A Hybrid Convolutional and Recurrent Deep Neural Network for Breast Cancer Pathological Image Classification

Rui Yan[1,2], Fei Ren[2], Zihao Wang[2,3], Lihua Wang[4], Yubo Ren[4], Yudong Liu[2],
Xiaosong Rao[4※], Chunhou Zheng[1※], Fa Zhang[2※]
[1]College of Computer Science and Technology, Anhui University, Hefei, China
[2]High Performance Computer Research Center, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, China
[3]University of Chinese Academy of Sciences, Beijing, China
[4]Department of Pathology, Peking University International Hospital, Beijing, China
raoxiaosong@pkuih.edu.cn, zhengch99@126.com, zhangfa@ict.ac.cn

***Abstract*—Hematoxylin and Eosin (H&E) stained breast tissue samples from biopsies are observed under microscopy for the gold standard diagnosis of breast cancer. However, a substantial workload increases and the complexity of the pathological images make this task time-consuming and may suffer from pathologist's subjectivity. Facing this problem, the development of automatic and precise diagnosis methods is challenging but also essential for the field. In this paper, we propose a new hybrid convolutional and recurrent deep neural network for breast cancer pathological image classification. Our method considers the short-term as well as the long-term spatial correlations between patches through RNN which is directly incorporated on top of a CNN feature extractor. Experimental results showed that our method obtained an average accuracy of 90.5% for 4-class classification task, which outperforms the state-of-the-art method. At the same time, we release a bigger dataset with 1568 breast cancer pathological images to the scientific community, which are now publicly available from http://ear.ict.ac.cn/?page_id=1576. In particular, our dataset covers as many different subclasses spanning different age groups as possible, thus alleviating the problem of relatively low classification accuracy of benign.**

Keywords—image classification, deep neural network, CNN, RNN, breast cancer pathological image, dataset

I. INTRODUCTION

Diagnosis from a pathological image is the gold standard in diagnosing of cancer. With the advent of precision medicine, accurate analysis of pathological image has become a necessary guarantee for medical diagnosis and treatment. At present, pathological image in clinical practice is mainly based on the manual qualitative analysis of pathologists. However, a large number of studies have shown that different pathologists have great inconsistencies in the manual analysis. The main reason for the inconsistency is artificial analysis method has strong subjectivity and vulnerable to environmental factors [1]. This inconsistency is very unfavorable for clinical diagnosis, so quantitative classification methods are urgently needed to solve the defects of artificial analysis.

Recently, deep learning methods have made great progresses and achieved remarkable performance in the field of computer vision and image processing. This has also inspired many scholars to apply this technique to pathological image analysis [2]. The automatic and accurate classification of high resolution pathological images is the cornerstone and bottleneck of other in-depth studies such as localization, detection and segmentation. Traditional machine learning methods have faced a lot of application limitations. It depends on the manual design features, while these features require a lot of professional knowledge and difficult to cover the whole characteristics of image. On the contrary, deep learning is an end-to-end learning model. It directly takes pathological images as input and can automatically learn and extract hidden information of disease feature from big image data, avoiding the process of artificially designing features. Meanwhile, the network structure has high robustness of the translation, rotation and scaling. These advantages provide conditions for effective quantitative analysis of pathological images.

In this paper, we propose a method that integrate the advantages of Convolution Neural Network (CNN) [3] and Recurrent Neural Network (RNN) [4] thus the short-term as well as the long-term spatial correlations between patches is preserved. Because the resolution of pathological image is too high, limited by hardware (GPU memory) conditions, it is inevitable to split it into small patches. Then, the CNN is used to extract the image features of each patch. Finally, the RNN is used to fuse the features of patches to make the final image classification. For 4-class classification task, we obtained an average accuracy of 90.5%, which outperforms the state-of-the-art method. At the same time, cooperated with Peking University International Hospital ,we released a dataset with 1568 breast cancer pathological images which led to an order of magnitude increase in dataset volume. The format of our dataset is completely consistent with the currently largest open dataset (249 images) which published by Bioimaging2015 [5]. Experimental results show that the average sensitivity of our method for normal, benign, *in situ* carcinoma and invasive carcinoma are improved 5.7%, 13.4%, 4.8% and 1.2% respectively, comparing with the results on Bioimaging2015 dataset. Especially worth emphasizing is that, due to our dataset covers as many different subsets spanning different age groups as possible, the classification sensitivity of benign is improved significantly from 66.7% to 80.1%. This indicates that high performance deep learning algorithm and large enough as

※Corresponding author.

well as diverse dataset are both essential to improve the ability of pathological image classification.

II. BACKGROUND

Although many studies have been conducted and a series of important advances have been made in the automatic classification of pathological image of breast cancer, the characteristics of pathological images themselves, such as the subtle differences between images, the phenomenon of cell overlapping and the uneven color distribution, have brought great difficulties to image classification. Moreover, labeled and public dataset is very scarce. These problems make automatic classification of breast cancer pathological image remains a challenge.

In the early pathological image classification of breast cancer, the studies focused primarily on the 2-class classification [6-7] of cancer and noncancer or a more complex 3-class classification [8] of normal, *in situ* carcinoma and invasive carcinoma. The most of the works have been carried on the nuclei extracted or entire image using textural, morphological and architectural features base on the approach of traditional machine learning. It is worth noting that most of the above classification approaches are carried out on different image magnification and different dataset size. There is no uniform comparison standard among these approaches. More importantly, these approaches used artificial-based feature extraction methods, which not only require enormous effort and professional domain knowledge, but also have certain difficulties in extracting distinguishing high-quality features. These seriously restricted the application of traditional machine learning approaches in the classification of pathological image of breast cancer.

Recently, the deep learning method has not only demonstrated excellent performance in computer vision tasks but also made breakthroughs in medical imaging tasks such as identification, detection and segmentation [9]. Spanhol et al. [10] disclosed the pathological image dataset of breast cancer which named BreaKHis. Based on the dataset, they used the AlexNet network and adopted different fusion strategies to classify, with a recognition rate of 6% higher than traditional machine learning algorithms. Bayramoglu et al. [11] also used the magnification-independent deep learning method on the BreaKHis dataset, with a recognition rate of approximately 83%. Araújo et al. [12] first considered 4-class classifications for breast cancer pathological image. They first extracts features based on CNN, and then uses a SVM to classify the extracted features.

Several methods for automatic classification of breast cancer pathological image have been developed and evaluated in ICIAR2018 challenge [13]. The basic ideas of these method are much the same. The high-resolution pathological image is first pre-processed and data-enhanced, and then divided into equal-sized patches, and each patch is classified or extracted the features by CNN. An image-wise classification is then made based on the vote of patch-wise classifications result or fusion of extracted features. Sulaiman Vesal et al. [14] proposed a transfer learning method. Based on the pre-training model of GoogleNet and ResNet, they first classify patches of one image, and then use the majority voting method to obtain the image-wise classification results. Yeeleng S. Vang et al. [15] first proposed using Inception-V3 to perform patch-wise classification. The patch-wise predictions are then passed through an ensemble fusion framework involving majority voting, gradient boosting machine and logistic regression to obtain the image-wise prediction. Alexander Rakhlin et al. [16] employed a different approach known as deep convolutional feature representation. To this end, pathological images are first encoded with the general purpose networks to obtain sparse descriptors of low dimensionality (1408 or 2048). Finally, they use gradient boosted trees for classification. Ruqayya Awa et al. [17] used ResNet to obtain twelve 8192-dimensional feature vectors which represented twelve non-overlapping patches of 512×512 pixels from the input image. To train a classifier with larger context, they then trained an SVM classifier with the flattened features of 2×2 overlapping blocks of patches which is equivalent to training the classifier with the features of patch 1024×1024 pixels in size. The image-wise label was then decided by majority voting on the labels of overlapping blocks.

Although all of the aforementioned methods focus on the result of patch-wise fusion to get the final image-wise classification results, these methods are either directly using majority voting and SVM, or just integrating short-distance patch dependencies. They ignore the important role of long-distance spatial dependence in classification.

III. DATASET

One major advantage of deep learning is that it can benefit from large amounts of training data. Breakthrough results in computer vision were obtained on the ILSVRC challenges based on the ImageNet dataset [18]. However, there are few public large scale image dataset in the medical image domain because of its large variation and the need for scarce and expensive medical experts to label high-quality medical image data. We can learn from the literature that most of the research on breast cancer pathological image analysis is done on small dataset, which are usually not available to the scientific community. Veta et al. [19] pointed out that the main obstacle in developing new pathological image analysis methods is the lack of large, open and labeled dataset. Moreover, conventional means of annotating natural images, e.g. crowd-sourcing, cannot be applied to medical images due to the fact that these tasks often require years of professional training and domain knowledges.

The grand challenge of medical image field have greatly contributed to the development of medical image analysis. Since 2007, medical imaging conferences and workshops, such as MICCAI, ISBI and ICIAR, published a large collection of medical dataset for benchmark research, available at http://www.grand-challenge.org. The obvious advantage of these public benchmark dataset is that they provide a precise definition of tasks and assessment metrics to facilitate fair and standardized comparison of the performance of various algorithms. In the field of breast cancer pathological image processing, one of the largest open datasets containing 249 images is released by "Bioimaging2015: 4th international symposium in applied bioimaging". The goal of this challenge is to provide an automatic classification for each input image. ICIAR2018 gland challenge on breast cancer pathological image provided 400 images which consistent with the format of the Bioimaging2015 dataset. The pathological images were divided into 4 categories, each with 100 pictures. To the best

of our knowledge, this is by far the largest dataset of breast cancer pathological image, but only can be available during the challenge opening. Although these open dataset have played a very significant role in improving the classification accuracy of breast cancer pathological images, the dataset volume of 249 and 400 images is still too small compared with the open dataset of natural images.

We cooperated with Peking University International Hospital to release a new pathological image dataset of breast cancer. The format of our increased breast cancer pathological image dataset is completely consistent with the dataset published by Bioimaging2015. Our image dataset consists of 1319 high resolution (2048 × 1536 pixels), uncompressed and annotated H&E stained images. All images are digital and have the same acquisition conditions: 200x magnification and 0.42m × 0.42m pixel size. The preparation procedure used in this work is the standard paraffin process, which is widely used in clinical routine. Each image is labelled as normal, benign, *in situ* carcinoma or invasive carcinoma according to the predominant cancer type in each image. The annotation was performed by two medical experts and images where there was disagreement were discarded. Table Ⅰ summarizes the image distribution. The initial is the dataset of Bioimaging2015 and the extended is our increased dataset, it can be regarded as the extension of the dataset in article [12]. Table Ⅱ describes the image format in our dataset. On the whole, based on the initial 249 images, we increased the amount of dataset to 1568. In particular, our dataset covers as many different subsets spanning different age groups as possible, which can fully reflect the morphology of breast tissue.

IV. METHODS

When input a pathological image with high resolution (2048 × 1536 pixels), our goal is to accurately classify it into four categories: normal, benign, *in situ* carcinoma and invasive carcinoma. We proposed a hybrid convolutional and recurrent deep neural network for breast cancer pathological image classification, the general workflow of our method is as follows (Fig. 2). First, the pathological images were pre-processed and enhanced to improve quantitative analysis. After pre-processing, one pathological image is divided into 12 small patches on average. Then, a fine-tuned Inception-V3 was used to extract the image features at patch-wise. Each of patch is extracted to a feature vector of 1 × 2048

TABLE I. SUMMARY OF OUR DATASET

Dataset	**Normal**	**Benign**	***In situ* carcinoma**	**Invasive carcinoma**	**Total**
Initial	55	69	63	62	249
Extended	273	451	339	256	1319
Overall	328	520	402	318	1568

TABLE II. DESCRIPTION OF PATHOLOGICAL IMAGES IN OUR DATASET

Color model	R(ed)G(reen)B(lue)
Size	2048×1536 pixels
Memory space	3-20 MB (approx.)
Type of label	image-wise

dimensions. That is, one pathological image can be extracted 12 feature vectors. Finally, the 12 feature vectors (12 × 1 × 2048) are input into a bidirectional LSTM to fuse the features of the 12 small patches to make the final complete image-wise classification. Since our method integrates the advantages of CNN and RNN, the short-term as well as the long-term spatial correlations between patches can be preserved.

A. Image pre-processing and augmentation

In order to improve quantitative analysis, we used the SVD geodesic method [20] to normalize the pathological images of H&E staining. In our study, we perform 50 random color augmentations for each image. In order to decompose the RGB color of the tissue into H&E color space, we used a color deconvolution algorithm [21] to adjust the amount of H&E by multiplying the magnitude of H&E of every pixel by two random uniform variables from the range (0.7,1.3) .

It is well known that deep learning approaches are heavily dependent on the volume of training dataset available, with models of higher complexity requiring more data to generalize well and avoid over-fitting to the training samples. In general, the breast pathological images provided are very large in size, spanning 2048 × 1536 pixels. In order to address the issues of insufficient data and large image sizes, we extracted patches from each image and augmented the data by applying varying degrees (90°, 180°, 270°) of rotation and flipping the extracted patches. This mode of data augmentation emulates a real-world scenario as there is no fixed orientation adopted by pathologists when analyzing pathological images. The label for each patch was inherited from the class assigned to the original image.

B. Patch-wise method

The application of CNNs pre-trained on large annotated image databases, such as ImageNet from different domains for various classification tasks, is referred to as transfer learning. With such an approach, the original network architecture is maintained and the pre-trained weights are used to initialize the network. The initialized weights are subsequently updated during the fine-tuning process, enabling the network to learn features specific to the task of interest. Recently, numerous studies have demonstrated that fine-tuning is efficient for a variety of classification tasks in the medical domain. In this paper, we use a well-known pre-trained CNN architecture, named, Google's Inception-V3 [22], which is fine-tuned to learn domain and modality specific features for classifying breast pathological images. Such a network is easier to optimize and consequently, enables training of deeper networks, which correspondingly leads to an overall improvement in network capacity and performance. The Google's Inception-V3 network employs factorized inception modules, allowing the network to choose suitable kernel sizes for the convolution layers. This enables the network to learn both low-level features with small convolutions and high-level features with larger ones.

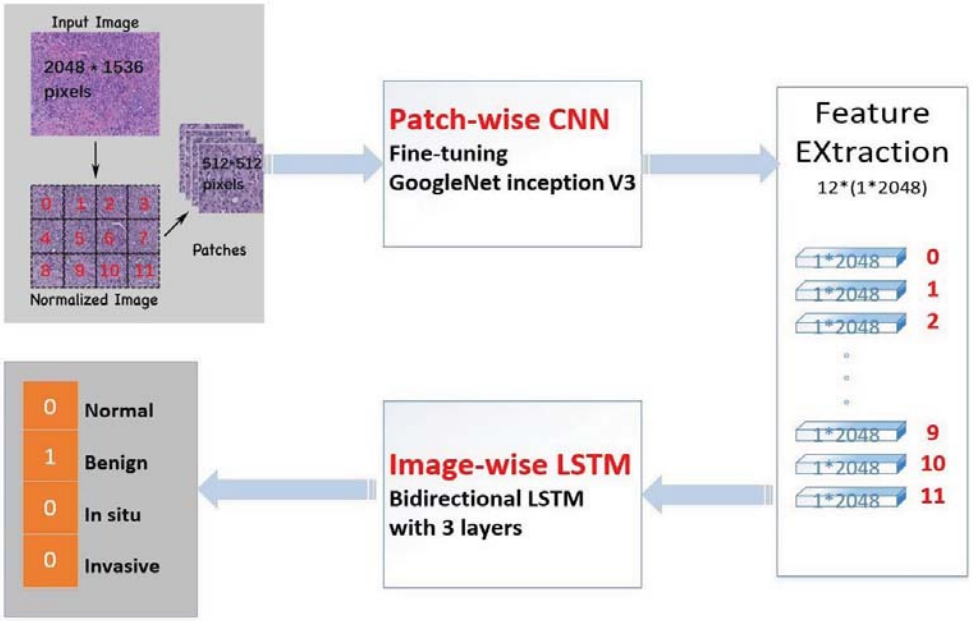


Fig. 2. An overview of the proposed workflow. First, we do pre-processing and data enhancement on pathological images. After this, a complete picture is divided into 12 small patches on average. Then, each pre-processed image is sent to a fine-tuned Inception-V3 to extract a feature of 1 × 2048 dimensions. In other words, one image gets 12 feature vectors. Finally, the 12 feature vectors (12 × 1 × 2048) are inputted into a bidirectional LSTM with 3 layers to fuse the features of the 12 small patches to make the final complete image-wise classification.

For features extraction, we used the standard pre-trained Inception-V3 from TensorFlow slim distribution. Here, we removed fully connected layers from model and converted the last convolutional layer consisting of 2048 channels via Global Average Pooling into a one-dimensional feature vector with a length of 2048.

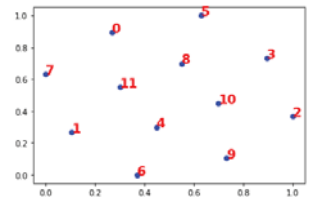
C. *Image-wise method*

When using the deep learning to process natural images, a complete image is directly used as input for end-to-end training. However, the size of the pathological image is too large, limited by hardware conditions, it is inevitable that the original picture is divided into several smaller patches. The accompanying problem is how to combine the results of each small patch and obtain the final image-wise classification result. The most common methods are majority voting and SVM. These two methods are simple and direct, but they also achieve good results. However, such a simple and direct method loses a lot of context information in the image. How to keep the context information of each small image is the focus of recent research. The two methods presented in the recently published paper by Ruqayya Awa et al. [5] and Kamyar Nazeri et al. [17] are also dedicated to retaining contextual information and have performed well. The method proposed by Ruqayya Awa et al. binds the four features extracted from the spatially close patches into one. But this simple flattened is difficult to fuse spatially close features together. Another method to preserve the contextual information of the picture is proposed by Kamyar Nazeri et al.. In their method, the first "patch-wise" CNN acts as an auto-encoder that extracts the most salient features of image patches while the second "image-wise" CNN performs classification of the whole image. However, they can only retain the information of the top, bottom, left and right of a picture, and the remote context information cannot be retained.

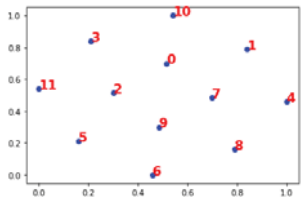
TABLE III. COMPARATION OF ACCURACY WITH PREVIOUS METHODS

Method	Patch-wise accuracy (%)	Image-wise accuracy (%)
Araujo et al.[12]	66.7	77.8
Le Hou et al.[23]	-	79.8
Rakhlin et al.[16]	-	87.2
Yeeleng S. Vang et al.[15]	-	87.5
Aditya Golatkar et al.[24]	79	85
Ruqayya Awan et al.[5]	-	83
Hongliu CAO et al.[25]	-	87.1
Our Proposed	82.1	90.5

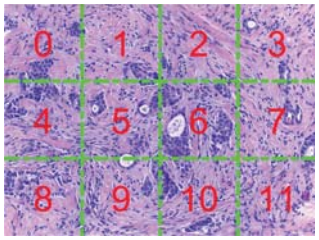
In response to the above deficiencies, we propose to use RNN to fuse the context information of features which is directly incorporated on top of a CNN feature extractor to make the final classification decision. In our model, the convolution layer captures patch features, while the recurrent layer captures short-term as well as long-term dependencies between the patch features in order to keep information of context. LSTM is one of the most commonly used variants of RNN, unlike a CNN, connections between units of an RNN form a directed cycle. This creates an internal state of the network that allows it to exhibit dynamic temporal or spatial behavior. A bi-directional long short-term memory network (BLSTM) is an extension of the LSTM that combines the outputs of two RNNs, one processing the input data from left to right, the other one from right to left. Instead of regular hidden units, the two RNNs contain LSTM blocks, which are smart network units that can remember a value for an arbitrary length of time. In our proposed method, one pathological image can be extracted 12 feature vectors (12 × 1 × 2048) by CNN. These 12 feature vectors are inputted into a 3-layer bidirectional LSTM. Finally, we add a fully connected layer in the last layer of the LSTM. Because we are doing 4 classifications, so the output of this fully connected layer has 4 nodes.



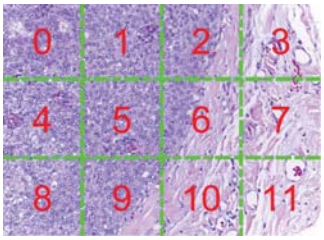
(a) Representation of invasive carcinoma image using t-SNE



(b) Representation of *in situ* carcinoma image using t-SNE



(c) Invasive carcinoma image



(d) *In situ* carcinoma image

Fig. 3. Visualization of two pathological images using t-SNE.

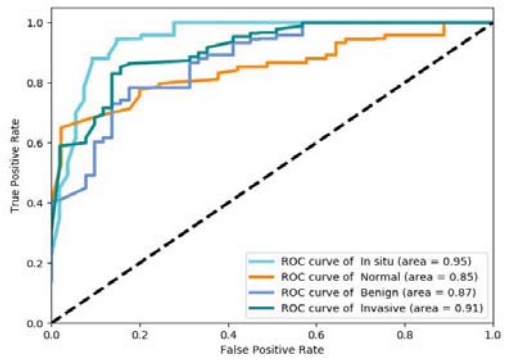
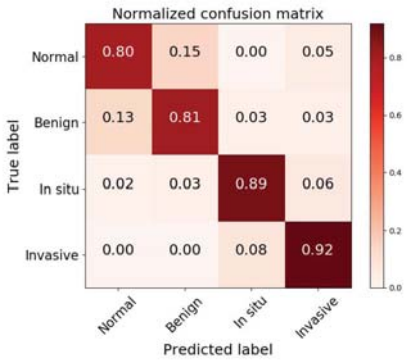


Fig. 4. Receiver Operating Characteristic (ROC) and AUC.

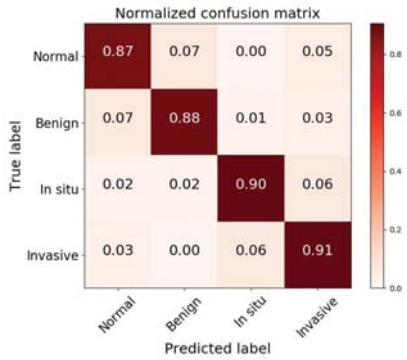
V. RESULTS

A. t-SNE visualization

T-distributed Stochastic Neighbor Embedding (t-SNE) [26] is a nonlinear dimensionality reduction technique well-suited for embedding high-dimensional data for visualization in a low-dimensional space of two or three dimensions. Specifically, it models each high-dimensional object by a two- or three-dimensional point in such a way that similar objects are modeled by nearby points and dissimilar objects are modeled by distant points with high probability. In short, the core idea behind t-SNE is to find a two-dimensional representation of the data, keeping the distance between the data points as much as possible. Fig. 3a and Fig. 3b show the two-dimensional (2D) representation of 12 feature vectors extracted from 12 patches from a breast cancer pathological image using t-SNE. Each data point in Fig. 3a and Fig. 3b represent the feature vector extracted from the corresponding patch in Fig. 3c and Fig. 3d. From the figure we can seen that the data points represent some spatially remote patches are very close on the final 2D representation. This phenomenon shows that there is also a close relationship between the distant patches. Hence, by joint considering the short-term as well as the long-term spatial correlations between patches, our proposed method can deep mine the pathologic features of breast cancer.



(a) Four class confusion matrix using dataset containing 400 images



(b) Four class confusion matrix using our released dataset

Fig. 5. Four class confusion matrix.

B. Accuracy comparison with previous methods

The performance of our proposed method on patch-wise and image-wise is shown in table Ⅲ. We compared the accuracy with state-of-the-art methods. Because some of the previously published papers used Bioimaging2015 dataset with 249 images and others used ICIAR2018 dataset with 400 images. Since the number of 249 and 400 images is not much different, for the convenience of comparison, we only compared the accuracy of the method in the case of 400 images. We randomly choose 400 images from our dataset and perform a 70%-20%-10% training-validation-test split. The training and validation sets are used for model development while the test set is held out and only used for evaluation.

For the four classifications of pathological images, our method achieved 83.1% average accuracy in patch-wise and 90.5% average accuracy in image-wise. The reason for the good performance in patch-wise is that we use a pre-training model that allows for better generalization on a smaller number of pathological image dataset. Moreover, different from the previous work which just using the original CNN, we use the more advanced Google's Inception-V3, which ensures the model's better learning ability. At the same time, we analyzed the reasons for achieving good performance on image-wise. Because we used an integration method of deep neural network, the short-term as well as the long-term spatial correlations between patches is preserved.

C. Confusion matrix and AUC

The confusion matrix of the predictions is presented in Fig. 5a using our method which trained on dataset contains a

total of 400 images. And Fig. 4 shows the mean Area Under Curve (AUC) of 89.5% corresponding to 95%, 85%, 87% and 91% for the four classes based on Receiver Operating Characteristic analysis.

From the results, we can see that the classification result of benign is relatively low. The reason for this phenomenon is that in the routine pathological diagnosis, the subclass of benign is not only diverse, but also closely related to the age of the patient. Therefore, in the case of a limited amount of data, it is difficult to cover enough features of pathological images of benign. So the final classification result is relatively low. To solve this problem, our dataset deliberately selected different subclasses of benign pathological images spanning different age groups. When the training dataset change to our expanded dataset (1568 images), the confusion matrix of the predictions is presented in Fig. 5b. It can be seen that the performance of benign is relatively improved, from 81% to 88%. However, with the increase in the amount of dataset and the use of our new method, the classification accuracy of *in situ* and invasive are not much improved. One possible reason is that their characteristics are relatively clear and the diversity of features is very low, while general models and small data volumes can already learn enough feature representations.

VI. CONCLUSIONS

In this article, we proposed a new method for breast cancer pathological image classification using a hybrid convolutional and recurrent deep neural network. Our method integrated the advantages of CNN and RNN, the short-term as well as the long-term spatial correlations between patches were both preserved. Through extensive assessments and comparisons, it was shown that our new method outperforms the state-of-the-art method. Meanwhile, we released a bigger dataset of breast cancer pathological images to the scientific community. We hope that the dataset can serve as a benchmark to facilitate a broader study of deep learning in the field of breast cancer pathologic images.

ACKNOWLEDGMENT

This research is supported by the National Key Research and Development Program of China (No. 2017YFE0103900 and 2017YFA0504702), the NSFC projects Grant (No. U1611263, U1611261, 61472397, 61502455 and 61672493), Peking University International Hospital Research Grant (No. YN2018ZD05) and Special Program for Applied Research on Super Computation of the NSFC-Guangdong Joint Fund (the second phase).

REFERENCES

[1] M. Veta, J. P. Pluim, P. J. van Diest, and M. A. Viergever, "Breast cancer histopathology image analysis: a review," IEEE transactions on bio-medical engineering, vol. 61, pp. 1400-11, 2014.

[2] G. Litjens, T. Kooi, B. E. Bejnordi, S. Aaa, F. Ciompi, M. Ghafoorian, et al., "A survey on deep learning in medical image analysis," Medical Image Analysis, vol. 42, pp. 60-88, 2017.

[3] Y. Lecun, Y. Bengio, and G. Hinton, "Deep learning," Nature, vol. 521, p. 436, 2015.

[4] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," Neural Computation, vol. 9, pp. 1735-1780, 1997.

[5] Available:https://rdm.inesctec.pt/dataset/nis-2017-003/resource/df04ea95-36a7-49a8-9b70-605798460c35.

[6] S. Doyle, S. Agner, A. Madabhushi, M. Feldman, and J. Tomaszewski, "Automated grading of breast cancer histopathology using spectral clustering with textural and architectural image features," in IEEE International Symposium on Biomedical Imaging: From Nano To Macro, pp. 496-499, 2008.

[7] C. Bilgin, C. Demir, C. Nagi, and B. Yener, "Cell-graph mining for breast tissue modeling and classification," Conf Proc IEEE Eng Med Biol Soc, vol. 2007, pp. 5311-5314, 2007.

[8] A. Brook, E. Y. Ran, E. Isler, R. Kimmel, R. Meir, and D. Peleg, "Breast Cancer Diagnosis From Biopsy Images Using Generic Features and SVMs," IEEE Transactions on Information Technology in Biomedicine, 2007.

[9] B. Zhang, "Breast cancer diagnosis from biopsy images by serial fusion of Random Subspace ensembles," in International Conference on Biomedical Engineering and Informatics, pp. 180-186, 2011.

[10] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte, "Breast cancer histopathological image classification using Convolutional Neural Networks," in International Joint Conference on Neural Networks, 2016.

[11] N. Bayramoglu, J. Kannala, and J. Heikkilä, "Deep learning for magnification independent breast cancer histopathology image classification," in International Conference on Pattern Recognition, pp. 2440-2445, 2017.

[12] T. Araújo, G. Aresta, E. Castro, J. Rouco, P. Aguiar, C. Eloy, et al., "Classification of breast cancer histology images using Convolutional Neural Networks," Plos One, vol. 12, p. e0177544, 2017.

[13] Available:https://iciar2018-challenge.grand-challenge.org/Home/.

[14] S. Vesal, N. Ravikumar, A. A. Davari, S. Ellmann, and A. Maier, "Classification of Breast Cancer Histology Images Using Transfer Learning," in International Conference Image Analysis and Recognition, pp. 812-819, 2018.

[15] Y. S. Vang, Z. Chen, and X. Xie, "Deep Learning Framework for Multi-class Breast Cancer Histology Image Classification," in International Conference Image Analysis and Recognition, pp. 914-922, 2018.

[16] A. Rakhlin, A. Shvets, V. Iglovikov, and A. A. Kalinin, "Deep Convolutional Neural Networks for Breast Cancer Histology Image Analysis," pp. 737-744, 2018.

[17] R. Awan, N. A. Koohbanani, M. Shaban, A. Lisowska, and N. Rajpoot, "Context-Aware Learning using Transferable Features for Classification of Breast Cancer Histology Images," 2018.

[18] O. Russakovsky, J. Deng, H. Su, J. Krause, S. Satheesh, S. Ma, et al., "ImageNet Large Scale Visual Recognition Challenge," International Journal of Computer Vision, vol. 115, pp. 211-252, 2015.

[19] Veta M, Pluim J P, van Diest P J, et al. Breast cancer histopathology image analysis: a review.[J]. IEEE Trans Biomed Eng, 2014, 61(5):1400-1411.

[20] M. Macenko, M. Niethammer, J. S. Marron, D. Borland, J. T. Woosley, X. Guan, et al., "A method for normalizing histology slides for quantitative analysis," in IEEE International Symposium on Biomedical Imaging: From Nano To Macro, pp. 1107-1110, 2009.

[21] A. C. Ruifrok and D. A. Johnston, "Quantification of histochemical staining by color deconvolution," Anal Quant Cytol Histol, vol. 23, pp. 291-299, 2001.

[22] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," pp. 2818-2826, 2015.

[23] L. Hou, D. Samaras, T. M. Kurc, Y. Gao, J. E. Davis, and J. H. Saltz, "Patch-Based Convolutional Neural Network for Whole Slide Tissue Image Classification," in Computer Vision and Pattern Recognition, pp. 2424-2433, 2016.

[24] A. Golatkar, D. Anand, and A. Sethi, "Classification of Breast Cancer Histology using Deep Learning," 2018.

[25] H. Cao, S. Bernard, L. Heutte, and R. Sabourin, "Improve the performance of transfer learning without fine-tuning using dissimilarity-based multi-view learning for breast cancer histology images," 2018.

[26] A. Pérezgonzález, M. Vergara, J. L. Sanchobru, D. M. Van, L J P, G. E. Hinton, D. Shanmugapriya, et al., "Visualizing Data using t-SNE," 2015.