Deep Learning-based Histopathological Segmentation for Whole Slide Images of Colorectal Cancer in a Compressed Domain

Hyeongsub Kim
Pohang University of 18 Science and Technology (POSTECH)
Hongjoon Yoon
Deepnoid Inc
Nishant Thakur
The Catholic University of Korea, Uijeongbu St. Mary's Hospital
Gyoyeon Hwang
The Catholic University of Korea, Yeouido St. Mary's Hospital
Eun Jung Lee
The Catholic University of Korea, Yeouido St. Mary's Hospital
Chulhong Kim (chulhong@postech.edu)
Pohang University of 18 Science and Technology (POSTECH)
Yosep Chong
The Catholic University of Korea, Uijeongbu St. Mary's Hospital

Research Article

Keywords: Deep Learning, Histopathological Segmentation, Colorectal Cancer, Compressed Domain

Posted Date: June 2nd, 2021

DOI: https://doi.org/10.21203/rs.3.rs-534311/v1

Version of Record: A version of this preprint was published at Scientific Reports on November 18th, 2021.
See the published version at https://doi.org/10.1038/s41598-021-01905-z.

Title

- Title: Deep Learning-based Histopathological Segmentation for Whole Slide Images of Colorectal Cancer in a Compressed Domain
- Running title: Pathological segmentation in a compressed domain

Authors

Hyeongsub Kim[1, 2]: hskim001@postech.ac.kr

Hongjoon Yoon[2]: hyoon@deepnoid.com

Nishant Thakur[3]: nishantbiotech2014@gmail.com

Gyoyeon Hwang[4]: hgy1206@gmail.com

Eun Jung Lee[4, 5]: ejlpath@catholic.ac.kr

Chulhong Kim[1*]: chulhong@postech.edu

Yosep Chong [3*]: ychong@catholic.ac.kr

Affiliations

[1]Departments of Electrical Engineering, Creative IT Engineering, Mechanical Engineering, School of Interdisciplinary Bioscience and Bioengineering, Medical Device Innovation Center, and Graduate School of Artificial Intelligence, Pohang University of Science and Technology (POSTECH), Pohang 37674, South Korea

[2]Deepnoid Inc., Seoul 08376, South Korea

[3]Department of Hospital Pathology, The Catholic University of Korea, College of Medicine, Uijeongbu St. Mary's Hospital, Seoul, South Korea

[4]Department of Hospital Pathology, The Catholic University of Korea, College of Medicine, Yeouido St. Mary's Hospital, Seoul, South Korea

[5]Department of Pathology, Shinwon Medical Foundation, Gwangmyeong-si, Gyeonggi-do, South Korea

*Corresponding author

Chulhong Kim

Tel: 82-54-279-8892

Fax: 82-54-279-8899

Email: chulhong@postech.edu

Yosep Chong

Tel: 82-31-820-3160

Fax: 82-31-820-3877

Email: ychong@catholic.ac.kr

Abstract

Automatic pattern recognition using deep learning techniques has become increasingly important. Unfortunately, due to limited system memory, general preprocessing methods for high-resolution images in the spatial domain can lose important data information such as high-frequency information and the region of interest. To overcome these limitations, we propose an image segmentation approach in the compressed domain based on principal component analysis (PCA) and discrete wavelet transform (DWT). After inference for each tile using neural networks, a whole prediction image was reconstructed by wavelet weighted ensemble (WWE) based on inverse discrete wavelet transform (IDWT). The training and validation were performed using 351 colorectal biopsy specimens, which were pathologically confirmed by two pathologists. For 39 test datasets, the average Dice score was 0.852 ± 0.086 and the pixel accuracy was 0.962 ± 0.027. We can train the networks for the high-resolution image (magnification x20) compared to the result in the spatial domain (magnification x10) in same the region of interest ($6.25 \times 10^{-2}\ \mu m^{2}$). The average Dice score and pixel accuracy are significantly increased by 6.4 % and 1.6 %, respectively. We believe that our approach has great potential for accurate diagnosis in pathology.

MAIN TEXT

Introduction

The large number of inspections for pathologists is exposed to the risk of misdiagnosis. This leads to a rapid increase in medical expenses, an increase in the false diagnosis rate, a decrease in medical productivity, and the risk of a cancer diagnosis. Automatic analyses of pathological images can mitigate human effort, save time, and provide a confident foundation for surgery and treatment. Convolutional neural networks (CNNs) are especially popular for the automatic diagnosis of many diseases in pathology [1–13]. However, despite the continued increase in the speed and memory capacity of central processing units (CPUs) and graphical processing units (GPUs), technological advances in pathological image analysis are still hampered by large image sizes [14].

For high-resolution and large scale images, a general preprocessing method to relieve memory limitation can induce important information loss. Several methods have been explored to reduce image sizes, such as decimation, cropping, and compression [15–17]. Decimation is the major process for down sampling large images, and it can also reduce noise power and improve signal-to-noise ratios (SNRs), thanks to an anti-aliasing filter. However, decimation can cause a loss of high-frequency information, resulting in low resolution due to the reduced signal bandwidth [18,19]. As another widely used method, cropping extracts the wanted areas from whole slide images (WSIs) into tiles. Although no information is missed with respect to a single tile, the spatial relationships between tiles may be lose, which is critical because object judgments depend on the relative size and color of each cell in the pathological image.

Compression is widely used both to minimize the size of an image file without degradation in the image quality and to reduce irrelevance and redundancy of data in the image. Thus, compression is mostly preferred to process large-scale images. For example, detecting ships in satellite images is difficult due to their high resolution and correspondingly large data volume. A compression technique called discrete wavelet transform (DWT) resolves the difficulty in high-resolution ship detection and performs better than conventional computer vision algorithms [20]. In addition, DWT is also useful for texture classification, because its finite duration provides both the frequency and spatial locality. In pathology, DWT analysis has been applied to classify tumors by using texture analysis [21].

In this work, we propose a pathological image segmentation method in the compressed domain. To compress large pathological images, we utilized not only DWT but also principal component analysis (PCA) according to hematoxylin and eosin (H&E) staining characteristics to reduce 3-channel RGB data to one channel [22]. We tested this inference method in the compressed domain on colorectal cancer pathologic images from the Catholic University of Korea Yeouido St. Mary's Hospital.

Our results imply that the method using the compressed domain is more useful for pathologic segmentation than the method using the spatial domain, for three reasons: 1) The average Dice score and pixel accuracy are significantly improved, by 6.4% and 1.6%, respectively. 2) Using DWT, neural networks can be trained not only by spatial information but also by texture information. 3) The performance can be more robust because of the large ROI in training after compression; the size of the input image is reduced by 8%. This new segmentation technique in the compressed domain can be potentially useful in applications where large scale data and texture information are important, such as remote sensing [23] and microscopy [24].

Results

Data Distribution

We used 390 WSIs of colorectal biopsy specimens. The average size of WSIs was 43,443 by 28,645 pixels. We split the dataset into three groups: 274 train data, 77 validation data, and 39 test data (Supplementary Table I). We used this dataset to implement a pipeline to achieve binary segmentation of normal and abnormal areas in colorectal cancer (CRC) tissue images.

Overall Result according to Each Method

Table I compares the average Dice and Acc according to each method. As it shows, for the model using compressed data, the average Dice and Acc results for the LL sub-band increased by 5.6 % and 1.6 %, respectively, compared to those of the model in spatial domain. The reason why LL's results improve is the impact of PCA. Channels are reduced and background is removed, reducing input complexity and improving performance. However, the average Dice and Acc results of the LH (-7.6 % for Dice and -2.5% for Acc), HL (-4.0% for Dice and -2.2% for Acc), and HH (-6.4% for Dice and -1.5% for Acc) sub-bands carrying high-frequency information decreased compared to those before compression.

The Trend for Dice and Acc according to Each Class

Fig. 1 shows distribution Dice and Acc according to all classes. In the case of all tumor classes, the average results for the LL sub-band are relatively high. Further, the average results of the LH, HL, HH sub-bands carrying high-frequency components are relatively high in ADENOCA and TAH (Fig. 1), which are relatively easy to detect due to advanced disease progression and consequent pathological modifications. However, the results of the LH, HL, HH sub-bands are less predictive for TAL, CARCINOID, and HYPERP (Fig. 1). TAL (relatively less advanced), CARCINOID (malignant tumors but occurring in the submucosa), and HYPERP (benign tumors) are difficult to accurately predict with only high-frequency components. Based on these results, we propose an ensemble method that can improve the results using both low-frequency and high-frequency information. In the

case of WAE, a widely used ensemble method, even though the weight is given to the result of the LL sub-band, the average Dice and Acc are lower than those of the result of the LL sub-band (-5.7% for Dice and -2.3 % for Acc). Compared to the no compression results, ADENOCA and TAH show good performance after WAE because the Dice and Acc in the high-frequency sub-bands such as LH, HL, and HH sub-band are higher than these of the no compression case (ADENOCA: +4.9% for Dice, +0.0% for Acc; TAH: +3.7% for Dice, +0.2% for Acc). However, in TAL, CARCINOID, and HYEPERP, which show low performance in the high-frequency sub-bands, the Dice and Acc after WAE are lower than those of no compression (TAL: +0.2% for Dice, -0.4% for Acc; CARCINOID: +4.3% for Dice, +2.3% for Acc; HYPERP: +11.1% for Dice, +4.1% for Acc;). On the other hand, after WWE, the average Dice and Acc increase by about 0.6% and 0.2%, respectively, compared to LL. For each class, the results of ADENOCA (+0.4% for Dice and +0.1% for Acc), TAH (+0.7% for Dice and +0.2% for Acc), TAL (+0.5% for Dice and +0.1% for Acc), CARCINOID (+1.5% for Dice and +0.1% for Acc), and HYPER (+0.8% for Dice and +1.8%for Acc) gradually increase. In particular, unlike WAE, the WWE also shows high predictive results for TAL (+9.2 % for Dice and +1.6% for Acc), CARCINOID (+8.7% for Dice and +3.7% for Acc), and HYPERP (+14% for Dice and +4.9% for Acc), and each shows higher performance than WAE.

Change in Dice in All Classes according to Low- (W_1) and High-frequency Weight (W_2, W_3, and W_4)

We checked change of Dice score in all classes according to low-frequency weight (W_1) and high-frequency weight (W_2, W_3, and W_4) to optimize each weight by conducting the empirical test. The best weights in the WWE are determined by the average Dice scores, as shown in Supplementary Table II. Fig 2 describes the change in Dice score with respect to various low-frequency weights (W_1) in all tumor classes (ADENOCA, TAH, TAL, CARCINOID, and HYPERP). From 0.6 to 0.9, the Dice scores of all the classes increase relatively steeply. Particularly, the increasing rates in the Dice scores of HYPERP and ADENOCA are relatively high. Beyond the W_1 value of 1.2, the Dice scores start being saturated in all classes. Further, we changed the values of the high-frequency weights (W_2, W_3, and W_4), but the changes in Dice scores are negligible as shown in Supplementary Fig 1.

Comparison of the Heat Map and Line Profiles between Annotation, the Result in the Spatial Domain, WAE, and WWE

Using a heat map and line profiles for tumor probability, we compared the segmentation prediction for annotation, the result in spatial domain, WAE, and WWE (Figs. 3a–d). The color bar indicates the tumor probability for each pixel. The heat map is overlaid on the original histology image, and a magnified image of the area in the colored border is located on top of the main image. The line profiles of the tumor probability cut along the red dotted dashed lines are located below the main image. Fig. 3(a) is the ground truth, annotated by a

pathologist. The pixel value in the annotation is 1, and the value in the other regions is 0. Fig. 3(b) shows the segmentation result in spatial domain. There is a slight loss of high-frequency information after decimation for efficient training, but the ROI used in single training is the same as for the other methods. The magnified image in Fig. 3(b) predicts a broader region than in the annotation, and the tumor probability in each pixel is relatively low. The segmentation result for WAE, shown in Fig. 3(c), clearly is qualitatively better than that in spatial domain. The final segmentation result with WWE has accurate edges as well as a high probability in each pixel, compared to the other methods. The tumor probability line profile processed with WWE is most similar to the original annotation profile, proving the accuracy of our method.

Average Dice for Each Method according to the Threshold

The Dice scores for WWE for the wavelet sub-bands, WAE for the wavelet sub-bands, and the result in spatial domain are compared across a range of threshold tumor probability values (Fig. 4). Between threshold values of 0.1 and 0.4, the Dice score of the result in spatial domain is slightly higher than that of WAE. However, beyond a threshold of 0.5, the Dice score for this method drops sharply, compared to those of the other methods. WAE and WWE continue perform robustly for all thresholds, and the Dice score of WWE is consistently higher than that of WAE, thanks to the high-frequency information.

Final Prediction Result of Five Different Tumor Classes using WWE

Finally, we compared our WWE predicted image with the image annotated by a pathologist. Fig. 5a-e shows tissue histology images from five different tumor categories. The pathologist's annotations are shown in Fig. 5f-j. The corresponding predicted probability map using WWE are shown in Fig. 5k-o and final overlaid tissue images are shown in Fig. 5p-t. The proposed WWE method generally segmented an afflicted area that corresponded well to the ground truth images. The average Dice and Acc of the WWE are 0.852 ± 0.086 and 0.962 ± 0.027, respectively. The best Dice (0.887 ± 0.101) is achieved in TAH, where the high-frequency information is important. On the other hand, the worst Dice (0.830 ± 0.057) is in TAL, where the low-frequency information is important. As shown in the yellow dotted boxes in the case of HYPERP (Figs. 5o, t), we often observed that the normal region where dead nuclei are gathered is abnormally predicted. Possibly, these abnormal predictions are caused by artifacts, such as tissue folds, ink, dust, and air bubbles, and further artifact removal may be required. Despite these abnormalities, the overall prediction of colorectal cancer using WWE was not biased to any one class: it performed well for all.

Discussion

The goal of this study is to increase diagnostic accuracy (e.g., Dice, Acc) by using a compressed domain to reduce high-frequency information loss. The compressed domain approach was employed in previous studies [20,25,26] showing good performance in

pathology classification not segmentation because there was no appropriate ensemble method for results for each sub-band (e.g. LL, LH, HL, HH sub-bands results) [21,27–30]. In this paper, we proposed the WWE method, which learns each low-frequency component and high-frequency component in the compression domain and then combines them. With the NVIDIA TITAN X 12Gb GPU used in this experiment, the U-net++ model can be trained on a maximum tile size of 512 by 512 at once. Therefore, in order to learn our experimental ROI size of $0.625 \times 10^{-1}\ \mu m^{2}$without compression, the resolution of the original image (×20 magnification) would have to be lowered (×10 magnification) (Table II). In this process, the loss of high-frequency components cannot be avoided. On the other hand, our proposed method can handle a tile size of 1024 by 1024 before compression. Thus, it is not necessary to lower the resolution to learn the same ROI size, and learning is possible with ×20 magnification. In addition, compared to the result in spatial domain, our proposed method can learn a tile that is four times larger than the limit of the hardware. However, our method requires four times more the number of GPUs (Table II) at the same time. From the perspective of time resources, in the case of a general CNN based on 2D convolution, the amount of computation increases exponentially as the input size increases. Therefore, it is faster to learn by separating one image into four images than to learn an image that is 4 times larger at a single time. This case is similar to the principle of the Cooley–Tukey FFT algorithm [31,32], and we believe that subsequent studies will also meaningfully to reduce time consuming.

We have conducted a study to prevent the loss of high-frequency information that occurs in the process of having to resize the image due to the limitation of the hardware and to increase the accuracy of the final result by using protected high-frequency information. Using a wavelet-weighted ensemble method, we found that accuracy was improved over that of images in spatial domain. The overall accuracy was determined by the low-frequency component, and the high-frequency component affected the margin. The disadvantage is that it requires a relatively large amount of GPU resources. However, we expect to reduce time-consuming compared to the result in spatial domain when the same as the initial tile size. To the best of our knowledge, this is the first study to do WWE in the compressed domain. We applied this processing method to colorectal cancer pathology images, and we believe that it can also be applied in general pathology images and show a similar increase in accuracy. Our proposed wavelet-weighted ensemble method can also be applied in other fields that process large-scale images (e.g., astronomy and satellite imagery) and that is important to margin (e.g., radiation therapy).

Methods

Data preparation

This study was reviewed and approved by the Institutional Review Board of the Catholic University of Korea College of Medicine (SC18RNSI0116). All experiments were conducted in accordance with relevant guidelines/regulations in the Catholic University of

Korea College of Medicine. Informed consent prior to the surgical procedures, all patients had given their informed consent to use tissue samples and pathological diagnostic reports for research purposes. We used a dataset using H&E stained-WSIs of colorectal biopsy specimens at the Yeouido St. Mary's Hospital. The WSIs were ×20 magnified images taken using a digital whole-slide camera (Aperio AT2, Leica biosystems, USA). The Whole slide images (WSIs) were manually annotated by the three trained pathologists supervised by the expert and performed routine histopathological examination by drawing the region of interest in the slides that corresponded to one of the five labels: adenocarcinoma (ADENOCA), high-grade adenoma with dysplasia (TAH), and low-grade adenoma with dysplasia (TAL), carcinoid (CARCINOID), and hyperplastic polyp (HYPERP). The average annotation time per WSI took 5-10 minutes. Next, annotations carried out by the trained pathologists were reviewed by the three senior pathologists and if necessary then modified and verified with the final checking verification by the one senior professors. Cases that had discrepancies in the annotation labels resolved the issue through further discussions. The images were excluded, when it was not possible to reach a consensus on a lesion type for an image. Most of the WSI contained multiple annotation labels. Therefore, a single WSI label of major diagnosis was assigned to a given WSI.

Compressed image analysis

In this study, we applied a compressed domain based on the wavelet transform used in JPEG2000 for the segmentation of pathologic images. The pipeline is as follows: tile extraction, z-axis compression, training and prediction in the compressed-domain using CNNs, prediction from one tile to the whole image, and wavelet-weighted ensemble (WWE) (Fig. 6). Each process is detailed in the following subsections.

Tile extraction based on a sliding window algorithm (Figs. 6(a) and Supplementary Fig. 2)

When the tiles are extracted from one WSI, the information about location and adjacent tiles is lost due to the limited fields-of-view. However, morphological information between adjacent areas is crucial for diagnostic decisions. Two typical tile extraction methods, the multiple ROI and sliding window methods, have been widely used to overcome this problem [12]. Although the multiple ROI method is faster than the sliding window because of its low redundancy, the sliding window method has the following advantages. First, the redundancy in the sliding window method assists data augmentation, an essential pre-processing step in a deep learning approach. Second, this method can overcome the limited field-of-view problem indirectly because the overlapping area depends on adjacent tiles. Finally, the overall accuracy can increase because the probability in the overlapping area is averaged during summation from the tile to the whole image. In this work, we choose the sliding window manner as the tile extraction method. Although the receptable maximum tile size is 512 × 512 pixels due to the limitation of our GPU memory size, we extracted a tile that

is 1,024 × 1,024 pixels in size before the compression step. The stride is set to 256 pixels, horizontally and vertically.

Z-axis compression based on principal component analysis (PCA)

Pathologic images have three red (R), green (G), and blue (B) channels (Fig. 7(a)). The correlation is high among each color (Fig. 7(c)). Color variation in the pathologic image is given by H&E staining, which dyes the cell nuclei blue, and dyes the extracellular matrix and cytoplasm pink. Therefore, z-axis compression was applied only to the R and B channels in the tissue region. First, Otsu algorithm were applied to extract the RGB values at tissue region, and then the G values were removed [33]. PCA was applied to maximize the variation between the R and B values and to minimize the mean squared error (Figs. 7(d)) [22]. This process reduces the image dimensionality and results in background reduction, widely used in histopathology (Fig. 7(b)). The PCA algorithm is described in detail in Supplementary Table III.

Training neural networks in the compressed-domain (x- and y-axis compression)

After the image depth compression (z-axis), discrete wavelet transform (DWT) was performed on each tile to compress the information along the x- and y- axis [35]. Haar wavelet is usually used to extract texture feature[28,30,36]. So, we decided 2D DWT based on Haar wavelet and its sub-band was calculated using the following Equations (1) – (4):

$$W_{\psi}^{A}(j,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \psi_{j,m,n}^{A}, \quad (1)$$

$$W_{\phi}^{V}(j,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \phi_{j,m,n}^{V}, \quad (2)$$

$$W_{\phi}^{H}(j,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \phi_{j,m,n}^{H}, \quad (3)$$

$$W_{\phi}^{D}(j,m,n) = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x,y) \phi_{j,m,n}^{D}, \quad (4)$$

where (x, y) is the coordinate of the input tile, (m, n) is the coordinate of the output sub-band, $\psi_{j,m,n}^{A}(x,y)$ and $\phi_{j,m,n}^{i}(x,y)$ represent the 2D wavelet basis function of level j, W_{ψ}^{A} describes an approximation of the original image called the LL (low-low) sub-band, and W_{ϕ}^{V}, W_{ϕ}^{H}, and W_{ϕ}^{D} are high-frequency components whose directions are vertical, horizontal, and diagonal. We call this transformed domain a compressed domain. [20]. These components are called the LH (low-high) sub-band, HL (high-low) sub-band, and HH (high-high) sub-band, respectively. Our proposed method using these compressed domain analyses has the following benefits. First, the image size is reduced (e.g., from 1,024 × 1,024 pixels to 512

× 512 pixels), but all needed information is retained to perfectly reconstruct the original image. After reconstruction, the ROI can be increased without losing information, which is proportional to the generalization performance. Second, the method is useful for classifying texture because the result of the 2D grey-level co-occurrence matrix (GLCM) in the wavelet domain can capture texture information from the wavelet sub-band according to the cancer grading [28]. We input all four DWT sub-bands in parallel to each separate segmentation model, U-Net++ [37]. We used the DiceCE loss function combined the Dice coefficients and the cross- entropies [38]. Each sub-band model took two NVidia Titan X GPUs. The total batch size was six for each GPU.

Prediction from tiles to whole images using wavelet weighted ensemble (WWE)

The reconstruction process is described here. After producing a whole probability map for each sub-band, as shown in Fig. 6(e), we applied ensemble learning based on wavelet weighted ensemble (WWE) to four trained neural networks for each sub-band (Fig. 8). Initially, a binary mask image (Fig. 8(b)) is obtained from the original image by using an Otsu algorithm (Fig. 8(a)) [33]. After a 2D wavelet transform based on the Haar wavelet, four wavelet sub-bands for the binary tissue mask were generated (Fig. 8(c)). We defined them as the wavelet weights, namely the LL weight, LH weight, HL weight, and HH weight. We added a small value, ε, to each wavelet weight, then multiplied it by their assigned weights (Fig. 8(d)). Lastly, we multiplied the weights by the corresponding probability map (Fig. 8(e)), and then applied an inverse discrete wavelet transform that also used the Haar wavelet to obtain a final probability map and overlay image (Fig. 8f-g). Parameters such as $W_1, W_2, W_3,$ and W_4 were empirically determined. Ideally, if the same region of each sub-band has a probability of 1, the reconstruction probability of that region should also be 1 without those parameters. However, we gave the LL sub-band more weight (i.e., 1.8) because the LL sub-band has a basic characteristic of the original image. Then, ε was added to remove the zero terms. The ensemble method is expressed by the following Equation (5):

$$R_{WWE} = \frac{1}{\sqrt{MN}} \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} W_1 (Y_{\psi}^{A} + \varepsilon) R_A(m,n) \psi_{1,m,n}^{A} + \frac{1}{\sqrt{MN}} \sum_{i=H,V,D} \sum_{m=0}^{M-1} \sum_{n=0}^{N-1} W_i (Y_{\varphi}^{i} + \varepsilon) R_i(m,n) \varphi_{1,m,n}^{i}, \quad (5)$$

where $\psi_{1,m,n}^{A}(x,y)$ and $\phi_{1,m,n}^{i}(x,y)$ represent 2D wavelet basis functions of level 1, Y_{ψ}^{A} describes an approximation of the binary tissue mask (LL sub-band weight), and Y_{ϕ}^{i} are high-frequency components (LH, HL, and HH sub-band weights) for the binary tissue mask whose directions are horizontal, vertical, and diagonal, respectively. R_A and R_i describe the probability map for each sub-band. R_{WWE} is the final prediction result after wavelet weighted ensemble (WWE).

$$\arg\max_{W \in \Re^4} f_d(R_{WWE}(R_A, R_H, R_V, R_D; W)). \quad (6)$$

To optimize the weight parameters such as W_1, W_2, W_3, and W_4, we applied optimization that satisfied Equation (6), where $\boldsymbol{W} = (W_1, W_2, W_3, W_4)$ and $f_d(\mathrm{x})$ is the function that decides the average Dice score of x. The range of each parameter is from 0.3 to 3.0, with a step size of 0.3. For comparison, Supplementary Table II shows the average Dice scores for W_1, W_2, W_3, and W_4. We chose the parameter values as $W_1 = 1.8$, $W_2 = 0.9$, $W_3 = 0.9$, and $W_4 = 0.3$.

Experimental setup

The qualities of the predictions were quantified by using the Dice score (Dice) and pixel accuracy (Acc), as follows:

$$Dice = \frac{2 \times N_{TP}}{2 \times N_{TP} + N_{FP} + N_{FN}}, \quad (7)$$

$$Acc = \frac{N_{TP} + N_{TN}}{N_{TP} + N_{TN} + N_{FP} + N_{FN}}, \quad (8)$$

where N_{TP}, N_{TN}, N_{FP}, and N_{FN} are the number of pixels for true-positive, false-positive, true-negative, and false-negative.

For the 39 WSIs test dataset, our proposed ensemble method (i.e., WWE) was compared with the model in three ways: (1) using the data after decimation and applying sliding window tile extraction instead of compression (Spatial domain, pixel size, 512 by 512 pixels; strides of sliding window, 128 pixels; x10 magnification), (2) using compressed data such as the LL, LH, HL, and HH sub-bands (Compressed domain), and (3) using the weighted average ensemble (WAE) for each sub-band result. The WAE is expressed as follows:

$$R_{WAE} = \frac{W_1 R_A + W_2 R_H + W_3 R_V + W_4 R_D}{W_1 + W_2 + W_3 + W_4}, \quad (9)$$

where R_A, R_H, R_V, and R_D describe the probability maps for each sub-band, and R_{WAE} is the final prediction result after the weighted average ensemble. R_A , R_H, R_V, and R_D describe probability the maps for the LL, LH, HL, and HH sub-bands, respectively. We set the same weight values in WAE as those in WWE ($W_1 = 1.8$, $W_2 = 0.9$, $W_3 = 0.9$, and $W_4 = 0.3$).

In order to verify the excellence of the proposed method, we progressed experiments as follows: (1) To compare average Dice and Acc according to each method, (2) To observe distribution of Dice and Acc according to all classes, (3) Check dice change of all classes according to low-frequency weight (W_1) and high-frequency weight (W_2, W_3, and W_4), (4) To compare sample images and its line profiles according to each method, (5) To compare with Dice of WWE, WAE, and the result in spatial domain according to threshold for tumor probability.

Reference

1. Yoshida, H. *et al.* Automated histological classification of whole slide images of colorectal biopsy specimens. *Oncotarget* **8**, 90719–90729 (2017).

2. Gertych, A. *et al.* Convolutional neural networks can accurately distinguish four histologic growth patterns of lung adenocarcinoma in digital slides. *Sci. Rep.* **9**, (2019).

3. Bera, K., Schalper, K. A., Rimm, D. L., Velcheti, V. & Madabhushi, A. Artificial intelligence in digital pathology — new tools for diagnosis and precision oncology. *Nat. Rev. Clin. Oncol.* **16**, 703–715 (2019).

4. Wang, S., Yang, D. M., Rong, R., Zhan, X. & Xiao, G. Pathology Image Analysis Using Segmentation Deep Learning Algorithms. *Am. J. Pathol.* **189**, 1686–1698 (2019).

5. Nagtegaal, I. D. *et al.* The 2019 WHO classification of tumours of the digestive system. *Histopathology* 0–1 (2019) doi:10.1111/his.13975.

6. Saha, M., Chakraborty, C. & Racoceanu, D. Efficient deep learning model for mitosis detection using breast histopathology images. *Comput. Med. Imaging Graph.* **64**, 29–40 (2018).

7. Yoon, H. *et al.* Tumor Identification in Colorectal Histology Images Using a Convolutional Neural Network. *J. Digit. Imaging* **32**, 131–140 (2019).

8. Kainz, P., Pfeiffer, M. & Urschler, M. Segmentation and classification of colon glands with deep convolutional neural networks and total variation regularization. *PeerJ* **5**, e3874 (2017).

9. Ho, D. J. *et al.* Deep Multi-Magnification Networks for Multi-Class Breast Cancer Image Segmentation. 1–12 (2019).

10. Komura, D. & Ishikawa, S. Machine Learning Methods for Histopathological Image Analysis. *Comput. Struct. Biotechnol. J.* **16**, 34–42 (2018).

11. Tokunaga, H., Teramoto, Y., Yoshizawa, A. & Bise, R. Adaptive weighting multi-field-of-view CNN for semantic segmentation in pathology. *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.* **2019-June**, 12589–12598 (2019).

12. Chang, H. Y. *et al.* Artificial Intelligence in Pathology. *J. Pathol. Transl. Med.* **53**, 1–12 (2019).

13. Thakur, N., Yoon, H. & Chong, Y. Current trends of artificial intelligence for colorectal cancer pathology image analysis: A systematic review. *Cancers (Basel).* **12**, 1–19 (2020).

14. Kayid, A. M. Performance of CPUs/GPUs for Deep Learning workloads. 25 (2018) doi:10.13140/RG.2.2.22603.54563.

15. Crochiere, R. E. & Rabiner, L. R. Interpolation and Decimation of Digital Signals—A Tutorial Review. *Proc. IEEE* **69**, 300–331 (1981).

16. Franco, M., Ariza-Araújo, Y. & Mejía-Mantilla, J. H. Automatic Image Cropping : A Computational Complexity Study Jiansheng. *Imagen Diagnostica* **6**, 49–56 (2015).

17. Brunton, S. L. & Kutz, J. N. Data Driven Science & Engineering - Machine Learning, Dynamical Systems, and Control. 572 (2017).

18. Carrillo-De-Gea, J. M., García-Mateos, G., Fernández-Alemán, J. L. & Hernández-Hernández, J. L. A Computer-Aided Detection System for Digital Chest Radiographs. *J. Healthc. Eng.* **2016**, (2016).

19. Liang, Y., Kong, J., Vo, H. & Wang, F. ISPEED: an Efficient In-Memory Based Spatial Query System for Large-Scale 3D Data with Complex Structures. *GIS Proc. ACM Int. Symp. Adv. Geogr. Inf. Syst.* **2017**-**Novem**, (2017).

20. Tang, J., Deng, C., Huang, G. Bin & Zhao, B. Compressed-domain ship detection on spaceborne optical image using deep neural network and extreme learning machine. *IEEE Trans. Geosci. Remote Sens.* **53**, 1174–1185 (2015).

21. Wang, J. Z., Nguyen, J., Lo, K. K., Law, C. & Regula, D. Multiresolution browsing of pathology images using wavelets. *Proc. AMIA Symp.* 430–434 (1999).

22. Zou, H., Hastie, T. & Tibshirani, R. Sparse principal component analysis. *J. Comput. Graph. Stat.* **15**, 265–286 (2006).

23. Ma, L. *et al.* Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS J. Photogramm. Remote Sens.* **152**, 166–177 (2019).

24. Falk, T. *et al.* U-Net: deep learning for cell counting, detection, and morphometry. *Nat. Methods* **16**, 67–70 (2019).

25. Williams, T. & Li, R. An Ensemble of Convolutional Neural Networks Using Wavelets for Image Classification. *J. Softw. Eng. Appl.* **11**, 69–88 (2018).

26. Liu, P., Zhang, H., Lian, W. & Zuo, W. Multi-Level Wavelet Convolutional Neural Networks. *IEEE Access* **7**, 74973–74985 (2019).

27. Jafari-Khouzani, K. & Soltanian-Zadeh, H. Multiwavelet grading of pathological images of prostate. *IEEE Trans. Biomed. Eng.* **50**, 697–704 (2003).

28. Bhattacharjee, S. *et al.* Multi-features classification of prostate carcinoma observed in histological sections: Analysis of wavelet-based texture and colour features. *Cancers (Basel).* **11**, 1–20 (2019).

29. Niwas, S. I., Palanisamy, P. & Sujathan, K. Wavelet based feature extraction method for Breast cancer cytology images. *ISIEA 2010 - 2010 IEEE Symp. Ind. Electron. Appl.* 686–690 (2010) doi:10.1109/ISIEA.2010.5679377.

30. Shaukat, A. *et al.* Automatic Cancerous Tissue Classification using Discrete Wavelet Transformation and Support Vector Machine. *J. Basic. Appl. Sci. Res* **6**, 1–1 (2016).

31. Cooley, J. W. & Tukey, J. W. An Algorithm for the Machine Calculation of Complex Fourier Series. *Math. Comput.* **19**, 297 (1965).

32. Sorensen, H. V., Jones, D. L., Heideman, M. T. & Burrus, C. S. Real-Valued Fast Fourier Transform Algorithms. *IEEE Trans. Acoust.* **35**, 849–863 (1987).

33. Nobuyuki Otsu. A Threshold Selection Method from Gray-Level Histograms. *IEEE Trans. Syst. Man Cybern* **9**, 62–66 (1979).

34. Tellez, D. *et al.* Quantifying the effects of data augmentation and stain color normalization in convolutional neural networks for computational pathology. *Med. Image Anal.* **58**, 101544 (2019).

35. Rabbani, M. & Joshi, R. *An overview of the JPEG 2000 still image compression standard*. *Signal Processing: Image Communication* vol. 17 (2002).

36. Lee, D., Choi, S. & Kim, H. J. High quality imaging from sparsely sampled computed tomography data with deep learning and wavelet transform in various domains. *Med. Phys.* **46**, 104–115 (2019).

37. Zhou, Z., Rahman Siddiquee, M. M., Tajbakhsh, N. & Liang, J. Unet++: A nested u-net architecture for medical image segmentation. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)* **11045 LNCS**, 3–11 (2018).

38. Isensee, F. *et al.* nnU-Net: Self-adapting framework for u-net-based medical image segmentation. *arXiv* (2018).

Acknowledgements

Funding: This work was supported by the National Research Foundation (NRF) grant (NRF-2019R1A2C2006269 and 2020M3H2A1078045) funded by Ministry of Science and ICT (MSIT), Institute of Information & communications Technology Planning & Evaluation (IITP) grant (No. 2019-0-01906, Artificial Intelligence Graduate School Program) funded by MSIT, Basic Science Research Program through the NRF grant (2018R1D1A1A02050922 and 2020R1A6A1A03047902) funded by the Ministry of Education, and BK21 Four project, Republic of Korea.

Author contributions

C. K. and Y. C. supervised the project. H. K. and H. Y. conceptualized and led the analysis. N. T., G. H., and E.J.L. obtained and annotated the data. All authors contributed to writing the manuscript and have approved the submission version.

Competing interests statement

The authors declare no competing interests.

Figure legends

Fig. 1 Comparison of Dice and Acc for each class between the result in spatial domain, the result in compressed domain (such as the LL, LH, HL, and HH sub-bands), the weighted average ensemble (WAE) model using the wavelet sub-bands, and the wavelet weighted ensemble (WWE) model using the wavelet sub-bands.

Fig. 2 Change in Dice in all tumor classes according to the low-frequency weight (W_1) in the WWE.

Fig. 3 (a) Annotated image, corresponding magnified image (above), and Dashed red line profiles (below). (b) Prediction result in spatial domain. (c) Weighted average ensemble (WAE) result for wavelet sub-band. (d) Wavelet weighted ensemble (WWE) result for wavelet sub-band.

Fig. 4 Comparative Dice of wavelet weighted ensemble (WWE) for wavelet sub-bands, weighted average ensemble (WAE) for wavelet sub-bands, and the result in spatial domain according to the threshold for tumor probability (95% confidence interval).

Fig. 5 Prediction results of the five different tumor classes. (a)-(e) are tissue histology images. (f)-(j) are annotation by a pathologist (i.e., ground truth). (k)-(o) are predicted probability map (WWE). (p)-(t) are overlaid tissue histology image and prediction heatmap. Yellow dotted boxes in (o) and (t) show misprediction due to dead nuclei.

Fig. 6 Overall flow chart of the proposed method. (a) Tile extraction based on a sliding window. (b) Image depth compression. (c) Forward transform to the compressed domain. (d) Training and

prediction using convolutional neural networks. (e) Prediction from one tile to the whole image. (f) Wavelet-weighted ensemble (WWE).

Fig. 7 The principal component analysis (PCA) for z-axis compression. (a) The original tile composed of RGB-channels. (b) The resulting tile composed of single channel. (c) RGB value distribution in a tissue region. (d) The PCA for the R and B channels.

Fig. 8 Schematic of wavelet weighted ensemble (WWE) for one low pass sub-band (LL sub-band) and three high pass sub-bands (LH, HL, and HH) based on discrete wavelet transform (DWT) and inverse discrete wavelet transform (IDWT). (a) Original image. (b) Binary tissue mask. (c) Initial wavelet weights. (d) Additional parameters. $\varepsilon = 0.1$, W1 = 1.8, W2= 0.9, W3= 0.9, and W4= 0.9. (e) Raw probability map for each sub-band. (f) Final probability map. (g) Overlaid image.

Tables

Method	Dice	Acc
w/o compression	0.788 ± 0.101	0.946 ± 0.032
LL Sub-band	0.844 ± 0.089	0.960 ± 0.028
LH Sub-band	0.712 ± 0.189	0.921 ± 0.053
HL Sub-band	0.748 ± 0.165	0.924 ± 0.057
HH Sub-band	0.724 ± 0.193	0.931 ± 0.045
WAE	0.787 ± 0.141	0.937 ± 0.039
WWE (proposed)	0.852 ± 0.086	0.962 ± 0.027

Table I Average Dice and Acc values for the result in the spatial domain, the model using compressed data (such as LL, LH, HL, and HH sub-band), the weighted average ensemble (WAE) model using the wavelet sub-bands, and the wavelet weighted ensemble (WWE) model using the wavelet sub-bands.

	Spatial domain	Compressed domain
Magnification	×10	×20
Initial tile size	512 by 512	1024 by 1024
ROI size	$6.25 \times 10^{-2}\ \mu m^2$	$6.25 \times 10^{-2}\ \mu m^2$
Number of GPUs	1	4

Table II Training conditions for magnification, initial tile size, ROI size, and the number of GPUs.

Figures

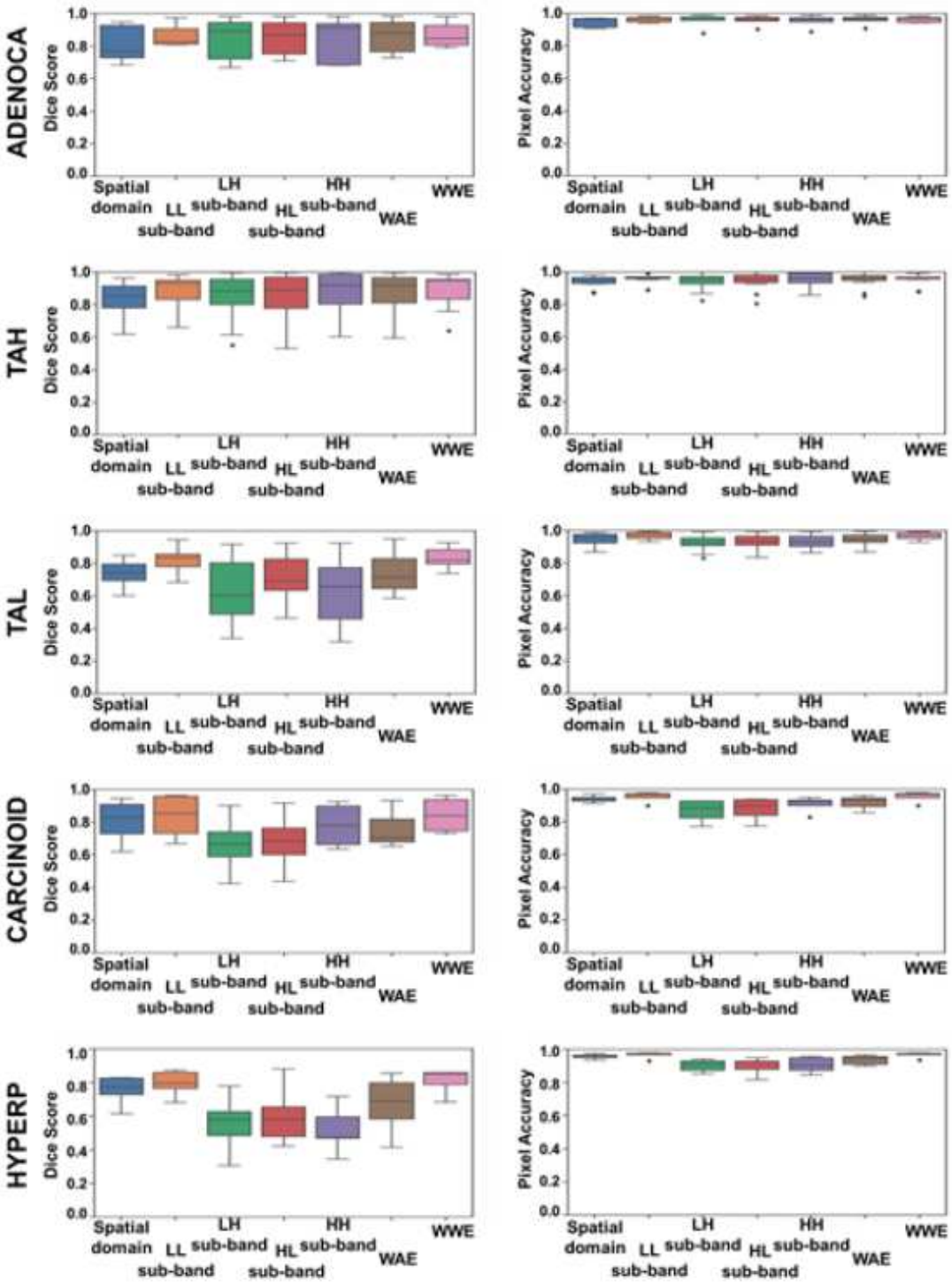

Figure 1

Comparison of Dice and Acc for each class between the result in spatial domain, the result in compressed domain (such as the LL, LH, HL, and HH sub-bands), the weighted average ensemble (WAE) model using the wavelet sub-bands, and the wavelet weighted ensemble (WWE) model using the wavelet sub-bands.

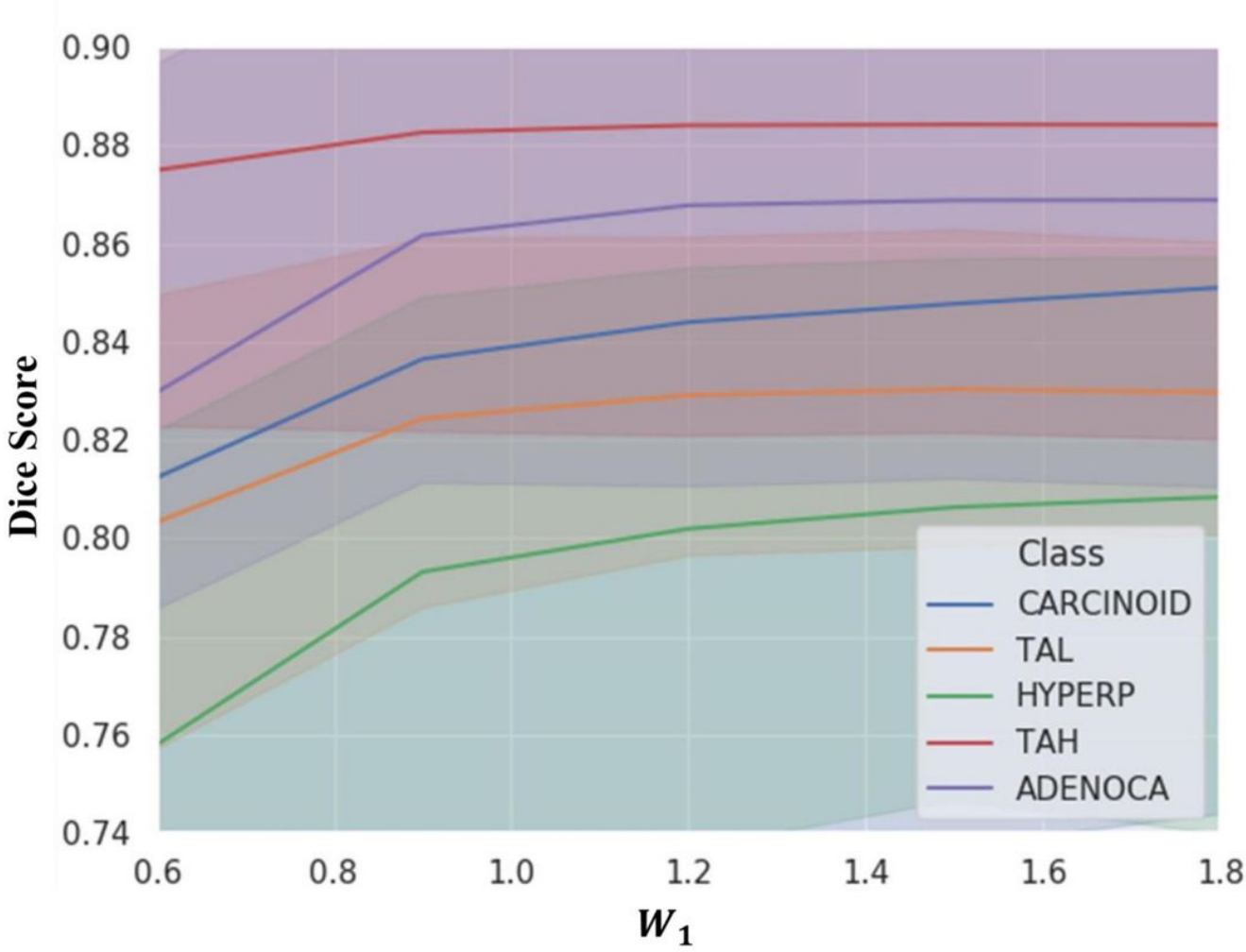

Figure 2

Change in Dice in all tumor classes according to the low-frequency weight (1) in the WWE.

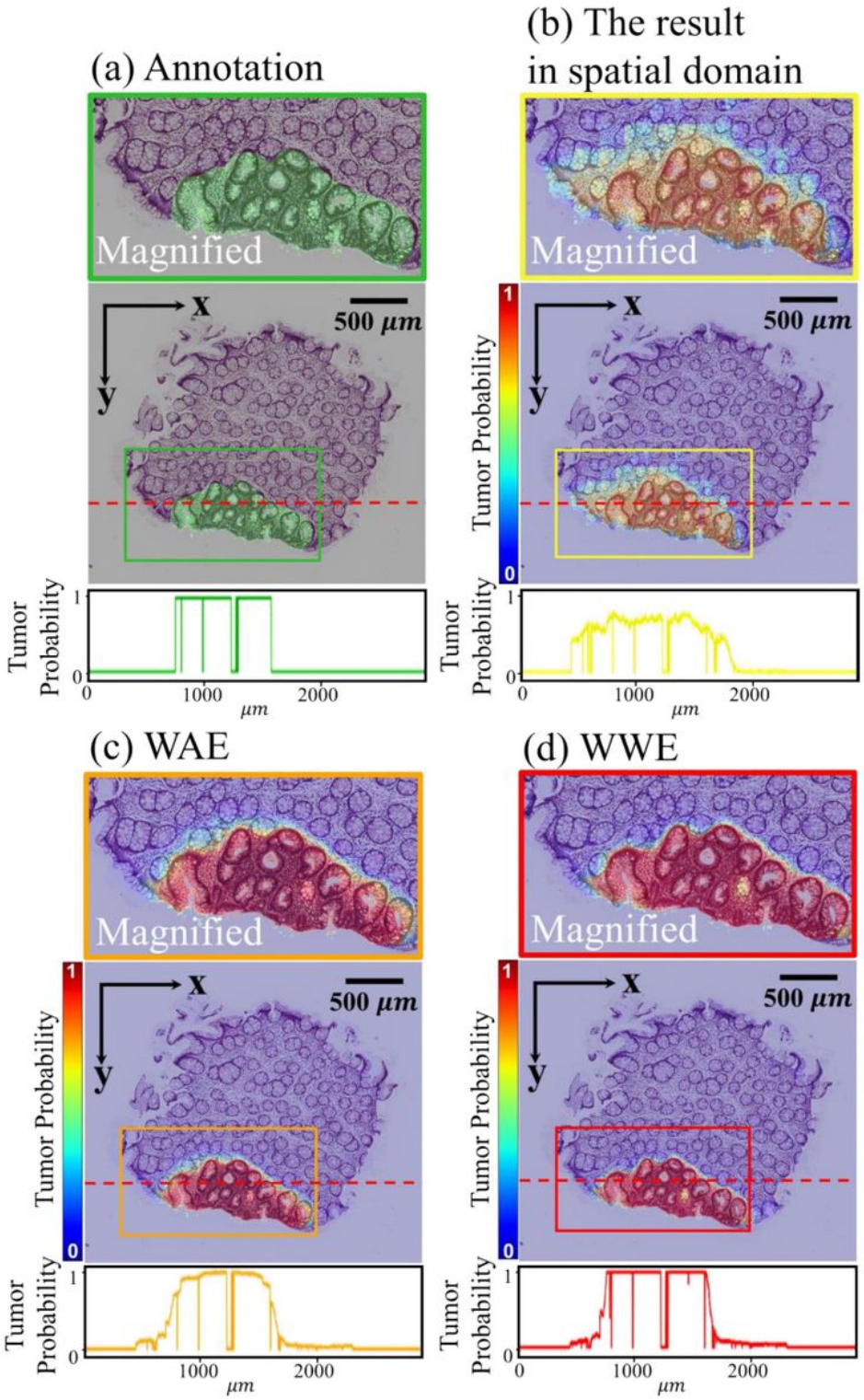

Figure 3

(a) Annotated image, corresponding magnified image (above), and Dashed red line profiles (below). (b) Prediction result in spatial domain. (c) Weighted average ensemble (WAE) result for wavelet sub-band. (d) Wavelet weighted ensemble (WWE) result for wavelet sub-band.

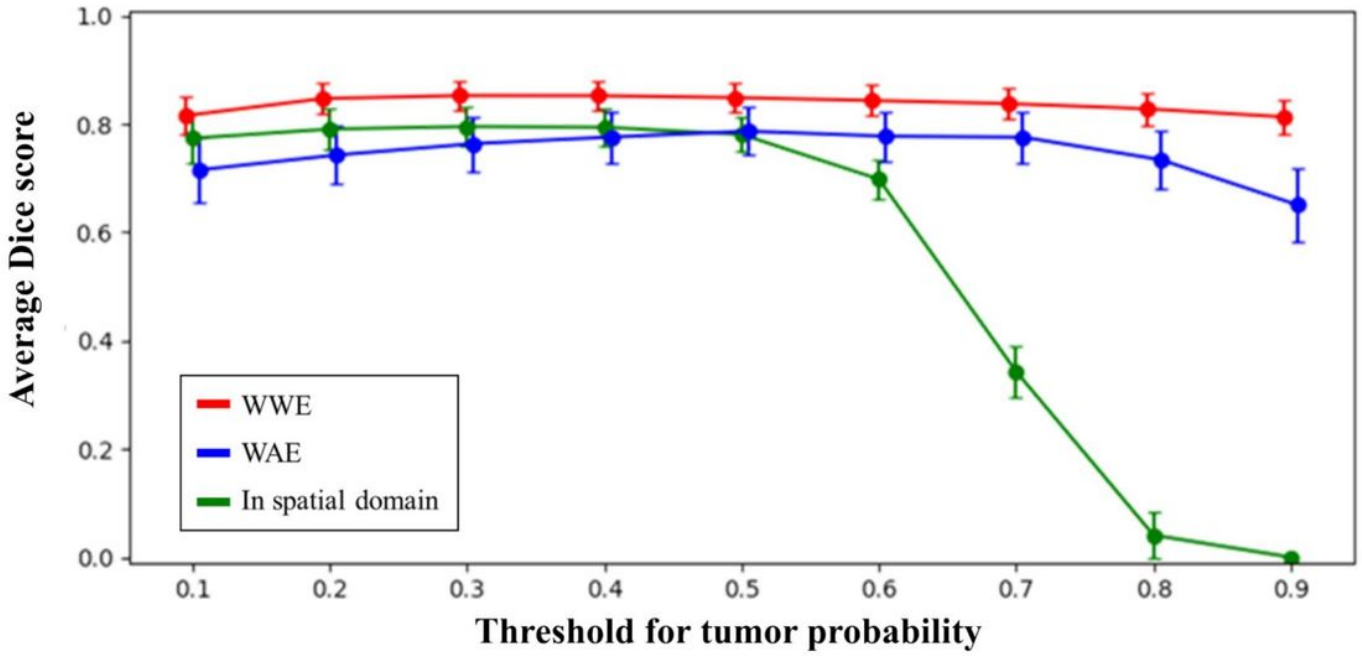

Figure 4

Comparative Dice of wavelet weighted ensemble (WWE) for wavelet sub-bands, weighted average ensemble (WAE) for wavelet sub-bands, and the result in spatial domain according to the threshold for tumor probability (95% confidence interval).

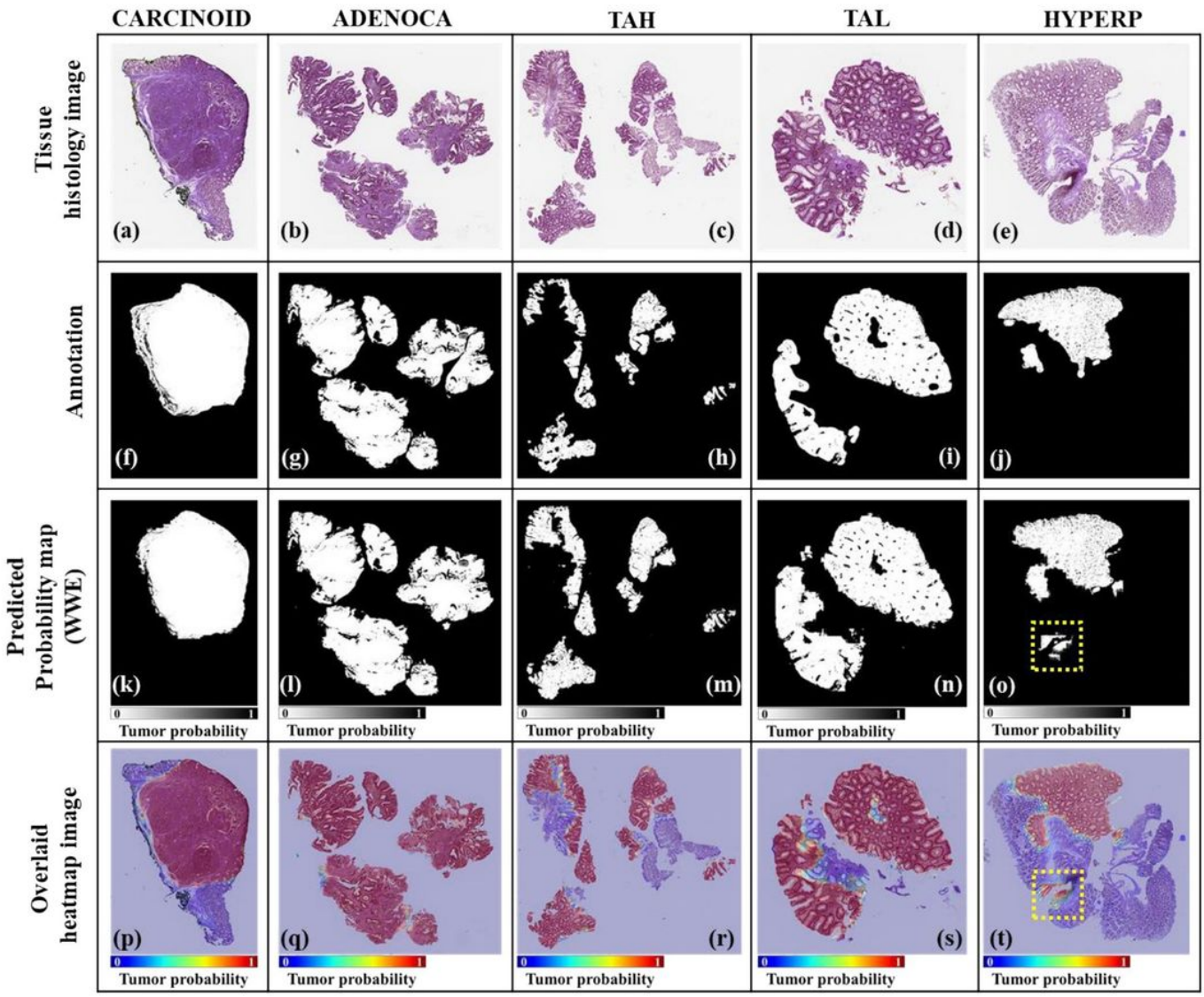

Figure 5

Prediction results of the five different tumor classes. (a)-(e) are tissue histology images. (f)-(j) are annotation by a pathologist (i.e., ground truth). (k)-(o) are predicted probability map (WWE). (p)-(t) are overlaid tissue histology image and prediction heatmap. Yellow dotted boxes in (o) and (t) show misprediction due to dead nuclei

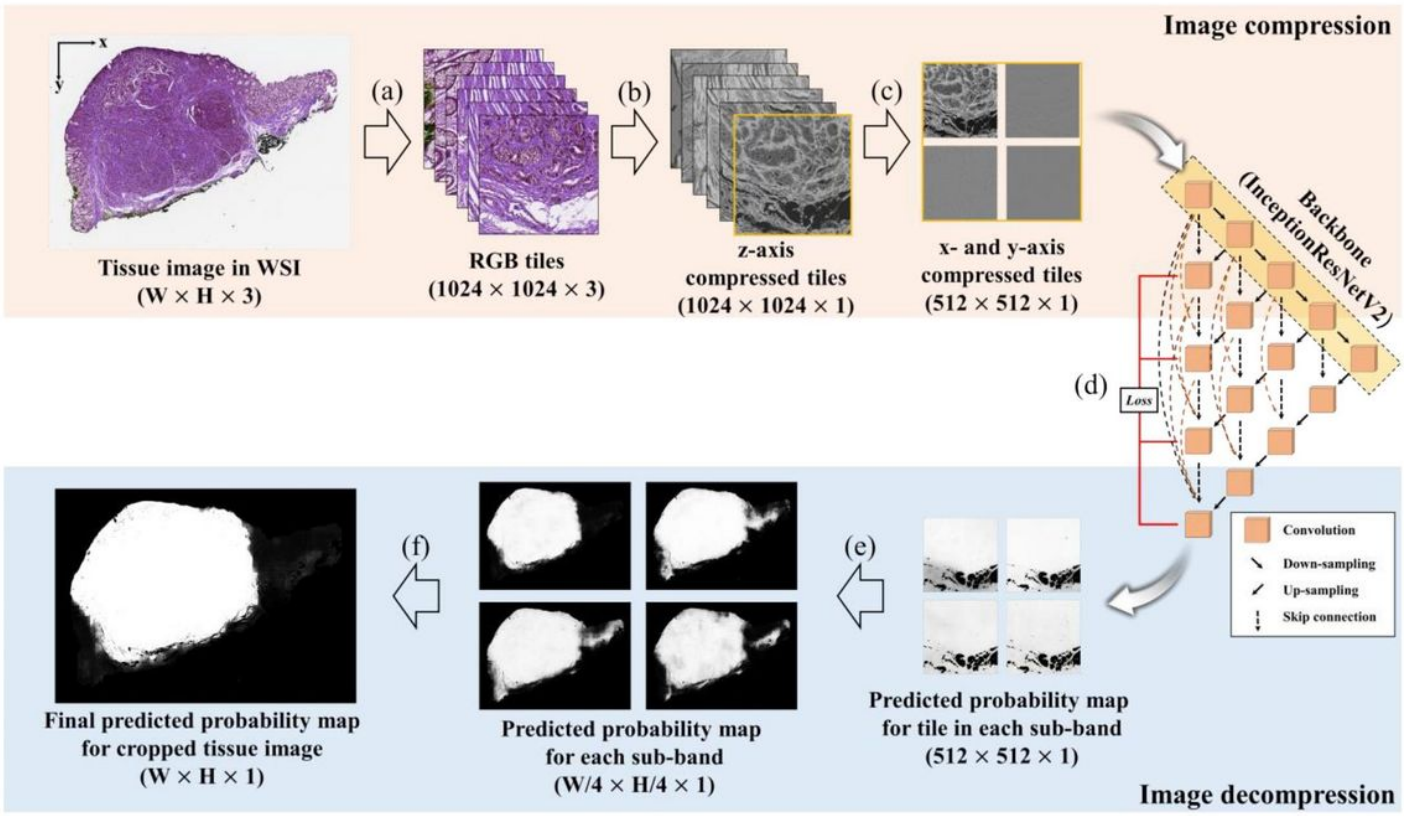

Figure 6

Overall flow chart of the proposed method. (a) Tile extraction based on a sliding window. (b) Image depth compression. (c) Forward transform to the compressed domain. (d) Training and prediction using convolutional neural networks. (e) Prediction from one tile to the whole image. (f) Wavelet-weighted ensemble (WWE)

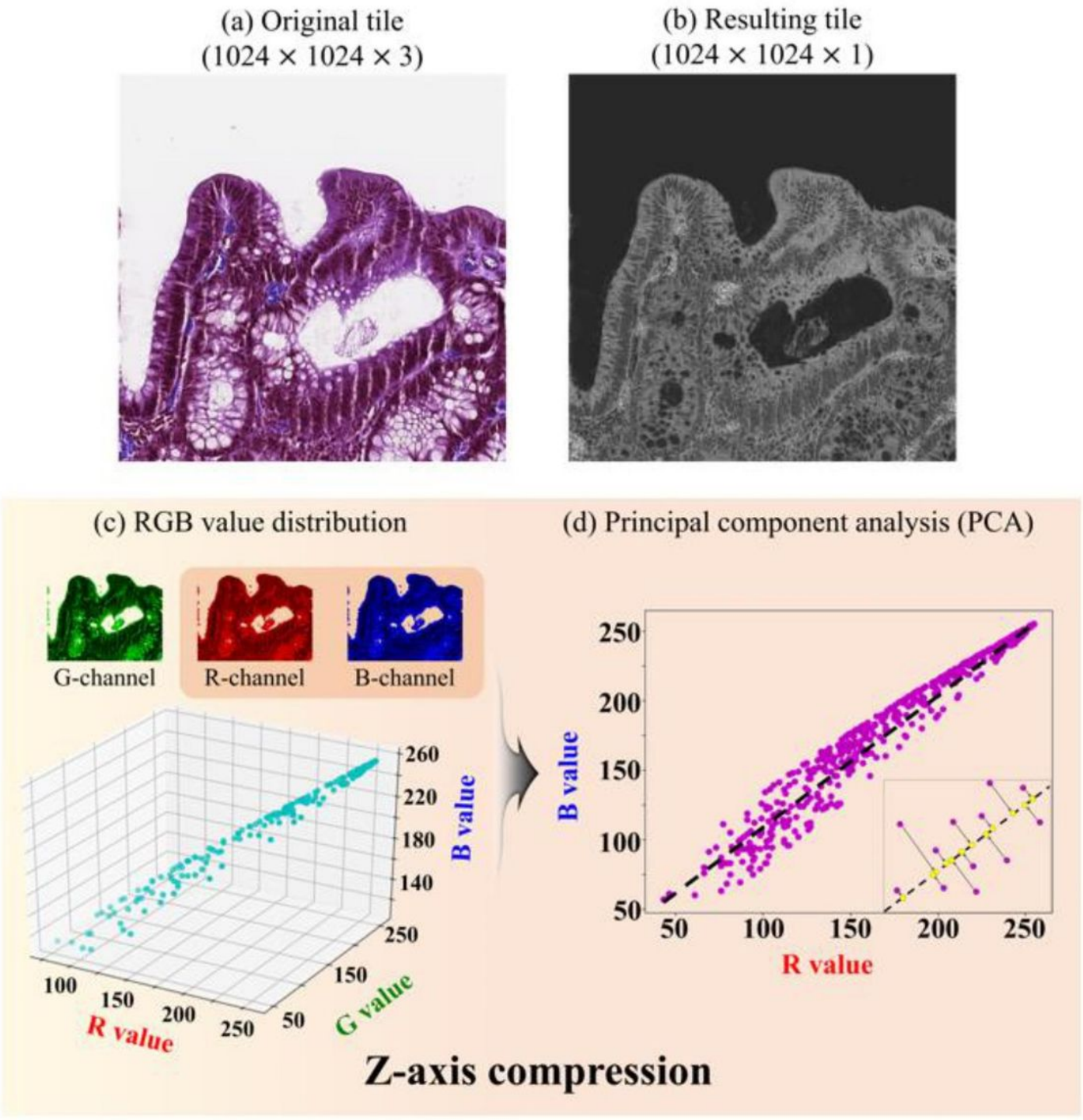

Figure 7

The principal component analysis (PCA) for z-axis compression. (a) The original tile composed of RGB-channels. (b) The resulting tile composed of single channel. (c) RGB value distribution in a tissue region. (d) The PCA for the R and B channels

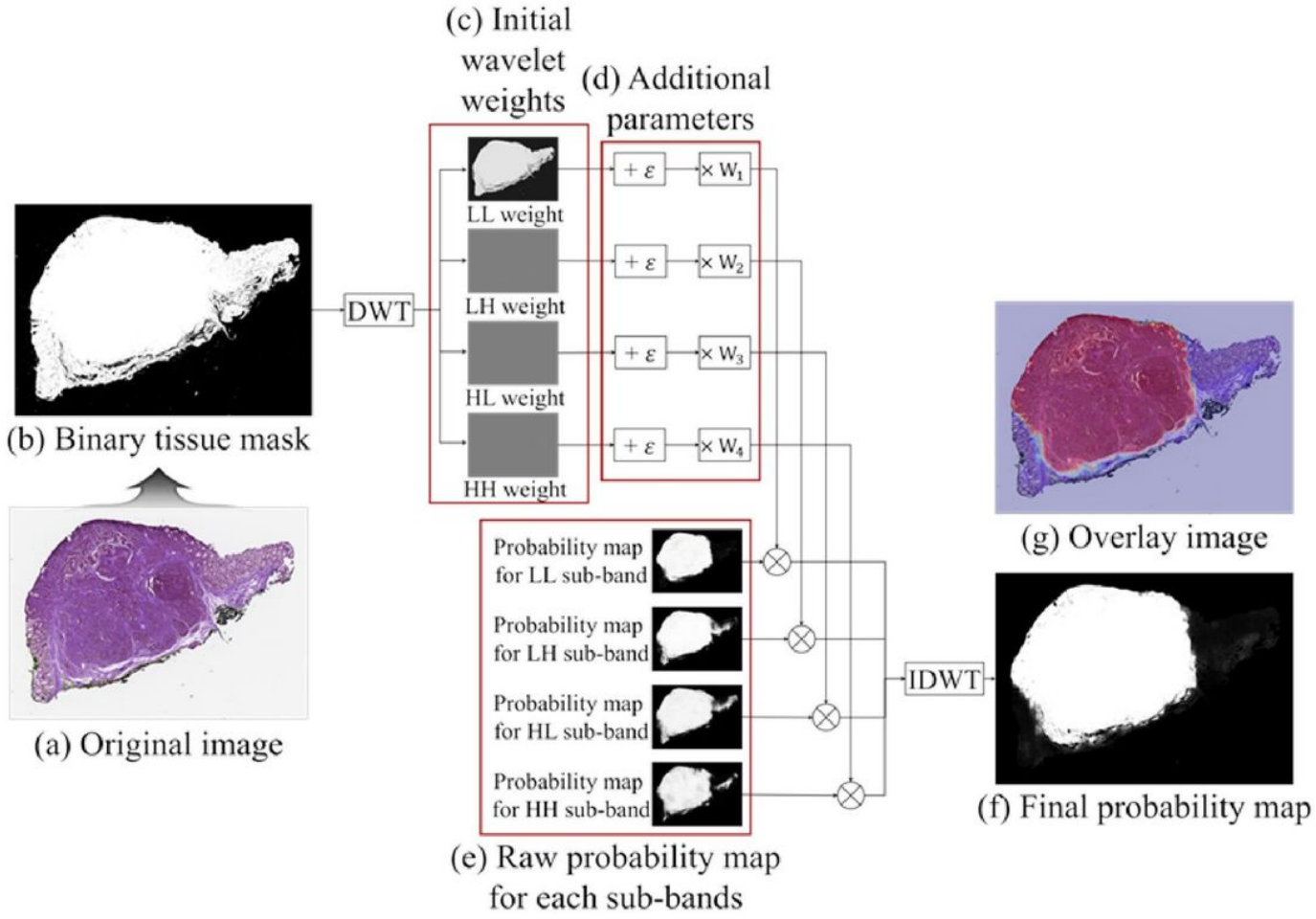

Figure 8

Schematic of wavelet weighted ensemble (WWE) for one low pass sub-band (LL subband) and three high pass sub-bands (LH, HL, and HH) based on discrete wavelet transform (DWT) and inverse discrete wavelet transform (IDWT). (a) Original image. (b) Binary tissue mask. (c) Initial wavelet weights. (d) Additional parameters. ε = 0.1, W1 = 1.8, W2= 0.9, W3= 0.9, and W4= 0.9. (e) Raw probability map for each sub-band. (f) Final probability map. (g) Overlaid image

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- 20210531SupplementaryinfoSR.pdf