



# HIC-net: A deep convolutional neural network model for classification of histopathological breast images<sup>☆</sup>



Şaban Öztürk<sup>a,\*</sup>, Bayram Akdemir<sup>b</sup>

<sup>a</sup> Department of Electrical and Electronics Engineering, Amasya University, Amasya, 05001, Turkey

<sup>b</sup> Department of Electrical and Electronics Engineering, Selçuk University, Konya, 42030, Turkey

## ARTICLE INFO

### Article history:

Received 13 August 2018

Revised 11 April 2019

Accepted 12 April 2019

### Keywords:

Cancer classification

Convolutional neural networks

CNN

Histopathological image

Whole-slide

## ABSTRACT

In this study, a convolutional neural network (CNN) model is presented to automatically identify cancerous areas on whole-slide histopathological images (WSI). The proposed WSI classification network (HIC-net) architecture performs window-based classification by dividing the WSI into a certain plane. In our method, an effective pre-processing step has been added for WSI for better predictability of image parts and faster training. A large dataset containing 30,656 images is used for the evaluation of the HIC-net algorithm. Of these images, 23,040 are used for training, 2560 are used for validation and 5056 are used for testing. HIC-net has more successful results than other state-of-art CNN algorithms with AUC score of 97.7%. If we evaluate the classification results of HIC-net using softmax function, HIC-net success rates have 96.71% sensitivity, 95.7% specificity, 96.21% accuracy, and are more successful than other state-of-the-art techniques which are used in cancer research.

© 2019 Elsevier Ltd. All rights reserved.

## 1. Introduction

Histopathologic imaging is a very useful method which is used by pathologists to examine cell behavior and diagnosis. However, the evaluation process becomes elongated and difficult because of the high number of cell and cell types in these images. In order to undertake this problem, where early diagnosis is extremely important, a lot of effort is exerted to produce fast examination process. Today, many improvements have been made in digital imaging devices to accelerate the diagnostic process. In addition, important developments in artificial intelligence methods are followed and used in the health field. Histopathological images consist of many cells and tissues. The condition of these cells and tissue fragments gives information about the disease. In some cases, unusual cell division and histopathological manifestations are known as tumor growth, but this is not always cancer. Tumor is a mass when it is generally defined, and each tumor does not cause cancer. However, when cells divide uncontrollably and spread (malignant tumor), they cause cancer. Thousands of cells and tissue areas in whole-slide images must be carefully examined to interpret these tumor cells. In some cases, misdiagnosis can be caused by negative factors such as fatigue and inexperience for pathologist. In order to prevent this negative situation, machine learning methods are used as an adjunct to the decision making process or a main diagnostic method. Given the advantages of the machines, the collaboration of experts with these methods is increasing the success in many cases [1].

<sup>☆</sup> This paper is for CAEE special section SI-mip. Reviews processed and recommended for publication to the Editor-in-Chief by Guest Editor Dr. Li He.

\* Corresponding author.

E-mail address: [saban.ozturk@amasya.edu.tr](mailto:saban.ozturk@amasya.edu.tr) (Ş. Öztürk).

Whole-slide images have a very high resolution and a complex background. For this reason, they contain a lot of cells and a misleading background texture. This complexity is a very difficult event to solve by simple image processing methods. Therefore, many noises and unwanted factors in the image are eliminated using image preprocessing techniques. Pre-processing methods are very useful for hand-crafted property extraction methods and have been preferred over many years. However, the emergence of automatic feature extraction algorithms and their success have reduced the use of hand-crafted feature extraction methods. Image pre-processing algorithms can be used with automatic feature extraction methods, but in this case, there is a problem. Because automatic feature extraction algorithms automatically generate features, every information in the image is important. The noise and information eliminated by pre-processing methods can reduce success in some cases. Therefore, the use of pre-processing algorithms with automatic feature extraction methods requires experience. Pre-processing is usually useful for images that contain important factors, such as a complex background pattern and number of target objects [2]. CNN is used as the most effective automatic feature generation technique for image processing applications. CNN became very popular after Krizhevsky et al. won the ImageNet competition [3]. The CNN architecture is based on core layers, so it is very suitable for image processing. These kernels are shifted over the image. As a result, it significantly reduces the number of parameters in the architecture. CNN architecture, which has made inspiring successes for various image processing problems, has entered the field of medical image processing fairly quickly. Many image processing and machine learning researchers who care about human health conduct studies to facilitate difficult situations and the decision-making process of pathologists. The analysis of histopathological images has long been a subject of considerable interest in the literature. Early studies have used various classification algorithms with hand-crafted features. Developments in histopathological image analysis have been observed following developments in machine learning methods. Sertel et al. [4] proposed computer-assisted grading method using the statistical framework. Recommended level co-occurrence matrix method based technique has the ability to make low and high grade classification. Al-Kadi [5] introduced optimum texture measure combination method to select the best channel by separating the channels. This method, which also includes hand-crafted features, is able to detect meningioma tumors. Raza et al. [6] have used the bag-of-features method to reveal medical image characteristics. They use scale and rotation invariant features to detect renal cell carcinoma. Zhang et al. [7] proposed Gaussian-based hierarchical voting and repulsive balloon model for analysis of histopathological images. One of the strengths of the proposed method is that it can operate in real time. As a result of the impressive results of automatic feature extraction techniques, these techniques have been used in the field of histopathological image processing, as in many image processing areas. Among these methods, the most striking is the CNN architecture, which is tailored for image processing problems. CNN architecture can produce highly successful results and can work in real time. Vu et al. [8] have used a discriminative feature-oriented dictionary learning technique. Images are more versatile with dictionary learning method which is an automatic feature extraction method. Coudray et al. [9] conduct a study for the diagnosis of Adenocarcinoma (LUAD) and squamous cell carcinoma (LUSC) by training a deep convolutional neural network (inception v3). Average area under the curve (AUC) value of the proposed method is 0.97. Zhang et al. [10] proposed multi-scale non-negative sparse coding in order to fill the semantic gap problem between low level features and high level features. Cruz-Roa et al. [11] proposed a deep learning architecture for detection of basal-cell carcinoma cancer. They recommend a novel deep learning technique that emphasizes visual patterns in the histopathological image. Xu et al. [12] introduced a deep convolutional neural network model for segmenting and classifying epithelial and stromal regions. Zhang et al. [13] have used deep learning based feature representation for skin biopsy histopathological images. Their method is a mechanism that learns potential high-level properties for local regions in skin biopsy images. Bayramoglu et al. [14] propose a deep learning method for magnification independent breast cancer analysis. Their work includes two separate architectures, single task and multi-task CNN. Huang et al. [15] introduced convolutional neural networks and self-taught learning for Epithelium-stroma classification. Zheng et al. [16] proposed nucleus-guided convolutional neural network for feature extraction from histopathological images. Their three hierarchy structure is trained with the cell nuclei. Gummeson et al. [17] have used deep CNN to classify prostate cancer. Their method has an error rate of 7.3% for benign and Gleason grade 3, 4, 5. Sudharshan et al. [18] use multiple instance learning method with CNN and other classifiers. According to their research, their results are more successful than conventional methods. Eycke et al. [19] propose a deep learning technique to segment glandular epithelium in histological images. Saha et al. [20] suggest a deep learning architecture consists of five convolution layers, four pooling layers, four rectified linear units (ReLU), and two fully connected layers to detect mitosis.

When the methods in the literature are examined, the success and speed of classification of CNN architecture are quite remarkable compared to other methods. CNN architecture has been selected to benefit from the incentive advantages of CNN architecture. But in this study, unlike the studies in the literature, an architecture has been developed specifically for the texture of histopathological images. For this purpose, a preprocessing step based on the effect of H&E stained components and the behavior of cells is added as the first layer of CNN architecture. The recommended pre-processing method is normalizing the images for CNN input. Firstly, histopathological images with high resolution are divided into small pieces ( $128 \times 128$  pixels) in a specific order. Each piece of image is applied to the network input by applying the proposed medical image pre-processing algorithm. On this count, training becomes more convenient for the network. Proposed structure includes 6 convolution layers, 6 ReLU layers, 6 pooling layers, FCL with 5 hidden layers, L2 regularization layer and dropout layer. The proposed histopathological image pre-processing technique and CNN structure are combined and named HIC-net. This method removes annoying backgrounds and makes problem generalization and learning easier and faster.

The contributions of this study can be listed as follows;

**Algorithm 1**

Pre-processing process.

**Inputs:** Histopathological image patch ( $I$ )**Outputs:** Normalized image ( $h_5$ )Initialize variables:  $h_1(x,y)$ ,  $h_2(x,y)$ ,  $h_3(x,y)$ ,  $h_4(x,y)$ ,  $h_5(x,y)$ ;

Begin:

Read the input image  $I$ 

$$h_1(x,y) = I * \frac{e^{-\frac{(x^2+y^2)}{2\sigma^2}}}{\sum_x \sum_y e^{-\frac{(x^2+y^2)}{2\sigma^2}}}, \sigma = 2 \quad (\text{Smoothing with Gaussian filter})$$

$$h_2(x,y) = h_1(x,y) - \text{median}(h_1) \quad (\text{Remove median value from smoothened image})$$

$$h_3(x,y) = I(x,y) - h_2(x,y) \quad (\text{Obtained image removes the original image})$$

$$h_4(x,y) = h_3(x,y) - f_{\text{smooth}}(h_3(x,y)) \quad (\text{Image is sharpened})$$

$$h_5(x,y) = f_{\text{median}}[5,5](h_4(x,y)) \quad (5 \times 5 \text{ 2D Median filter})$$

Stop

- Through the histopathological pre-processing step, which is the most important contribution, the learning ability of the network is increased and the training process is accelerated.
- The proposed CNN architecture is specially developed for H&E stained whole-slide histopathological images.
- The recommended method can be applied to all whole-slide images. For this, the customization step must be accompanied by special small changes. These changes are related to H&E staining and type of cells.
- The pre-processing step can be fitted to other machine learning techniques.

The rest of this paper is organized as follows: In [Section 2](#), the proposed HIC-net structure and details are mentioned. Also, it contains details of HIC-net parameter selections and experiments. In [Section 3](#), experimental results are presented and compared. Finally, conclusions are discussed in [Section 4](#).

## 2. Materials and methods

The analysis of histopathological images is quite challenging for both pathologist and machine learning algorithms. For pathologists, even a single analysis takes hours to analyze whole-slide tissue. Although machine learning algorithms produce faster results, many algorithms cannot achieve the desired level of performance. The reason for this is generally high resolution and disruptive factors.

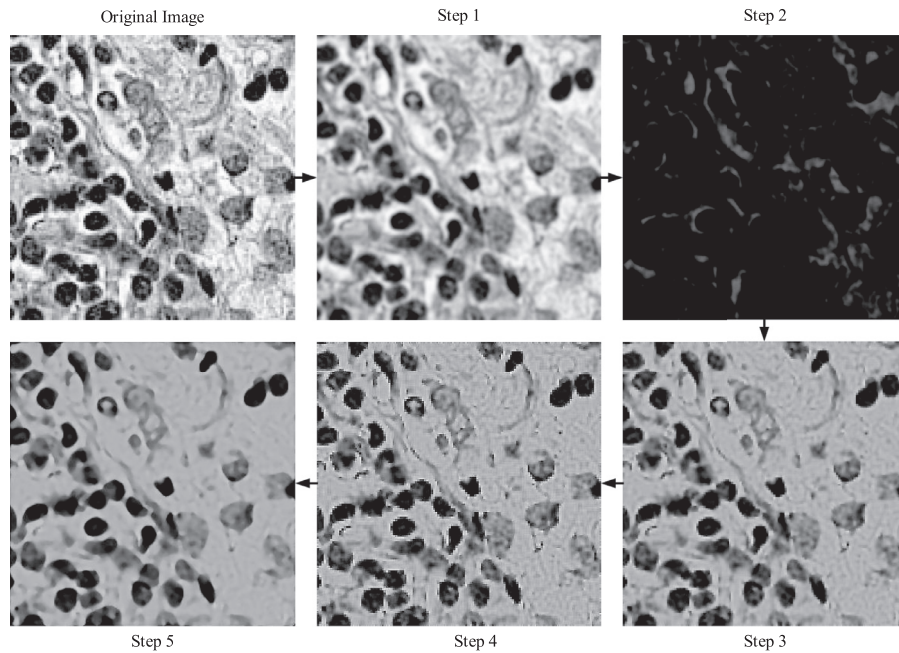
In the proposed algorithm, a highly effective pre-processing method has been introduced to remove the effect of disturbing factors and to facilitate the learning of the network. The intent of this method is to remove the fluctuations in the background gray level in the images and make the cells more visible. After pre-processing, non-cancerous cells are seen to be darker and smaller, and cancerous cells are found to be faded and larger. Another advantage of the pre-processing process is the suppression of gray value changes in the background and sudden white areas. These color changes are usually not directly linked to cancer. This fluctuation and balancing of the changes allow the CNN structure to focus only on the cells.

### 2.1. Histopathological image pre-processing and parameter selection

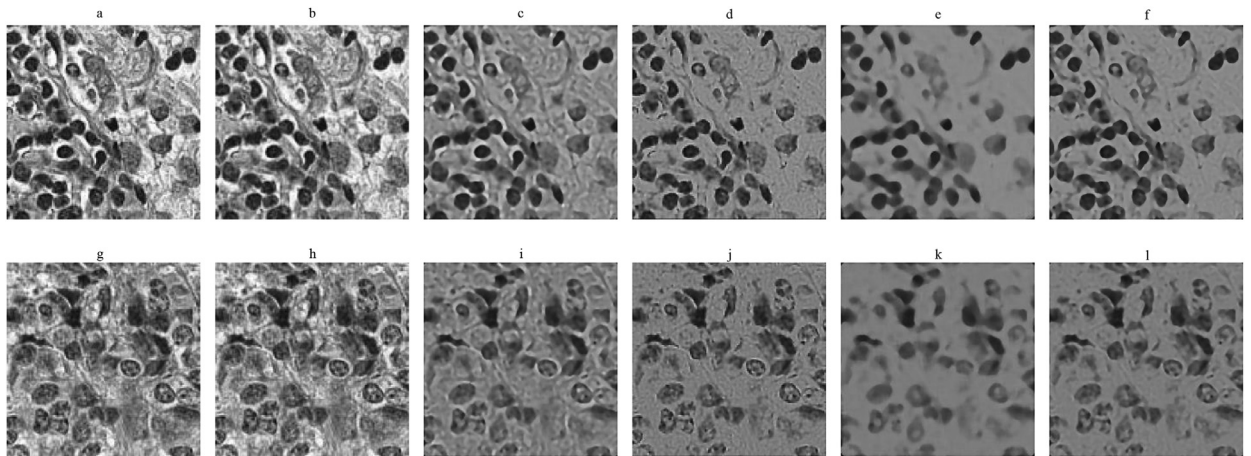
In the pre-processing process, the image is first smoothed by applying a Gaussian filter. The Gaussian filter helps to remove the association of the gray gradients in the background and the sharp transitions of the cells. Then, the average value of the pixels in the background image is calculated. The calculated average value is mathematically subtracted from all pixel values in the corrected image to remove background noise. All the regions where the effect of noises is high are determined. This new image is removed from the original image and the noise and sudden color glare in the image is compensated. Then, the noise is removed by a  $5 \times 5$  neighboring 2D median filter. However, the details are clarified so that it does not get lost before the 2D median filter. The proposed pre-processing method emphasizes the basic properties of the cells and is very effective in removing the disruptive factors such as fluctuations and noises in the background. In this regard, the learning of the CNN structure is accelerated. Because the CNN structure does not try to get meaningless features like the changes in the background. The algorithm of proposed histopathological image pre-processing method is as follow:

[Fig. 1](#) shows the visual responses of the operations in [Algorithm 1](#).

Pre-processing algorithms are usually used to improve performance in image segmentation and classification tasks. The intention of the pre-processing algorithms is to reduce the effect of noise and remove disturbing factors in the image. However, this situation may cause loss of useful information in detailed images. Pre-processing mostly can be used with handcrafted feature extraction algorithms. In contrast, deep learning algorithms can process raw input data and produce deep features. Automatic feature extraction algorithms can work without pre-processing, but the image must be normalized for the learning process to take place quickly. For this reason, the histopathological image pre-processing technique has been proposed in this study. First, the image is softened by applying a Gaussian filter with a sigma value of 2. If the sigma value is chosen to be a larger value, the intracellular gray values will affect the background total change. Small sigma values are limited to local regions only. Then, the median value is subtracted from the smoothed image. Because, the median value



**Fig. 1.** Histopathological image pre-processing steps in Algorithm 1.



**Fig. 2.** Histopathological image pre-processing, (a, g) original images, (b, h) small Gaussian neighborhood, (c, i) large Gaussian neighborhood, (d, j) small median neighborhood, (e, k) large median neighborhood, (f, l) our parameters.

includes overall recalculation of the intensity distribution of the image background. This effect is shown in Step 2 in Fig. 1. The resulting new image is removed from the original image. This means the background of the image is removed from the original image. Finally, the image is sharpened and a 2D median filter with  $5 \times 5$  adjacency is applied. If the median filter neighborhood value is increased, the cells appear adjacent to each other. If the median filter is reduced by the neighborhood value, the background details cannot be sufficiently resolved. The effects of parameter selection are shown in Fig. 2.

## 2.2. HIC-net structure and parameter selection

In the classification stage of histopathological images, the deep convolutional neural network structure is performed because of complex structure of histopathological images. The types, the numbers and the sequence of the cells in the image patches cannot be predicted. Each image has its own specific cell behavior and the structure changes for mitotic divisions. These make classification difficult. For this reason, there is a need for a classifier algorithm that automatically learns advanced features for the classification process and that can be achieved with high success. Our deep convolutional neural network contains convolutional neural network structure includes 6 convolution layers, 6 ReLU layers, 6 pooling layers, FCL with 5 hidden layers, L2 regularization layer and dropout layer. The proposed method is shown in Fig. 3.

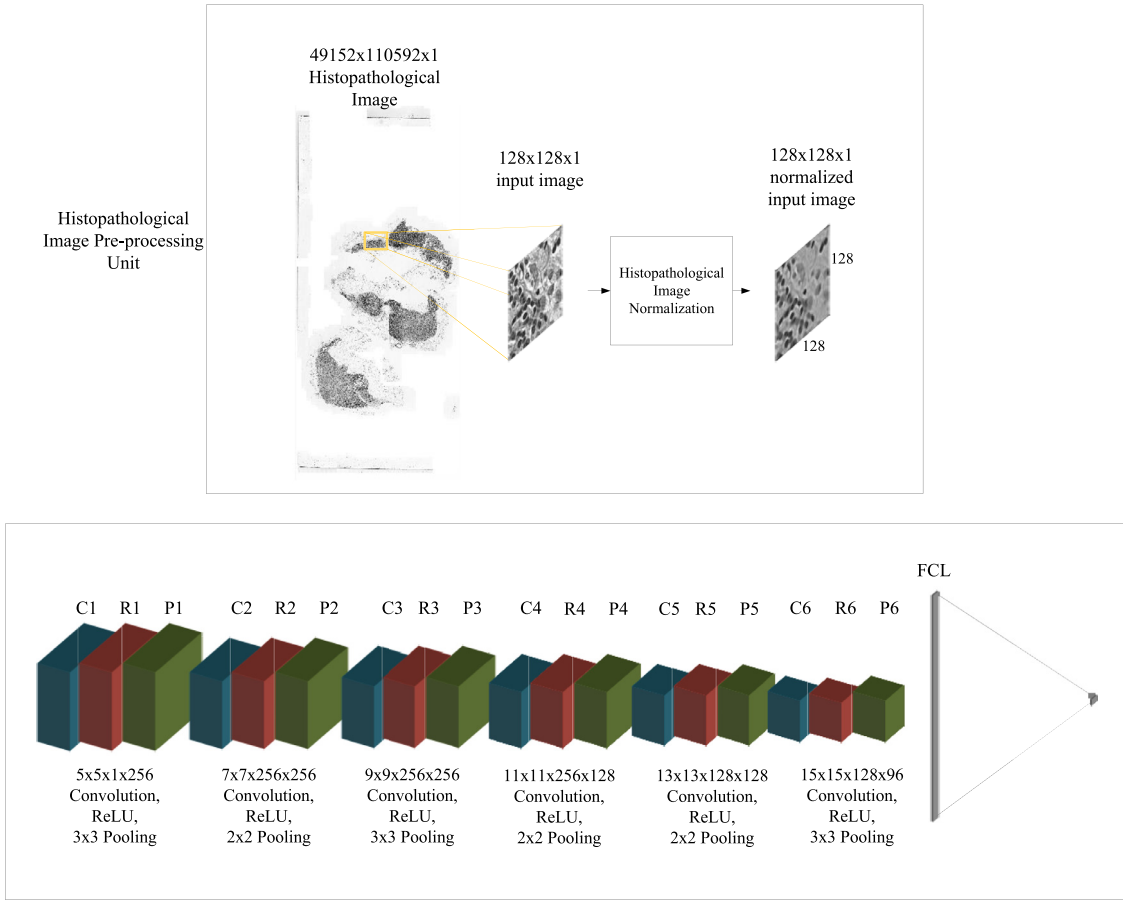


Fig. 3. Proposed HIC-net structure.

CNN structure is a very powerful algorithm for image processing problems. This is because of the advantages of the layers it contains. Convolution layer consists of feature matrices. Feature matrices contain important property information about the image. These feature matrices enhance success by updating after every iteration. In addition, feature matrices in the convolution layer are applied by convolution operation as in Eq. (1). The most important feature of this process is the sharing of features. In this case, the number of parameters in the network is reduced.

$$I_i^l = f \left( \sum_j I_j^{l-1} \otimes w_{ij}^l + b_i^l \right) \quad (1)$$

in which  $I_i^l$  represents the output matrix,  $w_{ij}$  represents convolutional kernel and  $b_i$  represents bias parameter.  $f$  is activation function.

Convolution layer is the basic layer used to automatically learn features and visualize learned features. However, it is inefficient to create a CNN network with only convolution layers. Because, the image size is large and the network is prone to memorize, they slow down the speed and reduces the total success. For this reason, additional layers are used. Rectified linear unit (ReLU) is used to degrade the linearity of network values. If the linearity does not deteriorate, the network parameters are memorized. Eq. (2) is used for this.

$$x = \begin{cases} 0, & x < 0 \\ x, & x > 0 \end{cases} \quad (2)$$

In order for the network to learn faster, the unimportant parameters need to be removed. A pooling layer is used for this. The pooling layer selects the maximum value according to the specified neighborhood value. Thus important features are transferred to the next layer. At the same time, the dimension is reduced because the unimportant matrix values are eliminated. This process is applied as in Eq. (3).

$$P_{jm} = \max_{k=1}^r (I_{j(m-1)n+k}) \quad (3)$$



Fully connected layer (FCL) is used to classify learned features. Therefore, hidden layers should be determined according to the problem. L2 regularization is a punishment function used to make the network learn better. Finally, the weights on the network change proportionately as they are updated. To avoid this, a dropout layer is used. Dropout layer ensures that some parts of the network are not used during training, according to the dropout value.

Using whole-slide images for network training will be quite inefficient, because there is no enough hardware for images of  $49,152 \times 110,592$  pixels. For this reason, the image is divided into  $128 \times 128$  pixels. These image patches are applied to the network in batches containing 64 images for network training. SGD is used to update the weights. Batch SGD has been preferred for this process. Batch SGD increases noise resistance and allows problem generalization. The total error value of images submitted in batches is calculated for each batch. Parameters of the net are updated with the stochastic gradient descent (SGD) algorithm according to the calculated error value. In the process of updating parameters with SGD, input batch data is processed by CNN architecture. After the last layer of architecture, the gradient of the error value is calculated. With this gradient value, the parameters in each layer are updated according to the chain rule. According to the chain rule, the error gradient value from the output layer of the CNN architecture to the input layer varies according to the layer type.

The HIC-net structure contains 6 convolution layers. The order of these layers is as follows:  $5 \times 5 \times 1 \times 256$ ,  $7 \times 7 \times 256 \times 256$ ,  $9 \times 9 \times 256 \times 256$ ,  $11 \times 11 \times 256 \times 128$ ,  $13 \times 13 \times 128 \times 128$ , and  $15 \times 15 \times 128 \times 96$ . This structure is represented as  $W \times W \times F \times N$ .  $W$  represents height and width values of features matrix.  $F$  is the number of the feature channels.  $N$  is the number of images in the convolution layer. The  $W$  values in the building are enlarged as the depth increased. The reason for this is to determine the interaction of the cells in the cancerous area with the other cells. The Xavier initialization method is used for weight initialization.

The pooling layer helps to reduce the image size while preserving the important features of the image. In this way, the transaction load is reduced without losing important data. Pooling layers with  $2 \times 2$  and  $3 \times 3$  neighborhood values are used in this study. At this point, the stride and padding coefficients in the layers gain importance. The stride value sets the property matrices to be slide over the image. At the end of this process, the image size changes. Eq. (4) is used to calculate the dimension of the image after stride value.

$$Dim = \left( \frac{I - w + 2P}{S} \right) + 1 \quad (4)$$

in which  $Dim$  represents image size,  $I$  represents input,  $w$  represents convolutional kernel,  $S$  represents stride parameter and  $P$  represents padding parameter. The padding parameter is the process of adding numbers to the edges of the view. At this point, image sizes can be kept at desirable levels. The formula used for its calculation is as in Eq. (5).

$$P_c = \frac{S(Dim - 1) + w - I}{2} \quad (5)$$

Our padding parameters are zero for every layer. Stride values are as follows: 1 for all layers except P1 layer, 2 for P1 layer. The image dimensions at each layer output up to the FCL input are as follows:  $124 \times 124$  pixels after C1,  $61 \times 61$  pixels after P1,  $55 \times 55$  pixels after C2,  $54 \times 54$  pixels after P2,  $46 \times 46$  pixels after C3,  $44 \times 44$  pixels after P3,  $34 \times 34$  pixels after C4,  $33 \times 33$  after P4,  $21 \times 21$  pixels after C5,  $20 \times 20$  pixels after P5,  $6 \times 6$  pixels after C6,  $4 \times 4$  pixels after P6 layers. Finally, 1536 values come to the FCL input.

The FCL consists of 1 input layer and 5 hidden layers. The FCL output layer consists of a single neuron. This neuron produces a value of '1' for cancerous tissue and '0' for normal tissue. The FCL input layer has 1536 neurons. These neurons depend on the output of the previous layer. The number of neurons is reduced as hidden layers deepen. There is only one neuron in the last layer.

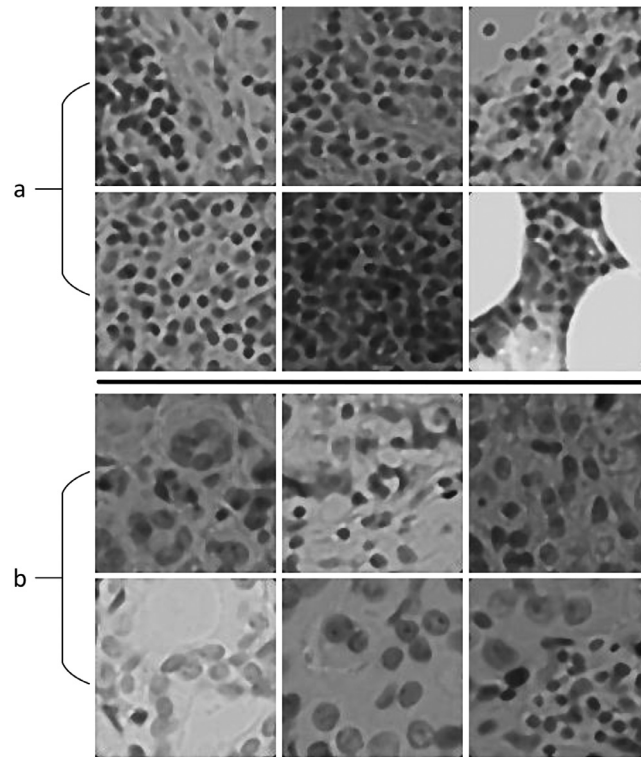
Deep convolutional neural networks and large neural networks face overfitting problems. Some parameters are overly compatible with the problem during training. Some parameters are updated proportionally to each other. This situation slows down the training and reduces the test success. Dropout layer is used to prevent these problems. Dropout layer randomly reduces network units. Thus, each unit can learn on its own. In this study, the dropout value is 0.25.

Other important parameters for network training are learning rate, weight decay and momentum. The learning rate for the proposed HIC-net is set to 0.00001. In every 50 iterations, this value is reduced by half. The weight decay parameter is 0.00005. The momentum value is 0.85.

### 3. Experimental results

#### 3.1. Dataset

The images on this dataset are used for breast cancer detection by examining sections of histological lymph nodes. Images from the dataset are reviewed by expert pathologists and a ground truth image is created for each image. The dataset contains 400 whole-slide images. 270 of these images are reserved for training and 130 for testing. Dataset images are stored in image pyramids that move from high resolution to low resolution. The lowest resolution image is  $512 \times 512$  pixels, the highest resolution image is  $97,792 \times 217,088$  pixels. A dataset containing 30,656 images is created using CAMELYON dataset [21] for the proposed network training. 14,500 image are breast cancerous tissue images and 16,156 images are normal tissue images. 25,600 images of dataset are randomly selected for training. These images contain tissues from more than



**Fig. 4.** Some images from our dataset. (a) normal tissue, (b) cancerous tissue.

one patient. To generate the data set, patches of size  $128 \times 128$  are cut at random positions from the whole-slide images. In order to classify the patches, we use only label information that are annotated as cancerous or normal. Label information for cancerous patches is '1' and label information for normal patches is '0'.

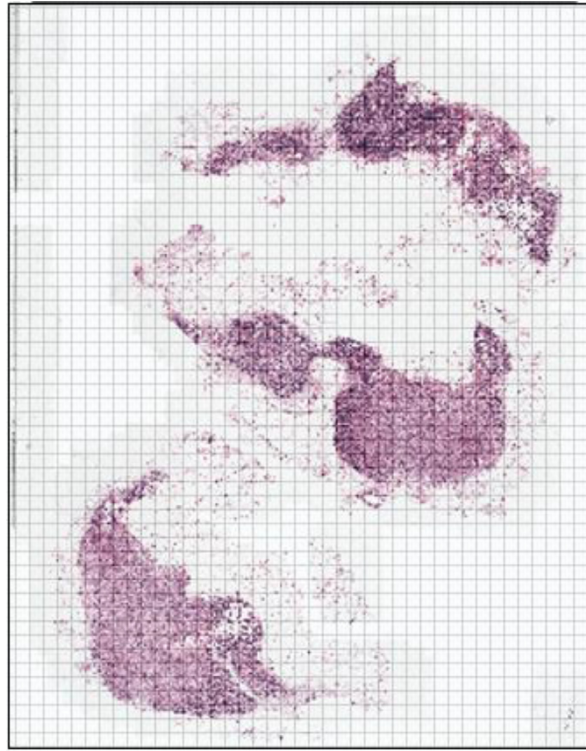
In order to train proposed network, 23,040 image patches are used for training, 2560 image patches are used for validation. The remaining 5056 images are used for the test. Some images in the data set are shown in Fig. 4.

When separating whole-slide images into patches with a size of  $128 \times 128$  pixels,  $128 \times 128$  size windows are used. The pixels remaining under these windows are saved as a patch. If the number of pixels remaining on the edges of the image is less than the  $128 \times 128$  window size, the patch mean value is assigned instead of the remaining pixel values. This process is shown in Fig. 5. If an image part does not contain any cell, we do not include it in the training or evaluation or test phase. Because it is not fair to evaluate a cell-free part. We mark such parts directly “without cancer”, without analyzing them with CNN.

### 3.2. Experiments

Proposed HIC-net is trained on a computer with Intel Core i7-7700K CPU (4.2 GHz), 32GB DDR4 RAM and NVIDIA GeForce GTX 1080 graphic card. The graphics card has a compute capability of 6.1. Thanks to this specificity, the GPU can operate much faster. It is especially suitable for batch operations. Processing images in mini-batches are quite useful for a stochastic gradient. This is useful for both GPU memory and random selection of instances. 600 epochs are selected for training the HIC-net on the GPU. The training process lasted about four days.

The feature extraction step is one of the most important steps in the classification process. Because it is the basic necessity for the perception of objects. The goal of machine learning algorithms is to be able to perceive scenarios as people perceive them. For this reason, it is necessary to obtain useful information about the scenario. The information obtained from the images is image features. These features can be obtained as handcrafted features using feature engineering. But this is a very challenging and time-consuming process. Nowadays, automatic feature extraction methods are used instead of hand-crafted. The features obtained from the images are examined as low-level features and high-level features. Low-level features are often simple features such as edges of objects, color components, Gabor filters. These features do not give information about the whole scenario when used alone. High-level features are about object-based or motion-based. They are more complex and contain clearer information. Usually, high-level features use low-level features for semantic-like operations. Convolutional neural network structure learns high-level features as the network deepens. As low level-level features are learned in the first layers of the network, high-level features are learned in the final convolution layers.



**Fig. 5.** Obtain a  $128 \times 128$  pixels patch from the whole-slide image.

The training-validation curves using pre-processing and without pre-processing, which is formed as a result of the network training with dataset composed of histopathological image patches, is shown in Fig. 6. It shows the training curves and the validation curves of our method with pre-processing and without pre-processing. As can be seen from the curves, the error curves fall down more quickly when the pre-processing method is used. The success rate at the end of 600 epochs without using the pre-processing layer is achieved in 100 epochs when the pre-processing layer is used. Image patches used for validation have not been used previously in training. For this reason, you can get tips about the success of your training by examining the situation of the two curves. If the difference between these two curves increases as the number of cycles advances, it is understood that the network has memorized training samples. For this reason, it is desirable that there is little difference between the two curves. Also, in order to prove the stability of the proposed method 3 fold cross validation is applied. 3 fold cross validation results are almost the same.

In the proposed method, the calculation of the error is based on the closeness value to the label value. That is, the value determined as 0.95 is evaluated according to the distance of the label value. Thus, a numeric error value comes out, not a label error. This process applies to all minibatches an error value is calculated for each minibatch. The success rate of the HIC-net architecture at the end of the training process is 95.2%. That is, the deviation rate from the label values is 4.8%. The HIC-net structure learned the features without interfering with the final layer results during the training phase. But it is pointless to use raw values during the test phase. For this reason, the resulting value is equal to the closest label value. Also, the midpoint of the two label values, 0.5, is set as the threshold value. Results below 0.5 are assigned as normal tissue. Results above 0.5 are assigned as cancerous tissue.

The area under on ROC curve is used to evaluate the test results of the proposed HIC-net algorithm. The ROC curve of HIC-net with pre-processing is shown in Fig. 7.

For the numerical evaluation of the success of the HIC-net algorithm, it is compared with well-known convolutional neural network algorithms. In the comparison phase, the same training set and the same test image patches are used for all CNN structures. Training samples include parts of the same whole-slide image and parts of different whole-slide images. The image patch dimensions are  $128 \times 128$  pixels for all methods. The input layers of other algorithms are set according to these values. AlexNet [22], FaceNet [23], VGG16 [24], and GoogLeNet [25] are used for comparison. AlexNet has 5 convolution layers. There are 3 hidden layers in the fully connected layer. It is suitable for images of  $227 \times 227 \times 3$  pixels as an input image. The VGG16 network consists of 16 layers. All convolution layers on the bottom are  $3 \times 3$  pixels. GoogLeNet consists of 27 layers. This network does not contain a fully connected layer. Thanks to this feature, it contains fewer parameter according to AlexNet.

Table 1 shows the comparison results. According to the comparison results, deeper networks produce more successful results. The results in Table 1 are not taken from the challenge page. Only the algorithms used here have been applied on



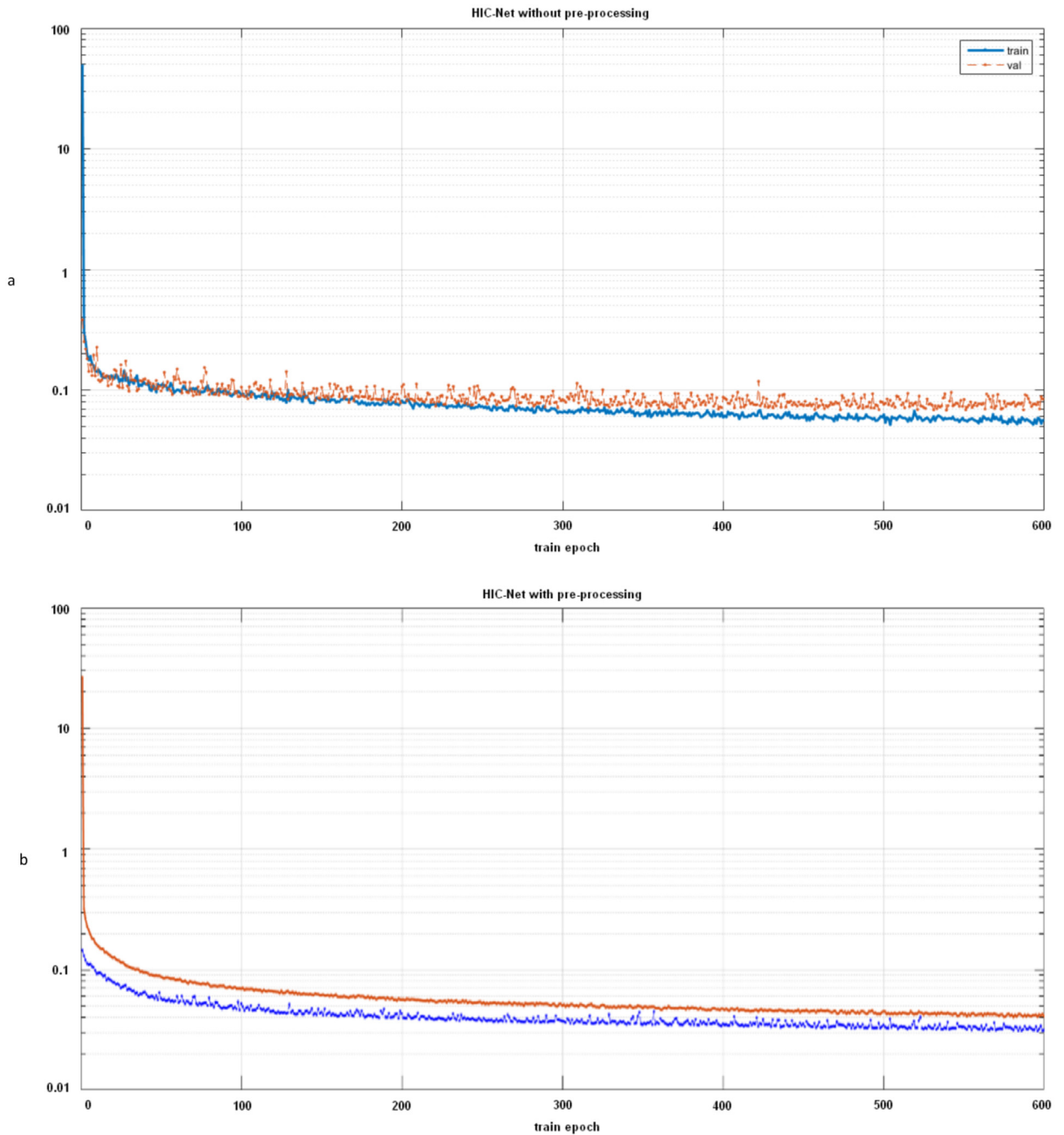


Fig. 6. Training curves of HIC-net, (a) without pre-processing, (b) with pre-processing.

**Table 1**

Comparison of CNN structures for patch based classification (training).

CNN structure	AUC (%)	Training time (days)
AlexNet	90.1	5
FaceNet	93.4	4
VGG16	93.9	4
GoogLeNet	95.3	4
HIC-net without pre-processing	92.05	4
HIC-net	<b>97.7</b>	4

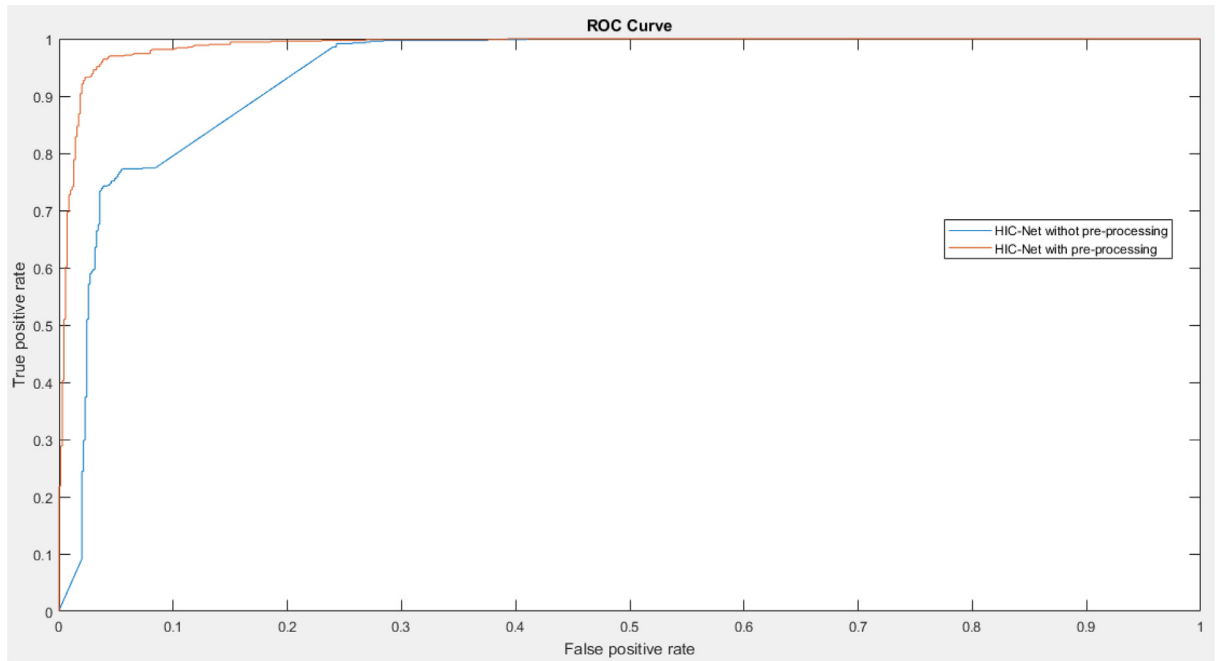


Fig. 7. ROC curve of HIC-net.

the new dataset we have created. The results are different from the results in the challenge because they are included in classification in the pixels without cell in the image patches. In this study, only patches containing cells are classified. Cell-free patches are not already cancerous. The proposed pre-processing method is used only for HIC-net. For other algorithms, the original states are used. All algorithms are compared on the same image patches. Also, Table 1 shows the training times for 600 epochs. The HIC-net algorithm produces better results than other well-known CNN algorithms.

Sensitivity, specificity and accuracy parameters are used to compare the proposed method with other state-of-art classification studies in the literature. Calculation of these parameters is as in Eqs. (6)–(8).

$$\text{Sensitivity} = \frac{TP}{TP + FN} \quad (6)$$

$$\text{Specificity} = \frac{TN}{TN + FP} \quad (7)$$

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (8)$$

in which TP is correctly identified white pixel number, FP is incorrectly identified white pixel number, TN is correctly identified black pixel number, FN is incorrectly identified black pixel number. Table 1 shows that state-of-art histopathological image classification algorithms.

The sensitivity parameter indicates the proportion of correctly identified cancerous sites. The specificity parameter represents a successful normal tissue. The accuracy parameter represents the total correct estimate. When evaluated according to these parameters, the histopathological image classification performance of the proposed algorithm is higher than the other algorithms. One of the most important factors in this success is the pre-processing technique used at the network entrance. The other factor is the network architecture.

In Table 2, the classification achievements of the state-of-art algorithms in the literature have been tested on the generated dataset. By comparing the results of these algorithms with the HIC-net algorithm, the HIC-net algorithm and pre-processing appear to be important.

The recommended method is a technique designed to classify histopathological images. For this purpose, a CNN architecture is designed for automatic learning and classification of image properties. When CNN architecture is used alone, training takes a long time because of the complexity and high resolution of histopathological images. In addition, success cannot be achieved at the desired level. For this reason, CNN architecture has been supported by an effective histopathological pre-processing method. In this study, a preprocessing technique has been developed, taking into account the H&E coloring of the image, the tactile state and the differences between the cells. Examining Fig. 6 clearly shows the effect of the pre-processing algorithm. When only CNN architecture is used, the training and validation curves are quite wavy. In addition,

**Table 2**  
Comparison of state-of-art histopathological classification algorithms.

Method	Sensitivity (%)	Specificity (%)	Accuracy (%)
Random Forest [26]	92.6	93.3	93
SVM [27]	85.9	90.6	88.3
C-RBH-PCA-net-SVM [28]	94.7	<b>97.36</b>	94.85
RBH-PCANet-LRBC [29]	71	83.23	78.46
N-CNN [30]	95.5	96.4	95.9
AlexNet	93.2	93.5	93.3
FaceNet	95.7	93.1	93.9
VGG16	96.4	95.8	96
GoogLeNet	96.4	96	96.1
HIC-net without pre-processing	95.3	94.5	94.9
<b>HIC-net</b>	<b>96.71</b>	95.7	<b>96.21</b>

educational success does not change much. When used in conjunction with the proposed preprocessing technique, it is seen that the training and validation curves are approaching zero faster and a continuous learning is experienced. The proposed method can be examined in two parts; pre-processing part and CNN architecture. Pre-processing is completely designed for problem-specific. The advantage of this part is that it is easy to find an important feature. Also, it can be used with other machine learning algorithms. The disadvantage is that it is necessary to design a separate perceptual perception for each problem. This process requires experience and expertise. Otherwise, the learning process is rather slow.

#### 4. Conclusion

Examination of the histopathological images provides important information about the cancer to the researchers. However, the difficulties of these images prolong the analysis period. In this study, an effective solution for histopathological image classification problem is presented. A special pre-processing method has been used to remove the complex structure of histopathological images and the effect of the disruptive factors in the background. In the proposed histopathological image pre-processing method, background irregularity is removed from the image, complexity is reduced and cell specificity is increased. At this stage, the distinction between cancerous cells and normal cells increases in the course of classification. The effect of the proposed pre-processing method has been examined on a linear CNN architecture. The proposed pre-processing layer increases the AUC score of the HIC-net architecture by 5.65%. In addition, validation and training error curves decrease more quickly when the pre-processing method is used. The proposed algorithm is compared with other state-of-art histopathological image classification methods. It has both cancerous tissue detection performance and normal tissue detection performance higher than other algorithms. The recommended pre-processing method produces the best results for  $128 \times 128$  pixels patches. Further studies will focus on early diagnosis and progress of cancer.

#### Funding

This study was funded by TUBITAK.

#### Conflict of interest

Authors declare that they have no conflict of interest.

#### Ethical approval

Histopathological images in this study were obtained from CAMELYON challenge (<https://camelyon17.grand-challenge.org/Home/>).

#### References

- [1] Fukuoka D, Hara T, Fujita H. Detection, characterization, and visualization of breast cancer using 3D ultrasound images. *Recent advances in breast imaging, mammography, and computer-aided diagnosis of breast cancer* 2006:557–67.
- [2] Zhang R, Zheng Y, Mak TWC, Yu R, Wong SH, Lau JY, Poon CC. Automatic detection and classification of colorectal polyps by transferring low-level CNN features from nonmedical domain. *IEEE J Biomed Health Inform* 2017;21(1):41–7.
- [3] Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*; 2012. p. 1097–105.
- [4] Sertel O, Kong J, Lozanski G, Shanaah A, Catalyurek U, Saltz J, Gurcan M. Texture classification using nonlinear color quantization: application to histopathological image analysis. 2008 IEEE international conference on acoustics, speech and signal processing; 2008.
- [5] Al-Kadi OS. Texture measures combination for improved meningioma classification of histopathological images. *Pattern Recognit* 2010;43(6):2043–53.
- [6] Raza SH, Parry RM, Moffitt RA, Young AN, Wang MD. An analysis of scale and rotation invariance in the bag-of-features method for histopathological image classification. In: *Lecture notes in computer science medical image computing and computer-assisted intervention – MICCAI* 2011; 2011. p. 66–74.

- [7] Zhang X, Xing F, Su H, Yang L, Zhang S. High-throughput histopathological image analysis via robust cell segmentation and hashing. *Med Image Anal* 2015;26(1):306–15.
- [8] Vu TH, Mousavi HS, Monga V, Rao G, Rao UA. Histopathological image classification using discriminative feature-oriented dictionary learning. *IEEE Trans Med Imaging* 2016;35(3):738–51.
- [9] Coudray N, Ocampo PS, Sakellaropoulos T, Narula N, Snuderl M, Fenyö D, Tsirigos A. Classification and mutation prediction from non-small cell lung cancer histopathology images using deep learning. *Nat. Med.* 2018;24:1559–67.
- [10] Zhang R, Shen J, Wei F, Li X, Sangaiah AK. Medical image classification based on multi-scale non-negative sparse coding. *Artif. Intell. Med.* 2017. doi:10.1016/j.artmed.2017.05.006.
- [11] Cruz-Roa AA, Ovalle JEA, Madabhushi A, Osorio FAG. A deep learning architecture for image representation, visual interpretability and automated basal-cell carcinoma cancer detection. In: *Medical image computing and computer-assisted intervention – MICCAI 2013 lecture notes in computer science*; 2013. p. 403–10.
- [12] Xu J, Luo X, Wang G, Gilmore H, Madabhushi A. A deep convolutional neural network for segmenting and classifying epithelial and stromal regions in histopathological images. *Neurocomputing* 2016;191:214–23.
- [13] Zhang G, Hsu C-HR, Lai H, Zheng X. Deep learning based feature representation for automated skin histopathological image annotation. *Multimedia Tools Appl* 2017. doi:10.1007/s11042-017-4788-5.
- [14] Bayramoglu N, Kannala J, Heikkilä J. Deep learning for magnification independent breast cancer histopathology image classification. 2016 23rd international conference on pattern recognition (ICPR); 2016.
- [15] Huang Y, Zheng H, Liu C, Rohde G, Zeng D, Wang J, Ding X. Epithelium-stroma classification in histopathological images via convolutional neural networks and self-taught learning. 2017 IEEE international conference on acoustics, speech and signal processing (ICASSP); 2017.
- [16] Zheng Y, Jiang Z, Xie F, Zhang H, Ma Y, Shi H, Zhao Y. Feature extraction from histopathological images based on nucleus-guided convolutional neural network for breast lesion classification. *Pattern Recognit* 2017;71:14–25.
- [17] Gummeson A, Arvidsson I, Ohlsson M, Overgaard NC, Krzyzanowska A, Heyden A, Bjartell A, Aström K. Automatic Gleason grading of H and E stained microscopic prostate images using deep convolutional neural networks. *Medical imaging 2017: digital pathology*, January; 2017.
- [18] Sudharshan PJ, Petitjean C, Spanhol F, Oliveira LE, Heutte L, Honeine P. Multiple instance learning for histopathological breast cancer image classification. *Expert Syst Appl* 2019;117:103–11.
- [19] Van Eycke YR, Balsat C, Verset L, Debeir O, Salmon I, Decaestecker C. Segmentation of glandular epithelium in colorectal tumours to automatically compartmentalise IHC biomarker quantification: a deep learning approach. *Med Image Anal* 2018;49:35–45.
- [20] Saha M, Chakraborty C, Racoceanu D. Efficient deep learning model for mitosis detection using breast histopathology images. *Comput Med Imaging Graph* 2018;64:29–40.
- [21] Camelyon Challenge, <https://camelyon17.grand-challenge.org/>.
- [22] Krizhevsky A, Sutskever I, Hinton GE. ImageNet classification with deep convolutional neural networks. *Commun ACM* 2017;60(6):84–90.
- [23] Ding H, Zhou SK, Chellappa R. FaceNet2ExpNet: regularizing a deep face recognition net for expression recognition. 2017 12th IEEE international conference on automatic face & gesture recognition (FG 2017); 2017.
- [24] Simonyan K, Zisserman A. "Very deep convolutional networks for large-scale image recognition," arXiv:1409.1556, 2014.
- [25] Szegedy C, Liu W, Jia Y, Sermanet P, Reed S, Anguelov D, Erhan D, Vanhoucke V, Rabinovich A. Going deeper with convolutions. 2015 IEEE conference on computer vision and pattern recognition (CVPR); 2015.
- [26] Valkonen M, Kartasalo K, Liimatainen K, Nykter M, Latonen L, Ruusuvaari P. Metastasis detection from whole slide images using local features and random forests. *Cytometry Part A* 2017;91(6):555–65.
- [27] Shibuya N, Nukala B, Rodriguez A, Tsay J, Nguyen TQ, Zupancic S, Lie D. A real-time fall detection system using a wearable gait analysis sensor and a Support Vector Machine (SVM) classifier. 2015 eighth international conference on mobile computing and ubiquitous networking (ICMU); 2015.
- [28] Shi J, Wu J, Li Y, Zhang Q, Ying S. Histopathological image classification with color pattern random binary hashing based PCANet and matrix-form classifier. *IEEE J Biomed Health Inform* 2017;21(5):1327–37.
- [29] Wu J, Shi J, Li Y, Suo J, Zhang Q. Histopathological image classification using random binary hashing based PCANet and bilinear classifier. 2016 24th European signal processing conference (EUSIPCO); 2016.
- [30] Zheng Y, Jiang Z, Xie F, Zhang H, Ma Y, Shi H, Zhao Y. Feature extraction from histopathological images based on nucleus-guided convolutional neural network for breast lesion classification. *Pattern Recognit* 2017;71:14–25.

**Şaban Öztürk** was born in İzmir, Turkey, in 1989. He received a B.S. degree in Electrical and Electronics engineering from Selçuk University, and M.S. degree in electrical & electronics engineering from Selçuk University in 2015. He is currently working toward the Ph.D. degree. His research interests include image and video processing, texture analysis, segmentation and classification algorithms.

**Bayram Akdemir** was born in Konya, Turkey in 1974. He received the B.S. degree in electrical & electronics engineering from Selçuk University, Turkey, in 1999, and the M.S. and Ph.D. degrees in electrical & electronics engineering from Selçuk University Konya, Turkey, in 2004 and 2009, respectively. His current research interests include electronic circuits, sensors, artificial intelligence, renewable energy sources.