Multi-path Convolutional Neural Network to Identify Tumorous Sub-classes for Breast Tissue from Histopathological Images

Rangan Das1, Utsav Bandyopadhyay Maulik2, Bikram Boote3, Sagnik Sen4, Saumik Bhattacharya5

Received: 22 February 2022 / Accepted: 23 June 2022
© The Author(s), under exclusive licence to Springer Nature Singapore Pte Ltd 2022

Abstract
Malignancy is one of the leading causes of death. It is on the rise in the developed and low-income countries with survival rates of less than 40%. However, early diagnosis may increase survival chances. Histopathology images acquired from the biopsy are a popular method for cancer diagnosis. In this work, we propose a deep convolutional neural network-based method that helps classify breast cancer tumor subtypes from histopathology images. The model is trained on the BreakHis dataset but is also tested on images from other datasets. The model is trained to recognize eight different tumor subtypes, and also to perform binary classification (malignant/non-malignant). The CNN model combines an encoder–decoder architecture and a parallel feed-forward network with attention mechanism. The proposed model provides state-of-the-art scores. Comparing with the other models, the accuracy of the proposed model is higher at different magnification and patient levels. The implementation is available at github.com/rangan2510/Residual-Unet

Keywords Breast histopathology · Deep learning · Convolutional neural networks · Image classification

Introduction
Cancer is one of the leading causes of death globally, particularly in developed countries [1, 2]. In the United States, breast cancer incidents increased by 0.3% per year between 2012 and 2016 [1]. The situation in low and middle-income countries is also alarming. For instance, in the Indian sub-continent, according to the National Institute of Cancer Prevention and Research (NICPR) under the Indian Council of Medical Research (ICMR) [3], around 2.25 million cases are being treated since 2018. Every year, around 750 thousand deaths are reported, and 11.6 million new cases are registered. However, the statistics on individual cancer types are also disturbing. Breast cancer cases make up 14% of the overall number of malignant cases. New cases registered in 2018 are 160 thousand. In rural areas, 1 in 60 women develop breast cancer respectively whereas, in urban areas, 1 in 22 do so. Early detection and quick diagnosis of breast cancer play an important role in increasing the survival rate. It has been observed that the survival rate of breast cancer is around 50% of the diagnosed cases. Internationally, non-invasive mammography is highly popularized for the early detection of breast cancer stages. It has also gained interest along with computational automation. Furthermore, the histopathological images [4] from invasive sampling through biopsy are extensively utilized to diagnose malignant cases including breast cancer. There are two major drawbacks. Firstly, diagnosed malignancy detecting affective zones are highly subject to expert interest. Moreover, it is also extremely time-consuming. Therefore machine learning techniques can be used to make the diagnosis procedure
much more efficient. Moreover, designing comprehensive machine learning techniques can reduce human error as well as the time is taken for manual diagnosis. Recently, deep learning-based methods have been used extensively for image classification tasks. Therefore, such models can help pathologists make precise diagnoses much faster. In this article, we have proposed a method based on deep learning [5] for diagnosing the tumorous sub-types of breast tissue. The proposed model implements a multi-path convolutional neural network that uses a U-net-like encoder–decoder architecture as well as a simple feed-forward CNN running in parallel. The encoder part makes use of residual blocks. Residual blocks have been extensively used in deep learning [6]. U-Net, on the other hand, is used for biomedical image segmentation. U-Net makes use of an encoder–decoder-like architecture to extract relevant features from the images. The proposed network is trained on the BreakHis data-set [7] that contains histopathological images of the benign and malignant tumors along with four sub-classes of each tumor type. Hence, the proposed tool can classify eight different types of tumors irrespective of the magnification levels of the microscopic images. The model has also been trained to do binary classification. Furthermore, a blind test has also been done to ensure that the model can classify the images from outside the training dataset.

Contributions

The notable contributions of this work are as follows:

- The proposed classifier is developed by incorporating a combination of multiple CNN architectures. The model has a feed-forward network constructed using a convolution block attention module (CBAM) that runs parallel to an encoder–decoder architecture. The encoder–decoder architecture is based on the U-Net. The encoder part is designed using residual connections. It outperforms multiple state-of-the-art CNN models in the task of classifying the different sub-types of breast tumors.
- The model makes use of skip connections and parallel paths that facilitate the gradient flow, enabling easier optimization.
- 1x1 convolutions are used to compute feature-map reductions before more expensive operations in the skip connections of the U-Net. This keeps the network much faster and reduces the number of trainable parameters.
- Transfer learning has been used to train the residual blocks for the encoder part of the U-net. This has significantly improved the training performance of the network.
- The generalization ability of the model has been tested extensively on various breast cancer histopathology datasets where it is trained on one dataset and is then evaluated on different datasets to simulate the real-life performance of the model. Here, the proposed method outperforms other state-of-the-art models by a significant margin.

Related Works

For more than forty years, automatic image processing for cancer diagnosis has been an active research topic [8]. Before deep learning, multiple attempts had been made for the classification and gradation of breast cancer using methods such as support vector machines (SVMs) [9, 10], random forest (RF) [11, 12], principal component analysis (PCA) [13] and so on. Feature extraction was done using scale invariant feature extraction (SIFT) [14], local phase quantization (LPQ), local binary patterning (LBP), and other similar methods [7]. Most methods require laborious feature engineering. Kowal et al. [15] have compared and tested different algorithms for nuclei segmentation and further classified the images as either benign or malignant. This case of binary classification was done on a dataset of 500 images and accuracy levels of 96.5% were reported. However, the feature selection was done manually and used only 699 images. Filipczuk et al. [16] proposed a method that makes use of four different classifiers and a 25-dimensional feature vector to perform binary classification on cytopathological images of fine needle biopsies. In this case, around 98.5% accuracy was achieved on a dataset containing only 737 images. Similarly, George et al. [17] also proposed a method that involves tedious manual feature extraction followed by classification using SVMs, probabilistic neural networks, multi-layer perceptron, and learning vector quantization on a dataset of 92 images only. Doyle et al. [18] makes use of textural features of images to develop an SVM for grading breast cancer. However, most of these methods were limited by the availability of large and open annotated datasets.

Over the last few decades, significant effort has been made in recognizing breast cancer from histopathological images using deep learning. Convolutional neural networks (CNNs) have been primarily used for the detection and classification of breast cancer [4]. Spanhol et al. [7] introduced a data-set called BreakHis that provides 7909 images, acquired from 82 patients. Not only is this dataset larger than the ones previously available, but it also comes with 8 different tumor subtypes: 4 for benign tumors, and 4 for malignant tumors. Multi-classification has more clinical value too. This dataset has enabled the development of multiple deep learning approaches. Teresa et al. [19] have combined CNN with SVM to achieve an accuracy of 77.8% for four classes and 83.3% for binary classification. Bayramoglu proposed [20] a CNN architecture with an average accuracy rate of 82.13%. Spanhol et al. [21] used the popular CNN architecture called LeNet and achieved around 72% accuracy. Han et al. [22]
proposed a class structure-based deep convolutional neural network (CSDCNN) which achieved around 93.2% accuracy on the BreakHis dataset. Using the same dataset, [23] proposed a deep manifold preserving autoencoder that performs binary classification and acquired an average of 89.51% on image-level classification whereas a more recent dependency-based lightweight convolutional neural network (DBLCNN) [24] achieves an image-level accuracy of 96.0%. However, none of these tasks verify the classification performance of the model on breast cancer histopathology images beyond the BreakHis dataset.

Previously, residual blocks have been used in U-Nets for segmentation tasks in other domains [25–27]. Newer models that use residual connections include the ResNeXt [28] model and Wide Residual Network [29, 30]. Variants of ResNeXt has also been used in classification of breast cancer histopathology images. Newer classification models include RegNet [31], EfficientNet [32] and Vision Transformers [33], all of which are used for comparison in this paper.

Data Description

The Breast Cancer Histopathological database or BreakHis [34] comprises images of benign and malignant breast tumors that were collected by open surgical biopsy (SOB). A key feature of the BreakHis dataset is the diverse classes of both malignant and benign tumors that are contained in the dataset. The histopathology study and labeling were done by pathologists of the P & D Lab, taking care of the anonymity of all the images. The digitized images for the breast tissue biopsy slides were obtained by an Olympus BX-50 system microscope having a relay lens with a magnification of 3.3x coupled with a Samsung digital color camera SCC-131AN. The dataset consists of 7909 images of breast cancer tissue, which were collected from 82 patients using 4 different magnification levels (40×, 100×, 200×, 400×). Table 1 contains details of how the images are distributed among the 4 magnification levels. It contains 2480 benign and 5429 malignant images in PNG format with 700 × 460 pixels, 3-channel RGB, and 8-bit depth in each channel. The 2 types are further subdivided into 4 categories each, i.e., four benign tumors and 4 malignant tumors. These are labelled as:

1. Benign: adenosis (A), fibroadenoma (F), phyllodes tumor (PT), tubular adenosis (TA)
2. Malignant: ductal carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC), papillary carcinoma (PC)

Figure 1 shows a sample from each type of tumor. The different magnification levels of a single sample is shown in Fig. 2. The details of the image distribution among these categories have been provided in Table 2.

| Magnification | Malignant | Benign | Total |
|---------------|-----------|--------|-------|
| 40×           | 1370      | 652    | 1995  |
| 100×          | 1437      | 644    | 2081  |
| 200×          | 1390      | 623    | 2013  |
| 400×          | 1232      | 588    | 1820  |
| Total number of images | 5429 | 2480 | 7909 |

Fig. 1 Breast histopathological images for each sub-classes from BreakHis Dataset. a Adenosis, b fibroadenoma, c phyllodes tumor, d tubular adenoma, e ductal carcinoma, f lobular carcinoma, g mucinous carcinoma, h papillary carcinoma
Methodology

The proposed method makes use of a deep learning [5] model that is built upon two popular CNN architectures: ResNet [35] and U-Net [36]. For the network, we have modified the U-net architecture and incorporated residual blocks for classification of breast cancer histopathological images. The following subsections explain the methodology.

Network Architecture

The proposed network architecture is made up of two primary components: a feed-forward network that uses an attention mechanism and an encoder–decoder architecture that is based on the U-Net model. These two feature extractors run in parallel and the final classification of the images is done from the aggregated feature maps of these two models. The model is illustrated in Fig. 4.

Attention-Based Feed-Forward Module

The feed-forward network uses sequential convolutions along with a convolution block attention module (CBAM) [37] with dilated convolutions. The CBAM module uses both spatial and channels attention mechanisms which boost the model performance. In the spatial attention module (SAM), a convolution operation is performed using a $3 \times 3$ kernel with a dilation rate $= 2$. The dilated kernel helps reduce the performance cost of the spatial attention mechanism. In the channel attention module (CAM), we use log-sum-exponent and a global average pooling of the input features and then pass each 1-D vector to a multi-layer perceptron (MLP). The output of the MLP for each vector is summed element-wise and is then passed through a ReLU function. We use the log-sum-exponent pooling instead of max pooling as it helps preserve the edge details and also helps deal with noise. The output of the CBAM is then stacked with the output of the encoder–decoder architecture that runs in parallel and is then passed through two more layers of convolution.

Encoder–Decoder Module

The encoder–decoder module is a modification of the standard U-Net architecture. In the standard U-net architecture, the encoder part uses successive convolutions, followed by $2 \times 2$ pooling for the down-sampling operation. In the proposed model, the successive convolution and pooling layers in each step of the encoder are replaced with a custom residual block. Every residual block, as illustrated in Fig. 3, has two convolution layers. These residual blocks have the same components as the ResNet18 architecture, allowing us to perform transfer learning. However, the forward pass through the layers is performed differently. This allows us to have improved performance while using the same weights as the original ResNet18 model. In the residual block, we use two skip connections. In the first convolution layer of every residual block, the convolution operation is implemented with a stride value of 2. This reduces the size of a $K \times K$ feature map to $K/2 \times K/2$. Hence, we need to perform identity downsampling. This is performed using two different methods for each of the skip connections. The first skip connection uses max pooling that halves the image dimensions. The second skip connection uses a strided convolution operation with batch normalization. This dense connectivity aids better information propagation through the network and removes dependence on just pooling operations for downsampling as pooling operations can be lossy [38]. Using these residual blocks enables us to simultaneously extract features and compress the image representation.

The encoder uses 4 residual blocks. The input image is down-sampled 5 times by a factor of 2. The number of stages was determined experimentally. The input to the
model is a 3-channel image with dimension 224 × 224. After the end of the contractive part, the dimension of the feature map is 7 × 7.

The residual blocks that are used in the residual U-Net follow the same architecture as those implemented in ResNet18. Each residual block contains two convolution operations each followed by batch normalization. So, for the first down-sampling step, convolution is performed using a 7×7 kernel with stride 2. For the second step of down-sampling, a simple max-pooling operation is done. From the next step onward, the residual blocks are used for down-sampling. The first residual block is slightly different from the rest as it does not reduce the spatial dimension of the feature maps. The convolution operation is performed with 3 × 3 kernels with stride 1 and padding 1 in the first residual block. In the following residual blocks, the stride is increased to 2 for the first convolution operation in the block. This is where the down-sampling takes place.

In the decoder part, up-sampling is done 5 times. After the feature maps of the decoding stage are concatenated from those of the encoding stage, a 3 × 3 convolution is performed. This convolution operation is also used to reduce the number of feature maps. After passing the feature maps through a ReLU activation layer, they are upsampled using bilinear interpolation. After the upsampling operation, another convolution operation is performed before the feature maps are again concatenated.

For the final classifier part, the output of the decoder part is passed through two layers of convolution, followed by average pooling and a fully connected layer. The number of nodes in the final layer is set based on the required number of classes.

Algorithm 1 Optimization of proposed network

```
Initialize dataloader D
Initialize model M
Loss l ← Infinity
Average Loss l_a ← 0
i ← 0
while i ≤ total number of epochs do
    for each Batch B ∈ D do
        Initialize L_B: List of loss values per batch.
        for each Images x, Labels y ∈ B do
            Compute loss l_b ← M(x, y)
            Add l_b to L_B
        end for
        l_a ← mean(L_B)
        if l_a ≤ l then
            l ← l_a
            Save model weights
        end if
    end for
end while
Load model weights for inference.
```

Ablation Study

In an encoder–decoder-based architecture, the performance of the model depends on the number of encoding and decoding steps used. In this case, the depth of the U-net structure needed to be assessed to see what number of steps provided the highest accuracy. An ablation study was done to find out the required number of encoding and decoding steps. Initially, 3 steps of encoding and three steps of decoding were used which provided an accuracy of 89.41% on 8 classes. Using 4 steps gave the maximum accuracy of 98.65% while using 5 steps reduced the accuracy again to 89.67%. Furthermore, tests were performed to see the effectiveness of the 1x1 convolution operations. These convolution operations were used before the concatenation operations to reduce the number of feature maps. The results of the study are summarized in Table 3.

![Implementation of the residual block in the proposed residual U-Net architecture](image-url)
Training and Testing

Training and testing were done separately in two phases. The BreakHis dataset itself comes with certain standard tests. These were done according to the 5-fold validation structure that is mentioned in [7]. For each of these folds, the set of 7909 images is split into two halves: a training set containing 5211 images and a testing set consisting of 2698 images. In the second case, the model was trained again from scratch and was compared to other methods for the 8-class classification. Furthermore, the performance for 2 class classification was also explored.

During training, the images were resized and augmented. The images were resized to 224 × 224 pixels directly without cropping. Augmentation techniques that were used include random rotation and flipping. Since the decoder part of the network makes use of the residual blocks that are similar in architecture to that of ResNet18, they were initialized with the ImageNet weights. This is shown in Fig. 4. For each of these folds, the model is trained for 200 epochs with a decaying learning rate and varying up-sampling. The

| Table 3 | Different network settings tested and their performance on the 8-class classification problem |
|---------|-----------------------------------------------------------------------------------|
|         | 1st | 2nd | 3rd | 4th | 5th | 6th |
| Number of encoding and decoding steps | 3   | 3   | 4   | 4   | 5   | 5   |
| Feature map reduction using 1–1 convolution | Yes | No  | Yes | No  | Yes | No  |
| Accuracy on test set | 89.41% | 93.24% | 98.65% | 98.47% | 95.20% | 89.67% |

The test set was set created from 20% of the images, randomly selected from the entire dataset.
learning rate gradually decreased from 0.001 to 0.0002 while the up-sampling technique was slowly refined. The learning rate was made to decay by 0.4% with every epoch. In the case of up-sampling, the first 50 epochs used Nearest Neighbor up-sampling. The next 50 epochs used Bilinear up-sampling. The final 100 epochs used Bicubic up-sampling. During training, the batch size was set to 16 and the Adamax optimizer was used.

To address the minor class imbalance, the loss function used during training is the Focal loss. This is given as

\[
FL(p_t) = -a_t (1 - p_t)^γ \log(p_t)
\]

where \(p_t = y \cdot p + (1 - y) \cdot (1 - p)\) and \(y\) is the binary class label indicator (0 or 1) and \(p\) is the predicted probability.

The mean loss for every batch of the epoch was monitored and the weights of the model were only updated after an epoch only if the mean loss improved after training. The best model that was saved during the training was used for testing. The entire process of acquiring the model weights for training is described in Algorithm 1.

The results of the 5-fold validation are provided in Table 4. We have used four different performance metrics: accuracy, precision, recall, and \(f_1\)-score.

A similar training strategy was adopted for the second case. In this case, the dataset was divided into a training set and a validation set. From Tables 1 and 2, it can be seen that the dataset suffers from a massive class imbalance problem. Here, the Ductal Carcinoma (DC) class has the largest number of images, which is 3451 images. Therefore, an augmented dataset was created where each class has 3451 images, totaling 27,608 images. This augmentation was done by simply creating copies of images in the other classes. The copies were also randomly rotated and flipped. During training, 80% of these image files were used as a training set, while the remaining 20% was used for validation of the trained model.

On this set of 27,608 images, the model was initially trained for 8-class classification. After training for 250 epochs, the accuracy level on the validation set reached its maximum value and plateaued. The varying up-sampling technique, the optimizer settings, and the learning rate value and decay were kept the same. Furthermore, other models were separately trained on this data-set for comparison with the same hyper-parameters. This allowed us to perform a direct comparison of the proposed model with the state-of-the-art models.

## Results

### Performance on BreakHis

We performed multiple evaluations to show the performance of the proposed model on the BreakHis dataset and the generalization ability of each of these models. Since the BreakHis dataset contains eight classes of images, which is the highest in any publicly available dataset, it allows us to train models that can be used on other breast cancer histopathology datasets as well. We evaluate the performance of the method on the 5-fold validation sets that are defined in [7]. The performance values are shown in Table 4. Figure 5 shows how the CBAM module helps in extracting relevant features from challenging images. The activation maps acquired by adding the attention module include more edge details that are that is otherwise absent from the corresponding feature maps.

We compare methods from literature that has been tested on the BreakHis dataset and also evaluate recent and popular computer vision models such as ResNet18 [35], U-Net [36], Densenet-121 [44], Xception Net [45], MnasNet [46], ResNeXt50 [28], ShuffleNet v2 [47], Wide ResNet [29, 30], RegNet [31], EfficientNet B-0 [32] and Vision Transformer (ViT) [33].

According to [4], we perform image-level classification for each magnification level, and we extend the evaluation to three more datasets. For image-level classification, we checked the classification performance of images belonging to different magnification levels. BreaKHIs contains four levels of magnification: 40x, 100x, 200x and 400x. For each magnification level, the classification accuracy is reported. Table 5 contains the accuracy of the eight-class classification problem for different datasets. Here, we can see that the proposed model either provides or surpasses the state-of-the-art.

| Fold | Accuracy (%) | F1-score | Precision | Recall |
|------|--------------|----------|-----------|--------|
| 1    | 99.1046      | 0.9929   | 0.9966    | 0.9896 |
| 2    | 99.9167      | 0.9992   | 0.9991    | 0.9993 |
| 3    | 98.1209      | 0.9686   | 0.9578    | 0.9883 |
| 4    | 98.5258      | 0.9897   | 0.999     | 0.9989 |
| 5    | 97.6590      | 0.9665   | 0.9535    | 0.9885 |
| Cumulative | 98.665 ± 0.786 | 0.983 ± 0.013 | 0.9812 ± 0.023 | 0.9929 ± 0.006 |
Blind Testing

To ensure the generalization capabilities of the proposed model, a blind test was performed. In blind testing, we trained the model on the BreakHis dataset and evaluated the model on other datasets. We created multiple binary classification datasets to ensure the model performs consistently in real-life scenarios. The data selection process is described in Table 6. Since the BreakHis dataset contains a diverse annotated classification of breast tumors (both benign and malignant), a model trained on the BreakHis dataset should perform well with other datasets too. This is demonstrated by three separate blind tests.

We used three separate datasets to test the model in real-life conditions. The datasets used are Breast Cancer Histology images (BACH) [40], an Invasive Ductal Carcinoma (IDC) Grading dataset [41] and a dataset with IDC patches that is aggregated from [42] and [43]. We devised three tests. The first test uses 200 samples from the BACH dataset. The second test uses 400 malignant samples from the IDC grading dataset and 100 benign samples from the BACH dataset. The third test uses 500 images from each class of the IDC patches dataset. The performance of each of these datasets is shown in Table 7.

Our model shows the best performance on the BACH dataset, which is the first blind test. On a total of 200 images, the proposed model gave an accuracy of 98.5%. In contrast, the other state-of-the-art models had a much lower score on the blind test, despite giving better results on the BreakHis dataset itself. This clearly shows that our model has better generalization power than others. On the BACH dataset, the ResNet18 model scored 76.0%. ResNeXt model used before scored 88.0% and DenseNet scored 83.5%. ShuffleNet got 72.5% while Mnasnet got a much lower score of 52.5%. The Vision Transformer model, the RegNet, and the Wide ResNet provided comparatively better performance. This shows that the proposed model is much more effective at the task of binary classification of breast cancer histopathological images in real life. Interestingly, these tests contain four sub-types of Breast tissue-associated benign tumor and three sub-types of Invasive Ductal Carcinoma (two of the rare tumor genesis cases). Despite that, the machine has outperformed in binary classification.

The tested models show a similar trend across the other two tests. In test 2, the model achieves a score of 86.0% whereas, in test 3, the model scores 75.3%. The lower score in the last test is possibly because of the low-resolution image patches (50 × 50 pixels) as well as the large size of the dataset. A higher score across all the blind tests corroborates the generalization ability of the model. Figure 6 shows the average accuracy of the proposed method across all the evaluations.

Conclusion

In this article, an improved deep learning model has been designed which is used to successfully classify breast cancer histopathology images. The model has been trained on the BreakHis dataset which allows it to classify images from...
eight types of tumors. The proposed model has been shown to have better performance while classifying the eight different classes of histopathology images. Furthermore, the model has been shown to successfully perform multi-class classification irrespective of the magnification levels of the images. Moreover, to simulate real-life scenarios, we also tested the trained model on other datasets to perform binary classification tasks. Multiple blind tests show that the model can work with images beyond the training dataset. Therefore, it is expected that the type, as well as subtypes of the tumor, can be separated unambiguously over a wide range of histopathology images. In the blind tests, the proposed model outperforms other state-of-the-art methods available in the literature. This ensures that the model can have real-life applications. The developed model will reduce the time and human error in breast cancer histopathological image analysis. We can infer that this model can be utilized to build an understanding of the sub-stages of malignancy in the future. Even though this paper shows the model for the classification of breast histopathological images only,

| Model          | Blind test 1 accuracy | Blind test 2 accuracy | Blind test 3 accuracy |
|----------------|-----------------------|-----------------------|-----------------------|
| ResNet18       | 76.0                  | 78.6                  | 70.1                  |
| ResNeXt        | 88.0                  | 82.2                  | 73.2                  |
| DenseNet       | 83.5                  | 80.8                  | 68.9                  |
| ShuffleNet     | 72.5                  | 71.4                  | 67.9                  |
| M NasNet       | 52.5                  | 58.6                  | 62.3                  |
| Wide ResNet    | 95.5                  | 81.6                  | 69.8                  |
| EfficientNet   | 87.0                  | 82.2                  | 70.6                  |
| RegNet         | 92.5                  | 77.8                  | 71.2                  |
| ViT            | 93.0                  | 82.2                  | 70.3                  |
| Proposed method| 98.5                  | 86.0                  | 75.3                  |

All the models are first trained on the BreakHis dataset for binary classification and are then tested on three breast cancer histopathology image sets acquired from different sources.

Fig. 5 Input samples (a) and the corresponding feature maps from the feed-forward module without the convolution block attention module (b) and with the convolution block attention module (c). Using CBAM allows the feed-forward module to be a more effective feature extractor when processing challenging samples.
we are assuming that this model can be utilized to build an understanding of the sub-stages of malignancy in the future.

Declarations

Conflict of interest The authors declare that no funds, grants, or other support were received during the preparation of this manuscript. The authors also declare that there are no conflicts of interest.

References

1. DeSantis CE, Ma J, Gaudet MM, Newman LA, Miller KD, Goding Sauer A, Jemal A, Siegel RL. Breast cancer statistics, 2019. CA Cancer J Clin. 2019;69(6):438–51.
2. Yadav S, Karam D, Bin Riaz I, Xie H, Durani U, Duma N, Giridhar KV, Hieken TJ, Boughey JC, Mutter RW, et al. Male breast cancer in the United States: treatment patterns and prognostic factors in the 21st century. Cancer. 2020;126(1):26–36.
3. Mallath MK, Taylor DG, Badwe RA, Rath GK, Shanta V, Pramesh C, Digumarti R, Sebastian P, Borthakur BB, Kalwar A, et al. The growing burden of cancer in India: epidemiology and social context. Lancet Oncol. 2014;15(6):e205–12.
4. Veta M, Plium JP, Van Dietel PJ, Viergever MA. Breast cancer histopathology image analysis: a review. IEEE Trans Biomed Eng. 2014;61(5):1400–11.
5. LeCun Y, Bengio Y, Hinton G. Deep learning. Nature. 2015;521(7553):436.
6. Szegedy C, Ioffe S, Vanhoucke V, Alemi AA. In: Thirty-First AAAI Conference on Artificial Intelligence, 2017.
7. Spanhol FA, Oliveira LS, Petitjean C, Heutte L. A dataset for breast cancer histopathological image classification. IEEE Trans Biomed Eng. 2015;63(7):1455–62.
8. Stenkvist B, Westman-Naeser S, Holmquist J, Nordin B, Bengtsson E, Vegelius J, Eriksson O, Fox CH. Computerized nuclear morphometry as an objective method for characterizing human cancer cell populations. Can Res. 1978;38(12):4688–97.
9. Rejani Y, Selvi ST. Early detection of breast cancer using SVM classifier technique. 2009. arXiv preprint arXiv:0912.2314.
10. Brook A, El-Yaniv R, Isler E, Kimmel R, Meir R, Peleg D. Breast cancer diagnosis from biopsy images using generic features and svms. Computer Science Department, Technion, Israel Institute of Technology: Tech. rep; 2008.
11. Chaurasia V, Pal S. Stacking-based ensemble framework and feature selection technique for the detection of breast cancer. SN Comput Sc. 2021;2(2):1–13.
12. Paul A, Dey A, Mukherjee DP, Sivaswamy J, Tourani V. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2015. p. 94–102.
13. Hasan H, Tahir NM. In: 2010 6th International Colloquium on Signal Processing and its Applications. IEEE; 2010. p. 1–4.
14. Irshad H, Jalali S, Roux L, Racoceanu D, Hwee LJ, Le Naour G, Capron F. Automated mitosis detection using texture, SIFT features and HMAX biologically inspired approach. J pathol inform. 2013;4(2):12.
15. Kowal M, Filipczuk P, Obuchowicz A, Korbicz J, Monczak R. Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. Comput Biol Med. 2013;43(10):1563–72.
16. Filipczuk P, Fevens T, Krżyżak A, Monczak R. Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies. IEEE Trans Med Imaging. 2013;32(12):2169–78.
17. George YM, Zayed HH, Roushdy MI, Elbagoury BM. Remote computer-aided breast cancer detection and diagnosis system based on cytological images. IEEE Syst J. 2013;8(3):949–64.
18. Doyle S, Agner S, Madabhushi A, Feldman M, Tomaszewski JJ. In: 2008 5th IEEE International Symposium on Biomedical Imaging: From Nano to Macro. IEEE; 2008. p. 496–9.
19. Araújo T, Aresta G, Castro E, Rouco J, Aguiah P, Eloy C, Polónia A, Campilho A. Classification of breast cancer histology images using convolutional neural networks. PloS One. 2017;12(6):e0177544.
20. Bayramoglu N, Kannala J, Heikkilä J. In: 2016 23rd International conference on Pattern Recognition (ICPR). IEEE; 2016. p. 2440–5.
21. Spanhol FA, Oliveira LS, Petitjean C, Heutte L. In: 2016 International Joint Conference on Neural Networks (IJCNN). IEEE; 2016. p. 2560–7.
22. Han Z, Wei B, Zheng Y, Yin Y, Li K, Li S. Breast cancer multi-classification from histopathological images with structured deep learning model. Sci Rep. 2017;7(1):4172.
23. Feng Y, Zhang L, Mo J. Deep manifold preserving autoencoder for classifying breast cancer histopathological images. IEEE/ACM trans comput biol bioinform. 2018;17(1):91–101.
24. Wang C, Gong W, Cheng J, Qian Y. DBCNN: dependency-based lightweight convolutional neural network for multi-classification of breast histopathology images. Biomed Signal Process Control. 2022;73: 103451.
25. Zhang Z, Liu Q, Wang Y. Road extraction by deep residual U-Net. IEEE Geosci Remote Sens Lett. 2018;15(5):749–53.
26. Alom MZ, Yakopcic C, Hasan M, Taha TM, Asari VK. Recurrent residual u-net for medical image segmentation. J Med Imaging. 2019;6(1): 014006.

27. Lian S, Li L, Lian G, Xiao X, Luo Z, Li S. A global and local enhanced residual u-net for accurate retinal vessel segmentation. IEEE/ACM trans comput biol bioinform. 2019;18(3):852–862.

28. Xie S, Girshick R, Dollár P, Tu Z, He K. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017, p. 1492–500.

29. Zagoruyko S, Komodakis N. Wide residual networks. 2016. arXiv preprint arXiv:1605.07146.

30. Zerhouni E, Lányi D, Viana M, Gabrani M. In: 2017 IEEE 14th International Symposium on Biomedical Imaging (ISBI 2017). IEEE; 2017. p. 924–8.

31. Radosavovic I, Kosaraju RP, Girshick R, He K, Dollár P. Designing network design spaces. 2020. arXiv:2003.13678 [cs.CV].

32. Tan M, Le Q. In: International Conference on Machine Learning. PMLR; 2019. p. 6105–14.

33. Dosovitskiy A, Beyer L, Kolesnikov A, Weissenborn D, Zhai X, Unterthiner T, Dehghani M, Minderer M, Gelly S, et al. An image is worth 16x16 words: transformers for image recognition at scale. 2020. arXiv preprint arXiv:2010.11929.

34. Benhammou Y, Achchab B, Herrera F, Tabik S. BreakHis based breast cancer automatic diagnosis using deep learning: Taxonomy, survey and insights. Neurocomputing. 2020;375:9–24.

35. He K, Zhang X, Ren S, Sun J. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016. p. 770–8.

36. Ronneberger O, Fischer P, Brox T. In: International Conference on Medical Image Computing and Computer-Assisted Intervention. Springer; 2015. p. 234–41.

37. Woo S, Park J, Lee JY, Kweon IS. In: Proceedings of the European Conference on Computer Vision (ECCV), 2018. p. 3–19.

38. Yu D, Wang H, Chen P, Wei Z. In: International Conference on Rough Sets and Knowledge Technology. Springer; 2014. p. 364–75.

39. Alom MZ, Yakopcic C, Nasrin M, Taha TM, Asari VK. Breast cancer classification from histopathological images with inception recurrent convolutional neural network. J digit imaging. 2019;32(4):605–617.

40. Aresta G, Aratijó T, Kwok S, Chennamsetty SS, Safwan M, Alex V, Marami B, Prastawa M, Chan M, Donovan M, et al. Bach: grand challenge on breast cancer histology images. Med Image Anal. 2019;56:122–39.

41. Bolhasani H, Amjadi E, Tabatabaeian M, Jassbi SJ. A histopathological image dataset for grading breast invasive ductal carcinomas. Inform Med Unlocked. 2020;19: 100341.

42. Cruz-Roa A, Basavanahally A, González F, Gilmore H, Feldman M, Ganesan S, Shih N, Tomaszewski J, Madabhushi A. In: Medical imaging 2014: digital pathology, vol. 9041. SPIE; 2014. p. 904103.

43. Janowczyk A, Madabhushi A. Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. J pathol inform. 2016;7(1):29.

44. Singh D, Kumar V, Kaur M. Densely connected convolutional networks-based Covid-19 screening model. Appl Intell. 2021;51(5):3044–51.

45. Chollet F. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017. p. 1251–8.

46. Tan M, Chen B, Pang R, Vasudevan V, Sandler M, Howard A, Le QV. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2019. p. 2820–8.

47. Ma N, Zhang X, Zheng HT, Sun J. In: Proceedings of the European Conference on Computer Vision (ECCV), 2018. p. 116–31.

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