

Deep Learning for Breast Cancer Classification Using Convolutional Neural Networks

Sérgio Barreto
Centro de Informática (UFPE)
Recife, Brasil
slbp@cin.ufpe.br

Yves Emmanuel
Centro de Informática (UFPE)
Recife, Brasil
yefo@cin.ufpe.br

Abstract—Breast cancer is one of the leading causes of death among women worldwide. Early detection is crucial for successful treatment, and the use of artificial intelligence techniques, such as Deep Learning, can provide a promising approach to this task. This project focused on the development and evaluation of deep learning models for breast cancer classification using mammogram images. Three pre-trained models, VGG16, EfficientNet, and ResNet, were fine-tuned and compared in terms of their ability to distinguish between benign and malignant cases. The models were evaluated using various performance metrics, including KS test, accuracy, and AUC. The results showed that all three models performed well, with the ResNet model demonstrating the highest performance. Interpretability of the models was also considered for future work. The project highlights the potential of deep learning models for breast cancer classification and provides insights into the effectiveness of different architectures for this task.

Index Terms—Breast cancer, Convolutional Neural Networks, Deep Learning, Histopathological images, Microscopic images, Image classification.

I. INTRODUCTION

Breast cancer is one of the most common cancers among women worldwide, and early detection is critical for successful treatment. The use of medical imaging techniques, such as mammography, can aid in the diagnosis of breast cancer by detecting suspicious lesions. However, accurate interpretation of mammography images can be a challenging task in the daily routine of a medical professional. Therefore, the development of advanced computer-aided diagnosis (CAD) systems can assist radiologists in accurately identifying potential signs of breast cancer. These CAD systems use algorithms and machine learning techniques to analyze mammography images and provide a second opinion to radiologists, improving diagnostic accuracy and reducing the chances of missed diagnoses.

In recent years, Deep Learning, a subfield of AI, has shown remarkable results in various image classification tasks, including medical imaging. Convolutional Neural Networks (CNNs) are a popular Deep Learning architecture for image classification and have demonstrated promising results in several medical image analysis tasks, including breast cancer diagnosis [1].

In this work, we propose the use of CNNs to classify microscopic images of breast tumor tissue into malignant and benign categories. Specifically, we use high-quality image data of breast tumor tissue and preprocess the data to fit pre-trained

models. We trained different CNN architectures: VGG-16, EfficientNet, and ResNet.

II. THE DATASET

The Breast Cancer Histopathological Database (BreakHis) [2] is a publicly available dataset containing high-resolution microscopic biopsy images of benign and malignant breast tumors. The images were collected through a clinical study conducted between January and December 2014 at the R&D Laboratory in Brazil. All patients referred to the laboratory with a clinical indication of breast cancer were invited to participate in the study, and the institutional ethics committee approved the study, obtaining written informed consent from all participants. To protect patients' privacy, all data was anonymized before being made available for use.

The database is already divided into 5 folds. For our purpose, we used the fold number 2, which had the highest number of training samples as shown in the graph in Figure 7. Additionally, based on readings of articles that used the BreakHis dataset [2], we found that images with a magnification level of 200X resulted in the highest accuracy, and thus we used this magnification level for all images.

To mitigate the class imbalance issue, we undertook an undersampling approach since the malignant class was over-represented in the dataset. Additionally, we implemented data augmentation layers, such as random flipping, random brightness, and random rotation, across all models to enhance performance on different image formats. We preprocessed the images utilizing TensorFlow, adapting to each model's specific requirements. By taking these steps, we aimed to enhance the models' robustness and improve their overall performance.

CNN ARCHITECTURES

TensorFlow was utilized to implement different CNN architectures in this study. Specifically, the VGG-16, EfficientNet, and ResNet architectures, pre-trained with Imagenet, were explored. These architectures are widely utilized in computer vision tasks and were selected for their demonstrated effectiveness.

Also, the Adam optimizer was used due to its strong performance and rapid convergence. For the loss function, BinaryCrossentropy() was used, which is commonly utilized

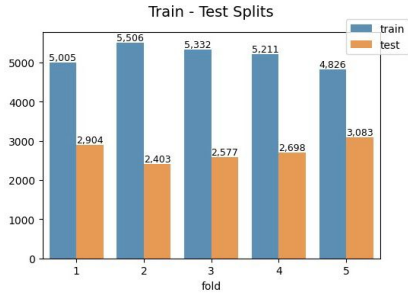


Fig. 1. Bar chart comparing the 5 data splits

in binary classification tasks. Each model was trained with a learning rate of 0.01, a batch size of 28, and for 100 epochs.

To prevent overfitting, dropout layers were included after the pre-trained model's output to facilitate knowledge sharing throughout the network during training. Dropout regularization has demonstrated efficacy in enhancing deep learning models' generalization abilities, and it proved to be especially valuable in this study. Additionally, to avoid overfitting, early stopping techniques and reduce plateau were employed while validating the model on training data. These measures were taken to ensure that the models would not overfit and would maintain high performance on unseen data.

The KS-test was the primary evaluation metric employed, with additional metrics such as ROC-AUC curve and the confusion matrix utilized to validate the model's ability to differentiate between classes. By employing a combination of these evaluation metrics, the model's performance in distinguishing between benign and malignant breast tumors was thoroughly assessed. The KS-test, in particular, was selected as the primary evaluation metric due to its robustness in detecting differences between two distributions and its effectiveness in classification tasks.

RESULTS

VGG-16 based

The VGG16-based model was chosen for its effectiveness in image recognition tasks [3] and its suitability for our dataset. The VGG16 architecture, proposed by Simonyan and Zisserman in 2014, is composed of 13 convolutional layers and 3 fully connected layers, resulting in a total of 138 million parameters. The VGG16 architecture is known for its strong performance in large-scale image classification tasks, and it has been utilized in various computer vision applications. For instance, it has been used for object detection, image segmentation, and face recognition, among other tasks.

The model's large number of parameters and the architecture's depth enabled it to learn complex image features effectively. Additionally, transfer learning was employed by utilizing the pre-trained weights of the VGG16 model, fine-tuning the last few layers for our dataset, resulting in improved accuracy. Through extensive experimentation and validation, the VGG16-based model proved to be a robust and effective approach for breast cancer classification in our study.

The model's performance was relatively good, as indicated by the KS test value of 0.56 on the test dataset. The AUC of 0.84 also suggests that the model has strong discriminative power. Additionally, upon examining the confusion matrix, it can be observed that the model's errors were within acceptable limits for a good performance. The overall accuracy of the model was 0.80.

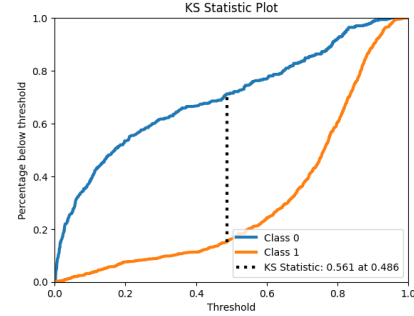


Fig. 2. KS-test of VGG-16 based

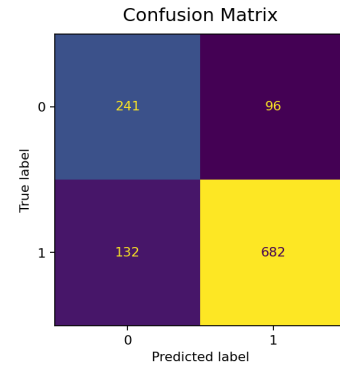


Fig. 3. Confusion Matrix of VGG-16 based

EfficientNet based

The EfficientNet architecture was chosen for its superior performance in image recognition tasks and its efficiency in terms of the number of parameters required [4]. The EfficientNet architecture was proposed by Tan et al. in 2019 and is known for achieving state-of-the-art performance in various computer vision tasks while requiring fewer parameters than other architectures. The EfficientNet architecture achieves this efficiency by using a compound scaling method that simultaneously scales up the depth, width, and resolution of the model. This approach results in a highly efficient architecture that performs well on a wide range of image recognition tasks.

In our study, we used the EfficientNet-B0 architecture, which has 5.3 million parameters and achieved state-of-the-art performance on the ImageNet dataset. Transfer learning was also employed by utilizing the pre-trained weights of the EfficientNet-B0 model, fine-tuning the last few layers for our dataset, resulting in further improved accuracy. Through

extensive experimentation and validation, the EfficientNet-based model proved to be an effective approach for breast cancer classification while requiring fewer parameters than other architectures.

The EfficientNet model also exhibited strong performance in breast cancer classification, with a KS value of 0.67 on the test dataset, an accuracy of 0.83, and an AUC of 0.9. These metrics suggest that the EfficientNet model was able to effectively distinguish between malignant and benign cases in our study.

When comparing the performance of the VGG16-based model and the EfficientNet-based model, the EfficientNet performed slightly better in terms of KS, AUC and accuracy.

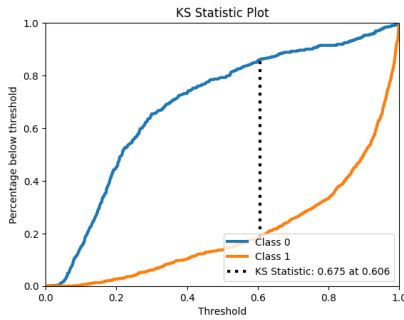


Fig. 4. KS-test of EfficientNet based

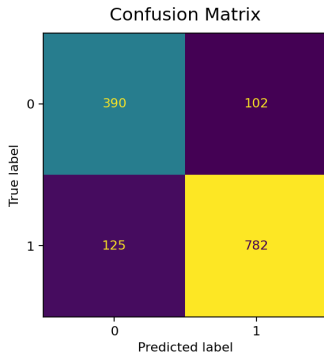


Fig. 5. Confusion Matrix of EfficientNet based

ResNet based

The ResNet model was chosen due to its well-established reputation for high performance in image recognition tasks. ResNet [5] is a deep residual neural network architecture that introduced a novel residual learning framework to address the issue of vanishing gradients in very deep neural networks. This architecture's ability to efficiently train very deep neural networks has resulted in its successful application in various image recognition tasks, including object detection, image segmentation, and classification.

The ResNet model also demonstrated strong performance in breast cancer classification, with a KS test value of 0.69, an accuracy of 0.83, and an AUC of 0.92. These results suggest

that the ResNet model was able to effectively distinguish between benign and malignant cases.

It performed slightly better than the other two models, with the highest KS test value, accuracy, and AUC. The VGG16 and EfficientNet models also demonstrated strong performance, with similar AUC and accuracy values, although their KS test values were slightly lower than the ResNet model. Overall, these results suggest that all three models are effective approaches for breast cancer classification, with the ResNet model demonstrating the highest performance among the three.

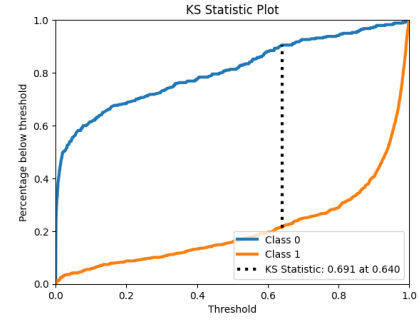


Fig. 6. KS-test of ResNet based

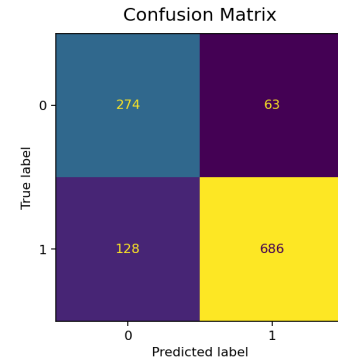


Fig. 7. Confusion Matrix of ResNet based

OVERALL METRICS

	VGG-16	EfficientNet	ResNet
Loss	0,4981	0,3879	0,4325
Accuracy	0,8019	0,8377	0,8340
KS-test	0,5610	0,6750	0,6910
ROC-AUC	0,8369	0,8990	0,9147

FUTURE DIRECTIONS

One crucial aspect of evaluating a model is its interpretability, which can aid in understanding the model's decision-making process. Unfortunately, in this study, interpretability was not taken into account until later stages, and it could not be implemented. However, the last pooling layer of the model can be utilized to extract the regions on the image that the model used to differentiate between the malignant and benign classes. Medical interpretation suggests that a malignant tumor

is identified when metastasis occurs, indicating that the tumor is spreading rapidly throughout the cells.

For future works, it is highly recommended to attempt to identify the regions of metastasis, providing important context for medical professionals to be aware of. Furthermore, stacking models can potentially improve the overall performance, and it is recommended to explore this approach in further studies.

REFERENCES

- [1] Spanhol, F. A., Oliveira, L. S., Petitjean, C., Heutte, L. (2016). "Breast cancer histopathological image classification using convolutional neural networks." In International Joint Conference on Neural Networks (IJCNN) (pp. 2560-2567). IEEE. doi: 10.1109/IJCNN.2016.7727519. Available at: <https://ieeexplore.ieee.org/document/7727519>.
- [2] Spanhol, F., Oliveira, L. S., Petitjean, C., Heutte, L. (2016). "A dataset for breast cancer histopathological image classification." IEEE Transactions on Biomedical Engineering.
- [3] Karen Simonyan and Andrew Zisserman, "Very Deep Convolutional Networks for Large-Scale Image Recognition," arXiv preprint arXiv:1409.1556, 2015, cs.CV. Available at: <https://arxiv.org/abs/1409.1556>.
- [4] Mingxing Tan and Quoc V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," CoRR, vol. abs/1905.11946, 2019. Available at: <http://arxiv.org/abs/1905.11946>.
- [5] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, "Identity Mappings in Deep Residual Networks," CoRR, vol. abs/1603.05027, 2016. Available at: <http://arxiv.org/abs/1603.05027>.