Enhancing IDC Histopathology Image Classification: A Comparative Study of Fine-Tuned and Pre-trained Models

Anusree Kanadath[(✉)], J. Angel Arul Jothi, and Siddhaling Urolagin

Department of Computer Science, Birla Institute of Technology and Science Pilani,
Dubai Campus, Dubai International Academic City, 345055 Dubai, UAE
{p20180904,angeljothi,siddhaling}@dubai.bits-pilani.ac.in

Abstract. Invasive ductal carcinoma (IDC) is a type of breast cancer that affects adult women all around the world. This cancer starts in the duct cells of the breast, spreads through the lymph system, and eventually affects nearby organs and bones. It's crucial for physicians to correctly identify the various forms of breast cancer. Instead of doing this manually, it's better to use computer programs because it saves time and reduces mistakes. This study presents a computer-assisted diagnosis method that utilizes deep convolutional neural networks to classify IDC histopathology images. These networks are trained using two forms of transfer learning: feature extraction and fine-tuning. In this study IDC classification is done using 4 well known deep learning networks, Xception, DenseNet169, ResNet101 and MobileNetV2. The dataset used is a publicly available IDC dataset containing 168 whole slide images. The evaluation results show that the fine-tuned models give better classification results than feature extractor models for IDC histopathology image classification.

Keywords: Deep learning · Transfer learning · Histopathology · Image classification · Fine-tuned models · Computer aided diagnosis

1 Introduction

Invasive ductal carcinoma (IDC) is the most prevalent form of breast cancer, accounting for approximately 80% of all breast cancer cases in women. The term 'invasive' signifies that the cancer has extended into the adjacent breast tissues. 'Ductal' implies that the cancer originated within the milk ducts, the channels responsible for transporting milk from the lobules to the nipple. Lastly, 'Carcinoma' denotes any cancer that initiates in the skin or other tissues covering internal organs, including breast tissues. IDC is characterized by the transformation of abnormal cells within the milk duct lining, leading to their invasion of breast tissue beyond the confines of the duct walls. Once that happens, the cancer cells can spread. In the United States, it is anticipated that 297,790 cases of invasive breast cancer and 55,720 cases of non-invasive breast cancer will be diagnosed in the year 2023 [2]. In the last few decades, the number of women

diagnosed with invasive breast cancer has increased by approximately 0.5% per year. For the year 2023, estimates suggest that 2800 men in the United States will also be diagnosed with invasive breast cancer. Tragically, it is anticipated that breast cancer will result in approximately 43700 deaths in the United States. Among these, the vast majority, around 43170, are expected to affect women, with a smaller number, approximately 530 affecting men.

Timely identification of invasive carcinoma plays a pivotal role in cancer treatment. Detecting invasive carcinoma at an early stage substantially enhances the prospects of successful treatment and long-term survival. As invasive carcinomas progress, it is typically more difficult to treat, necessitating more aggressive therapies and posing greater health risks for the patient. Not only does early detection increases the likelihood of complete tumor removal, but it also allows for a wider range of treatment options that may be less invasive and have fewer adverse effects.

Researchers have made significant advances in utilizing deep learning (DL) algorithms to aid in cancer screening, including breast cancer. These DL algorithms aid pathologists in rapidly analyzing histopathology images and diagnosing cancer. Despite the remarkable success of DL algorithms, their seamless integration into digital pathology faces significant obstacles. Some of these challenges include the absence of the necessary labeled data for complex deep learning models, the texture variation of the tissue types and the vast dimensionality of whole slide images (WSIs) with common image sizes exceeding 50000×50000 pixels.

Transfer learning is one of the effective solutions for histopathology image processing [4,14,15]. Transfer learning is the process of applying a model that has been trained on a large dataset for a specific task to a similar task, despite having a smaller dataset. It offers numerous benefits, including the reduction of training time, the improvement of output accuracy, and the requirement for less training data. Negative transfer and overfitting are the two major disadvantages of transfer learning.

This paper focuses on transfer learning-based approaches for classifying histopathology image regions as IDC+ve or IDC-ve. The pretrained deep models are reused for feature extraction approach and fine-tuned approach. 4 deep learning models, XceptionNet, DenseNet169, ResNet101 and MobileNetV2 pretrained on the ImageNet dataset were used. An IDC dataset with 168 WSIs that is publicly available was used for this study. The detailed comparative analysis helps the researchers to identify the best transfer learning models for IDC histopathology image classification.

The remainder of this paper is structured as follows: Sect. 2 provides an overview of the relevant literature. Section 3 describes the methodology used in this study. Section 4 describes in detail the experimental setup. The results and analysis are presented in Sect. 5. Section 6 is the conclusion of the paper.

2 Literature Survey

In this section, we review several recent papers related to the automated detection of IDC in breast cancer using deep learning techniques. These papers demonstrate the advancements in this field and the ongoing efforts to enhance IDC classification accuracy and efficiency.

Andrew janowczyk and Anant madabhushi [12] have developed and implemented a deep learning model for a variety of digital pathology tasks, including segmentation, detection, and classification. The study achieved an F1 score of 0.7648 in IDC detection task. Aiza and Alexander [18] presented an improved convolutional neural network (CNN) architecture for predicting IDC. Their model demonstrated remarkable performance with an F1 score of 85.28% and a balanced accuracy of 85.41%, surpassing previous deep learning approaches. This paper emphasized the significance of CNN enhancements for accurate IDC detection.

Jianfei zhang et al. [22] proposed a method that merge a multi-scale residual CNN (MSRCNN) and support vector machine (SVM) for IDC detection. The approach demonstrated an average accuracy of 87.45%, average balanced accuracy of 85.7%, and an average F1 score of 79.89% after 5-fold cross-validation. Avishek and Sunanda [7] implemented a CNN model for breast cancer classification. The model demonstrated a classification accuracy of 78.4% when evaluated on the IDC dataset. Mohammad et al. [3] investigated two approaches for IDC classification: a baseline CNN model and transfer learning using the VGG16 CNN model. The baseline model achieved an F1 score of 83% and an accuracy of 85%. Notably, transfer learning through feature extraction produced superior classification results compared to the baseline model.

Justin et al. [20] investigated a range of CNN architectures for automated breast cancer detection. They assessed four different architectures using a substantial dataset and achieved remarkable results with one particular fine-tuned CNN architecture. This model yielded impressive results, including an F1 score of 92%, a balanced accuracy of 87%, and an accuracy of 89%. The study identified a finely tuned CNN architecture that consistently delivered outstanding performance. Érika et al. [5] emphasized the advantages of using deep learning for IDC detection. Their 3-hidden-layer CNN, with data balancing, achieved both accuracy and an F1-Score of 0.85.

In conclusion, the literature survey has provided a comprehensive overview of the state of research in IDC image classification using deep learning. While significant advancements have been achieved in attaining elevated accuracy and F1-scores, there remains an ongoing need for deeper exploration into the most efficient approaches like fine-tuning and feature extractor models. This encompasses a comprehensive examination of various model architectures and transfer learning techniques to enhance their efficacy.

3   Materials and Methods

3.1   Transfer Learning

Machine learning and deep learning have revolutionized computer vision, natural language processing, and speech recognition by achieving success at complex tasks. However, these models often need huge, high-quality datasets and significant computational resources. The performance of DL models is critically dependent on the availability of an adequate volume of accurately labeled training data [17]. While there is an abundance of labelled data for natural images, a lack of annotated medical images presents a significant challenge in the domain of medical image analysis. This lack of training data has the potential to hinder the effectiveness of deep learning models. Therefore, transfer learning has emerged as a viable alternative to conventional DL approaches, providing a valuable solution to enhance model performance and overcome data limitations [9].
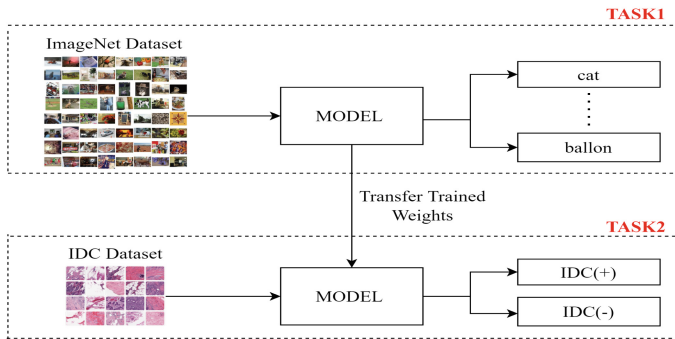

Fig. 1. General diagram of transfer learning for image classification.

Pre-trained models are DL models that have been used to solve one problem using a large dataset and then reused to solve another similar problem with a smaller dataset. Transfer learning is the process of transferring a pre-trained model's weights to solve another problem. Transfer learning saves training time, improves neural network performance, and reduces the demand for data. These advantages collectively contribute to the widespread popularity of transfer learning as a powerful machine learning method. The general diagram of transfer learning is given in the Fig. 1. The deep learning model is initially trained using the ImageNet visual database, which contains more than 14 million images, with the objective to classify images into 1000 distinct classifications. After training, the model's weights are set to their optimal values, resulting in a model that has been effectively learned. This pre-trained model is then used within a context for transfer learning. The pre-trained model is repurposed for binary classification in order to solve the IDC histopathology image classification problem. The
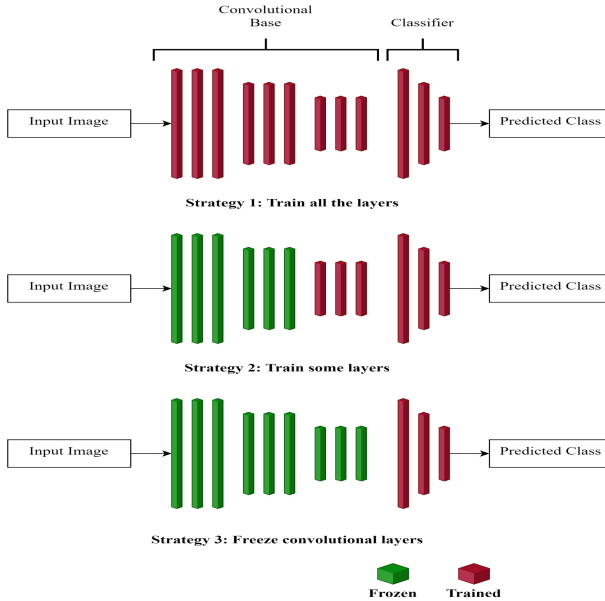
Fig. 2. Fine-tuning strategies.

model is modified according to the binary classification problem. After this modification, the model is trained with the IDC dataset. Finally, the model classifies histopathological images into IDC+ve or IDC-ve classes by using the knowledge it gained during its original training.

The pre-trained model can be used in three different ways [21], as shown in Fig. 2. The pre-trained model consists of a pre-trained convolutional base followed by a classifier. The convolutional base is comprised of a series of convolutional and pooling layers designed to extract image features. In contrast, the classifier is comprised of fully connected layers whose primary function is to classify images based on the extracted features [13]. The first approach is to use the IDC dataset to train the whole model. In this method, the architecture of the model that has already been trained is preserved, while training is tailored to suit the specific requirements of the IDC dataset. Nonetheless, this method requires a large dataset and substantial computational resources. The second approach is to freeze some layers of the convolutional base while training others. The lower layers capture general characteristics, whereas the upper layers focus on problem-specific characteristics. During the training process, certain layers can be "frozen" and kept unchanged by altering their layer weights. This method is especially useful when working with limited datasets or models with numerous parameters. This method helps the model to acquire both general and task-specific features, which may result in improved performance. This adaptability is a significant advantage, but it requires more computational resources. The third approach involves freezing all the layers of the convolutional base,

in its original form. This strategy is useful when computational resources are scarce, dataset size is small, or the pre-trained model has already demonstrated proficiency in solving a problem closely related to the target task. This approach can be computationally efficient, as it avoids the need to train all layers from scratch, making it more feasible in resource-limited situations.

3.2   Pre-trained Models

Several image classification models are training on extensive image datasets, including the widely known ImageNet. Some of the most well-known pre-trained classification models are AlexNet, VGG, GoogLeNet, ResNet, DenseNet, MobileNet, EfficientNet, Xception, NASNet, SqueezeNet, ShuffleNet, etc. [19]. 4 popular pre-trained models that have shown promising results for medical image classification such as DenseNet, ResNet, MobileNet, and XceptionNet are used in this work. This next section provides an overview of the characteristics of these pre-trained models.

ResNet: ResNet, commonly referred to as residual networks, represents a notable breakthrough in the domains of deep learning and computer vision [8]. As networks grow deeper with more layers, they often encounter the vanishing gradient problem, affecting effective training [16]. ResNet proposed a solution with the implementation of residual blocks, which are alternatively referred to as skip connections or shortcut connections. These connections provide alternative pathways for data and gradients to flow thus making training possible.

Figure 3 depicts the fundamental building block of the resnet, known as the 'residual block.' A residual block consists of two convolutional layers (Conv) accompanied by batch normalization (BN) and ReLU activation functions. The input feature map is denoted as X, while $F(X)$ represents the output obtained after passing through the two convolutional layers followed by BN and ReLU layers. Then the final output $H(X)$ from the residual block is defined by Eq. 1.
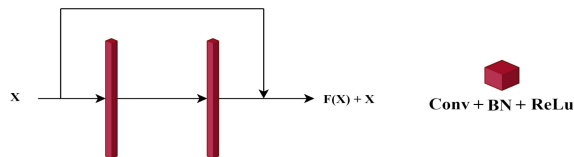
$$H(X) = F(X) + X \tag{1}$$
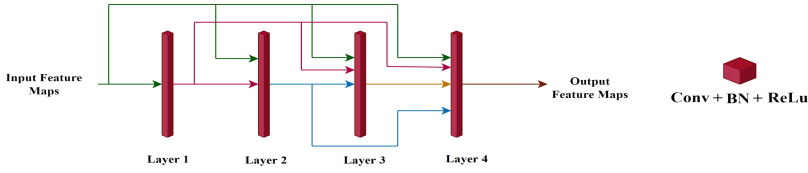

Fig. 3. Architecture of a residual block.

Fig. 4. Architecture of a DenseNet Block.

DenseNet: The DenseNet model developed by Huang et al. shown remarkable classification performance in 2017 when applied to publically available image datasets such as CIFAR-10 and ImageNet [11]. In the DenseNet architecture, every layer is connected to the successive layers within the network. This means that the features acquired by any layer are readily shared throughout the entire network, creating an enhanced information flow. Consequently, this architecture significantly improves the efficiency of training deep networks, all the while enhancing model performance. Furthermore, the presence of dense connections plays a role in mitigating overfitting, particularly on tasks involving smaller datasets. Figure 4 depicts a fundamental building block in the DenseNet architecture, referred to as a DenseNet block. In this block, the output of each convolutional layer is not only passed forward to the next immediate layer but also serves as input to every subsequent convolutional layer within the same block.

MobileNet: MobileNet is a family of lightweight deep neural network architectures designed for fast and effective deployment on mobile and embedded devices [10]. MobileNet's efficiency is based on the concept of depthwise separable convolution, which divides the standard convolution operation into two distinct steps: depthwise convolution and pointwise convolution. In depthwise convolution, a single filter is applied per input channel, resulting in a substantial reduction in computational load. The subsequent pointwise convolution combines the outcomes of the depthwise convolution to produce feature maps. This technique significantly reduces the computational overhead while preserving model accuracy. MobileNet architectures are renowned for their parameter efficiency. The use of depthwise separable convolution and reduced model size means they have significantly fewer parameters compared to traditional deep neural networks while maintaining competitive accuracy. This efficiency is crucial for deployment on resource-constrained devices. MobileNet also has a variety of model versions, from MobileNetV1 to MobileNetV3, each of which is made to meet different needs in terms of model size, speed, and accuracy.

XceptionNet: XceptionNet, often known as "Extreme Inception," represents a significant breakthrough in the field of deep learning and computer vision [6]. The basic concept behind XceptionNet is depthwise separable convolutions, similar to the approach used in MobileNet. The application of depthwise separable convolutions reduces model parameters and improves computational efficiency. The

design of XceptionNet was inspired by the multi-scale feature extraction capabilities of the Inception architecture. However, Xception takes this concept to an extreme by implementing depthwise separable convolutions across all Inception modules. This deep and efficient architecture enables XceptionNet to capture intricate patterns and features in data while achieving impressive computational efficiency.
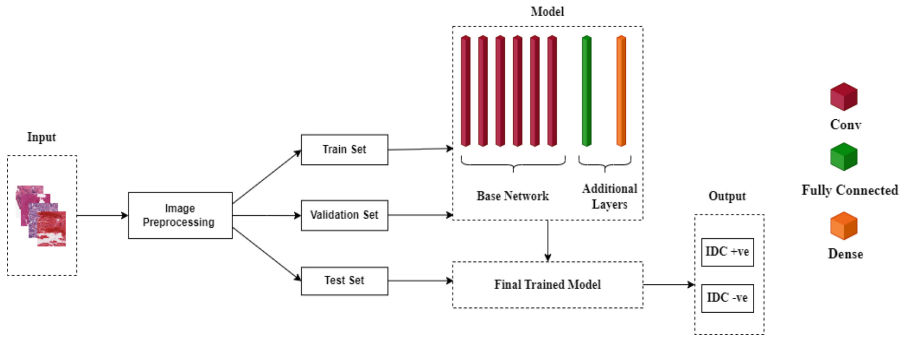
3.3   Proposed Model

Fig. 5. An illustration of Proposed model for IDC image classification

Figure 5 provides a detailed overview of the IDC image classification process. The process begins by acquiring images from the IDC dataset, which are then subjected to image preprocessing. Each image is resized to 48×48 dimensions during this preprocessing phase. To address class imbalances in the dataset, oversampling techniques such as SMOTE are applied to the images. After achieving class balancing, the dataset is divided into three sets: the training set, the test set, and the validation set. The classification model is trained using images from the training and validation sets as well as their respective class labels. These classification models are pre-trained models that consist of a base network followed by additional layers tailored for the specific classification task. After the training phase, the trained model is prepared for testing. During the testing phase, images from the test set are fed into the trained model, which predicts whether each image is IDC+ve or IDC-ve.

The top layers of each pre-trained model is removed treating the remaining architecture as the base network for the proposed model. The specific number of top layers removed was determined through experimentation and analysis of the model's architectures. After that, a series of batch normalization, dropout, and fully connected layers are added to this network. These additions were thoroughly selected to optimize the performance of the model. In our experiments on the IDC dataset, we investigate two distinct approaches. The first approach is the

feature extraction approach, which involves freezing all layers of the base network while training the remaining layers using a balanced IDC dataset. It is decided to freeze all layers so that the general feature representations that the pre-trained model had learned would remain the same. In contrast, the second approach is the fine-tuning approach, which entails freezing specific layers of the base network, enabling the remaining layers to undergo training using the balanced IDC dataset. It also enables the capture of IDC specific features while retaining general knowledge from the pre-trained layers.

4   Experimental Setup

The IDC dataset employed in this study comprises digitized histopathology slides collected from 162 individuals diagnosed with IDC at the Hospital of the University of Pennsylvania and the Cancer Institute of New Jersey [1]. From this dataset, a total of 277,524 patches, each measuring 50×50×3 (RGB), were extracted. Among these patches, 198,738 were identified as IDC-ve, while 78,786 were classified as IDC+ve. Sample images from each class is given in the Fig. 6. First 6 images belongs to IDC -ve classes and last 6 images belongs to IDC +ve classes.
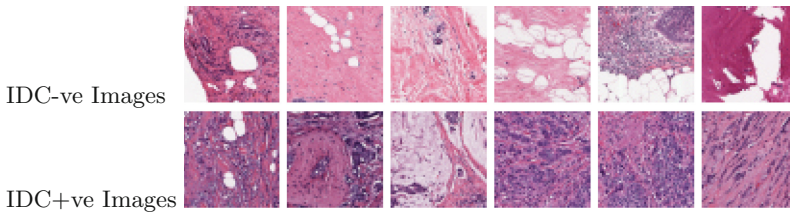

Fig. 6. Sample images from IDC dataset.

During the training phase, the IDC dataset was split 80:10:10 into training, testing, and validation sets. The model was trained for 100 epochs using the Adam optimizer. To evaluate the performance of the model, the binary cross-entropy loss function was used, and a batch size of 128 was chosen to process the 48×48 histopathology images efficiently. Accuracy, precision, recall, and F1 score were employed as evaluation metrics. The experiments were conducted on a computer with an Intel(R) Xeon(R) W-2123 CPU, 16 GB of RAM, and a 1 TB hard drive. The TensorFlow framework was employed for implementing the model, and the code was developed using the Python programming language.

5   Results and Discussion

In this section, we present the experimental results and analyze the outcomes of two distinct methods applied to the IDC dataset: the feature extraction approach

and the fine-tuning approach. We selected four deep classification models for our experiments and evaluated their performance on a balanced IDC dataset. A comparative analysis was conducted by considering both the number of parameters to be trained and the classification accuracy.

5.1 Parameter Analysis of Deep Learning Models

Table 1 provides a comprehensive overview of the trainable and non-trainable parameters associated with the deep models used in our experiments. The feature extractor models were directly applied to the dataset without any structural modifications, with the only alteration being the addition of fully connected layers at the end. Consequently, feature extractor models tend to have a high count of non-trainable parameters while keeping the number of trainable parameters relatively low. Among the various feature extractor models, MobileNet stands out with an exceptionally low number of parameters. This is attributed to its design, which prioritizes efficiency for lightweight devices by employing depthwise separable convolution operations to reduce parameter count. In contrast, ResNet and its variants tend to have a notably higher number of parameters due to their utilization of residual networks within their architecture.

Similarly fine-tuning involves taking an existing deep model and adjusting it to better suit a specific task or dataset. Unlike feature extractor models, fine-tuning allows for more flexibility in modifying the model's architecture, including unfreezing and retraining certain layers. By keeping most of the model's parameters frozen, fine-tuning requires training only a fraction of the total parameters, making the process more efficient and faster compared to training an entirely new model. Among the various fine-tuned models, ResNet101 has a high number of parameters to train. When comparing fine-tuning with feature extraction approach, one notable difference is that fine-tuning typically involves training more parameters than feature extraction approach. This allows the model to adapt its representations and features for a specific task, making it more suitable for the target application.

Table 1. Details of trainable and non-trainable parameters (in million).

Models	Feature extraction		Fine-tuning	
	Non-trainable	Trainable	Non-trainable	Trainable
Xception	20.89	0.0369	14.07	6.82
DenseNet169	12.64	0.0033	11.66	0.9788
ResNet101	42.67	0.0164	28.20	14.46
MobileNetV2	2.26	0.0102	0.0387	2.22

5.2   Classification Results and Analysis

The classification results of the developed models are detailed in Table 2 and Table 3. Table 2 showcases the classification results of feature extractor models when applied to the IDC dataset, whereas Table 3 illustrates the classification outcomes of fine-tuned models on the same IDC dataset

Among the feature extractor models, Xception and DenseNet169 demonstrate the highest levels of efficiency. The Xception model achieved a training accuracy of 0.85 and a testing accuracy of 0.80, demonstrating a strong ability to match the training data and effectively adapt to new, unseen data. The model achieved 0.81 precision and 0.81 recall, resulting in an F1 score of 0.81. DenseNet169 exhibits excellent results, with a training accuracy of 0.84 and a testing accuracy of 0.83, respectively. Both its precision and recall are high at 0.84, yielding an F1 score of 0.84. On the other hand, ResNet101 and MobileNetV2 exhibit slightly lower levels of performance. The ResNet101 model achieved a training accuracy of 0.74 and a testing accuracy of 0.74. The precision and recall values of the model are both 0.75, leading to an F1 score of 0.74. The MobileNetV2 model achieved a training accuracy of 0.79 and a testing accuracy of 0.76. Additionally, the precision, recall and F1 score are 0.77.

Table 2. Classification results of feature extractor models on balanced dataset.

Models	Train_acc	Test_acc	Precision	Recall	F1
Xception	0.85	0.80	0.81	0.81	0.81
DenseNet169	0.84	0.83	0.84	0.84	0.84
ResNet101	0.74	0.74	0.75	0.75	0.74
MobileNetV2	0.79	0.76	0.77	0.77	0.77

Among the fine-tuned models, Xception and DenseNet169 exhibit the highest levels of performance across all metrics. DenseNet169 achieved a training accuracy of 0.99 and a testing accuracy of 0.91. It maintains a remarkable precision of 0.91 and a high recall of 0.90, resulting in an F1 score of 0.90, indicating its effectiveness in both accuracy and the balance between precision and recall. Xception achieved a training accuracy of 0.99, indicating an excellent fit to the training data, while maintaining a testing accuracy of 0.87. Its precision, recall, and F1 score are 0.88, showcasing a well-balanced performance. In contrast, ResNet101 and MobileNetV2, while still achieving high training accuracy, demonstrate slightly lower testing accuracies of 0.81 and 0.78, respectively. Their precision and recall values are also lower than those of Xception and DenseNet169. ResNet101 has a precision, recall, and F1 score of 0.83, while MobileNetV2 has a precision of 0.81, a recall of 0.79, and an F1 score of 0.79.

The comparison between fine-tuned models and feature extractor models on the IDC dataset clearly states the superiority of fine-tuned models in terms of the

Table 3. Classification results of fine-tuned models on balanced dataset.

Models	Train_acc	Test_acc	Precision	Recall	F1
Xception	0.99	0.87	0.88	0.88	0.88
DenseNet169	0.99	0.91	0.91	0.90	0.90
ResNet101	0.99	0.81	0.83	0.83	0.83
MobileNetV2	0.99	0.78	0.81	0.79	0.79

accuracy of classification. Fine-tuning feature extractor models allows architecture changes, especially in the top layers that provide task-specific predictions. This modification substantially improves classification accuracy and other performance metrics. Fine-tuned models integrate the large amount of information acquired from a feature extractor model with the domain-specific knowledge obtained from the target dataset. This combination increases the model's ability to generalize to new, unknown data and decreases the possibility of overfitting. Finely-tuned models have a greater number of parameters to train than feature extractor models, resulting in a modest increase in training time. As part of future work, the application of additional deep learning models to the IDC classification is planned. Furthermore, the models will be extended to address multiclass classification challenges within other histopathology datasets.

6   Conclusion

In this study, we conducted an automated IDC image classification task by utilizing transfer learning techniques. Four deep learning models, Xception, ResNet, MobileNet, and DenseNet, were employed in the IDC classification. Among these models, the fine-tuned DenseNet169 outperformed, achieving a test accuracy of 91% and an F-score of 90%. This research highlights the significance of fine-tuning deep learning models to improve the accuracy of IDC diagnosis. The future scope of this research includes a broader exploration of deep learning models in IDC classification and a focus on multiclass classification challenges in various histopathology datasets.

References

1. Invasive ductal carcinoma (IDC) histology image dataset. http://www.andrewjanowczyk.com/use-case-6-invasive-ductal-carcinoma-idc-segmentation/
2. Breast cancer facts and statistics (2023). https://www.breastcancer.org/facts-statistics
3. Abdolahi, M., Salehi, M., Showkatian, E., Reiazi, R.: Artificial intelligence in automatic classification of invasive ductal carcinoma breast cancer in digital pathology images. Med. J. Islamic Repub. Iran **34**, 140 (2020)
4. Ahmed, S., et al.: Transfer learning approach for classification of histopathology whole slide images. Sensors (Basel, Switzerland) **21**, 5361 (2021)

5. de Assis, É. G., do Patrocinio, Z.K., Nobre, C.N.: The use of convolutional neural networks in the prediction of invasive ductal carcinoma in histological images of breast cancer. Stud. Health Technol. Inform. **290**, 587-591 (2022)
6. Chollet, F.: XCeption: deep learning with depthwise separable convolutions. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (2017)
7. Choudhury, A., Perumalla, S.: Detecting breast cancer using artificial intelligence: convolutional neural network. Technol. Health Care **29**, 33–43 (2020)
8. He, K., Zhang, X., Ren, S., Sun, J.: Deep residual learning for image recognition. In: Proceedings of 2016 IEEE Conference on Computer Vision and Pattern Recognition, CVPR '16. IEEE (2016)
9. Hee, K., Cosa, A., Santhanam, N., Jannesari, M., Maros, M., Ganslandt, T.: Transfer learning for medical image classification: a literature review. BMC Med. Imaging **22**, 69 (2022)
10. Howard, A.G., et al.: MobileNets: efficient convolutional neural networks for mobile vision applications. CoRR (2017)
11. Huang, G., Liu, Z., van der Maaten, L., Weinberger, K.Q.: Densely connected convolutional networks. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (2017)
12. Janowczyk, A., Madabhushi, A.: Deep learning for digital pathology image analysis: a comprehensive tutorial with selected use cases. J. Pathol. Inf. **7**, 29 (2016)
13. Kandel, I., Castelli, M.: How deeply to fine-tune a convolutional neural network: a case study using a histopathology dataset. Appl. Sci. **10**, 3359 (2020)
14. Ikromjanov, K., Bhattacharjee, S., Hwang, Y.B., Kim, H.C., Choi, H.K.: Multiclass classification of histopathology images using fine-tuning techniques of transfer learning. J. Korea Multimedia Soc. **24**, 849–859 (2021)
15. Mormont, R., Geurts, P., Marée, R.: Comparison of deep transfer learning strategies for digital pathology. In: 2018 IEEE/CVF CVPRW (2018)
16. Pascanu, R., Mikolov, T., Bengio, Y.: On the difficulty of training recurrent neural networks. In: ICML'13, JMLR.org (2013)
17. Rashmi, R., Prasad, K., Udupa, C.: Breast histopathological image analysis using image processing techniques for diagnostic purposes: a methodological review. J. Med. Syst. **46**, 1–24 (2021)
18. Romano, A.M., Hernandez, A.A.: Enhanced deep learning approach for predicting invasive ductal carcinoma from histopathology images. In: 2019 2nd International Conference on Artificial Intelligence and Big Data (2019)
19. Wang, J., Zhu, H., Wang, S., Zhang, Y.: A review of deep learning on medical image analysis. Mob. Netw. Appl. **26**, 351–380 (2021)
20. Wang, J.L., Ibrahim, A.K., Zhuang, H., Muhamed Ali, A., Li, A.Y., Wu, A.: A study on automatic detection of idc breast cancer with convolutional neural networks. In: 2018 International Conference on Computational Science and Computational Intelligence (CSCI) (2018)
21. Yamashita, R., Nishio, M., Do, R.K.G., Togashi, K.: Convolutional neural networks: an overview and application in radiology. Insights Imaging **9**, 611–629 (2018)
22. Zhang, J., Guo, X., Wang, B., Cui, W.: Automatic detection of invasive ductal carcinoma based on the fusion of multi-scale residual convolutional neural network and SVM. IEEE Access **9**, 40308–40317 (2021)