB recurrence. Therefore, the RGB of an H&E-stained image is first color-space-converted to HLS, and an LLL image consisting of the L components of HLS is created.

While inputting the RGB and LLL images, we construct the DAB model including feature extractors A and B using CNNs, as shown in Fig. 2. Fig. 3 shows the architecture of the building block shown in Fig. 2. We adopt batch normalization,[10] a rectified linear unit (ReLU),[11] and global average pooling[12] for both feature extractors A and B. The architecture in Fig. 3 is similar to the residual block and shortcut connection in Resnet.[13] By creating both feature extractors A and B containing the building block shown in Fig. 3, we suppress the disappearance of the gradient when training a model that identifies the presence or absence of LAIB recurrence.

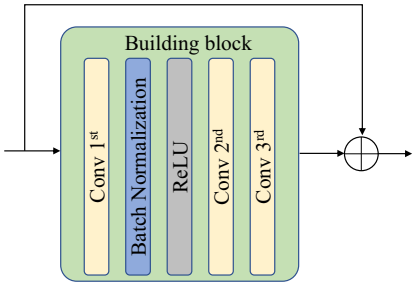

Fig. 3. Building block architecture.

The DAB model also includes one FC layer. Feature extractors A and B each consist of three convolutional (Conv) layers, one max pool layer, and five building block layers (one building block layer consists of three convolution layers, i.e., Conv 1st, Conv 2nd, Conv 3rd, batch normalization, and ReLU). Feature extractor A convolves the image in the Conv1 layer to extract features that represent the partial shape of non-recurrence tumors or recurrence tumors in H&E-stained RGB images. In the max pool layer, the features of the presence or absence of recurrence are extracted to absorb small displacements in the H&E-stained image. After the building block 0 layer in Fig. 2, the image is convolved to extract features that express a larger partial shape of non-recurrence or recurrence tumors in the H&E-stained image compared with the previous layer.

Feature extractor B then extracts the features of the partial shapes of different sizes for non-recurrence or tumors that will recur, which cannot be captured only from RGB images, from H&E-stained LLL images and the features of the presence or absence of recurrence to absorb the displacement in the LLL image. In the final FC layer, the results of the FC layers of feature extractors A and B are combined and aligned into a one-dimensional array (see Fig. 8 for details). The DAB model, which consists of the above architecture, identifies whether LAIB will or will not recur. The weights and biases of each Conv layer, FC layer, and final FC layer are obtained by machine learning using multiple H&E-stained pathological images.

DA model

The DA model consists only of Feature extractor A, as shown in Fig. 4.

SD model

As shown in Fig. 5, the SD model switches between DA and DAB models in accordance with the image. We constructed the DA and DAB models by

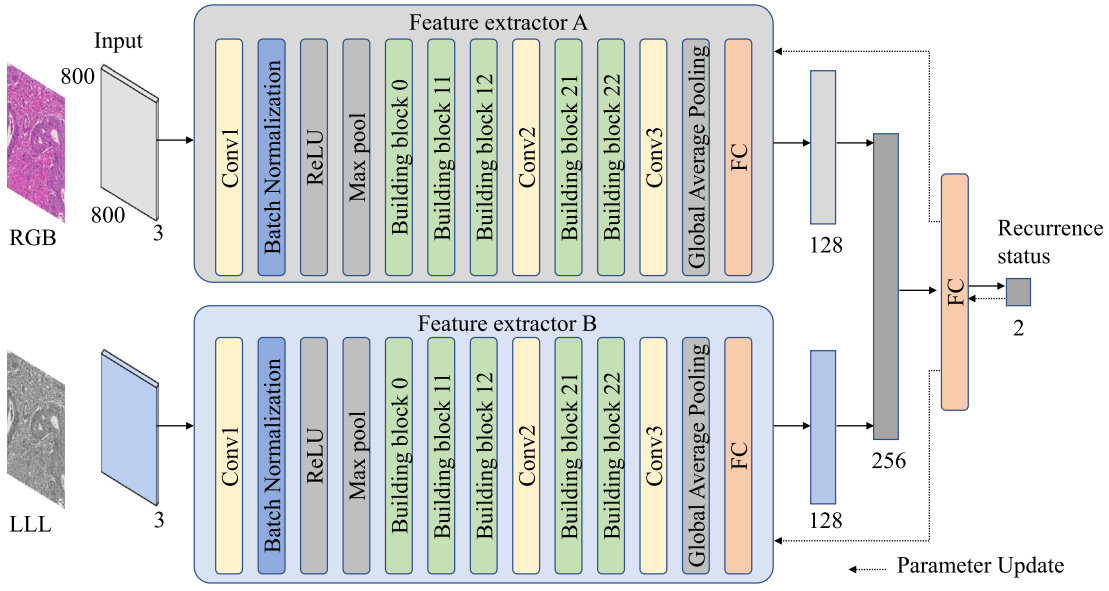

Fig. 2. Network architecture of DAB model consisting of feature extractors A and B: input of model is RGB and LLL image pair of size 800 × 800 × 3.

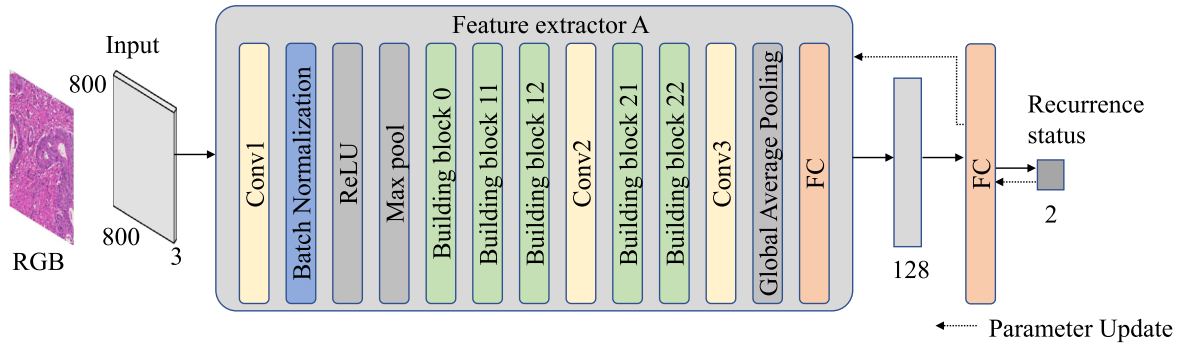

Fig. 4. DA model consisting of feature extractor A.

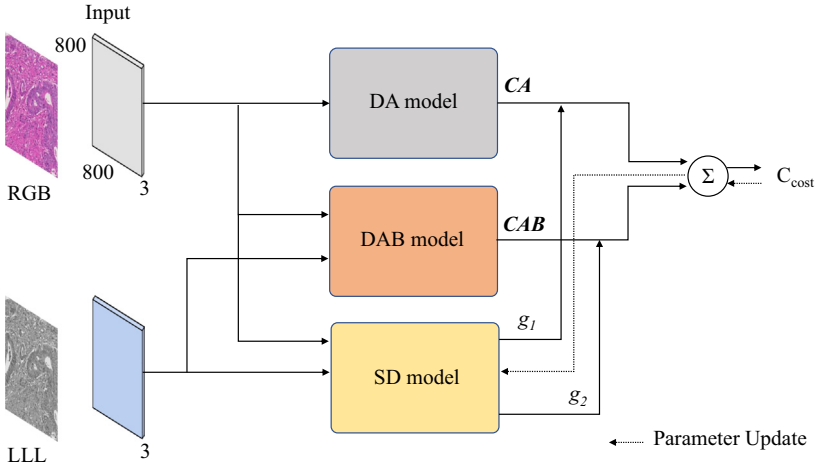

Fig. 5. Architecture to switching between DA and DAB models.

machine learning in advance. The RGB and LLL images are input to the SD and DAB models. However, only RGB images are input to the DA model. The identification results of both DA and DAB models are obtained. The SD model was constructed so that when training it, the cost C_{cost} decreases by using Eq. (2). On the basis of the obtained C_{cost}, the parameter of the SD model is updated using Adam.[14] In Eq. (2), where ***CA*** and ***CAB*** represent the output of the DA and DAB models, respectively, *nll* represents the negative log likelihood and ***label*** represents teaching data. The coefficients g_k represent the coefficients that select each model, and $\Sigma_{k=1}^{2} g_k = 1$.

$$\boldsymbol{G_v} = g_1 \times \boldsymbol{CA} + g_2 \times \boldsymbol{CAB} \tag{2}$$

$$\boldsymbol{C_{cost}} = nll(\boldsymbol{G_v}, \boldsymbol{label})$$

The architecture of the SD model is shown in Fig. 6. This model has the same architecture as that of the DAB model. However, the output of the model is the result of model selection to select the DAB or DA model.

The DA or DAB model is automatically selected for each image to identify the presence or absence of LAIB recurrence using the SD model.

Convolution layer

In the convolution layers of the DA, DAB, and SD models, as shown in Fig. 7, multiple filters are prepared for feature extractors A and B. The convolution of each filter is executed for the input data $\boldsymbol{x}_{\mathbf{conv}}$, then the output $\boldsymbol{y}$ shown in Eq. (3) is calculated for each feature extractor. The $\boldsymbol{x}_{\mathrm{conv}}$ is an input image (three channels of RGB) for Conv1 of feature extractor A, input image (three channels of LLL) for Conv1 of feature extractor B, output data of the max pool layer for building block 0, and output data of the previous layer for building block 11 and later. The bold italics in each equation represent a matrix.

$$\boldsymbol{y} = \mathrm{convolution}(\boldsymbol{x}_{\mathrm{conv}}, \boldsymbol{W}) + \boldsymbol{b} \tag{3}$$

$$\boldsymbol{z}_{\mathrm{conv}} = \mathrm{ReLU}(\boldsymbol{y})$$

Next, the nonlinear function ReLU[11] is applied to $\boldsymbol{y}$ to calculate the output data $\boldsymbol{z}_{\mathrm{conv}}$ for each feature extractor. Using multiple pathological images for training, the filter weight $\boldsymbol{W}$ and bias $\boldsymbol{b}$ are calculated in accordance with Eq. (3) by machine learning so that a tumor that will not recur is determined to be such a tumor, and a tumor that will recur is determined to be such a tumor. During machine learning, it is desirable to set the filter size to a size that includes a part of the cell and set the number of filters to the number of shape types that express the partial shape that constitutes one cell.

Pooling layer

In the max pool layers of the DA, DAB, and SD models, the horizontal and vertical information is thinned out to extract the features robust to the displacement of the object in the pathological image. If the $\boldsymbol{z}_{\mathrm{conv}}$ of the convolution layer is set as the input data $\boldsymbol{x}_{\mathrm{pool}}$ of the pooling layer, the pooling size[8] should be set to a size that suppresses the position fluctuation of cells in the pathological image and extracts tumor information regarding the presence or absence of LAIB recurrence.

Fully connected layer

In the final FC layers of the DA and DAB models, the predicted value for identifying the presence or absence of tumor recurrence is calculated as

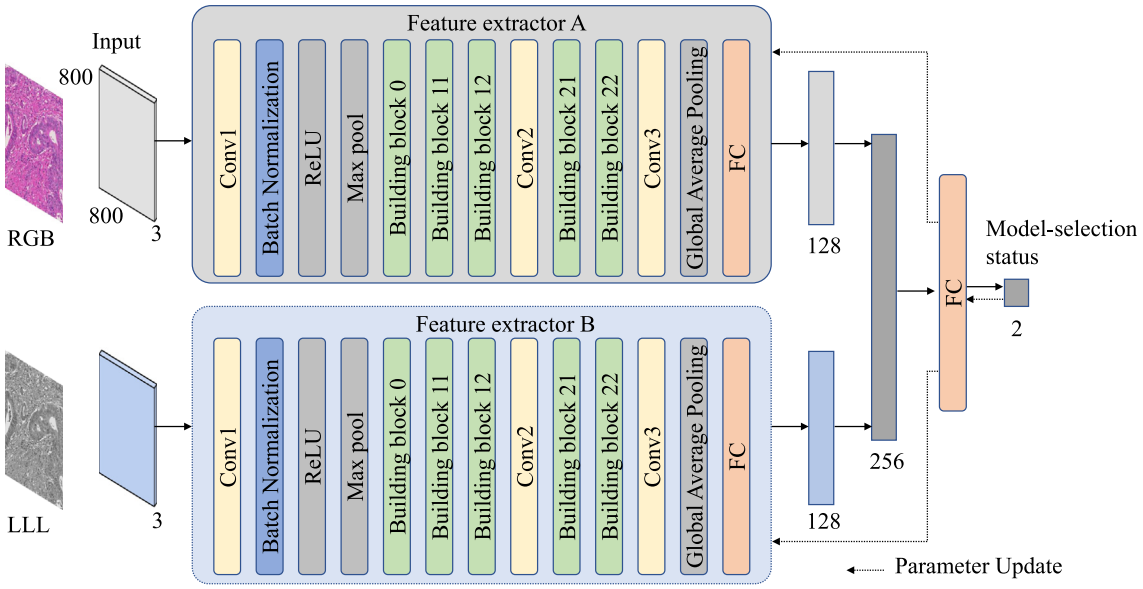

Fig. 6. Architecture of SD model.

between 0 and 1. As shown in Fig. 8, the data obtained by combining the output data $\boldsymbol{zA}_{\mathrm{pool}}$ of the FC layer of feature extractor A and the output data $\boldsymbol{zB}_{\mathrm{pool}}$ of the FC layer of the feature extractor B are aligned into the output data $\boldsymbol{z}_{\mathrm{pool}}$ of the one-dimensional array, and $\boldsymbol{z}_{\mathrm{pool}}$ is set as the input data $\boldsymbol{x}_{\mathrm{LR}}$ of the final FC layer. In the DA and DAB models, a unit for determining a tumor that does not recur and one for determining a tumor that will recur are arranged, and these units are fully combined with $\boldsymbol{x}_{\mathrm{LR}}$. In the SD model, a unit that selects the DA model and one that selects the DAB model are arranged, and these units are fully combined as $\boldsymbol{x}_{\mathrm{LR}}$. As shown in Eq. (4), the bias $\boldsymbol{b}_{\mathrm{LR}}$ is added to obtain the output $\boldsymbol{yv}$ after calculating the matrix product of $\boldsymbol{x}_{\mathrm{LR}}$ and weight $\boldsymbol{W}_{\mathrm{LR}}$. Finally, the DA and DAB models apply the softmax function to $\boldsymbol{yv}$ to calculate the identification prediction value $\boldsymbol{z}_{\mathrm{LR}}$ (range 0–1.0) for both the lung adenocarcinoma that does not recur and that will recur. In the SD model, the softmax function is applied to $\boldsymbol{yv}$ to calculate $\boldsymbol{z}_{\mathrm{LR}}$ (range is 0–1.0) that selects the DA and DAB models.

$$\boldsymbol{x}_{\mathrm{LR}}\left(\boldsymbol{z}_{\mathrm{pool}}\right) = \mathrm{concat}\left(\boldsymbol{zA}_{\mathrm{pool}}, \boldsymbol{zB}_{\mathrm{pool}}\right) \tag{4}$$

$$\boldsymbol{yv} = \boldsymbol{x}_{\mathrm{LR}} \times \boldsymbol{W}_{\mathrm{LR}} + \boldsymbol{b}_{\mathrm{LR}}$$

$$\boldsymbol{z}_{\mathrm{LR}} = \mathrm{softmax}(\boldsymbol{yv})$$

For the parameter update process of the DA and DAB models, the cost of these models is calculated using the $\boldsymbol{z}_{\mathrm{LR}}$ of the tumor that does not recur and the tumor that will recur, then $\boldsymbol{W}$ and $\boldsymbol{b}$ of the convolution layer and $\boldsymbol{W}_{\mathrm{LR}}$ and $\boldsymbol{b}_{\mathrm{LR}}$ of the FC layer are updated so that the cost is minimized using Adam.[14] However, C_{cost} is calculated using the $\boldsymbol{z}_{\mathrm{LR}}$ for selecting the DA and DAB models, then $\boldsymbol{W}$ and $\boldsymbol{b}$ are updated so that C_{cost} is minimized using Adam in the parameter-update process of the SD model.

The DA and DAB models, which identify the presence or absence of LAIB recurrence and consist of feature extractor A, feature extractor B, and a final FC layer, were constructed using CNNs.

The SD model, which automatically selects the DA and DAB models and consists of feature extractor A, feature extractor B, and a final FC layer, was also constructed using CNNs. The DA and DAB models are automatically selected for each image using the SD model, and the presence or absence of LAIB recurrence in the input H&E-stained image is identified.

The parameters of the models we constructed for the proposed method, i.e., the SD and DAB models, model-training parameters, and computational environment and framework are summarized in Tables 1, 2, and 3, respectively. We also set the parameter of "Number of filters" so that features that express some of the shapes, sizes, and arrangement relationships of cells, cell nuclei, etc. are captured in the first layer of our models. We also set the parameters of these models so that the features that express these complex shapes are captured in building block 0 and subsequent layers.

Table 4 shows the classification of the identification results. Sensitivity, specificity, and accuracy rate (accuracy) are calculated as true positive (TP)/(TP + false negative (FN)), true negative (TN)/(false positive (FP) + TN), and (TP + TN)/(TP + FP + TN +FN), respectively, and were set as the evaluation index.

Results

We used 19986 H&E-stained partial images and 4000 for training and validation. We also used the validation images for validation during training to suppress overfitting of the models to be constructed and improve generalization performance. The training and validation images were extracted from the same samples, but the partial image containing the same area as

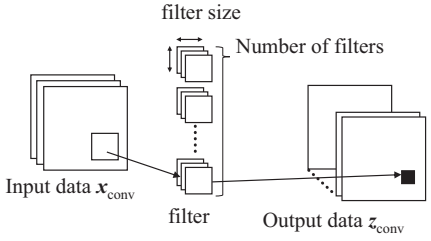

Fig. 7. Convolution-layer processing.

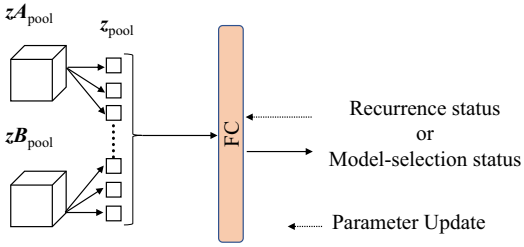

Fig. 8. Final FC layer processing.

Table 1
Parameters of the SD and DAB models.

Parameter	Filter size	Number of filters
Conv1	7x7,	32
Building block 0(Conv 1st,2nd,3rd)	1x1,3x3,1x1,	32,64,64
Building block 11(Conv 1st,2nd,3rd)	1x1,3x3,1x1,	64,64,64
Building block 12(Conv 1st,2nd,3rd)	1x1,3x3,1x1,	64,64,64
Conv2	1x1,	128
Building block 21(Conv 1st,2nd,3rd)	1x1,3x3,1x1,	128,128,128
Building block 22(Conv 1st,2nd,3rd)	1x1,3x3,1x1,	128,128,128
Conv3	1x1,	256

Table 2
Model-training parameters.

Parameter	Value
Batch size	6
Optimizer (Learning rate)	0.001
Optimizer (Adam) beta1	0.9
Optimizer (Adam) beta2	0.999
Optimizer (Adam) epsilon	1e-8

Table 3
Computational environment and framework.

OS	Ubuntu 16.04 64bits
CPU	3.6GHz 8cores
Memory	32GB
GPU	Geforce RTX 2080 SUPER
Framework	Pytorch

Table 4
Classification of identification results.

		Ground Truth	
		LAIB recurrence	No LAIB recurrence
Prediction with proposed method	LAIB recurrence	TP	FP
	No LAIB recurrence	FN	TN

the training data was not included in the validation images. We also created 7106 evaluation images randomly selected from samples different from the training and validation samples. Using SD (RGB+LLL) and DAB (RGB+LLL) for predicting tumor recurrence, we conducted an identification experiment to predict recurrence of LAIB. The results are shown in Table 5. However, the training, validation, and evaluation images did not include partially overlapping of partial images.

We first compared SD (RGB+LLL) and DAB (RGB+LLL) with various DA and DAB models using images in various color spaces, i.e., DA (RGB), DA (HSV), DA (HLS), DA2 (RGB), DA2 (LLL), DAB (RGB + HLS), DAB (RGB + L), DAB (RGB + HSV), and DAB (RGB + Lab), where DA2 denotes a model consisting of double the number of filters of feature extractor A, then compared our models with the conventional VGG16 (RGB), Resnet18 (RGB), Inception v3 (RGB), and Densenet121 (RGB)[15] models, as shown in Table 5. The type of image used is shown in parentheses.

The sensitivity, specificity, and accuracy of DA (RGB) constructed using only the RGB image and feature extractor A in Fig. 4 were respectively 93.6, 83.3, and 87.1%, and those of DAB (RGB + LLL) using the RGB and LLL images and feature extractors A and B in Fig. 2 were respectively 91.2, 89.6, and 90.2%, and those of SD (RGB + LLL) in Fig. 6, which switches between DA (RGB) and DAB (RGB+LLL) in accordance with the image were respectively 91.7, 90.2, and 90.9%.

Discussion

Partial image-based analysis results

By creating a feature extractor using not only the RGB component but also three L components, i.e., in DAB (RGB + LLL), our method can extract the morphological features regarding the presence or absence of LAIB recurrence in the image that cannot be captured by the R, G, and B components alone. Fig. 9 shows examples of LAIB images from which DAB (RGB + LLL) correctly identified the presence or absence of LAIB recurrence and that DA (RGB) misidentified.

Table 5
Identification results of presence or absence of LAIB recurrence in partial image units.

Identification model (Image type used)	Sensitivity [%]	Specificity [%]	Accuracy [%]
SD (RGB+LLL): proposed	91.7	90.2	90.9
DAB (RGB+LLL)	91.2	89.6	90.2
DA (RGB)	93.6	83.3	87.1
DA (HSV)	91.2	83.4	86.3
DA (HLS)	91.7	78.7	83.5
DA2 (RGB)	93.6	79.4	84.6
DA2 (LLL)	93.7	78.7	84.3
DAB (RGB+HLS)	93.3	77.8	83.5
DAB (RGB+L)	91.6	86.0	88.1
DAB (RGB+HSV)	91.0	82.6	85.7
DAB (RGB+Lab)	91.6	79.9	84.2
VGG16 (RGB)	88.7	84.4	86.0
Inceptionv3 (RGB)	92.4	80.7	85.0
Resnet18 (RGB)	95.0	81.8	86.7
Densenet121 (RGB)	88.7	79.4	84.0

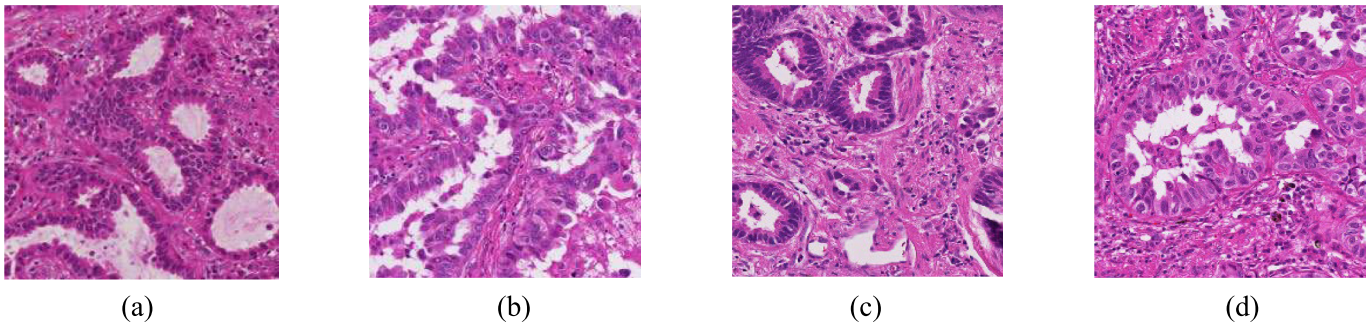

Fig. 9. Example images that were misidentified with DA (RGB) and correctly identified with DAB (RGB + LLL). (a) Acinar subtype with recurrence correctly identified with DAB (RGB + LLL), (b) papillary subtype with recurrence correctly identified with DAB (RGB + LLL), (c) acinar subtype without recurrence correctly identified with DAB (RGB + LLL), and (d) papillary subtype without recurrence correctly identified with DAB (RGB + LLL).

As shown in Fig. 9, DAB (RGB + LLL) could correctly identify the presence or absence of LAIB recurrence for acinar and papillary subtypes of LAIB by using RGB and LLL components together. In this experiment, the acinar and papillary subtypes accounted for 60.3% of all LAIBs for which DAB (RGB + LLL) could correctly identify compared with DA (RGB). The acinar subtype forms glandular structures with a central luminal space surrounded by tumor cells. Tumor cells in the papillary subtype characteristically grow in a papillary fashion. These lung adenocarcinomas contain many white areas. By using the model with a parallel structure of feature extractors and the features of RGB and LLL images, i.e., DAB (RGB+LLL), features of tubular structures surrounded by tumor cells are extracted in the acinar subtype, while in the papillary subtype, features of papillary growth of tumor cells within the glandular structure are extracted. Therefore, DAB (RGB+LLL) enables more accurate distinction of the presence or absence of LAIB recurrence.

Fig. 10 shows examples of images that were misidentified with DAB (RGB + L) but correctly identified with DAB (RGB + LLL). The specificity of DAB (RGB + LLL) improved by 3.6% and its accuracy improved by 2.1%. When training a model that configures feature extractors A and B in parallel, it is possible that the model better reflects the features of L by giving the same amount of image components.

From the results of DA2 (RGB) and DA2 (LLL), which are configured by doubling the number of filters in feature extractor A, the presence or absence of LAIB recurrence could not be accurately identified simply by increasing the number of filters in the model. Therefore, we believe it is effective to combine the features of both RGB and LLL to identify the presence or absence of LAIB recurrence.

Comparing the results of conventional VGG16 (RGB), Inceptionv3 (RGB), Resnet18 (RGB), and Densenet121 (RGB), the accuracy of Resnet18 (RGB) was 86.7%, which is the most accurate result. Comparing the results of DA (RGB), DA (HSV), DA (HLS), and Resnet18 (RGB) using one type of image, the accuracy of DA (RGB) was 87.1%. Although the specificity and sensitivity of DA(RGB) were respectively 1.5% higher and 1.4% lower than Resnet 18(RGB), the accuracy of DA(RGB) was 0.4% higher than Resnet 18 (RGB). However, DA (RGB) tended to misidentify non-recurrence LAIB as recurrence LAIB. Comparing the results of DAB (RGB + LLL), DAB (RGB + HLS), DAB (RGB + HSV), and DAB (RGB + Lab)

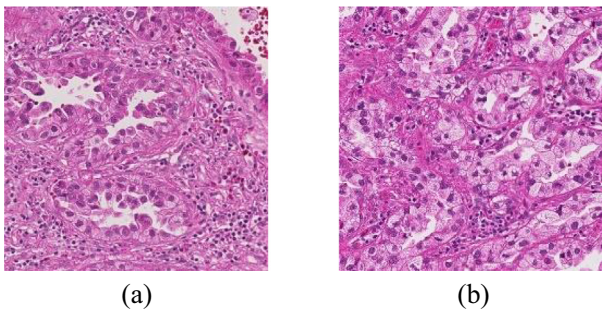

Fig. 10. Example images that were misidentified with DAB (RGB + L) and correctly identified with DAB (RGB + LLL). (a) Cancer cells showing papillary growth correctly identified with DAB (RGB + LLL), (b) cancer cells showing glandular structure correctly identified with DAB (RGB + LLL).

Table 6
Prediction results of presence or absence of LAIB recurrence in image units.

Model (Image type used)	Prediction rate [%]
SD (RGB+LLL): proposed	84.8
DAB (RGB+LLL)	78.8
DA (RGB)	72.7

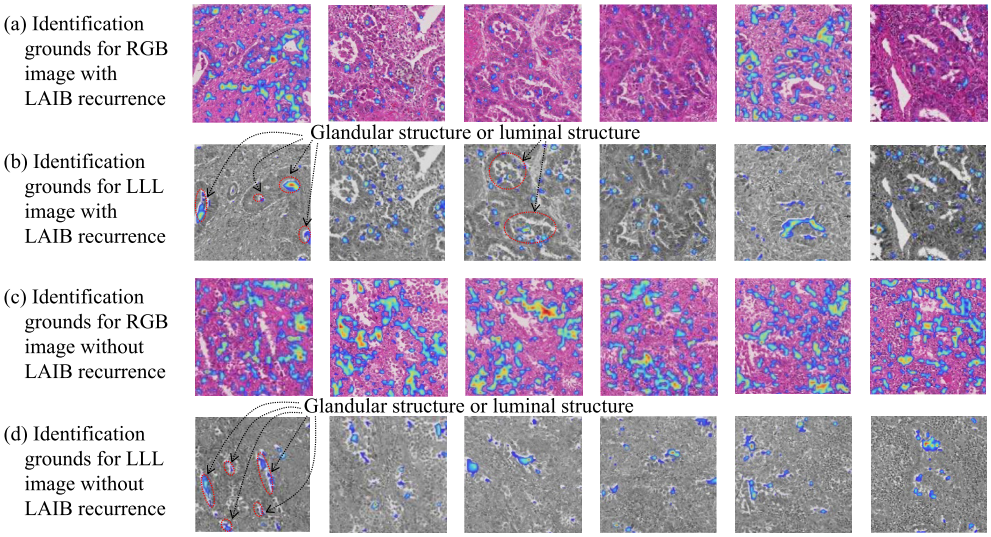

Fig. 11. Display examples of identification grounds of feature extractors A and B. (a) Identification grounds for RGB image with LAIB recurrence that feature extractor A predicted, (b) identification grounds for LLL image with LAIB recurrence that feature extractor B predicted, (c) identification grounds for RGB image without LAIB recurrence that feature extractor A predicted, and (d) identification grounds for LLL image without LAIB recurrence that feature extractor B predicted.

using two types of images in the model architecture of the proposal method, the accuracy of DAB (RGB + LLL) was 90.2%, which is the most accurate result. Although the sensitivity of DAB (RGB + LLL) was 2.4% lower than DA (RGB), the specificity and accuracy of DAB (RGB + LLL) improved by 6.3 and 3.1%, respectively, compared with DA (RGB). Thus, DAB (RGB + LLL) can identify between non-recurrence and recurrence LAIB in a well-balanced manner. Comparing the results of SD (RGB+LLL) and DAB (RGB+LLL), the sensitivity, specificity, and accuracy of SD (RGB+LLL) respectively improved by 0.5, 0.6, and 0.7%.

Image-based analysis results

Table 6 shows the results of recurrence prediction of LAIB by image unit of SD (RGB + LLL), which automatically switches between DAB (RGB+LLL) and DA (RGB) for each image; DAB (RGB + LLL) only; and DA (RGB) only. When the presence or absence of LAIB recurrence was correctly identified for 70% or more of the partial images in an image, the image was regarded as a correct prediction, and the prediction rate was calculated by dividing the number of images predicted correctly by the total number of images. DAB (RGB + LLL) had a 6.1% improvement in the prediction rate for images compared with DA (RGB). Thus, DAB (RGB + LLL) can determine the presence or absence of LAIB recurrence more accurately than DA (RGB). Furthermore, SD (RGB + LLL) had a 6.0% improvement in the prediction rate for images compared with only DAB (RGB + LLL). It is therefore possible that SD (RGB + LLL) can determine the presence or absence of LAIB recurrence more accurately than only DAB (RGB + LLL). Thus, SD (RGB + LLL) makes it possible to appropriately select a model that determines the presence or absence of LAIB recurrence for each image.

Identification-grounds-based analysis results

As shown in Fig. 11, the identification grounds for feature extractors A and B of SD (RGB + LLL) and DAB (RGB + LLL) are displayed on the image using Grad-CAM.[16] The numerical values (0.0–1.0) indicating the identification grounds are displayed in gradation from blue (0.0) to red (1.0) (the display of the identification grounds of 0.0 was omitted).

Fig. 11 (a) shows the identification grounds when feature extractor A predicted that LAIB in RGB images will recur, and Fig. 11 (b) shows the identification grounds when feature extractor B predicted that LAIB in LLL images will recur. Fig. 11 (c) shows the identification grounds when feature extractor A predicted that LAIB in RGB images will not recur, and Fig. 11 (d) shows the identification grounds when feature extractor B predicted that LAIB in LLL images will not recur.

Comparing Fig. 11 (a) with (b), and (c) with Fig. 11 (d), the features captured by feature extractor A using RGB images differed from those captured by feature extractor B using LLL images. It is thus possible that feature extractor B using LLL images can capture the features related to the presence or absence of LAIB recurrence, such as in the glandular structures or luminal structures, which cannot be captured only with conventional models or feature extractor A using only RGB images. Therefore, the presence or absence of LAIB recurrence for lung adenocarcinoma including glandular structures or luminal structures can be accurately predicted

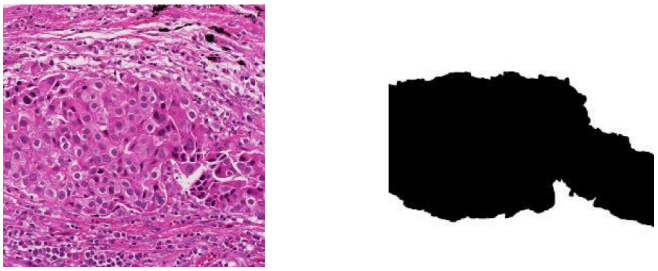

Fig. 12. Example of mask image.

Table 7
Comparison of regions of identification grounds for cancer-cell and non-cancer-cell regions. Numbers indicate number of blocks (400 × 400 pixels) of identification grounds. Numbers in parentheses indicate percentage of total number of pixels of cancer and non-cancer identification grounds. We calculated *P*-values for number of blocks of identification grounds using chi-square test.

Region of identification grounds	Recurrence	Non-recurrence	P-value
RGB			
Region of RGB identification grounds for cancer	68 (76.7%)	162 (74.0%)	P=0.656
Region of RGB identification grounds for non-cancer	21 (23.3%)	57 (26.0%)	
LLL			
Region of LLL identification grounds for cancer	24 (58.3%)	28 (85.1%)	P=0.027
Region of LLL identification grounds for non-cancer	17 (41.7%)	5 (14.9%)	

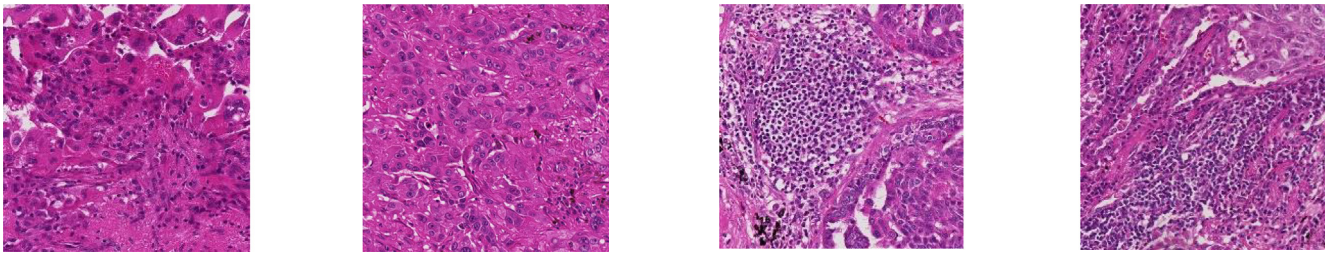

Fig. 13. Example image that was misidentified with DAB (RGB + LLL) and correctly identified with SD (RGB + LLL). (a) Cancer cells with recurrence correctly identified with SD (RGB + LLL), (b) cancer cells without recurrence in solid subtype correctly identified with SD (RGB + LLL).

Fig. 14. Example of images misidentified with SD (RGB + LLL). (a) Cancer cells at upper- and lower right, (b) cancer cells at upper right.

using the features that feature extractor A extracts from RGB images and those that feature extractor B extracts from LLL images.

Next, we analyzed the percentage of identification grounds contained in the cancer-cell and non-cancer-cell regions. As shown in Fig. 12, we created a mask image in which the cancer-cell region was black and the non-cancer cell region was white. Comparing the mask image with the identification grounds, we calculated the ratio of the identification grounds contained in each region for the RGB image of feature extractor A and the LLL image of feature extractor B by using Eq. (5). Table 7 lists the results. In Eq. (5), *CNC* represents the number of pixels of the identification grounds for the non-cancer cell or the cancer-cell region, and RIB_{all} represents the total number of identification grounds.

$$RIB_{\mathrm{ratio}} = CNC/RIB_{\mathrm{all}} \tag{5}$$

As shown in Table 7, when comparing the percentage of identification grounds for LLL images indicated by feature extractor B in the cancer-cell region for with and without LAIB recurrence, it is clear that the region of the identification grounds for LAIB without recurrence is 26.8% larger than that of the identification grounds for LAIB with recurrence. The region of LLL identification grounds in the cancer-cell region for without LAIB recurrence is statistically different from that for with LAIB recurrence (*P*-value = .027). This study has shown that LAIB consisting of glandular structures or luminal structures that have a large area tends not to recur. Therefore, SD (RGB + LLL) switching between DAB (RGB + LLL) and DA (RGB) in accordance with the image, and DAB (RGB + LLL) consisting of feature extractors A and B can more accurately capture the features regarding the presence or absence of LAIB recurrence including in glandular structures or luminal structures.

DAB (RGB + LLL) is suitable for identifying the presence or absence of LAIB recurrence in acinar, lepidic, and papillary subtypes, including white areas such as glandular structures or luminal structures. However, DA (RGB) is suitable for identifying the presence or absence of LAIB recurrence of the solid subtype, which does not contain many white areas. As shown in Fig. 13, SD (RGB + LLL) can further improve the identification accuracy of cancer cells by appropriately switching DAB (RGB + LLL) and DA (RGB) for each image and accurately identifying the recurrence or no recurrence of LAIB.

Fig. 14 shows example images misidentified with SD (RGB + LLL). If the partial image contains almost no features of the presence or absence of LAIB recurrence, SD (RGB + LLL) may misidentify.

Conclusions

The proposed tumor-identification method is capable of extracting features related to the presence or absence of LAIB recurrence.

Since the proposed method can improve the accuracy of predicting the presence or absence of LAIB recurrence in pathological images, it had been shown to be effective as an identification method for predicting LAIB recurrence. Therefore, our method should be effective in formulating clinical treatment plans.

Compliance with ethical standards

The data of this study were examined and obtained in accordance with the ethical examination standards established by Keio University (approval number: 2021-52) and National Cancer Center (IRB number:2020-158).

Funding

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

Conflict of interest

The authors declare that they have no conflict of interest.

Acknowledgments

In carrying out this study, we would like to express our deep gratitude to Dr. Yutaro Tamiya of the Department of Respiratory Medicine for providing pathological image data and compiling clinical data.

References

1. Yu K, Zhang C, Berry GJ, et al. Predicting non-small cell lung cancer prognosis by fully automated microscopic pathology image features. Nat Commun 2016;7:1-10.
2. Consonni D, Pierobon M, Gail HM, et al. Lung cancer prognosis before and after recurrence in a population-based setting. J Natl Cancer Inst 2015;107(6):1-12.
3. Amin MB, Edge SB, Greene FL, et al. *American Joint Committee on Cancer. AJCC Cancer Staging Manual. 8th ed.* NY: Springer, New York. 2017.
4. Wang X, Janowczyk A, Zhou Y, et al. Prediction of recurrence in early stage non-small cell lung cancer using computer extracted nuclear features from Digital H&E images. Scient Rep 2017;7(1):1-10.
5. Szegedy C, Vanhoucke V, Ioffe S, Shlens J, Wojna Z. Rethinking the inception architecture for computer vision. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas; 2016. p. 2818–2826.
6. Wang S, Chen A, Yang L, et al. Comprehensive analysis of lung cancer pathology images to discover tumor shape and boundary features that predict survival outcome. Scient Rep 2018;8:1–9.
7. Diaz-Cely J, Arce-Lopera C, Mena JC, Quintero L. The effect of color channel representations on the transferability of convolutional neural networks. In: *Arai K, Kapoor S, eds.* CVC 2019. AISC. Cham: Springer; 2020. p. 27–38. https://doi.org/10.1007/978-3-030-17795-9_3.
8. LeCun Y, Bottou L, Bengio Y, Haffner P. Gradient-based learning applied to document recognition. Proc IEEE 1998;86(11):2278–2324.
9. Otsu N. Threshold selection method from gray-level histograms. IEEE Trans Syst Man Cybern 1979;9:62–66.
10. Ioffe S, Szegedy C. Batch normalization: accelerating deep network training by reducing internal covariate shift. ICML; 2015.
11. Nair V, Hinton GE. Rectified linear units improve restricted Boltzmann machines. ICML; 2010.
12. Lin M, Chen Q, Yan S. Network in network. Proc. of ICLR; 2014.
13. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas; 2016. p. 770–778.
14. Kingma DP, Ba J. Adam: a method for stochastic optimization. International Conference on Learning Representations (ICLR), San Diego; 2015. p. 1-13.
15. Huang G, Liu Z, Van der Maaten L, Weinberger KQ. Densely connected convolutional networks. Proceedings of IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu; 2017. p. 4700–4708.
16. Selvaraju RR, Cogswell M, Das A, Vedantam R, Parikh D, Batra D. Grad-CAM: visual explanations from deep networks via gradient-based localization. IEEE International Conference on Computer Vision (ICCV), Venice; 2017. p. 618–626.