KICS The Korean Institute of Communications and Information Sciences

Available online at www.sciencedirect.com
ScienceDirect

ICT Express 4 (2018) 247–254

www.elsevier.com/locate/icte

Breast cancer histology images classification: Training from scratch or transfer learning?

Shallu*, Rajesh Mehra

Department of Electronics & Communication, National Institute of Technical Teachers' Training & Research, Chandigarh, 160019, India

Received 1 October 2018; accepted 15 October 2018
Available online 3 November 2018

Abstract

We demonstrated the ability of transfer learning in comparison with the fully-trained network on the histopathological imaging modality by considering three pre-trained networks: VGG16, VGG19, and ResNet50 and analyzed their behavior for magnification independent breast cancer classification. Concurrently, we examined the effect of training–testing data size on the performance of considered networks. A fine-tuned pre-trained VGG16 with logistic regression classifier yielded the best performance with 92.60% accuracy, 95.65% area under ROC curve (AUC), and 95.95% accuracy precision score (APS) for 90%–10% training–testing data splitting. Layer-wise fine-tuning and different weight initialization schemes can be a future aspect of this study.

Keywords: Breast cancer; Histopathological images; Convolutional neural network; Full training; Transfer learning

1. Introduction

The pink ribbon in the world of medicine is a symbol to bring awareness among people regarding breast cancer. Breast cancer is one of the leading causes of the high mortality rate in women [1]. BRCA1, BRCA2 (name of two genes), obesity, inhaling of birth control pills, irregular menstrual cycles, greater exposure to radiation therapy and estrogen hormone are some high-risk factors to cause breast cancer [2–4]. These factors are responsible for inducing mutations in cells and resulted in the uncontrollable growth of the cells. Soreness in the breast is the primary symptom occurred during breast cancer which becomes life-threatening if not diagnosed in the early stage [https://www.nationalbreastcancer.org/breast-pain]. Skin irritation, redness, pain, and swelling are some other symptoms occur at the initial stage which becomes ominous with the erosion of nipples or sudden watery discharge from the nipples [3,4]. Therefore, early detection is essential to lead an easy treatment and to increase the chances of survival as well [5].

The techniques used in screening and monitoring of the breast cancer, including mammography, magnetic resonance imaging (MRI), ultrasound, positron emission tomography (PET), thermography, and surgical incision [4,6–8]. Interpretation of the data obtained from these techniques is very complicated due to the similar clinical manifestations of different types of cancer, which make further analysis of the data very difficult. Analysis of the data is a time-consuming and labor-intensive process but also a very critical step to provide a differential diagnosis. So, it has become necessary to automate some of the tasks in the diagnostic workflow to reduce the burden on the radiologist and pathologist. In this context, the machine learning techniques emerged as a dominant paradigm which offers reliable solutions and can perform some diagnostic task automatically and intelligently [9].

Among the various machine learning techniques, the deep neural networks are attracting remarkable interest due to their automatic feature extraction and representation learning ability [10–12]. Advances in computational power are another

* Corresponding author.
E-mail address: shallu.ece@nitttrchd.ac.in (Shallu).
Peer review under responsibility of The Korean Institute of Communications and Information Sciences (KICS).

https://doi.org/10.1016/j.icte.2018.10.007

cause for its popularity [13]. A common variant of the deep neural network, i.e., convolutional neural network (CNN) has broad applications in computer vision and known for their weight sharing and local connectivity characteristics [14–17]. These two characteristics permit the CNN to act like local filters and to detect the same pattern at the entire image with less number of trainable parameters. The ability of representation learning makes the CNN to combine low-level features into high-level features, successively, until the obtainment of the class label for the image. However, full-training (training from scratch) of CNN may not be very easy as CNN requires plenty of training data for better performance [18,19]. High-performance graphical processing units (GPUs) also needed in training of CNN for fast processing because the training with such a big collection of data is a time-consuming process [18]. Besides, the process of training in CNN is very complicated and required constant adjustments of parameters to ensure equivalent learning of all the layers due to issues of convergence and overfitting [20].

Transfer learning may be a good substitute for training from scratch. In transfer learning, a network trained on one task is fine-tuned and applied to a different but related task. Transfer learning can be applied in two ways, as a baseline or a feature generator [21–24]. The parameters in a pre-trained network are adjusted according to the required task when used as a baseline [21]. However, the features are extracted from the input and then used to train a new classifier when employed as a feature generator [22]. Herein, the following question is addressed in the context of histopathological images of breast cancer tissues classification: Whether a fully trained or fine-tuned pre-trained network performs well for magnification independent classification of breast cancer histopathology images? In order to answer this vital question, the following experiments were performed: (a) magnification independent classification of histopathological images related to breast cancer obtained from the BreakHis dataset (https://web.inf.ufpr.br/vri/databases/breast-cancer/histopathological-database-breakhis/) and, (b) the study of network performance for three different training–testing data splitting. For every set of experiments, the performance of the pre-trained fine-tuned network is compared with the network that trained from scratch. The motivation behind this work is the research contribution given by Tajbaksh et al. as they were not able to perform a comparative analysis of full-training vs. fine-tuning of pre-trained CNNs over histopathological and magnetic resonance imaging modality [25]. Hence, our motive is to extend their work for histopathological imaging modality through this paper. The aim of this paper is to analyze the ability of transfer learning as compared to the fully-trained network for the application of breast cancer classification using histopathological imaging and MR modality and to determine which pre-trained network performed better for this application. As best of authors' knowledge concern, the study is performed very first time in the histopathological imaging modality.

2. Material and method used

In this study, the three well recognized pre-trained deep CNN models: VGG16, VGG19, and ResNet50 are used, for both full training and transfer learning. Since the task to classify breast cancer histology images is very complex so very deep architecture required in solving this problem. The VGG16, VGG19, and ResNet50 are very popular pre-trained CNN models due to their more in-depth architecture. Additionally, these models have shown relatively high performance for challenging computer vision problem multifariously. For example, VGG16 and VGG19 hold the first position in localization and the second position in classification for the ImageNet Large Scale Visual Recognition Challenge (ILSVRC-2014). Moreover, VGG16 and VGG19 have also shown a good generalization on the other datasets like Caltech-101, Caltech-256, PASCAL VOC 2007, and 2012 [26,27]. Similarly, ResNet50 (Residual Network) was also a winner in ILSVRC-2015 and surpassed even the human performance in the task of classification. Easy training is the significant advantage to use the ResNet50 model as it learns residuals from the image instead of some features [28]. In the current study, the entire above listed pre-trained network employed as a feature generator by considering all the activations earlier to the last fully-connected layer in the network. Further, these activations are used as features vector to learn a new classifier for the classification purpose. In this regard, logistic regression (LR) is utilized as a new classifier to make the final decision. Therefore, the current approach is different from that used by [25] in terms of fine-tuning wherein fine-tuning of pre-trained was performed in a layer-wise manner using only one pre-trained CNN model, the AlexNet. However, the current approach also considered the influence of training–testing data size on the performance of fine-tuned and fully-trained network as performed in [25].

2.1. Dataset

Generally, in the medical domain, the dataset consists of limited samples due to a complicated and expensive procedure of data collection. So it is very challenging to access a relevant and large-scale well-annotated dataset. A well-annotated dataset provides a common platform for the comparison and validation of purposed models. Thus, to develop a robust model, availability of large-scale well-annotated dataset is the foremost requirement. In this work, a publicly available *BreakHis* dataset is used built in collaboration with the Prognostics and Diagnostics (P&D) Laboratory, Parana, Brazil [14,35]. The dataset contains a total of 7909 breast cancer histopathology image samples collected from 82 patients under four different magnification levels. The samples in the dataset are divided into two major groups: benign and malignant wherein benign and malignant group consists of 2480 and 5429 samples, respectively. A glimpse of the dataset is provided in Fig. 1.

2.2. Data augmentation

Unbalanced and limited data size is the major challenge in the development of robust computer-aided diagnosis (CAD) system. Data augmentation is an approach used in deep models to enlarge the dataset in order to alleviate the problem of limited data size. Some popular data augmentation techniques like

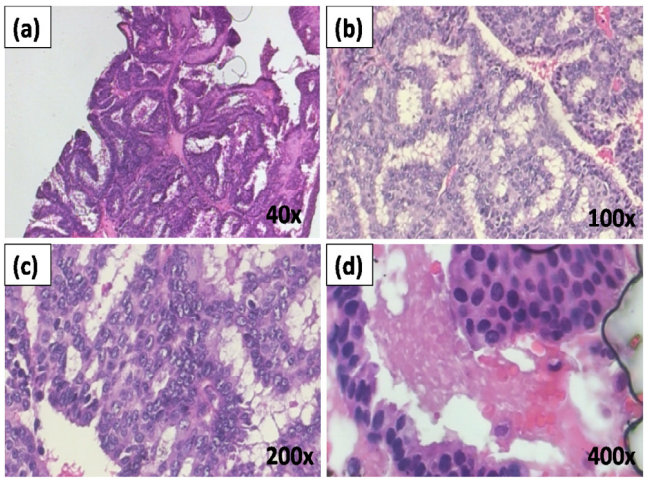

Fig. 1. Breast Cancer Histopathological Images from BreakHis Dataset of a Patient Suffered from Papillary carcinoma (Malignant) with four magnification levels (a) 40x, (b) 100x (c) 200x and (d) 400x. [10].

flipping, cropping, scaling, rotation, interpolation, translation, and noise insertion have already been applied in many previous studies [15,29,30]. But all augmentation approaches used for natural images would not work over medical images, as a lot of medical images are a top-down solved problem, not the bottom-up like natural images. Further, intensities play a pivotal role in many medical imaging modalities as compared to natural images. Therefore, the selection of a data augmentation approach should be performed wisely, based on the dataset. Since the histopathological images possess rotation and reflection symmetry [31], so there is always a possibility to eliminate some inherent properties from the image if other augmentation techniques are used which may cause loss of discriminating features or information. Consequently, for both the full training and transfer learning, only rotation is employed as data augmentation technique. Rotations of the images are carried out about their center with the three angles: 90°, 180°, and 270°. Another advantage in using the augmentation process lies in the minimization of overfitting [32], which is a major issue in the successful implementation of a machine learning model.

2.3. Magnification independent classification

Magnification factor plays a crucial role in the analysis of histology images. In magnification, the size of an object is altered so that the object can be seen with comfort [33]. The histology images consist of the variety of tissues, but the analysis of these tissues becomes complex at low magnification. Since the acquiring of an image at different magnification introduces the diversity in the background. So, it becomes challenging for an automatic CAD system to learn or extract distinct features from images with different magnification to make a differential diagnosis. Utilization of images with the same magnification can be a way to reduce the variation in the background, as preferred in most of the previous studies. They performed classification for a particular magnification level [34,35]. In some previous studies, images with multiple magnification factors have also been considered and employed different classifiers for the specific magnification factor [36,37]. In the implementation of such approaches, multiple stages of training are required in conjunction with the prior knowledge of the magnification factor. Also, an image with a new magnification factor also put a significant limit on the performance of the network which is a major weakness in the execution of magnification dependent approaches. Thus, there is a need to develop a magnification independent automatic diagnosis system that has the adaptability to the new magnification factor as well.

3. Results & Discussion

In this study, we have demonstrated coherent results for breast cancer classification application in the histopathological imaging modality. A balanced dataset is used for both full training and fine-tuning CNNs. To get a balanced dataset, the class (malignant) with a large number of images is down-sampled equal to the image samples present in another class (benign). All the experiments are performed on the same machine with configuration: Intel(R) Core(TM) i7-7500U @ 2.90 GHz, NVIDIA GeForce 940MX, Window 10, 8 GB memory, using Tensorflow and Keras library. To evaluate the classification performance of the fine-tuned pre-trained network and full trained network, the whole dataset is divided into training and testing data. Splitting of the data into the training–testing data is a common practice in neural networks that are used for performance analysis. To determine the influence of training–testing data size on the network's performance, three splitting fashions for training–testing data (90%–10%, 80%–20%, and 70%–30%) are utilized.

The set of experiments are performed for both fine-tuned and fully-trained network using these three splitting fashions. The elapsed time for each experiment was around 1–2 hours. The results for these experiments are computed in terms of precision, recall and f1 score for each class separately. Then, for each experiment, an average of results obtained for both the classes is computed in order to perform an easy comparison. In addition, classification performance is studied using a receiver operating characteristics (ROC) analysis, and area under the curve (AUC) to validate the results. In conjunction with this, average precision score (APS) is also computed for performance evaluation.

In this study, for full-training of the network weights are initialized randomly to train the model on the BreakHis dataset from scratch. The random initialization of weight performed to utilize only the architecture of the pre-trained model, not the weights trained on previous data. However, for transfer learning the weights of the pre-trained network are kept as it is, based on the assumption that the pre-trained networks have already been trained pretty well. Tables 1 and 2 are summarizing the results obtained from the transfer learning and full-training of VGG16, VGG19, and ResNet50 on BreakHis dataset. The performance of all the models studied in the context of classification of breast histopathological images into two classes: benign (B) and malignant (M). From Table 1,

Table 1
Performance Analysis for Histopathological Image Classification using Fine-tuned Pre-trained Network (VGG16, VGG19, and ResNet50).

Classifier	Training–Testing Data Splitting	Class Type	Precision	Recall	F_1 Score	Accuracy	AUC	APS
VGG 16 + LR	90%–10%	B	0.93	0.92	0.93	**92.60%**	**95.65%**	**95.95%**
		M	0.93	0.93	0.93			
		Avg/Total	**0.93**	**0.93**	**0.93**			
	80%–20%	B	0.92	0.93	0.92	92.20%	93.95%	93.89%
		M	0.93	0.92	0.92			
		Avg/Total	0.92	0.92	0.92			
	70%–30%	B	0.92	0.92	0.92	91.73%	93.49%	93.29%
		M	0.92	0.91	0.92			
		Avg/Total	0.92	0.92	0.92			
VGG 19 + LR	90%–10%	B	0.88	0.93	0.90	90.00%	91.85%	91.27%
		M	0.93	0.88	0.90			
		Avg/Total	0.90	0.90	0.90			
	80%–20%	B	0.89	0.90	0.89	89.50%	91.76%	91.13%
		M	0.90	0.89	0.90			
		Avg/Total	0.90	0.90	0.90			
	70%–30%	B	0.90	0.91	0.91	90.40%	91.14%	90.38%
		M	0.90	0.90	0.90			
		Avg/Total	0.90	0.90	0.90			
ResNet50 + LR	90%–10%	B	0.77	0.81	0.79	79.40%	79.39%	82.03%
		M	0.81	0.78	0.80			
		Avg/Total	0.79	0.79	0.79			
	80%–20%	B	0.80	0.78	0.79	78.90%	79.23%	80.56%
		M	0.81	0.78	0.80			
		Avg/Total	0.79	0.79	0.79			
	70%–30%	B	0.78	0.81	0.79	78.73%	79.12%	79.09%
		M	0.80	0.77	0.78			
		Avg/Total	0.79	0.79	0.79			

it can be observed that the fine-tuned pre-trained VGG16 network significantly outperformed the ResNet50 whereas the performance of VGG16 and VGG19 is comparable. In transfer learning, ResNet50 provided unsatisfactory results over the BreakHis dataset and indicated their inability to generalize over the new problem as shown in Table 1. Overfitting was the only reason for its inadequate performance which arose due to the excessively large capacity of the network. Freezing of more layers in the network can be a solution to reduce the effective capacity of the network and overfitting as well [25]. This solution is not taken into account in this work due to space constraints, but we assured that the solution will be considered in the extended version of this paper. Further, in performance evaluation for the fully trained network, ResNet50 showed a great performance over VGG16 and VGG19 network. From Table 2, it has been found that VGG16 and VGG 19, both the networks are biased to one particular class. The value of recall (also termed as sensitivity) can be considered as strong evidence for this statement. The value of recall is very high for one particular class and simultaneously very low for other class in both fully trained VGG16 and VGG19, but the ResNet50 is equally sensitive to both the classes which are going to support its better working in comparison of VGG16 and VGG19.

Further, to study the performance of the model with respect to the size of the training–testing data, three different splitting of training–testing data (90%–10%, 80%–20%, and 70%–30%) are utilized because the size of the dataset has a significant impact on the performance of CNNs. In this context, ROC analysis and AUC is used to compare the performance of all the networks, as illustrated in Figs. 2–4. In Fig. 2, ROC curve and AUC of the pre-trained and fully-trained network are compared for 90%–10% training–testing data splitting. It is found that, pre-trained VGG 16 (AUC-95.65%) and VGG19 (AUC-91.85%) outperformed the fully-trained VGG16 (AUC-75.00%) and VGG19 (AUC-65.40%) by a significant amount while the pre-trained ResNet50 (AUC-79.23%) is outperformed by fully-trained ResNet50 (AUC-82.88%) but with a small value. The same trend is observed for the remaining splitting fashion of training–testing data which are shown in Figs. 3 and 4. From Tables 1 and 2, it has been observed that the degradation in performance of fine-tuned VGG16, VGG19, and the ResNet50 network is insignificant with respect to reduction in the size of training data. All the three networks performed well when the training data is 90% of the total dataset, but the situation is somewhat different for full training. VGG19 performed well for the 80%–20% training–testing data splitting. However, the performance of ResNet50 and VGG16 network for the entire splitting fashion was almost similar. The deviation of VGG19 from the usual trend is due to its more sensitivity towards benign and malignant during 90%–10% and 70%–30% splitting, respectively. In summary, it has been demonstrated that the employment of the transfer learning approach results in a noteworthy performance with respect to CNN with full training over the histopathological imaging modality, even when training dataset is limited in size.

Table 2
Performance Analysis for Histopathological Image Classification using Full-trained Network (VGG16, VGG19, and ResNet50).

Classifier	Training–Testing Data Splitting	Class Type	Precision	Recall	F_1 Score	Accuracy	AUC	APS
VGG 16	90%–10%	B	0.63	0.64	0.64	64.40%	**75.00%**	**76.79%**
		M	0.66	0.64	0.65			
		Avg/Total	0.64	0.64	0.64			
	80%–20%	B	0.63	0.63	0.63	63.20%	66.77%	65.23%
		M	0.63	0.64	0.63			
		Avg/Total	0.63	0.63	0.63			
	70%–30%	B	0.66	0.60	0.63	63.93%	75.18%	73.23%
		M	0.62	0.68	0.65			
		Avg/Total	0.64	0.64	0.64			
VGG 19	90%–10%	B	0.50	0.87	0.64	51.80%	60.40%	59.88%
		M	0.60	0.19	0.29			
		Avg/Total	0.55	0.52	0.46			
	80%–20%	B	0.57	0.76	0.65	59.70%	65.61%	62.00%
		M	0.65	0.43	0.52			
		Avg/Total	0.61	0.60	0.59			
	70%–30%	B	0.81	0.02	0.04	50.27%	61.34%	55.44%
		M	0.50	0.99	0.66			
		Avg/Total	0.66	0.50	0.35			
ResNet50	90%–10%	B	0.77	0.78	0.79	78.80%	82.71%	79.34%
		M	0.80	0.78	0.79			
		Avg/Total	0.79	0.79	0.79			
	80%–20%	B	0.79	0.80	0.79	79.30%	82.88%	79.67%
		M	0.80	0.78	0.79			
		Avg/Total	0.79	0.79	0.79			
	70%–30%	B	0.70	0.82	0.80	**79.93%**	84.35%	80.76%
		M	0.81	0.78	0.79			
		Avg/Total	**0.80**	**0.80**	**0.80**			

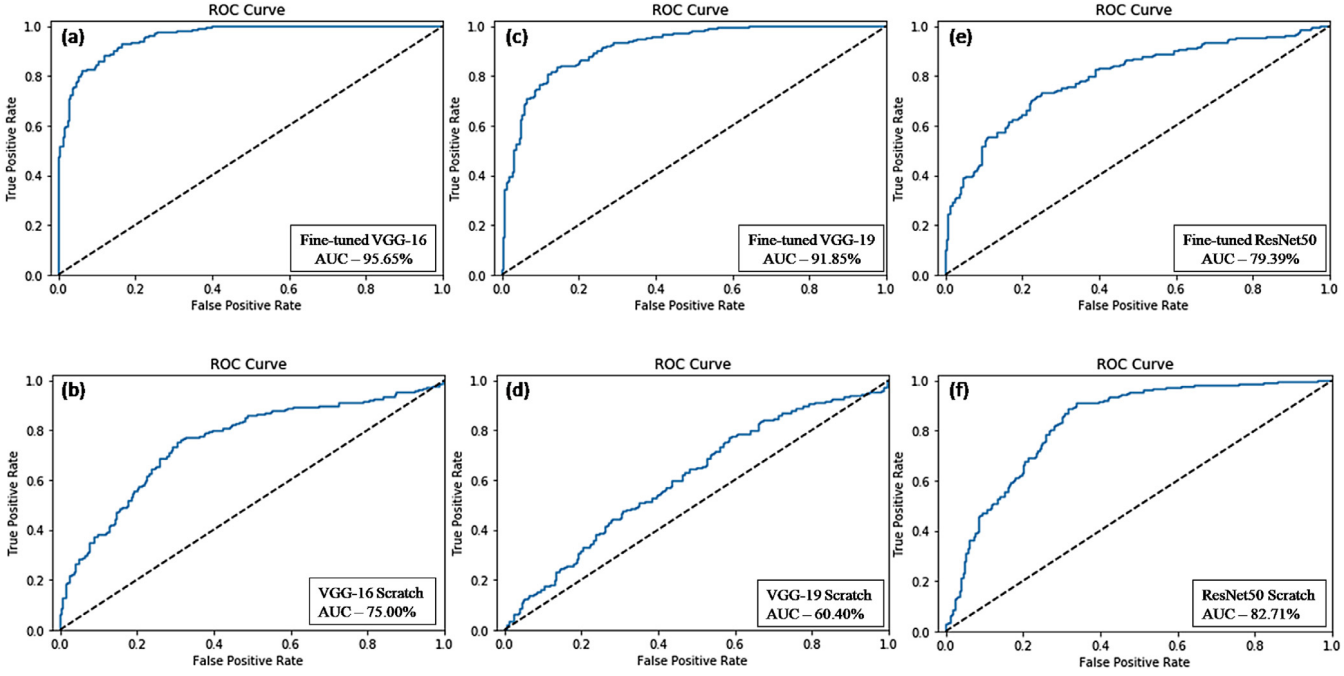

Fig. 2. ROC analysis for breast cancer classification with 90%–10% training and testing set splitting in (a) Fine-tuned pre-trained VGG16, (b), Fine-tuned pre-trained VGG19, (c) Fine-tuned pre-trained ResNet50, (d) Fully-trained VGG16, (e) Fully-trained VGG19, and (f) Fully-trained ResNet50.

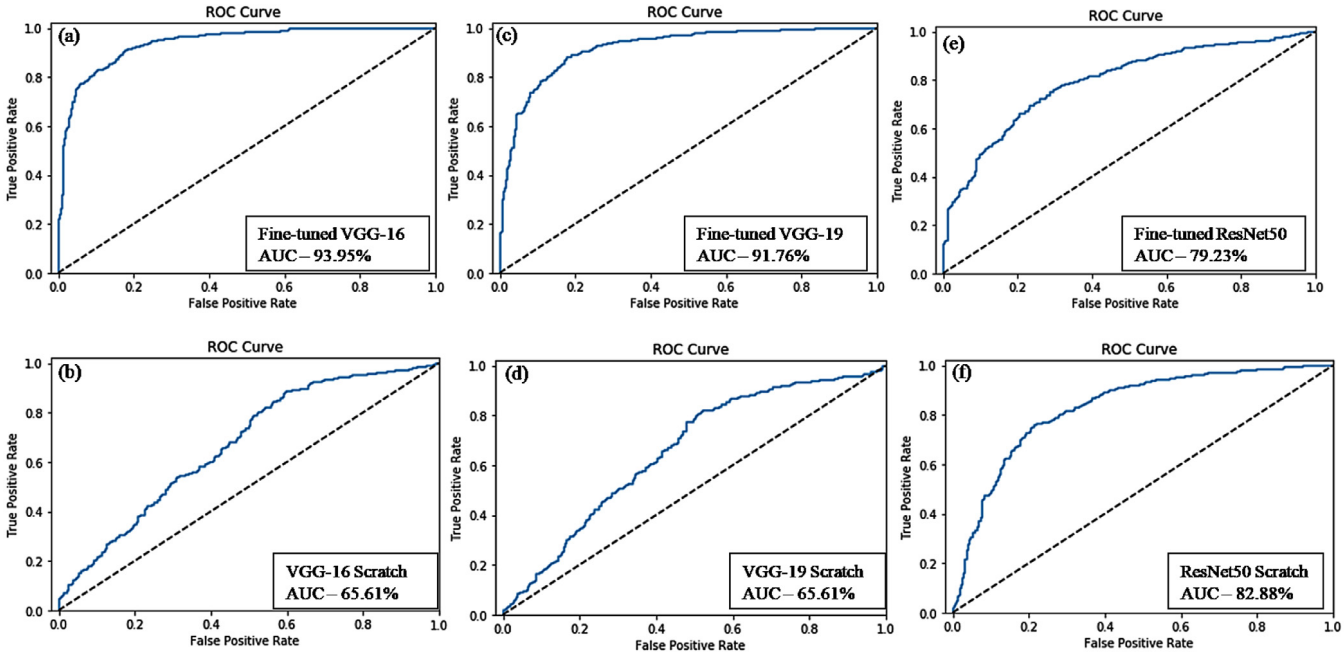

Fig. 3. ROC analysis for breast cancer classification with 80%–20% training and testing set splitting in (a) Fine-tuned pre-trained VGG16, (b), Fine-tuned pre-trained VGG19, (c) Fine-tuned pre-trained ResNet50, (d) Fully-trained VGG16, (e) Fully-trained VGG19, and (f) Fully-trained ResNet50.

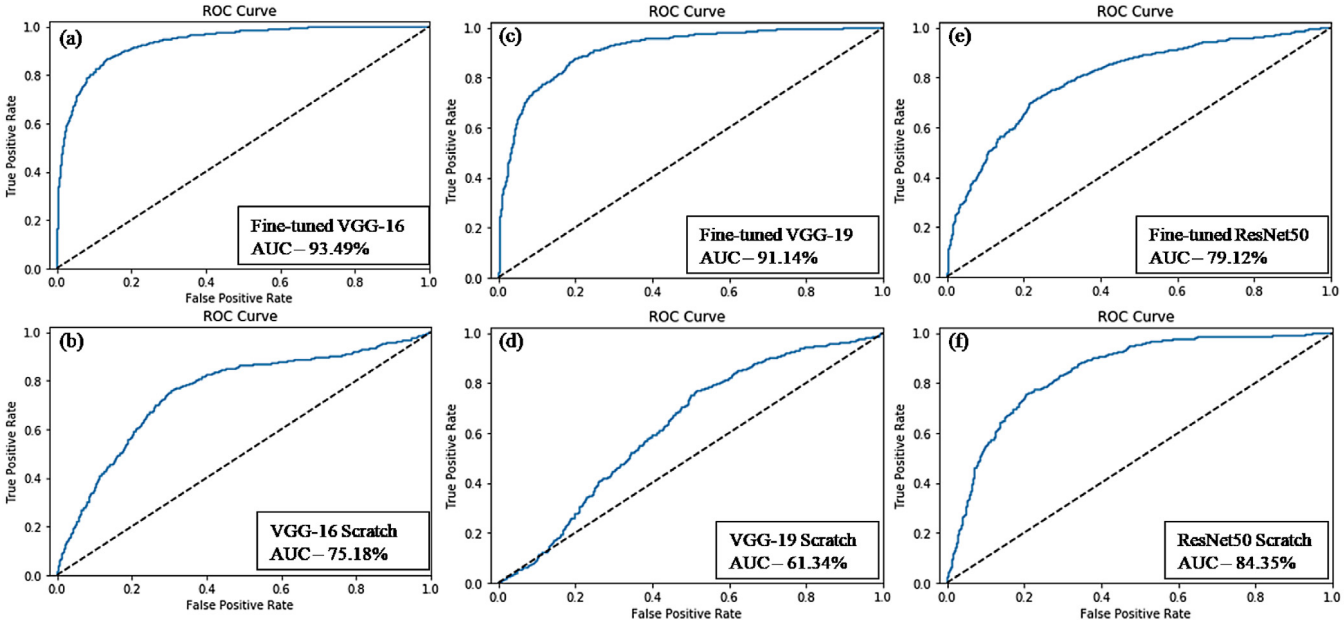

Fig. 4. ROC analysis for breast cancer classification with 70%–30% training and testing set splitting in (a) Fine-tuned pre-trained VGG16, (b), Fine-tuned pre-trained VGG19, (c) Fine-tuned pre-trained ResNet50, (d) Fully-trained VGG16, (e) Fully-trained VGG19, and (f) Fully-trained ResNet50.

4. Conclusions and future directions

This work determines the possibility of knowledge transfer from natural to histopathological images, by employing three pre-trained networks (VGG16, VGG19, and ResNet50) for fine-tuning and full-training. For transfer learning, these pre-trained networks have been utilized as a feature generator and extracted features used to train logistic regression classifier. Some significant findings of this work are:

- In comparison with the full-trained network, the fine-tuned pre-trained VGG16 with logistic regression classifier yielded the best performance with 92.60% accuracy, 95.65% area under ROC curve (AUC), and 95.95% accuracy precision score (APS) for 90–10 training–testing data splitting.
- Fine-tuned pre-trained network is more robust to the size of training data over a full-trained network as the

performance of fine-tuned network not degraded much even when the size of training data is reduced.
- Biasing towards one particular class in the task of classification degrades the network performance considerably. Therefore, a network should be unbiased or equally sensitive to all the classes for better performance.
- The capacity of the network also has a significant influence on the network performance as it gives rise to the problem of over-fitting and under-fitting when too large or too small; respectively. Therefore, the optimal capacity of a network largely depends on the application to handle.

Layer-wise fine-tuning, dataset with increased size, advanced data augmentation technique (like conditional generative adversarial networks (GANs) and deep photo style transfer) and the different weight initialization techniques (such as Xavier, He, MSRA, Gaussian distribution) in full training of network, are some methods to pursue as future aspects of this study.

Conflict of interest

The authors declare that they have no conflict of interests in this paper.

References

[1] R.L. Siegel, K.D. Miller, S.A. Fedewa, D.J. Ahnen, R.G. Meester, A. Barzi, A. Jemal, Colorectal cancer statistics, 2017, CA Cancer J. Clin. 67 (3) (2017) 177–193.

[2] D. Ford, D.F. Easton, M. Stratton, S. Narod, D. Goldgar, P. Devilee, D.T. Bishop, B. Weber, G. Lenoir, J. Chang-Claude, H. Sobol, M.D. Teare, J. Struewing, A. Arason, S. Scherneck, J. Peto, T.R. Rebbeck, P. Tonin, S. Neuhausen, R. Barkardottir, J. Eyfjord, H. Lynch, B.A.J. Ponder, S.A. Gayther, J.M. Birch, A. Lindblom, D. Stoppa-Lyonnet, Y. Bignon, A. Borg, U. Hamann, N. Haites, R.J. Scott, C.M. Maugard, H. Vasen, S. Seitz, L.A. Cannon-Albright, A. Schofield, M. Zelada-Hedman, Genetic heterogeneity and penetrance analysis of the BRCA1 and BRCA2 genes in breast cancer families, Am. J. Hum. Genet. 62 (1998) 676–689.

[3] M. Kallenberg, K. Petersen, M. Nielsen, A.Y. Ng, P. Diao, C. Igel, C.M. Vachon, K. Holland, R.R. Winkel, N. Karssemeijer, M. Lillholm, Unsupervised deep learning applied to breast density segmentation and mammographic risk scoring, IEEE Trans. Med. Imaging 35 (5) (2016) 1322–1331.

[4] K. Ganesan, U.R. Acharya, C.K. Chua, L.C. Min, K.T. Abraham, K.H. Ng, Computer-aided breast cancer detection using mammograms: A review, IEEE Trans. Biomed. Eng. 6 (2012) 77–98.

[5] Early detection: a long road ahead, Nature Review-Editorial. 18 (2018) 401.

[6] Y. Zheng, Breast cancer detection with Gabor features from digital mammograms, Algorithms 3 (1) (2010) 44–62.

[7] B.D. Fleet, J. Yan, D.B. Knoester, M. Yao, J.R. Deller, E.D. Goodman, Breast cancer detection using haralick features of images reconstructed from ultra wideband microwave scans, Transl. Res. Med. Imaging (2014) 9–16.

[8] M. Tan, B. Zheng, J.K. Leader, D. Gur, Association between changes in mammographic image features and risk for near-term breast cancer development, IEEE Trans. Med. Imaging 35 (7) (2016) 1719–1728.

[9] D. Shen, G. Wu, H.-Il Suk, Deep learning in medical image analysis, Annu. Rev. Biomed. Eng. 19 (2017) 221–248.

[10] Y. LeCun, Y. Bengio, G. Hinton, Deep learning, Nature 521 (7553) (2015) 436–444.

[11] R. Mehra Shallu, S. Kumar, An insight into the convolutional neural network for the analysis of medical images, in: IEEE International Conference on Nanotechnology for Instrumentation and Measurement Workshop (NANOfIM), 2017, in press.

[12] E. Hoffer, N. Ailon, Deep metric learning using triplet network, in: International Workshop on Similarity-Based Pattern Recognition, (2015) 84–92.

[13] A. Voulodimos, N. Doulamis, A. Doulamis, E. Protopapadakis, Deep learning for computer vision: A brief review, Comput. Intel. Neurosc. 2018 (2017) 1–13.

[14] F.A. Spanhol, P.R. Cavalin, L.S. Oliveira, C. Petitjean, L. Heutte, Breast cancer histopathological image classification using convolutional neural networks, in: International Joint Conference on Neural Networks, 2016, pp. 2560–2567.

[15] Z. Han, B. Wei, Y. Zheng, Y. Yin, K. Li, S. Li, Breast cancer multi-classification from histopathological images with structured deep learning model, Nature, Sci. Rep. (2017) 1–10.

[16] M. Anthimopoulos, S. Christodoulidis, L. Ebner, A. Christe, S. Mougiakakou, Lung pattern classification for interstitial lung disease using a deep convolutional neural network, IEEE Trans. Med. Imaging 35 (5) (2016) 1207–1216.

[17] M. Aubreville, C. Knipfr, N. Oetter, C. Jaremenko, E. Rodner, J. Denzler, C. Bohr, H. Neumann, F. Stellzle, A. Maier, Automatic classification of cancerous tissue in laser endo-microscopy images of the oral cavity using deep learning, Nature, Sci. Rep. (2017) 1–10.

[18] Lung Pattern Classification for Interstitial Lung Diseases Using Deep Convolutional Neural Network. 35, (5) (2016) 1207–1216.

[19] S. Hoo-Chang, H.R. Roth, M. Gao, L. Lu, Z. Xu, I. Nogues, J. Yao, D. Mollura, R.M. Summers, Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characterristics and transfer learning, IEEE Trans. Med. Imaging 35 (5) (2016) 1285–1297.

[20] O. İrsoy, E. Alpaydın, Continuously constructive deep neural networks, 2018 arXiv preprint arXiv:1804.02491.

[21] A.S. Razavian, H. Azizpour, J. Sullivan, S. Carlsson, CNN features off-the-shelf: an astounding baseline for recognition, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2014, pp. 512–519.

[22] H. Azizpour, A.S. Razavian, J. Sullivan, A. Maki, S. Carlsson, From generic to specific deep representations for visual recognition, 2014 arXiv preprint arXiv:1406.5774.

[23] K. Gopalakrishnan, S.K. Khaitan, A. Choudhary, A. Agrawal, Deep convolutional neural networks with transfer learning for computer vision-based data-driven pavement distress detection, Constr. Buil. Mater. 157 (2017) 322–330.

[24] O. Penatti, K. Nogueira, J. Santos, Do deep features generalize from everyday objects to remote sensing and aerial scenes domains?, in: IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2015, pp. 44–51.

[25] N. Tajbakhsh, J.Y. Shin, S.R. Gurudu, R.T. Hurst, C.B. Kendall, M.B. Gotway, J. Liang, Convolutional neural networks for medical image analysis: Full training or fine tuning?, IEEE Trans. Med. Imaging 35 (5) (2016) 1299–1312.

[26] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition, in: International Conference on Learning Representations, 2015, pp. 1–14.

[27] P. Wang, G.W. Cottrell, Basic level categorization facilitates visual object recognition, in: International Conference on Learning Representations, 2016, pp. 1–13.

[28] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, 2015. arXiv:151203385v1.

[29] J. Shijie, W. Ping, J. Peiyi, H. Siping, Research on data augmentation for image classification based on convolution neural networks, in: Chinese Automation Congress (CAC), 2017, pp. 4165–4170.

[30] J. Wang, L. Perez, The effectiveness of data augmentation in image Classification using deep learning, in: IEEE Conference on Computer Vision and Pattern Recognition, 2017, pp. 1–8.

[31] B.S. Veeling, J. Linmans, J. Winkens, T. Cohen, M. Welling,.
[32] Deep Neural network based recognition of plant diseases by leaf image classification, Comput. Intell. Neurosci.
[33] T.L. Sellaro, R. Filkins, C. Hoffman, J.L. Fine, J. Ho, A.V. Parwani, L. Pantanowitz, M. Montalto, Relationship between magnification and resolution in digital pathology systems, J. Pathol. Inform. 4 (2013).
[34] M. Peikari, M.J. Gangeh, J. Zubovits, G. Clarke, A.L. Martel, Triaging diagnostically relevant regions from pathology whole slides of breast cancer: A texture based approach, IEEE Trans. Med. Imaging 35 (1) (2016) 307–315.
[35] C. Loukas, S. Kostopoulos, A. Tanoglidi, D. Glotsos, C. Sfikas, D. Cavouras, Breast cancer characterization based on image classification of tissue sections visualized under low magnification, in: Computational and Mathematical Methods in Medicine, 2013.
[36] F. Spanhol, L. Oliveria, C. Petitjean, L. Heutte, A dataset for breast cancer histopathological image classification, IEEE Trans. Biomed. Eng. 63 (7) (2015) 1455–1462.
[37] F. Keskin, A. Suhre, K. Kose, T. Erashin, A.E. Cetin, R. Cetin-Atalay, Image classification of human carcinoma cells using complex wavelet based covariance descriptors, PLoS One 8 (1) (2013).