International Journal of Innovation Studies

ENHANCED BREAST CANCER DETECTION USING TRANSFER LEARNING: A COMPARATIVE STUDY OF VGG19, RESNET50, XCEPTION, AND INTEGRATED CLASSIFIER APPROACHES

S. Leena Nesamani[1], M. Gracy[2], S.Nirmala Sugirtha Rajini[3], P.Abhirami[4], T.Chandrakala[5]

[1]Department of Computer Applications, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, Tamilnadu – 603 203, India

[2]Department of Computer Applications, Faculty of Science and Humanities, SRM Institute of Science and Technology, Kattankulathur, Tamilnadu – 603 203, India

[3]Department of Computer Applications, Dr.M.G.R.Educational and Research Institute, Maduravoyal, Chennai, India

[4]Department of Computer Science, Ethiraj College for Women, Egmore, Chennai,India

[5]Department of Computer Science and Applications, Jawahar Science College, Neyveli, India

Abstract

Breast cancer is distinguished by the abnormal and uncontrolled proliferation of cells within breast tissues. Certain irregularities may be missed or improperly detected because of the complicated patterns of breast abnormalities and the limitations of human visual judgment. In recent years, computer vision has developed as an important tool in healthcare, enabling applications such as disease diagnosis, tumor detection, medical imaging, and patient monitoring. While Convolutional Neural Networks (CNNs) have shown effective in image processing applications, their success is often dependent on vast amounts of training data. However, a lack of labeled medical imaging data makes it difficult to train CNNs from scratch for clinical applications. To overcome this, transfer learning is used, which allows pre-trained models to be adapted for effective medical picture categorization on little datasets.

This study uses three well-known pre-trained CNN models to categorize mammographic pictures as normal or abnormal: VGG19, Xception, and ResNet50. The collected features from these models were identified using two distinct machine learning techniques: Logistic Regression and a custom-designed Neural Network classifier. Each model's performance was measured using common measures such as , F1-score, accuracy, recall, precision, and confusion matrix analysis. Among the examined models, VGG19 in conjunction with the neural network classifier produced the greatest classification accuracy of 94%, surpassing ResNet50 and Xception, which had accuracies of 90% and 88%, respectively. VGG19's outstanding performance across various evaluation measures demonstrates its applicability for breast cancer diagnosis in the analyzed dataset.

Keywords: Breast Cancer, Convolutional Neural Networks, Benign, Malignant, Mammogram, Machine Learning, Transfer Learning, Computer Vision,

I INTRODUCTION

Cancer is a big concern in every country. It is a disease that is often fatal, that has taken numerous lives and is expected to continue doing so in the future [1]. Among the most prevalent types of cancer is breast cancer, which is caused by uncontrolled proliferation of abnormal cells in the breast [2]. It is inevitable that, due to the features of breast anomalies and the nature of human visual perception, abnormalities are occasionally ignored or misclassified. As a result, biopsies are taken that aren't essential.

To address this issue, a system for computer-aided diagnosis (CAD) has been developed. Using image processing techniques and machine learning algorithms, a CAD system is proposed and implemented as an integrated system in this work. CAD seeks to detect and localize anomalies at an early stage, preventing the anomaly from spreading further [3]. Mammography is the screening method with the lowest radiation exposure, proven to reduce breast cancer-related mortality [4]. Mammography can detect breast tumors that are as small as 1mm even before they become palpable or clinically detectable through physical examinations [4].

Mammography testing reduces breast cancer death by 19% overall, with less effect for women in their 40s (15%) and larger benefit for those in their 60s (32%) [5]. As a result, the American Cancer Society recommends initiating screening mammography at age 45, with the possibility of starting earlier based on personal choice [6].

II LITERATURE

Deep learning enhances the detection accuracy of breast cancer in comparison to other methods. The detection procedure uses an end-to-end training methodology. Deep learning techniques employ automatic feature extraction which avoids human feature engineering, a time-consuming and cumbersome process which makes it more desirable in the classification process as compared to Machine Learning algorithms. The below mentioned research articles discuss how Deep learning methods can be effectively used in the identification of breast cancer. Kayla Mendel et al. highlight the importance of accurately assessing mitotic count as a key factor in breast cancer classification and diagnosis. Their study explores the potential of transfer learning for mitosis detection. The proposed approach utilizes the pre-trained convolutional neural network that has been modified by combining the completely connected layers with a random forest classifier. This combination helps in distinguishing characteristics in the nuclear regions to precisely determine the cell nuclei's classification label. By carefully fine-tuning the pre-trained model and preprocessing the extracted features, the proposed method achieves enhanced classification accuracy.

Leilei Zhou et al. study the influence of learning algorithms on CT images for the categorization of malignant and benign kidney cancers and create patient-level algorithms to increase the accuracy of classification. Within 15 days of improved CT examination, 192 cases of kidney tumors were gathered and diagnosed by pathologic diagnosis. To achieve this categorization, images from the ImageNet database was used to cross-train the Inception V3 network.

A computer-aided diagnosis approach utilizing convolutional neural networks is recommended by Woo Kyung Moon et al. to support radiologists in classifying mammographic mass of lesions. It is known that voluminous data are required to train networks from the ground up in deep learning. However, transfer learning serves as an effective strategy for handling relatively smaller datasets, particularly in the context of medical imaging. The results obtained demonstrate that the suggested approach is quite successful and can be used to identify benign or malignant tumors. Convolutional neural networks are used in this computer-aided diagnosis technique, which helps the radiologist classify mammography mass lesions. It is commonly known that in order to create networks with a certain depth from the base level, deep learning typically requires large datasets. Furthermore, transfer learning is a useful method for working with somewhat little datasets when it comes to medical imaging. The results obtained unequivocally demonstrate the efficacy of the suggested methodology and its potential for use in determining the benignity or malignancy of the heap lesions.

Ghulam Murtaza et al. observed that breast cancer classification involved the frequent analysis of mammograms and histopathologic images. Nearly half of the studies involved used public datasets. Normalization, scaling, and image augmentation were performed as part of preprocessing. CNN was employed to perform the classification process, while some studies performed well by employing transfer learning techniques from pre-trained models. Finally, the paper gave out some open research challenges for future researchers.

Nirmala Sugirtha Rajini, Leena Nesamani, and Abirami extracted the salient features based on their shape, position, and surface. The results proved to be most promising when compared to the other algorithms. Leena Nesamani et al. experimented with various computer-aided diagnosis experiments on MRI images of breast tumors and have identified that neural network classifiers perform much better than machine learning classifiers in the classification process of identifying whether the given MRI image belongs to malignant or benign. Moreover, it was observed that feature selection process plays a vital role in deciding the model performance, which leads to a careful selection of efficient features for an efficient diagnostic process.

The research work carried out by Sana Ullah Zhou et al. to classify breast tumors as either malignant or benign, extracted features from the breast cytology images using pre-trained CNN architectures like the Res Net, GoogLe Net, and VGG Net. The extracted features were passed through a fully connected layer for classification, which is carried out by the average pooling layer. The performance was observed on various benchmark datasets and was concluded that the proposed model produced higher accuracy of detection.

Sara Hosseinzadeh Kassani et al., in categorizing histological breast cancer images stained with hematoxylin and eosin suggest a deep learning-based solution that is completely automated that uses features derived from models of deep convolutional neural networks and a pooling procedure. The average performance of the pre-trained Xception model that has been pre-trained is 92.50%. Leilei Zhou et al. to enhance the reliability of classifiers, the author recommends using cross-modal transfer learning. Using a network trained on mammography pictures, identified tumors in breast MRI scans. Transfer learning may provide standard pre-trained shared models with a 94% accuracy rate.

III METHODOLOGY

Transfer learning is a machine learning approach that transfers the learned weights and biases of a neural network of one task to a neural network on a different target task. Two strategies can be used to improve the diagnostic success of computer vision tasks using transfer learning. The first uses feature extraction, which involves freezing the neural network's convolutional bases. The second strategy involves training a dense layer (dense → dropout → classifier) [7], [8]. The advantages of transfer learning are: firstly, it reduces training time; secondly, it provides better performance for neural networks; and finally, it lowers the need for a large amount of target domain data in the process of constructing a target learner. Pre-trained networks have been trained on more than 1.2 million images and are able to categorize pictures into a thousand different categories, like table, horse, airplane, and many more [7]. Compared to training a network from scratch, using a pre-trained network with transfer learning is far more effective and quicker.

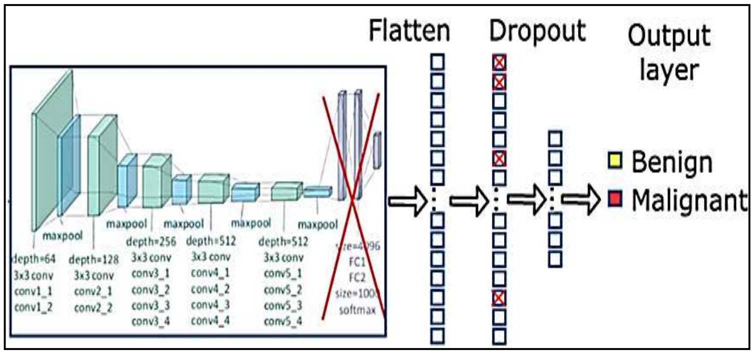

Fig. 1. VGG19 for Transfer Learning

Breast cancer classification task performance was enhanced using the concept of same-domain transfer learning. To identify mammography images as either benign or malignant, a benchmark pre-trained model is created and trained on a collection of radiology images [8]. The VGG19 for transfer learning is displayed in Fig.1.

The dataset employed for this research consists of both normal and pathological breast mammogram images are utilized in this research.. Breast mammogram images were collected from various hospitals. The right and left breasts were analyzed using standard imaging perspectives, including mediolateral oblique (MLO) and bilateral craniocaudal (CC) views. A total of 734 images, out of which 367 were Benign and 422 were Malignant in nature. Additionally, data was also considered from the publicly available Miniddsm2 dataset, where the images were categorized into benign and malignant and images.

One type of convolutional neural network (CNN) is the Visual Geometry Group (VGG 19) architecture that consists of 19 layers. Sixteen convolutional layers, three fully connected layers, five max pooling layers, and one SoftMax layer make up this conventional convolutional neural network (CNN) architecture. It uses alternating structure of multiple convolutional and non-linear activation functions, Maxpooling for down sampling to reduction in the spatial resolution and the activation function used was ReLU. The Max Pooling sampling function could be expressed as:

$$\chi_{p_j}^{(n)} = f\left(\tau_j^n \text{ down}\left(\chi_j^{(n-1)}\right) + b_j^{(n)}\right) \quad \text{… Eqn. (1)}$$

where,

f – Activation function (ReLu)

τ_j^n - Coefficient corresponding to the jth feature map of the nth layer

$b_j^{(n)}$ – Bias for the jth featuremap of the nth layer

The problem of vanishing/exploding gradients is a problem with the multilayer training networks. This problem was addressed by introducing a deep residual learning framework where shortcut connections that simply performed identity mappings were chosen. In a ResNet model, the layers are designed to learn a residual mapping, represented as H(x). A function F(x) is approximated by the non-linear layers, where F(x) = H(x) - x. Consequently, the initial mapping is formulated as H(x) = F(x) + x. For the ResNet 50, a minor alteration was made where the shortcut connections skipped two tiers.
Xception, is the short form for "extreme Inception," pushes the principles of the Inception architecture to further levels. 1x1 convolutions of the Inception network are initially applied to reduce the input's dimensionality, followed by the application of various filters to each depth-wise feature space. But Xception flips this process around. It compresses the input space using 1x1 convolutions across the depth after applying filters to each depth-wise feature map. This approach closely resembles the depthwise separable convolution, a technique that has been incorporated into neural network architectures since 2014.

The way non-linearity is handled following the initial operation is another significant difference between Xception and Inception. A ReLU activation function follows both procedures in the Inception model, however, Xception skips this step and introduces non-linearity instead. The initial investigations consisted of six experiments conducted with the pretrained architectures. These experiments can be divided into two groups, based on the final classifier employed for the classification process. Two types of classifiers were employed for the final classification. The first group employed a Logistic regressor as the classifier. The second group employed a custom-made neural network classifier.

One popular machine learning method for categorization challenges is logistic regression. The relationship between the independent and dependent variables is modeled using a logistic function. The dependent variable in this approach is binary, meaning it can only have two possible outcomes (e.g., classifying cancer as either malignant or benign). Due to this binary nature, logistic regression is particularly suitable for handling such classification problems. To convert predicted values into probabilities is done by using the sigmoid function; any real-valued number can be converted into a range between 0 and 1. The hypothesis is given by the function:

$$h=(X\theta) \qquad \ldots\text{Eqn. (2)}$$

where,

- X is the input feature matrix
- θ is the vector of parameters (weights) that the model learns and
- g- Sigmoid function

A custom-made neural network architecture, which is utilized for the ultimate classification of breast mammogram images, consists of a flatten_1 layer, which is the first input layer that gets the data from the feature extraction module of the retrained network. The next dense_1 _1 layer consists of 128 neurons, which is followed by a dropout layer to handle overfitting during training, and finally, the dense_2 layer is the output layer with 2 neurons. The structure of the uniquely designed neural network is shown below in Fig. 2.

```
_________________________________________________________________
Layer (type)                 Output Shape              Param #
=================================================================
flatten_1 (Flatten)          (None, 4608)              0
_________________________________________________________________
dense_1 (Dense)              (None, 128)               589952
_________________________________________________________________
dropout_1 (Dropout)          (None, 128)               0
_________________________________________________________________
dense_2 (Dense)              (None, 2)                 258
=================================================================
Total params: 590,210
Trainable params: 590,210
Non-trainable params: 0
_________________________________________________________________
```

Fig. 2. Custom-made Neural Network classifier

Each layer computes a linear transformation given by the formula

$$Z^{(l)} = W^{(l)} . A^{(l-1)} + b^{(l)} \quad \text{...Eqn. (3)}$$

where,

- l represents the layer number
- $Z^{(l)}$ is the pre-activation output of layer.
- $W^{(l)}$ *and* $b^{(l)}$ represent the weight matrix and bias vector of the layers, respectively.
- $A^{(l-1)}$ is the activation of layer.

The **loss function** (categorical cross-entropy) is given by:

$$L = \sum_{i=1}^{N} \sum_{j=1}^{C} y_i j \, log(\hat{y}_i j) \quad \text{...Eqn. (4)}$$

where,

- N represents the quantity of samples.
- CC represents the total number of classes.
- Yi denotes the actual label, which is one-hot encoded.
- $\hat{y}_i j$ signifies the predicted probability for class j.

In order to improve the model's performance, the categorical cross-entropy loss function assesses the difference between the true labels and the predicted probabilities.

IV RESULT AND DISCUSSION

The initial investigations were carried out with the Target dataset, which is the breast mammogram image dataset. The experiment began with the preparation of the raw data, where the images were compressed with Lossless compression (LJPEG encoding), followed by the bi-filtering to remove noise, and finally cropping the region of interest (ROI). As part of the feature extraction process, label encoding was performed in addition to the automatic feature extraction done by the convolutional base models. Classification using Logistic Regression Classifier and a Neural Network Classifier was performed and the results were captured.

A summary of the performance metrics for experiments conducted with the Logistic Regression classifier is shown below in Table I.

TABLE I EVALUATION METRICS FOR EXPERIMENTS CONDUCTED WITH THE LOGISTIC REGRESSION CLASSIFIER

Classifier	Model	Precision	recall	F1-score	Accuracy %
Logistic Regression	**VGG 19**	**0.83**	**0.80**	**0.82**	**72%**
	Res Net50	**0.83**	**0.76**	**0.79**	**70%**
	Xception	**0.83**	**0.76**	**0.79**	**70%**

From Table I, we can infer that VGG19 with LR classifier performs better than the other models. The performance of the NN classifier for the classification of breast mammogram images as benign and malignant is shown in Fig. 3. Summary of the performance metrics for experiments conducted with the Neural Network classifier as referred the Table II.

TABLE II EVALUATION METRICS FOR EXPERIMENTS CONDUCTED WITH THE NEURAL NETWORK-BASED CLASSIFICATION MODEL

Classifier	Model	Precision	recall	F1-score	Accuracy %
Neural Network	VGG19	0.95	0.94	0.95	94%
	Res Net50	0.91	0.9	0.92	90%
	Xception	0.86	0.85	0.89	88%

From the Table II, we can infer that VGG19 with a NN classifier performs better than the other models. Performance of the pretrained models employing the transfer learning technique in terms of the learning curves is in Fig 3 and the comparison of the various performance metrics for the three models is shown in Fig 4. Both the figures show that VGG19 performs better than the other two models using transfer learning in classification of mammogram images as benign or malignant.

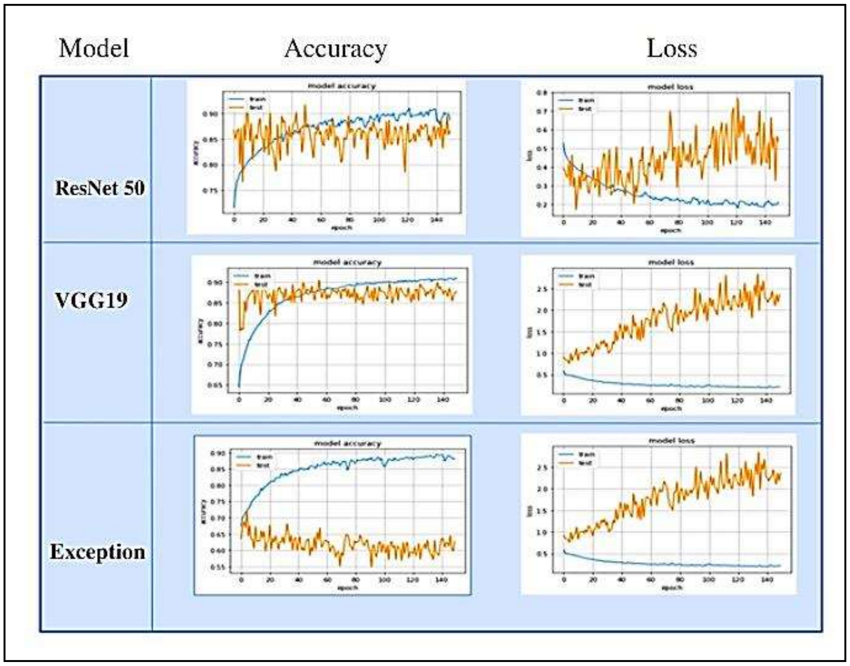

Fig. 3. Performance of pre-trained models using NN classifier

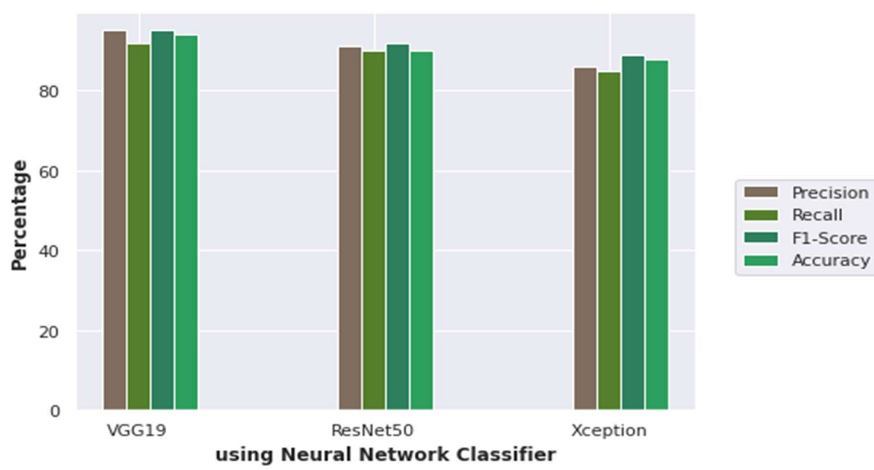

Fig. 4. Comparison of performance metrics for VGG19, ResNet50, and Xception Architectures

CONCLUSION

Transfer learning has revolutionized deep learning, particularly in medical image classification, where handling limited data while ensuring high accuracy is crucial. From the experiments conducted, it can be concluded that VGG 19 architecture with a neural network classifier has performed very well while classifying images from breast mammograms. The Convolutional base models are trained on the very popular ImageNet dataset which is a different domain dataset. The structure and shapes of the Radiology images are completely different from the images in the ImageNet dataset. This could be considered as a future scope for the research where the convolutional base is also trained from the scratch on a medical image dataset to improve more domain knowledge and increase the degree of accurate predictions.

REFERENCES

1. H. Sung *et al*., "Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries," *CA: A Cancer Journal for Clinicians*, vol. 71, no. 3, pp. 209–249, 2021, doi: 10.3322/caac.21660.
2. C. E. DeSantis *et al*., "Breast cancer statistics, 2019," *CA: A Cancer Journal for Clinicians*, vol. 69, no. 6, pp. 438–451, 2019, doi: 10.3322/caac.21583.
3. K. Doi, "Computer-aided diagnosis in medical imaging: Historical review, current status and future potential," *Comput. Med. Imaging Graph.*, vol. 31, no. 4–5, pp. 198–211, 2007, doi: 10.1016/j.compmedimag.2007.02.002.
4. C. D. Lehman *et al*., "Diagnostic accuracy of digital screening mammography with and without computer-aided detection," *JAMA Intern. Med.*, vol. 175, no. 11, pp. 1828–1837, 2015, doi: 10.1001/jamainternmed.2015.5231.
5. E. R. Myers *et al*., "Benefits and harms of breast cancer screening: A systematic review," *JAMA*, vol. 314, no. 15, pp. 1615–1634, 2015, doi: 10.1001/jama.2015.13183.
6. K. C. Oeffinger *et al*., "Breast cancer screening for women at average risk: 2015 guideline update from the American Cancer Society," *JAMA*, vol. 314, no. 15, pp. 1599–1614, 2015, doi: 10.1001/jama.2015.12783.
7. S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Transactions on Knowledge and Data Engineering*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010, doi: 10.1109/TKDE.2009.191.
8. T. Y. Kong, A. G. Abdul Jabar, and R. A. Salam, "Transfer learning for breast cancer classification using deep convolutional neural networks," *IEEE Access*, vol. 7, pp. 124677–124685, 2019, doi: 10.1109/ACCESS.2019.2938315.
9. Breast Cancer Screening in Average-Risk Women. Ann Intern Med. 2019 Apr 16;170(8). doi: 10.7326/P19-0007. Epub 2019 Apr 9. PMID: 30959524.
10. Bredemeyer M. ACS Releases Guideline on Breast Cancer Screening. Am Fam Physician. 2016 Apr 15;93(8):711-2. PMID: 27175846.
11. Kayla Mendel, Hui Li, Deepa Sheth & Maryellen Giger 2019, 'Transfer learning from convolutional neural networks for computer-aided diagnosis: a comparison of digital breast tomosynthesis and full-field digital mammography', *Academic Radiology*, vol, 26, no. 6, pp. 735-743
12. Leilei Zhou, Zuoheng Zhang, Yu-Chen Chen, Zhen-Yu Zhao, Xin-Dao Yin & Hong-Bing Jiang 2019, 'A deep learning-based radiomics model for differentiating benign and malignant renal tumors', *Translational Oncology,* vol. 12, no. 2, pp. 292-300
13. Woo Kyung Moon, Yan-Wei Lee, Hao-Hsiang Ke, Su Hyun Lee, Chiun-Sheng Huang & Ruey-Feng Chang 2020, 'Computer-aided diagnosis of breast ultrasound images using ensemble learning from convolutional neural networks', *Computer Methods and Programs in Biomedicine,* vol. 190, doi: 10.1016/j.cmpb.2020.105361

14. Ghulam Murtaza, LiyanaShuib, Ainuddin Wahid Abdul Wahab, Ghulam Mujtaba, Ghulam Mujtaba, Henry Friday Nweke et al. 2020, 'Deep learning-based breast cancer classification through medical imaging modalities: state of the art and research challenges', *Artificial Intelligence Review*, vol. 53, no. 3, pp. 1655-1720
15. Nirmala Sugirtha Rajini S, Leena Nesamani, S & Abirami, P 2021, 'Brain Tumor Segmentation from MRI Images using Deep Learning-based CNN with SVM Classifier', *International Journal of Grid and Distributed Computing,* vol. 14, no. 1, pp. 1557-1564
16. Leena Nesamani, S, Nirmala Sugirtha Rajini, S, Josphine, MS & Jacinth Salome, J 2021, 'Deep Learning-Based Mammogram Classification for Breast Cancer Diagnosis Using Multi-level Support Vector Machine', *Advances in Automation, Signal Processing, Instrumentation, and Control, Lecture Notes in Electrical Engineering 700,* pp. 371-383
17. Sana Ullah Zhou, Naveed Islam, Zahoor Jan, Ikram Ud Din & Joel Rodrigues, JPC 2019, 'A novel deep learning based framework for the detection and classification of breast cancer using transfer learning', *Pattern Recognition Letters,* vol. 125, pp. 1–6
18. Sara Hosseinzadeh Kassani, Peyman Hosseinzadeh Kassani, Michal Wesolowski, J, Kevin Schneider, A & Ralph Deters 2019, *IEEE International Conference on Information and Communication Technology Convergence (ICTC), 16-18 Oct,* 'Breast cancer diagnosis with transfer learning and global pooling', Jeju-si Jeju-do, South Kores, pp. 519-524
19. Leilei Zhou, Zuoheng Zhang, Yu-Chen Chen, Zhen-Yu Zhao, Xin-Dao Yin & Hong-Bing Jiang 2020, 'Transfer learning-based DCE-MRI method for identifying differentiation between benign and malignant breast tumors', *IEEE Access,* vol. 8, pp. 1752