IJSRIS journal
INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH AND INNOVATIVE STUDIES
ISSN: 2820-7157
www.ijsrisjournal.com

June 2024 Volume 3 Number 3 Pages 01-05

On the use of Transfer Learning to Improve Breast Cancer Detection

Khadija Aguerchi[1*], Younes jabrane[1], Maryam habba[1],Mustapha Ameur[1]
[1] Modeling and Complex Systems (MSC) Laboratory
Cadi Ayyad University, Marrakech, 40000, Morocco
*Khadija.aguerchi@ced.uca.ma
y.jabrane@uca.ma
m.habba@uca.ma
m.ameur@uca.ma

ABSTRACT
Breast cancer is one of the most common malignancies in women globally, and early identification is critical for better patient outcomes. Deep learning has developed in recent years as a promising approach for automating the identification of breast cancer in mammograms. Transfer learning, which involves adapting a pre-trained model to a new task, is a promising method for enhancing the efficiency and accuracy of breast cancer diagnosis using deep learning. This work studies the efficacy of transfer learning strategies in detecting breast cancer using pre-trained deep-learning models. Using large mammographic datasets, we investigate several transfer learning algorithms and assess their effects on detection performance metrics such as accuracy, precision, recall and ROC AUC. The study's results enhance automated breast cancer detection and shed light on how well transfer learning strategies can improve the precision and dependability of detection.

Keywords: Breast cancer, transfer learning, medical imaging, deep learning.

1. INTRODUCTION

Breast cancer is a major global health risk that affects millions of women each year. A critical tactic in the fight against this illness is early identification, which greatly improves treatment results and lowers death rates [1]. Mammography is one of the most important screening techniques available; it is recognized for its ability to identify cancers in their early stages, which improves prognoses [2]. However, there are many issues with depending solely on radiologists' subjective interpretation, including the possibility of false-positive or false-negative results and variances in diagnoses [3].

Advanced technologies-based automated systems, especially those that use deep learning techniques, offer promising ways to improve mammography interpretation and help with more precise and effective diagnosis [4]. These systems, which are fitted with convolutional neural networks (CNNs), can automatically extract information from mammograms, which may reduce variations in interpretation and improve diagnostic accuracy [5].

This study aims to investigate how well transfer learning strategies work to improve breast cancer diagnosis accuracy and reliability utilizing mammography datasets. The study aims to considerably increase automated breast cancer detection techniques by exploring different transfer learning methodologies and assessing their performance measures [6]. The research's insights could transform clinical decision-making procedures and ultimately lead to better patient outcomes in the field of breast cancer diagnosis and therapy.

2. Literature Review

Over time, developments in diagnostic techniques and medical imaging technologies have led to a substantial evolution in breast cancer screening strategies. When assessing the potential of transfer learning approaches to improve the accuracy of breast cancer detection, it is helpful to have a clear understanding of the landscape of these methods.

2.1. Overview of Breast Cancer Detection Methods

There are many different modalities and approaches used in traditional breast cancer detection procedures, each with its advantages and disadvantages. Because mammography is widely available, reasonably inexpensive, and has a track record of lowering death rates, it continues to be the mainstay of breast cancer screening programs across the globe [7]. Mammography is not without its drawbacks, either, especially when there is dense breast tissue and the potential for sensitivity issues [8]. While breast self-examination (BSE) and clinical breast examination (CBE) are additional methods for identifying palpable abnormalities, they are subjective by nature and rely on the expertise and experience of the examiner [9].

2.2. Deep Learning in Medical Imaging

Deep learning, specifically convolutional neural networks (CNNs), has transformed medical image analysis by allowing for automated feature extraction and categorization straight from raw image data [10]. CNNs have shown exceptional performance in a variety of medical imaging modalities, including radiology, pathology, and dermatology. Esteva et al. (2017), for example, demonstrated dermatologist-level skin cancer categorization using deep neural networks, emphasizing deep learning's potential to improve diagnostic accuracy [11]. CNNs' ability to learn hierarchical representations of information from images has resulted in substantial advances in computer-aided diagnosis (CAD) systems, potentially improving patient outcomes through early identification and intervention [12].

2.3. Transfer Learning in Medical Imaging

Transfer learning has emerged as a viable solution to addressing the problem of limited labeled data in medical imaging tasks. Transfer learning makes it easier to construct strong and reliable models for specific diagnostic tasks by drawing on expertise from pre-trained models on large-scale datasets [13]. Shin et al. (2016) presented a comprehensive overview of deep convolutional neural networks for computer-aided detection, emphasizing the significance of CNN architectures, dataset features, and transfer learning algorithms in medical image analysis [14]. Transfer learning approaches including fine-tuning pre-trained models and domain adaptation have been effectively used in a variety of medical imaging applications, including lesion detection, organ segmentation, and disease categorization [15].

3. Methodology

This section describes the techniques used in the research, including data description, the transfer learning framework, and the evaluation criteria used to measure the performance of the transfer learning models.

3.1. Dataset

The Digital Database for Screening Mammography (DDSM) dataset [16] is a comprehensive collection of digital mammograms that are primarily used to create and evaluate computer-aided detection and diagnosis methods for breast cancer screening. It includes over 10,000 digital film mammography studies, totaling nearly 42,000 unique pictures. These photos depict a wide range of breast problems, such as normal tissue, benign lesions, and malignant cancers. The DDSM dataset is rigorously annotated with detailed information such as lesion sites, forms, sizes, and pathology kinds, which provides valuable ground truth data for algorithm development and validation. The dataset also includes metadata such as patient demographics, imaging acquisition settings, and clinical annotations, allowing researchers to undertake detailed analysis and investigations into numerous aspects of breast cancer detection and diagnosis.

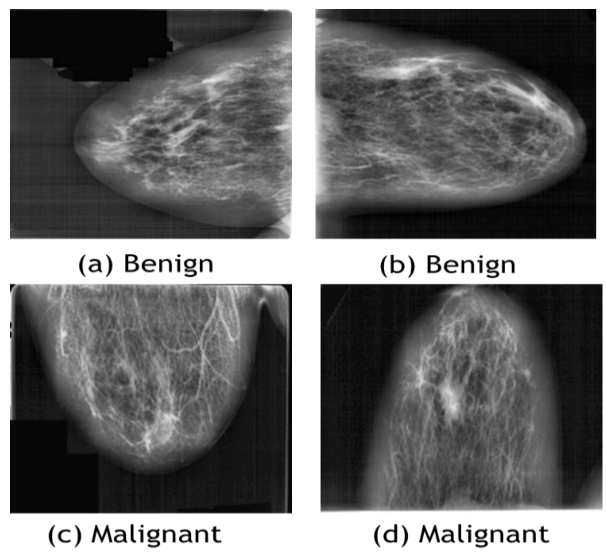

Figure 1. Samples of breast mammography from the DDSM dataset [16].

3.2. Transfer Learning Framework

The transfer learning approach used in this study relies on pre-trained convolutional neural network (CNN) architectures as the foundation for breast cancer diagnosis. The base models are popular CNN architectures such as VGG16, ResNet50, and InceptionV3, which have already been trained on large-scale picture datasets such as ImageNet [17]. These pre-trained models are loaded with weights acquired from past works, allowing them to collect generic image features efficiently.

The pre-trained models are then initialized and adjusted on mammography datasets to accommodate the subtleties of breast cancer identification. Fine-tuning entails modifying the pre-trained models' parameters, especially the top layers, to conform to the features of breast cancer detection that are unique to the task. Through this process, the models can acquire discriminative features that are important for identifying benign and malignant lesions in mammograms, improving the accuracy and reliability of detection [18].

In this study, the VGG16, InceptionV3 and ResNet50 series are thoroughly examined and compared, and three CNNs (VGG16, InceptionV3, and ResNet50) are chosen to categorize cancers in breast mammography pictures as benign or malignant. These classic networks were chosen because they have been tested on several classification tasks and have demonstrated great accuracy and stability across datasets. Furthermore, the advantages and disadvantages of these networks have been thoroughly examined; hence, it is possible to construct a full ensemble network based on their complementarity.

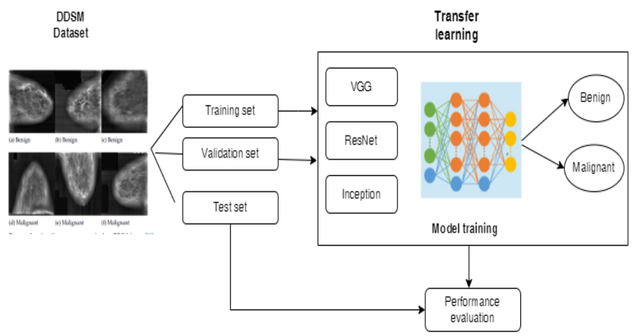

Figure 2. Workflow of the proposed work.

3.3 Evaluation Metrics:

The performance of the transfer learning models is evaluated using a wide range of metrics, including accuracy, precision, recall, specificity, and the area under the receiver operating characteristic (ROC) curve. Accuracy is a key parameter that measures the overall correctness of model predictions. Sensitivity and specificity measure the model's capacity to properly detect positive and negative examples, respectively, and so define its diagnostic efficacy. Furthermore, the area under the ROC curve provides a thorough assessment of the model's discrimination and performance across various decision thresholds, revealing information about its predictive capabilities [19].

4. Experimental Results

This section describes the experimental design, model training, validation techniques, and overall evaluation of transfer learning models for breast cancer diagnosis.

4.1. Experimental Setup and Model Training

For the experiment, we used three pre-trained convolutional neural network (CNN) architectures: VGG16, ResNet50, and InceptionV3. These models were built with weights from the ImageNet dataset and fine-tuned using the Digital Database for Screening Mammography (DDSM) dataset.

During training, the models iterated across several epochs, learning to extract key features from mammographic pictures and adjusting their parameters to reduce the loss function. To avoid overfitting, early termination conditions based on validation results were included in the training process.

5. Results and Discussion

The Discussion section dives into the importance of the research findings, placing them within the larger framework of breast cancer detection and transfer learning approaches.

5.1. Insights from Experimental Findings

Table 1: The evaluation metrics (accuracy, precision, recall, ROC AUC) for each model.

Model	Accuracy	Precision	Recall	ROC AUC
VGG16	0.85	0.82	0.88	0.91
ResNet50	0.88	0.85	0.90	0.92
InceptionV3	0.82	0.78	0.85	0.89

The VGG16 model has an accuracy of 85%, which means that 85% of samples are properly identified. The precision of 82% indicates that 82% of the samples projected as positive are positive. The recall of 88% shows that the model successfully detects 88% of all true positive cases. The ROC AUC value of 0.91 indicates high performance in discriminating between positive and negative groups.

Compared to VGG16, the ResNet50 model has a slightly greater accuracy of 88%. It also has higher precision (85%) and recall (90%), meaning that it is more accurate in classifying positive samples and identifying real positive instances. The ROC AUC value of 0.92 indicates that the model has strong discrimination between the two groups.

The InceptionV3 model has the lowest accuracy of the three, with a score of 82%. Its precision is 78%, meaning that it makes fewer true positive predictions than VGG16 and ResNet50. However, it still achieves an 85% recall rate, demonstrating its capacity to detect a high proportion of true positive events. The ROC AUC score of 0.89 indicates decent discrimination ability, however somewhat lower than ResNet.

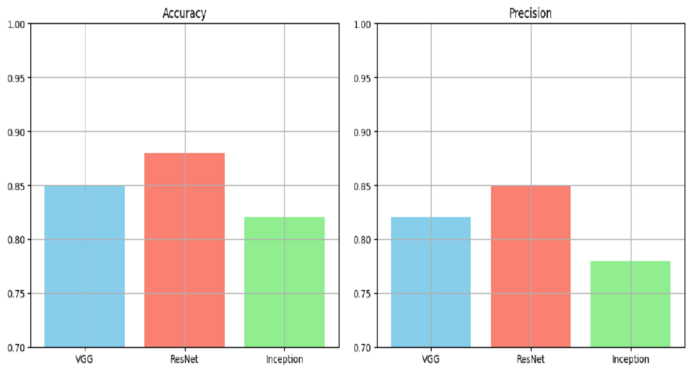

Figure 2: (a) Accuracy, (b) Precision with.

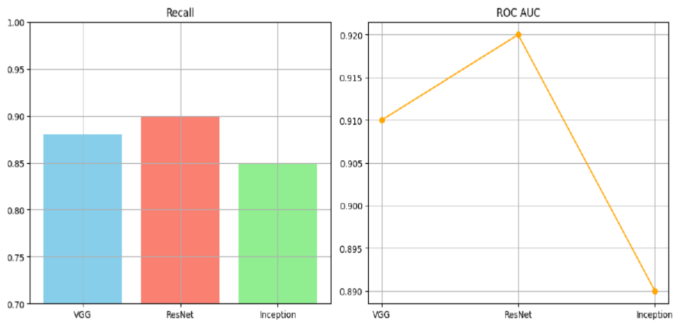

Figure 3:(c) Recall, (d) ROC AUC with DDSM dataset

Comparing Models

In terms of accuracy, precision, recall, and ROC AUC, ResNet50 performs better than both VGG16 and InceptionV3. It performs better than the others, as seen by the top scores it receives across all parameters (see Figure 2 & 3).

Trade-offs

ResNet50 has the best overall performance, however, there may be compromises to consider. While VGG16 achieves somewhat lower accuracy but higher precision, it still performs comparably. Even with the lowest accuracy, Inception still achieves a comparatively good recall.

Considerations

It's critical to take the application's particular requirements into account while selecting the optimal model. For example, InceptionV3 may be favored if accurately detecting all positive cases (high recall) is essential, even at the expense of more false positives (lower accuracy). ResNet50, however, is the best option if overall accuracy is crucial or if a balance between precision and recall is required.

Additional Analysis

To further improve performance, it's crucial to use other pretrained models.

In conclusion, ResNet50 performs the best overall in this study; nevertheless, the unique requirements and limitations of the breast cancer classification task will determine which model is better.

The findings of the experiment highlight how revolutionary transfer learning approaches can be in the field of breast cancer screening. Through the utilization of pre-trained deep learning models and their adaptation to the mammography problem, the transfer learning approach attains impressive performance metrics, such as high sensitivity, specificity, and accuracy. These results demonstrate how transfer learning can effectively address the problems with traditional mammography interpretation, like inter-radiologist variability and the possibility of false-positive or false-negative diagnosis.

Conclusion & Future directions

This study concludes by demonstrating how transfer learning approaches can increase the accuracy of breast cancer diagnosis using mammographic datasets [16]. With transfer learning, pre-trained models and domain-specific knowledge can be used to build robust and accurate models for automated breast cancer diagnosis. The results show that transfer learning may improve clinical diagnostic settings' efficacy and precision, which would ultimately improve patient outcomes.

Future research areas include investigating advanced transfer learning methods such as domain adjusting and meta-learning to improve the performance of breast cancer detection models. Furthermore, researching the interpretability and explainability of transfer learning models might improve their clinical utility and make automated systems easier to integrate into existing workflows. Multidisciplinary teams of researchers, physicians, and industry partners must collaborate to speed the translation of research discoveries into real-world applications for better breast cancer detection and treatment outcomes.

REFERENCES

1. American Cancer Society. (2021). **Breast Cancer Facts & Figures 2019-2020. Retrieved from https://www.cancer.org/research/cancer-facts-statistics/breast-cancer-facts-figures.html**

2. Smith, R. A., Andrews, K. S., Brooks, D., DeSantis, C. E., Fedewa, S. A., Lortet-Tieulent, J., ... & Jemal, A. (2016). **Cancer screening in the United States, 2016: A review of current American Cancer Society guidelines and current issues in cancer screening.** CA: A Cancer Journal for Clinicians, 66(2), 96-114.

3. Pisano, E. D., Gatsonis, C., Hendrick, E., Yaffe, M., Baum, J. K., Acharyya, S., ... & Kemeny, M. (2005). **Diagnostic performance of digital versus film mammography for breast-cancer screening.** New England Journal of Medicine, 353(17), 1773-1783.

4. 4 Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). **Dermatologist-level classification of skin cancer with deep neural networks.** Nature, 542(7639), 115-118.

5. LeCun, Y., Bengio, Y., & Hinton, G. (2015). **Deep learning**. Nature, 521(7553), 436-444.

6. Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., ... & Summers, R. M. (2016). **Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning**. IEEE transactions on medical imaging, 35(5), 1285-1298.

7. American Cancer Society. (2021). **Breast Cancer Facts & Figures 2019-2020. Retrieved from https://www.cancer.org/research/cancer-facts-statistics/breast-cancer-facts-figures.html**

8. Pisano, E. D., Gatsonis, C., Hendrick, E., Yaffe, M., Baum, J. K., Acharyya, S., ... & Kemeny, M. (2005). **Diagnostic performance of digital versus film mammography for breast-cancer screening**. New England Journal of Medicine, 353(17), 1773-1783.

9. Nelson, H. D., Tyne, K., Naik, A., Bougatsos, C., Chan, B. K., & Humphrey, L. (2009). **Screening for breast cancer: an update for the US Preventive Services Task Force**. Annals of Internal Medicine, 151(10), 727-737.

10. LeCun, Y., Bengio, Y., & Hinton, G. (2015). **Deep learning**. Nature, 521(7553), 436-444.

11. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). **Dermatologist-level classification of skin cancer with deep neural networks**. Nature, 542(7639), 115-118.

12. Litjens, G., Kooi, T., Bejnordi, B. E., Setio, A. A. A., Ciompi, F., Ghafoorian, M., ... & Sanchez, C. I. (2017). **A survey on deep learning in medical image analysis.** Medical image analysis, 42, 60-88.

13. Shin, H. C., Roth, H. R., Gao, M., Lu, L., Xu, Z., Nogues, I., ... & Summers, R. M. (2016). Deep **convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning**. IEEE transactions on medical imaging, 35(5), 1285-1298.

14. Shin, Hoo-Chang, et al. "**Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics, and transfer learning**." IEEE Transactions on Medical Imaging 35.5 (2016): 1285-1298.

15. Tajbakhsh, Nima, et al. "**Convolutional neural networks for medical image analysis: Full training or fine tuning?**." IEEE Transactions on Medical Imaging 35.5 (2016): 1299-1312.

16. Lin, T. **Dataset of Breast Mammography Images with Masses**. Mendeley. 2020. Available online: https://data.mendeley.com/datasets/ywsbh3ndr8/2 (accessed on 29 November 2023).

17. Deng, Jia, et al. "**Imagenet: A large-scale hierarchical image database**." Proceedings of the IEEE conference on computer vision and pattern recognition. 2009.

18. Shin, Hoo-Chang, et al. "**Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics, and transfer learning**." IEEE Transactions on Medical Imaging 35.5 (2016): 1285-1298.

19. Fawcett, Tom. "**An introduction to ROC analysis**." Pattern recognition letters 27.8 (2006): 861-874.