Automated Classification of Breast Cancer Histology Images Using Deep Learning Based Convolutional Neural Networks

Majid Ali Nawaz
Faculty of Computer and information, Assiut University

Adel A. Sewissy
Faculty of Computer and information, Assiut University

Taysir Hassan A. Soliman
Faculty of Computer and information, Assiut University

Summary
Automated classification of cancers using histopathological images is a challenging task of accurate detection of tumor sub-types. In this paper, we applied fine-tuned pre-trained deep neural networks classified on BreakHis datasets on eight distinct classes for benign has four sub-classes (adenosis, fibroadenoma, phyllodes tumor, and tubular adenoma) malignant has four sub-classes (ductal carcinoma, lobular carcinoma, mucinous carcinoma, and papillary carcinoma) all together on difference model on Inception (V1,V2) and ResNet V1 50. The confusion matrix showing high accuracy value 95% with less error rate 0.011 .

Keywords:
Medical imaging, Computer-aided diagnosis (CAD), Deep Learning, Medical image processing, Convolution Neural Network.

1. Introduction

Breast cancer is considered to be the most rampant and deadliest among various types of cancers. It is reported that around 523,000 women die of breast cancer (BC) [1]. It is increasing at an alarming rate. The following chart on the statistics of BC cases of Australia can drive home the point.

Figure 1 shows the year-wise comparisons between a number of new female patient and the number of female death since the year 2007 in Australia. The case of Australia with a population of 20-25 million may act as an indication to understand the alarming situation worldwide. Carcinomas are grouped into two class namely benign and malignant each group have four types of tumor. While benign types of breast cancer contain adenosis, fibroadenoma, phyllodes tumor and tubular adenoma, malignant types containductal carcinoma, lobular carcinoma, mucinous carcinoma, and papillary carcinoma. Histopathological classification of breast carcinoma is usually based on the morphological features of the tumors. On the basis of the morphological features of tumors, they are grouped as major tumor types and minor tumor subtypes.

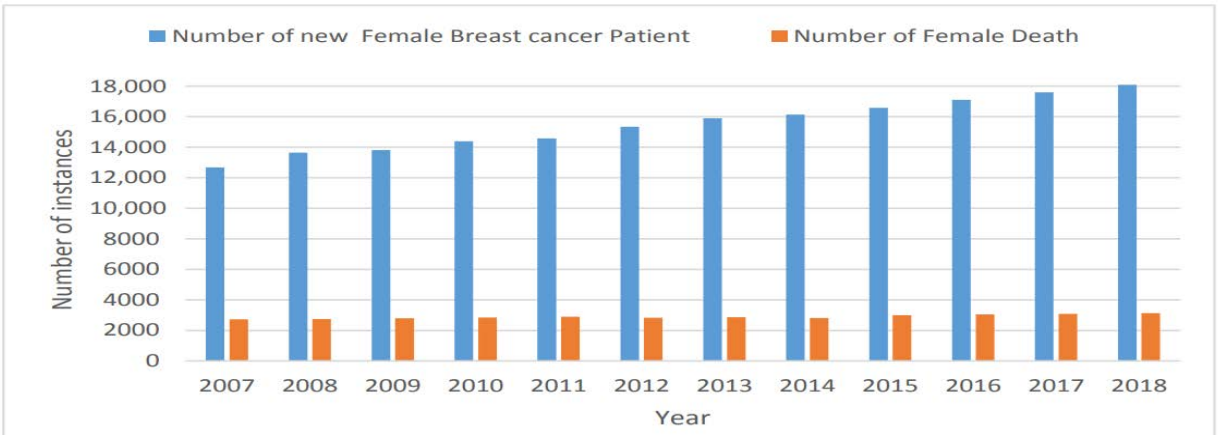

Fig. 1 New cases of breast cancer for women and number of women dying in the last twelve years.

There are 20 major tumor types and 18 minor tumor sub-types ([27]). Approximately, 70-80 percent of all breast cancers belong to either one of the two major histopathological classes, namely invasive ductal carcinoma (IDC) or invasive lobular carcinoma (ILC) [34]. The IDC class is divided into five different carcinoma sub-types including tubular, medullary, papillary, mucinous and cribriform carcinomas, while benign types of breast cancer contain adenosis, fibroadenoma, phyllodes tumor and tubular adenoma. More importantly, Identification of minor tumor sub-types known as special tumor types provides clinically useful information to

Manuscript received April 5, 2018
Manuscript revised April 20, 2018

determine an effective therapy. Many clinical studies reported a mismatch between immunohistochemically and molecular classification of breast cancer [3]. In 2011, StGallen International Expert Consensus validated the application of immunohistochemistry for identification breast cancer sub-types. The diversity in breast cancer types and the limited predictive power of the histopathological classification exceedingly urge to find a comprehensive approach for accurate evaluation of the morphological features of carcinomas. Ascertaining the type of tumor or carcinoma is complicated and time-consuming activity. It is also subject to human errors. For the detection of cancer and interpretation, pathologists have to study large numbers of tumor tissue slides. The processes of quantifying the cell or tumor on different parameters (e.g. miotic counts, surface area, and cell size) and evaluation of immunohistochemical molecular markers are multifarious and time-consuming. Manual inspection and interpretation are vulnerable to statistical, distributional and human errors due to improper images. These errors affect the accuracy in classification of cancers in conventional diagnosis. So there is a need to for an automated and reproducible diagnostic device to meet the requirements efficiently. Computer-aided diagnosis (CAD) has proven to be an excellent method of image-based medical examination. It also enables grading and staging of tumors. This method of diagnosis is also considered to be a cost-effective as reduces unwarranted expenses. Conventional image processing and machine learning techniques require extensive pre-processing, segmentation and manual extraction of specific visual features before classification. However, deep learning approaches have surpassed human performance in visual tasks by utilization of automated hierarchical feature extraction and classification by multi layers. These kinds of approaches can be applied for cancer diagnosis using tumor tissue images. The first application of the image processing on analytical pathology for cancer detection was introduced by True et al. [33]. This kind of application showed the implication of morphological features in diagnostic methods for malignant tumors. They used a series of morphological features including area fraction, shape, size and object counting to detect cell abnormalities. A large body of evidence has been published concerning cancer detection using various image processing and machine learning techniques. Application of these methods is limited due to manual feature extraction of the features. On the other hand, deep learning approach offers an automated, accurate and sensitive method to feature extraction from medical images. There emerged several deep learning based neural network systems that could identify different kinds of cancer. Incidentally, the Neighboring Ensemble Predictor (NEP) coupled with Constrained Convolutional Neural Network (CCNs) could lead to nucleus detection in colon cancer [8]. In agreement with this, four deep learning network architectures including GoogLeNet, AlexNet, VGG16 deep network [35] and ConvNet with 3, 4, and 6 layers were recently applied to identify breast cancer. Despite improvements in image analysis and interpretation, numerous questions related to the reliability and sensitivity of appropriate pathological diagnosis systems, particularly for breast cancer classification, have remained to be answered. In particular, there were no significant, comprehensive and promising solutions for discrimination of breast cancer subtypes. This study presents the accuracy comparison between two deep learning Inception (V1, V2) and ResNet-50 (InceptionV1) architectures to discriminate microscopic cancerous imaging. We demonstrate a highly accurate automatic framework for cancer detection and classification of its sub-types. Our framework employs data augmentation and advanced pre-processing. The paper is organized as follows: Section 2 describes related research, Section 3 describes the proposed approach, Section 4 describes materials and methods used in the present study, Section 5 describes the performance of our model on the BreakHis dataset as well as compare with the present findings, and we conclude our paper in Section 6.

2. Related work

Both Image Analysis and Machine Learning Methodologies include the CAD systems. As the name implies the Computer Aided Diagnosis systems have been designed to aid doctors to diagnose cases. They were attentively developed as second opinion systems and yield in such advantages as raising the efficiency level of the diagnostic process as well as the cost benefits and helping specialists save effort. Owing to the aforementioned advantages and in more particular terms, cancer diagnosis method has been the focus and several replications have been noted. The analysis of nuclei morphology, for example, can be clear enough to categorize a tissue malignant or benign. As a result, many research concentrated on the nuclei analysis malignant-benign categorization. Kowal et al. [11], however, employed a different classification of the algorithms for nuclei dissection on fine needle biopsy microscopic images. In order to reach a precision level that ranges from 84% to 93%, measured on 500 images of 50 patients, morphological, topological and texture traits were implemented. As for the Patient-wise clustering which has been performed by a great majority that voted on 10 images each, results indicated an accuracy level of 96–100%. Concomitantly, Filipczuk et al. [12] and George et al. [13] used the fine needle biopsies from which they excerpted nuclei-based traits. As a matter of fact, the

circular Hough transform was performed in order to identify nuclei subjects. Afterward, machine-learning and Otsu threshold were used to reduce false- positive variables. George et al. [13], on the other hand, the nuclei refinement was further segmented by watershed. In these studies, texture and shape traits were employed so as to train different classifiers. Filipczuk et al. [12] reached 98.51% accuracy by majority voting over 11 images for each of the 67 patients. However, George et al. [13] used 92 images results ranged from 71.9% and 97.15%in individual image classification. For the binary classification of more complex images, Belsare et al. [14] used tissue organization in addition to the nuclei. These researchers assessed 70 images extracted from a private 40× magnification breast histology H&E dataset. In order to divide the epithelial layer around the lumen of the cells, Spatio-color-texture graphs were employed while the statistical texture features were used in order to train the final classifiers. This resulted in accuracy between 70% and 100%. Many other scholars considered a more complex 3-classes classification of breast cancer histology images. Brook et al. [15] and Zhang et al. [16], for instance, divided the tissue images of breast cancer into normal, in situ carcinoma and invasive carcinoma. In order to reach such classification, these authors referred to the database from the Israel Institute of Technology [17]. Further to this, Brook et al. [15] dichotomized the images through the multiple threshold values and made use of related component statistics for the training of a support vector machine (SVM) classifier. They asserted that an average accuracy of 93.4% has been reached but can rise to 96.4 % in case 20% of the images were rejected. The cascade classification approach was used by Zhang et al. [16]. They randomly fed Subsets of Curvelet Transform and local binary pattern (LBP) features as the first set of parallel SVM classifiers. The images whose classifiers do not conform were disregarded and analyzed by another set of artificial neural networks (ANN) over other random feature subsets. Again, all images whose classifiers do not agree were rejected. This operation led to 97% accuracy with 0.8% rejection rate. In the recent years, the technology boom in computing power and dataset sizes helped the smooth application of Convolutional Neural Networks (CNNs) to visualize clustering problems. In contrast to the old approach of hand-crafted feature extraction methods, CNNs employed effective features directly from the training image patches by the optimization of the classification function loss. These deep learning models have resulted in achieved effective performance in image classification difficulties in different fields [18],[19], such as medical image analysis [20], and the histopathology images [21]. In addition, the CNNs helped reduce the field-knowledge that was required to develop a classification system. Accordingly, the methods' performance is less biased by the dataset used and thus similar network architectures can achieve more reliable results. Indeed and using multiple magnifications. Spanhol et al. [22] employed CNN architecture in the Imagenet network [18] in order to divide the samples of H&E breast tissue biopsy for benign and malignant tumors. In this study, 32 × 32 and 64 × 64 pixels patches were excerpted from the initial images and used to train the CNN. The last clustering was reached through the combination of the probabilities with the sum, the product or maximum rules. As a matter of fact, two patch extraction methods were investigated namely; sliding window and random extraction. The extraction of patches helped decrease the complexity of the model by reducing the input in the successive layers. The authors acknowledged the drop of the accuracy level for higher magnifications, which entails that their CNN architecture cannot extract relevant features for higher magnifications. In fact, only nuclei edge-related features are extracted for higher magnifications, as it will be discussed in this work. Other researchers, however, have successfully attuned the architecture of the CNN to breast histology related-problems. Ciresan et al. [19], for example, and in an attempt to train a CNN for mitosis detection in H&E stained breast biopsy slides, employed 101 × 101 patches. The architecture used helped to explore nuclei of different sizes and their districts. This method gained the ICPR 2012 Mitosis Detection Contest with an F1-score of 0.782. In order to identify invasive carcinoma regions in breast histology slides, Cruz-Roa et al. [23] trained a CNN on 100 × 100 pixels whole-slide patches, extracted using grid sampling. Owing to the global nature of the problem, however, the CNN feature-extraction scale ranges from nuclei to overall tissue organization. This method outperformed other state-of-the-art methods, achieving an F1-score of 0.780. For these last two works, the model was slipped through the image so as to reach a probability map and then the detection result was obtained via thresholding. In [19], the training, dataset size, and complexity were increased by using random rotations and mirroring the training occurrences.

2.1 Proposed Approach

In this study, we developed and introduced an accurate and reliable computer-based technique empowered with deep learning approaches to classify breast cancer sub-types from histopathological images derived from Hematoxylin and eosin stain (H&E). Our framework contains five steps: (a) Image acquisition and applying stain normalization, (b) Data augmentation (c) Deep learning pre-processing, (d) Transfer learning and fine-tuning pre-trained models, and (e) Hierarchical feature extraction and classification with Inception and ResNet networks. All steps have been illustrated in figure 2.

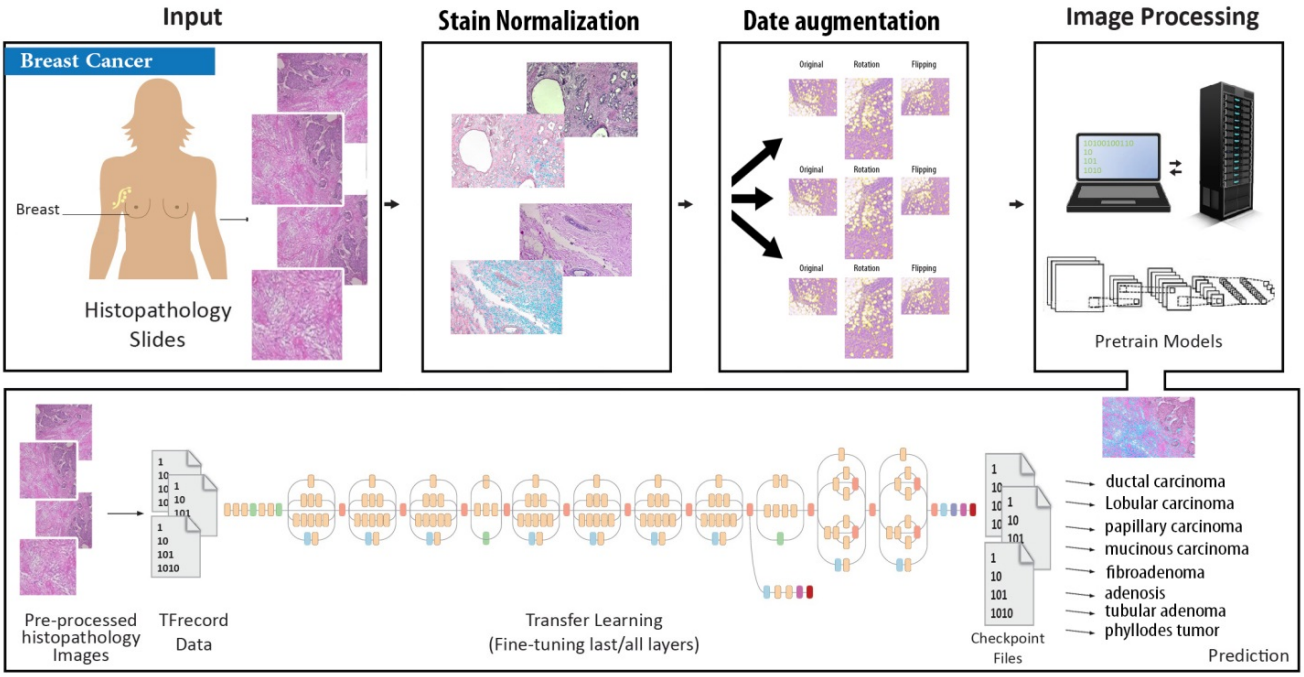

Fig. 2 Proposed approach

3. Materials and methods

3.1 Datasets

The dataset [22] used in this project was the Breast Cancer Histopathological Database or BreaKHis. The BreaKHis database is composed of 7,909 microscopic images of breast tumors. The images were collected from 82 different patients with 4 (40X, 100X, 200X, 400X) different magnifying factors. The database is separated into 2 super classes, benign and malignant, with 4 subclasses each. There are 2,480 benign and 5,429 malignant samples. The distribution of the sample is given in Table 1.

Table 1: BREAKHIS STRUCTURE

Magnification	***Benign***	***Malignant***	Total
40X	652	1370	1995
100X	644	1437	2081
200X	623	1390	2013
400X	588	1232	1820
Total of Images	2480	5429	7909

The benign(B) subclasses are as follows: adenosis (A), fibroadenoma (F), phyllodes tumor (PT), and tubular adenoma (TA); and the malignant(M) subclasses are as follows: carcinoma (DC), lobular carcinoma (LC), mucinous carcinoma (MC) and papillary carcinoma (PC), in Figure 3 represent different types of breast cancer tumors. The subclasses represent different types of breast cancer tumors. Different types of breast tumors are known to have different prognoses and need a different kind of classification. Each image filename contains information about the actual tissue itself. For example, SOB_B_TA-14-4659-40-001.png tells us that the biopsy procedure employed was SOB, the class B means that it was for the benign class, the TA means that the subclass of the tissue was a tubular adenoma, the patient identification number is 14-1659, magnification factor was 40X, and this was image 1.

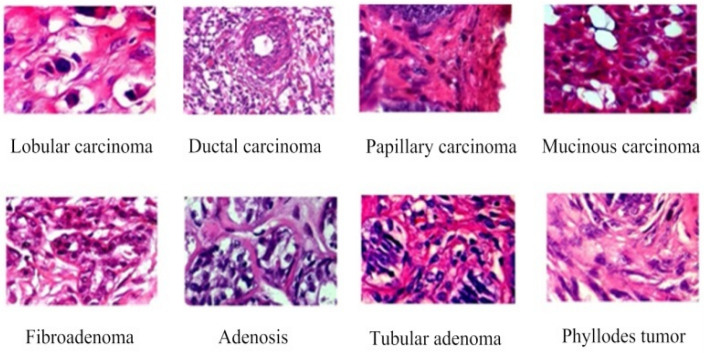

Fig. 3 Samples of BreaKHis subclasses.

3.2 Stain normalization.

Before analyzing them, images were normalized [37]. This method takes into account the staining technique used for the histology slides preparation. As a starting point and through the use of a logarithmic transformation, the colors of the images were converted into optical density (OD). Secondly, singular value decomposition (SVD) was implemented to the OD tuples in order to find the 2D projections with higher variance. The resulting color space transform is then applied to the original image. As a final step, the image histogram is overextended so as the dynamic range covers the lower 90% of the data. Figure 4 displays the two images before and after being normalized.

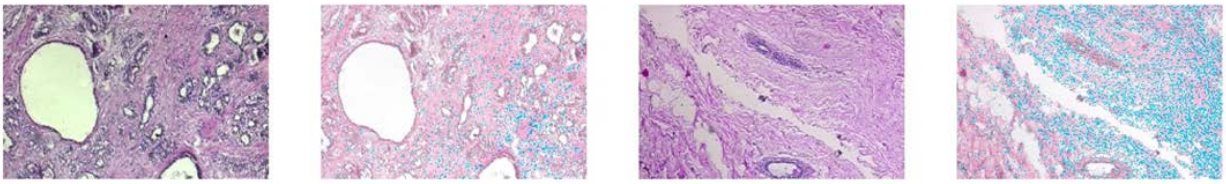

Fig. 4 Histology image normalization. A and C original images; B and D images after normalization.

3.3 Data Augmentation

Data augmentation is an essential step to have enough diverse samples which are required to train a deep network to learn from the images. Several studies investigated the role of data augmentation in deep learning [28].We considered data augmentation for breast cancer sub-types due to the existing differences in the number of images among different sub-type classes. Technically, data augmentation was accomplished on data acquired from Augmentor Python library by using and rotating and flipping methods (see Figure 5).In the Augmentor input images are rotated at 270 degrees, and then, the input images are flipped top-bottom to right with 0.6 probabilities and are recorded. Each output image was further processed on Protocol Buffers to convert to it a format called TFRecord.

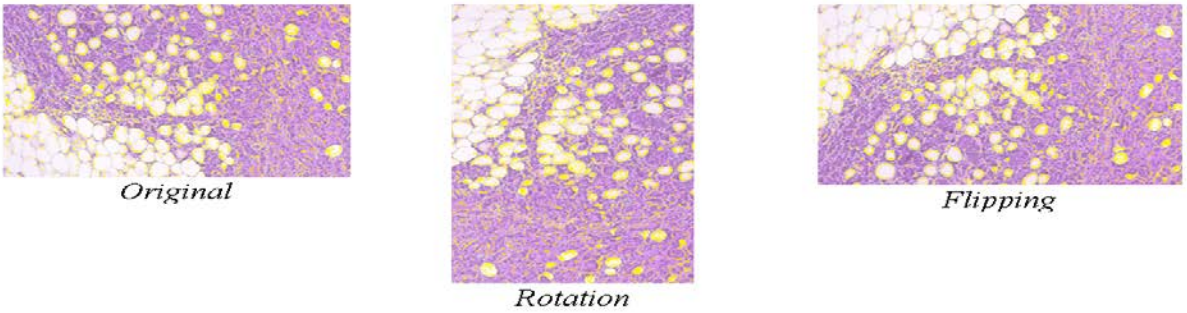

Fig. 5 Data augmentation techniques including rotating, flipping.

3.4 Pre-Processing Steps

Macenko normalization and data augmentation were followed by pre-processing steps as a preliminary recommended phase to prepare data for further feature extraction and analysis. Previous studies ([10],[11], [12], [13], [14]) proposed different pre-processing methods because of the nature of their data. This work proposed a series of calculation, divided into five steps. The first step focused on stain normalization and augmentation, followed by TFRecord [6] format conversion based on Protocol Buffers [7]. In the third step, TFRecords were normalized to [0, 1]. Afterward, whole image bounding box was re-sized to (299 × 299 × 3) or (224 × 224 × 3) according to the recommended model image size for Inception and ResNet architectures. Finally, as Inception and ResNet pre-processing, input training images were randomly flipped left to right horizontally and then cropped to create image summaries to display the different transformations on images. In order to improve the power of learning and to make the network invariant to aspects of the image that do not affect the label.

3.5 Transfer Learning

Transfer learning is defined as exporting knowledge from previously learned source to a target task ([24], [3]). Learning from clinical images from scratch is often not the most practical strategy due to its computational cost, convergence problem [31], and an insufficient number of high quality labeled samples. A growing body of experiments has investigated pre-trained models in the presence of limited learning samples [36].Pre-trained ConvNets alongside fine-tuning and transfer learning lead to faster convergence and outperform training from scratch [31]. Our target data-set consists of total 7909 breast cancer sub-types histopathological images) is obviously smaller than the used reference data-set (ImageNet; training data with 1.2M [32]. Therefore, we initialized weight of different layers of our proposed network by using ImageNet Inception and ResNetpre-trained models. Then, we employed last layer fine-tuning on cancer images data-set. Therefore, the ImageNet pre-trained weights were preserved while the last fully connected layer was updated continuously. Since the cancer data-sets analyzed here are large and very different from ImageNet, the full layer fine-tuning was applied to compare accurately classification of cancers with the last layer fine tuning[32].

3.6 Inception and ResNet Architectures

Among various deep learning methods, we considered Inceptions and ResNet architectures. It is known that Inception models have migrated from fully-to-sparsely-connected architectures. In order to add more non-linearity capability, Inception module technically included 1 × 1 factorized convolutional neural networks followed by the rectified linear unit (ReLU). Also, a 3×3 convolutional layer was employed. Auxiliary logits with a combination of average pool, convolutional 1 × 1, fully connected, and softmax activation was applied to preserve the low-level detail features and tackle vanishing gradient problem in last layers. ResNet permanently utilized shortcut connections between shallow and deep networks to control and adjust training error rate [30]. This study examined different frameworks of Inception (V1 and V2) and ResNet V1 50 ([30], [24]) on cancer digital images. Furthermore, RMSProp adaptive learning rate was applied with start- (0.001), decay- (0.9), and end-points (0.0001) settings. Because of an insufficient number of available histopathological cancer images (section III-A) compared to numerous model parameters (up to 5 million in Inception and 10 million in ResNet), dropout regularization and batch normalization [25]were applied with batch sizes of 32 in training with 4,000, 6,000 epochs and 100 in evaluation steps.

3.7 Computerized System Configuration

Deep learning training with an extreme number of network parameters, computational tasks, and large data-sets was significantly accelerated by a single computing platform with following specifications: Asus Intel i7 core, 10 GB RAM, 1 TB HDD, Nvidia GeForce GTX 950m 4GB with ubuntu 16.0.4 64-bit operating system and python 3.5. In addition, Tensorflow-GPU platform requires CUDA 8.0 toolkit and cuDNN 5.1. All GPU necessary settings and details were obtained from TensorFlow and TFslim documentations and NVIDIA GPUs support ([6]).

4. Experiment Results

4.1 Transfer learning the last

The results were classified into eight sub-types of BreakHis cancer. Several standard performance terms such as true positive (TP), false positive (FP), true negative (TN), false negative (FN), were isolated from the confusion matrix. Using an 80% training set and a 20% test set, an 8 × 8 confusion matrix was used to represent prediction results of eight cancer histological types. Statistical performance measurements were summarized in Tables 3 and 4. The result shows that ResNet V1 50 Inception (V1 and V2) with fine-tuning all layers gave higher accuracy (95 %) than Inception V1 and V2 Meanwhile, the each of the following Tables 2 and 3 is derived from data which were presented in Figure 6 and 7 successively.

Table 2: The Last Layers Fine-Tuning of Breast Cancer sub-classes

Model	***Epoch***	***TP***	***TN***	***FP***	***FN***	***Accuracy***	***Error rate***	selectivity
Inception V1	4000	2287	65	147	29	34%	0.069	0.9509
Inception V1	6000	2289	81	131	27	39%	0.062	0.9529
Inception V2	4000	2296	51	161	20	27%	0.071	0.9422
Inception V2	6000	2253	100	112	63	51%	0.069	0.9564
ResNet-50 V1	4000	2285	166	46	31	82%	0.030	0.9897

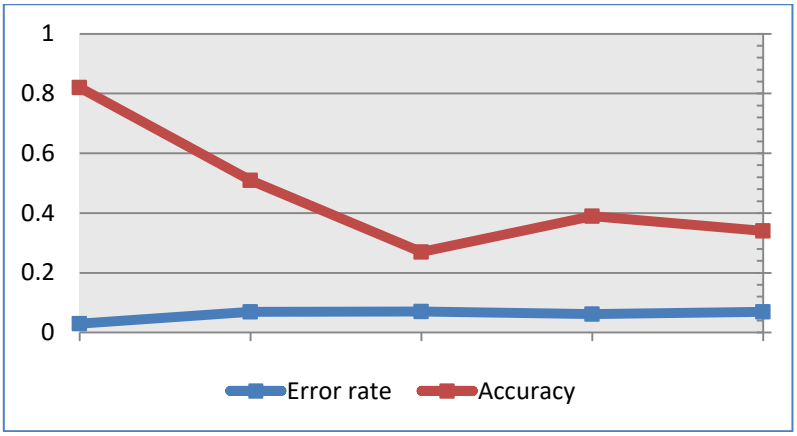

Fig. 6 The Last Layers Fine-Tuning of Breast Cancer sub-classes

Table 3: The All Layers Fine-Tuning of Breast Cancer sub-classes

Model	***Epoch***	***TP***	***TN***	***FP***	***FN***	***Accuracy***	***Error rate***	sensitivity
Inception V1	4000	2314	74	138	2	46%	0.055	0.9647
Inception V1	6000	2311	116	96	5	62%	0.039	0.9869
Inception V2	4000	2316	19	193	0	41%	0.076	0.9339
Inception V2	6000	2313	109	103	3	58%	0.041	0.9889
ResNet-50 V1	4000	2297	201	11	19	95%	0.011	0.9992

Fig. 7 The all Layers Fine-Tuning of Breast Cancer sub-classes

5. Discussion

This work examined data from BreakHis datasets to classify Carcinomas breast cancer sub-types. Previous studies ([26],[28],[29]) focused on binary benign-malignant classification and did not perform the further quantitative assessment. In this work, we introduced automated breast cancer multi-classification methods. We suggested a generic CAD framework based on deep networks for learning histopathology images to avoid the error committed by hand-crafted pathological features. In this study, we compared the performance of Inception and ResNet deep learning models using transfer learning strategy on several large image datasets. We found that deep ResNet models were more sensitive. In recent comparative studies, ([29], [26], [4]), conventional machine learning methods (SVM, KNN, QDA, ASSVM, SSVM- SCAD, etc) hand-crafted feature extraction were used. In these approaches, the results were evaluated at various magnifications (i.e. 40X, 100X, 200X and 400X). These methods could achieve 90 to 93% accuracy in the benign and malignant classification. For example, AlexNet deep learning approach which used many learning parameters to classify benign and malignant Carcinomas showed an accuracy rate of 90% ([28]). Besides, Han and his colleagues ([1]) reported about a deep learning-based system for multi-classification of breast cancers that showed an average accuracy rate of the 93.2%. Table 4 presents a comparison of the rate of accuracy between our proposed approach and some other architecture.

Table 4: study comparison

Approach	Accuracy
([4],[26],[29])	90% to 93%
[28]	90%
Proposed Approach	95%

6. Conclusion

This study suggests a modified approach to the existing ResNet approach for an effective and reliable strategy for the detection of Carcinomas subtypes. It suggests steps develop deep learning based convolutional neural network system which would reduce human errors in analyzing and classifying subtypes of cancer tumors in a manual study of histopathological images. In a comparative study between Inception V1, V2 and a modified deep learning based ResNet -50 V, it was established that the latter is highly efficient with 95% accuracy and low error rate 0.11 in detecting eight subtypes of cancer carcinomas. This also showed that the modified ResNet -50 Coffers trivial false positive average values (0.3 out of 900 for four cancer types, 6.3 out of 809 for all breast cancer, 5 out of 800 for benign and 0.3 out of 1000 for malignant). Thus this kind of efficient multi-classification system relieves the pathologists and medical experts workloads regarding analyze and interpretation of the Histopathological slides for assisting the doctors to choose more efficient therapeutic approaches.

References

[1] Global Burden of Disease Cancer Collaboration. Fitzmaurice C, Allen C, Barber RM, Barregard L, Bhutta ZA, Brenner H, Dicker DJ, Chimed-Orchir O, Dandona R, et al. Global, regional, and national cancer incidence, mortality, years of life lost, years lived with disability, and disability-adjusted life-years for 32 cancer groups, 1990 to 2015: a systematic analysis for the global burden of disease study. JAMA Oncol. 2017;3:524–548.

[2] M. T. Bahadori, Y. Liu, and D. Zhang.A general framework for scalable transductive transfer learning. Knowledge and information systems, 38(1):61–83, 2014.

[3] L. Carey, E. Winer, G. Viale, D. Cameron, and L. Gianni. Triple-negative breast cancer: disease entity or title of convenience? Nature reviews Clinical oncology, 7(12):683–692, 2010.

[4] A. Chan and J. A. Tuszynski.Automatic prediction of tumour malignancy in breast cancer with fractal dimension. Open Science, 3(12):160558, 2016.

[5] M. Colleoni, N. Rotmensz, P. Maisonneuve, M. Mastropasqua, A. Luini, P. Veronesi, M. Intra, E. Montagna, G. Cancello, A. Cardillo, et al. Outcome of special types of luminal breast cancer. Annals of oncology, 23(6):1428–1436, 2011.

[6] T. Google-Developers. Installing tensorflow on ubuntu. https://www. tensorflow.org/install/install linux, 2017.

[7] T. Google-Developers. Protocol buffers. https://developers.google.com/ protocol-buffers/?hl=en, 2017

[8] A. Esteva, B. Kuprel, R. A. Novoa, J. Ko, S. M. Swetter, H. M. Blau, and S. Thrun. Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639):115–118, 2017.

[9] L. Gui, R. Xu, O. Lu, J. Du, and Y. Zhou. Negative transfer detection in transductive transfer learning.International Journal of Machine Learning and Cybernetics, pages 1–13, 2017.

[10] Veta M, Pluim JPW, Van Diest PJ, Viergever MA. Breast cancer histopathology image analysis: A review. IEEE Transactions on Biomedical Engineering. 2014;61(5):1400–1411.

[11] Kowal M, Filipczuk P, Obuchowicz A, Korbicz J, Monczak R. Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. Computers in Biology and Medicine. 2013;43(10):1563–1572.

[12] Filipczuk P, Fevens T, Krzyzak A, Monczak R. Computer-aided breast cancer diagnosis based on the analysis of cytological images of fine needle biopsies. IEEE Transactions on Medical Imaging. 2013;32(12):2169–2178.

[13] George YM, Zayed HH, Roushdy MI, Elbagoury BM. Remote computer-aided breast cancer detection and diagnosis system based on cytological images. IEEE Systems Journal. 2014;8(3):949–964.

[14] Belsare AD, Mushrif MM, Pangarkar MA, Meshram N. Classification of breast cancer histopathology images using texture feature analysis. In: TENCON 2015—2015 IEEE Region 10 Conference. Macau: IEEE; 2015. p. 1–5.
[15] Brook A, El-Yaniv R, Issler E, Kimmel R, Meir R, Peleg D. Breast Cancer Diagnosis From Biopsy Images Using Generic Features and SVMs. 2007; p. 1–16.
[16] Zhang B. Breast cancer diagnosis from biopsy images by serial fusion of Random Subspace ensembles. In: 2011 4th International Conference on Biomedical Engineering and Informatics (BMEI). vol. 1. Shanghai: IEEE; 2011. p. 180–186.
[17] Israel Institute of Technology dataset;. Availablefrom: ftp.cs.technion.ac.il/pub/projects/medic-image.
[18] Krizhevsky A, Sutskever I, Hinton GE. ImageNet Classification with Deep Convolutional Neural Networks. In: Advances in Neural Information Processing Systems 25; 2012. p. 1106–1114.
[19] Ciresan DC, Giusti A, Gambardella LM, Schmidhuber J. Mitosis detection in breast cancer histology images with deep neural networks. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). 2013;8150 LNCS(PART 2):411–418.
[20] Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, et al. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Scientific Reports. 2016;6(January):26286
[21] Sirinukunwattana K, Raza SEA, Tsang YW, Snead DRJ, Cree IA, Rajpoot NM. Locality Sensitive Deep Learning for Detection and Classification of Nuclei in Routine Colon Cancer Histology Images. IEEE Transactions on Medical Imaging. 2016;35(5):1196–1206.
[22] Spanhol FA, Oliveira LS, Petitjean C, Heutte L. Breast Cancer Histopathological Image Classification using Convolutional Neural Networks. In: International Joint Conference on Neural Networks (IJCNN 2016). Vancouver; 2016.
[23] Cruz-Roa A, Basavanhally A, González F, Gilmore H, Feldman M, Ganesan S, et al. Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks. In: Proc. SPIE. vol. 9041.San Diego, California; 2014. p. 904103–904115.
[24] K. He, X. Zhang, S. Ren, and J. Sun. Deep residual learning for image recognition. In Proceedings of the IEEE conference on computer vision and pattern recognition, pages 770–778, 2016.
[25] S. Ioffe and C. Szegedy. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In International Conference on Machine Learning, pages 448–456, 2015.
[26] M. A. Kahya, W. Al-Hayani, and Z. Y. Algamal.Classification of breast cancer histopathology images based on adaptive sparse support vector machine. Journal of Applied Mathematics and Bioinformatics, 7(1):49, 2017.
[27] W. H. Organization et al. Tumours of the breast and female genital organs. World Health Organization Classification of Tumours: Pathology and Genetics of Tumours of the Breast and Female Genital Organs. Lyon: International Agency for Research in Cancer, 2003.
[28] A. A. A. Setio, F. Ciompi, G. Litjens, P. Gerke, C. Jacobs, S. J. van Riel, M. M. W. Wille, M. Naqibullah, C. I. Sanchez, and B. van Ginneken. ´ Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE transactions on medical imaging, 35(5):1160–1169, 2016.
[29] F. A. Spanhol, L. S. Oliveira, C. Petitjean, and L. Heutte.A dataset for breast cancer histopathological image classification. IEEE Transactions on Biomedical Engineering, 63(7):1455–1462, 2016.
[30] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens, and Z. Wojna. Rethinking the inception architecture for computer vision.In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 2818–2826, 2016.
[31] N. Tajbakhsh, J. Y. Shin, S. R. Gurudu, R. T. Hurst, C. B. Kendall, M. B. Gotway, and J. Liang. Convolutional neural networks for medical image analysis: Full training or fine tuning? IEEE transactions on medical imaging, 35(5):1299–1312, 2016.
[32] T. TensorFlow-Group. Tensorflow-slim image classification library. https://github.com/tensorflow/models/tree/master/slim, 2016. [Apache License, Version 2.0, Online; accessed 17-July-2017].
[33] L. D. True. Morphometric applications in anatomic pathology. Human Pathology, 27(5):450 – 467, 1996.
[34] G. Viale, N. Rotmensz, P. Maisonneuve, E. Orvieto, E. Maiorano, V. Galimberti, A. Luini, M. Colleoni, A. Goldhirsch, and A. S. Coates. Lack of prognostic significance of classic lobular breast carcinoma: a matched, single institution series. Breast cancer research and treatment, 117(1):211, 2009.
[35] D. Wang, A. Khosla, R. Gargeya, H. Irshad, and A. H. Beck. Deep learning for identifying metastatic breast cancer.arXiv preprint arXiv:1606.05718, 2016.
[36] P. Sermanet, D. Eigen, X. Zhang, M. Mathieu, R. Fergus, and Y. LeCun.Overfeat: Integrated recognition, localization and detection using convolutional networks. arXiv preprint arXiv:1312.6229, 2013.
[37] *M. Macenko et al., 'A method for normalizing histology slides for quantitative analysis', in 2009 IEEE International Symposium on Biomedical Imaging: From Nano to Macro, 2009, pp. 1107–1110.*