

A CAD-Like Tool Using Convolutional Neural Networks to Improve Radiologist Screening of Breast Cancer

Abdelhafidh KACHER, *Mathematics and Informatics research Lab. (LIM), University of Tiaret, Algeria*, abdelhafidh.kacher@univ-tiaret.dz

Keltoum LARBI, *University of Tiaret, Algeria*, keltoum.larbi@univ-tiaret.dz

Mourad HORRI, *Intelligence Artificielle et Systèmes Lab., University of Tiaret*, mourad.horri@univ-tiaret.dz

Mohamed HOCEINI, *Intelligence Artificielle et Systèmes Lab., University of Tiaret*, mohammed.hoceini@univ-tiaret.dz

Ibrahim SI LARBI, *Intelligence Artificielle et Systèmes Lab., University of Tiaret*, ibrahim.silarbi@univ-tiaret.dz

Abstract—Breast cancer remains a significant global health challenge, with early detection being the primary factor in reducing mortality. While mammography is the standard screening tool, the risk of human error in interpretation persists. This paper presents a Computer-Aided Detection (CAD) system based on a Convolutional Neural Network (CNN) to assist radiologists in identifying abnormal mammograms. Using a dataset of 2,000 mammographic images from DDSM dataset, the model was trained over 50 epochs. The experimental results obtained during a practical workshop demonstrate test accuracies of 90%, 83%, 95%, 91%, and 82% across multiple CNN architectures.

Index Terms—Breast Cancer, Mammography, Convolutional Neural Networks, Medical Imaging, Transfer Learning, Diagnostic Accuracy.

I. INTRODUCTION

Breast cancer is the most prevalent cancer among women worldwide, and early diagnosis is essential for effective treatment and improved survival rates. Mammography is an efficient screening technique to detect the breast cancer at its early stages, allowing doctors to provide timely treatment and surgical intervention, then reducing mortality rates. However, the growing of breast cancer cases produces increasing volume of mammographic examinations, placing excessive workload on radiologists. Additionally, interpreting mammograms manually is known to be challenging and time consuming due to the nature of cancerous lesions and density variation of breast tissue [1].

A CAD system in mammography is designed to assist radiologists by automatically identifying regions of interest (ROIs) that may correspond to masses, calcifications, or other suspicious lesions [2]. Unlike traditional systems, modern CAD tools increasingly rely on deep learning (DL) models,

which are capable of learning complex and non-linear representations directly from image data. Importantly, such systems are not intended to replace radiologists but rather to complement their expertise and improve diagnostic efficiency.

Mammographic images present several challenges for analysis. Images often exhibit low contrast, especially in dense breast tissue making lesion detection difficult [3]. Lesions may be small, subtle and visually similar to normal breast tissue. Additionally, acquisition variability, noise, and the presence of labels or artifacts further complicate processing. These challenges motivate the need for robust preprocessing techniques and powerful feature extraction models such as CNNs.

This paper details the development and obtained result of a CNN-based models designed to classify mammograms into Normal and Abnormal. The CNN architectures were implemented during the workshop entitled “CAD-Like Tool with a CNN Architecture to Improve Radiologist Screening of Breast Cancer” organized within the second edition of the scientific study days JIMC 2025.

This workshop aimed to provide participants with both theoretical foundations and practical experience in designing and implementing CNN models for mammography image classification and focused on the complete pipeline from dataset preparation and image preprocessing to model training, evaluation, and performance improvement using data augmentation and transfer learning (TL).

The results obtained by the participants were promising, highlighting the potential of CNN-based CAD systems to improve breast cancer screening accuracy and to support radiologists in the early detection of abnormalities.

II. OBJECTIVES OF THE WORKSHOP

The primary objective of this workshop was to introduce participants to the practical use of machine learning (ML) and DL techniques for medical image analysis, with a particular focus on mammography and CNNs. More specifically, the

workshop aimed to:

- Explaining the principle of CAD systems and breast cancer screening.
- Providing a deep understanding of CNNs and their suitability for image-based applications.
- Demonstrating the challenges in mammographic imaging such as low contrast, dense breast tissue and subtle lesion appearance.
- Assisting the participants through the construction of a complete CNN model for breast cancer classification using Python, TensorFlow, Keras and Google Colab covering all stages from data loading and preprocessing to model training, evaluation, and performance optimization.

By the end of the workshop, participants are expected to be capable of independently implementing and evaluating basic CNNs for binary classification of mammography images.

III. DATASET

The workshop introduced participants to the use of public mammography datasets, highlighting the importance of well-organized data splits for training, validation, and testing.

In this workshop we used a total of 2000 images from DDSM dataset [4]. The images were structured into separate directories corresponding to the training, validation and testing splits, with subfolders representing the class labels:

- Training Set: 1,400 images.
- Validation Set: 400 images.
- Test Set: 200 images.

This data preparation enabled the direct use of Keras data generators (ImageDataGenerator with a batch size of 32) for efficient data loading, normalization and augmentation.

IV. PREPROCESSING

Image preprocessing is a critical step to ensure consistent input to the CNN models and to improve the performance. The preprocessing pipeline implemented during the workshop included:

- Removal of irrelevant regions and labels where applicable.
- Resizing all images to a fixed resolution of 224×224 pixels, consistent with common CNN architectures.
- Intensity normalization through rescaling pixel values to the $[0, 1]$ range.
- Contrast enhancement using techniques such as CLAHE, discussed conceptually to address low-contrast mammographic images.

These steps aimed to reduce variability, enhance diagnostically relevant structures, and decrease computational complexity.

V. MODEL ARCHITECTURE

A CNN is a class of DL models designed for image analysis. CNNs extract spatial features by applying convolutional filters that slide over the image. Early convolutional layers typically learn simple features such as edges and blobs, while deeper layers capture more complex patterns, including textures and shapes.

Key CNN components discussed in the workshop included convolutional layers, activation functions (ReLU and Sigmoid), pooling layers, and fully connected layers and others. The rectified linear unit (ReLU) activation function was presented as the most commonly used activation function in hidden layers, while sigmoid and softmax activations were discussed for binary and multi-class classification, respectively.

Initially, the participants implemented a sequential CNN model composed of three convolutional blocks, each followed by Maxpooling layer. The convolutional layers used 3×3 kernels with an increasing number of filters (32, 64, and 128), enabling progressive feature abstraction. The extracted feature maps were then flattened and passed through fully connected layers and dropout layer to mitigate overfitting. The final output layer employed a sigmoid activation function for binary classification. The model was trained using the Adam optimizer (with a learning rate of 0.001) and binary cross-entropy loss function, reflecting standard practice in binary medical image classification tasks.

To further enhance performance, the workshop introduced transfer learning using pre-trained CNN architectures such as VGG [5], ResNet [6], and EfficientNet [7]. These models, pre-trained on the ImageNet dataset, provide rich feature representations that can be adapted to medical imaging tasks.

In the demonstrated approach, the convolutional base of the pre-trained model was frozen, and custom fully connected layers were added for mammography classification. This strategy significantly reduced training time and improved performance, particularly when the available dataset was relatively small. Participants were encouraged to experiment with different base models and fine-tuning strategies as future work.

VI. TRAINING STRATEGY AND EVALUATION

A. Data Augmentation

To address limited dataset size and reduce overfitting, data augmentation techniques were implemented into the training set. These techniques included random rotations, horizontal flips, and minor spatial transformations. Data augmentation effectively increased the diversity of the training data and improved the model's generalization capability.

B. Model Training

Model training was conducted in a Google Colab environment using a Tesla T4 GPU. The network was compiled with the Adam optimizer and Binary Cross-Entropy loss function. Training was performed for 50 epochs with a batch size of 32.

The training process was conducted using Google Colab, allowing participants to execute experiments without local hardware constraints or library dependencies. Model checkpoints were employed to save the best-performing weights based on validation accuracy. Training and validation curves for loss and accuracy were visualized to monitor convergence behavior and detect potential overfitting.

C. Evaluation Metrics

Model performance was assessed on a held-out test set using multiple metrics relevant to medical diagnosis. In addition to accuracy, precision, recall, specificity, and F1-score were computed. Confusion matrices were analyzed to better understand the distribution of true positives, false positives, true negatives, and false negatives. This comprehensive evaluation emphasized the importance of sensitivity and specificity in medical applications.

VII. RESULTS

The workshop adopted a hybrid pedagogical approach combining theoretical lectures with hands-on coding sessions. Participants actively followed code demonstrations, modified parameters, and observed their impact on model performance. Interactive discussions and a dedicated question-and-answer session facilitated deeper understanding and knowledge exchange.

This approach proved effective in bridging the gap between theory and practice, particularly for participants with limited prior experience in CNNs.

The results obtained during the workshop demonstrated the feasibility and effectiveness of CNN-based CAD-like tools for mammography classification. The best three results obtained by the participants are detailed in Tables I, II, III, IV and V.

The model proposed by Team-A achieved an overall accuracy of 95%, indicating strong classification capability. The high recall value of 0.97 for abnormal cases demonstrates the model's effectiveness in detecting potentially cancerous cases, which is critical in medical screening applications. Precision values of 0.93 for abnormal and 0.97 for normal cases show that the model maintains a low false-positive rate. The equal F1-score of 0.95 for both classes reflects balanced performance. Additionally, the macro and weighted averages confirm that the model does not favor one class over the other, ensuring reliable and unbiased predictions.

Team-b achieved an overall accuracy of 83% using sequential CNN model and 90% using MobileNet pre-trained CNN model, indicating the capabilities of using TL.

Finally, Team-C and Team-D achieved overall accuracies of 91% and 82% respectively.

From a research perspective, the workshop highlighted the potential of DL to support radiologists and improve breast cancer screening outcomes when carefully designed and validated.

TABLE I
CLASSIFICATION REPORT (TEAM-A)

Class	Precision	Recall	F1-Score
Abnormal	0.93	0.97	0.95
Normal	0.97	0.93	0.95
Accuracy	-	-	0.95
Macro Avg	0.95	0.95	0.95
Weighted Avg	0.95	0.95	0.95

TABLE II
CLASSIFICATION REPORT (TEAM-B) USING SEQUENTIAL CNN

Class	Precision	Recall	F1-Score
Abnormal	0.85	0.82	0.83
Normal	0.83	0.85	0.84
Accuracy	-	-	0.83
Macro Avg	0.84	0.83	0.83
Weighted Avg	0.84	0.83	0.83

TABLE III
CLASSIFICATION REPORT (TEAM-B) USING MOBILENET

Class	Precision	Recall	F1-Score
Abnormal	0.88	0.92	0.90
Normal	0.92	0.88	0.90
Accuracy	-	-	0.90
Macro Avg	0.90	0.90	0.90
Weighted Avg	0.90	0.90	0.90

TABLE IV
CLASSIFICATION REPORT (TEAM-C)

Class	Precision	Recall	F1-Score
Abnormal	0.89	0.92	0.91
Normal	0.92	0.89	0.90
Accuracy	-	-	0.91
Macro Avg	0.91	0.91	0.90
Weighted Avg	0.91	0.91	0.90

TABLE V
CLASSIFICATION REPORT (TEAM-D)

Class	Precision	Recall	F1-Score
Abnormal	0.81	0.85	0.83
Normal	0.84	0.80	0.82
Accuracy	-	-	0.82
Macro Avg	0.83	0.82	0.82
Weighted Avg	0.83	0.82	0.82

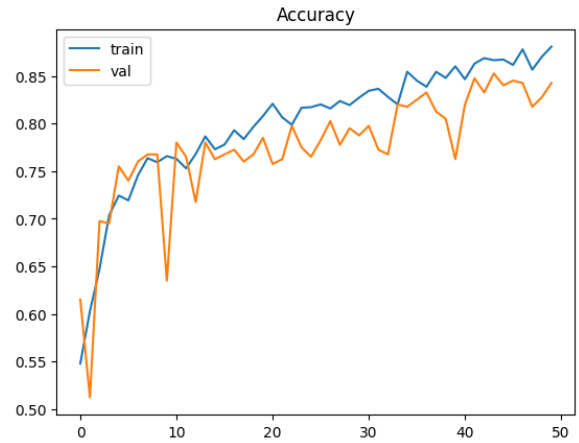


Fig. 2. Training accuracy using sequential CNN (Team-B)

TABLE VI
HYPERPARAMETERS

Hyperparameter	Value
Batch size	32
Loss function	Binary cross-entropy
optimizer	Adam
Learning rate	0.01-0.0001
Epochs	50

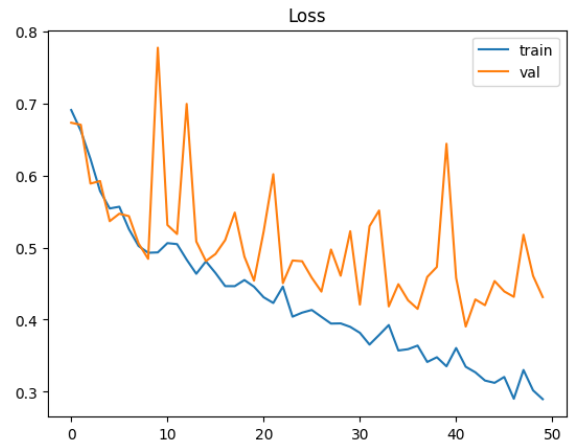


Fig. 3. Training loss using sequential CNN (Team-B)

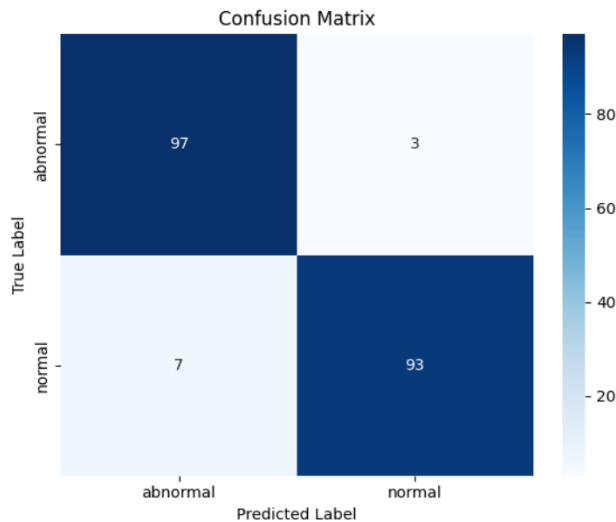


Fig. 1. Confusion matrix (Team-A)

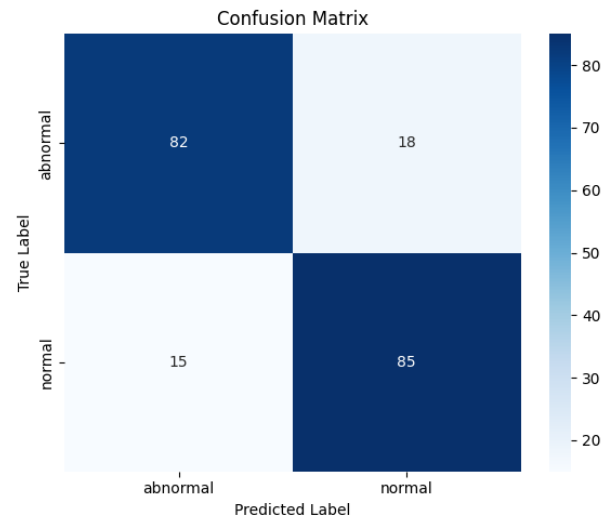


Fig. 4. Confusion matrix using sequential CNN (Team-B)

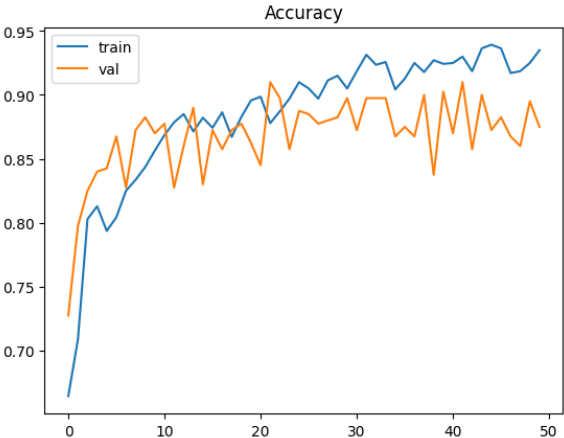


Fig. 5. Training accuracy using MobileNet (Team-B)

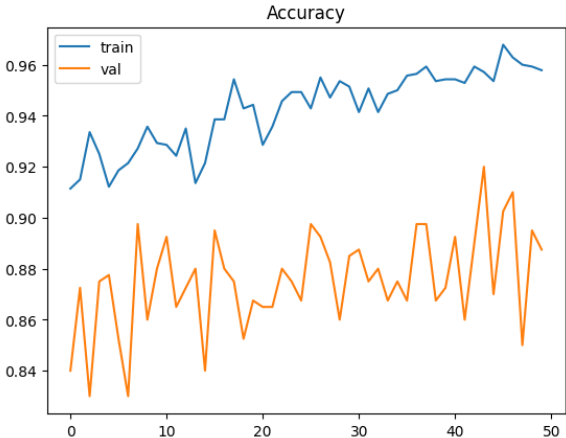


Fig. 8. Training accuracy (Team-C)

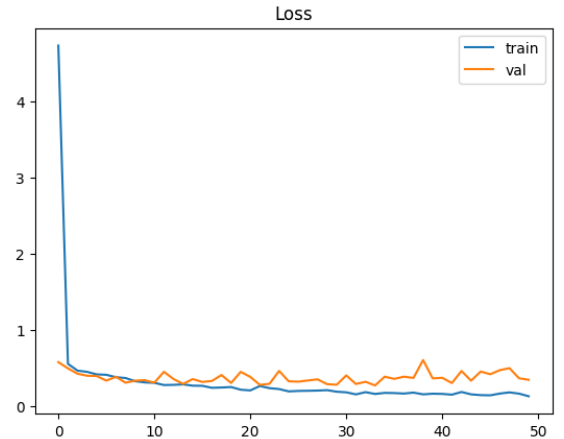


Fig. 6. Training loss using MobileNet (Team-B)

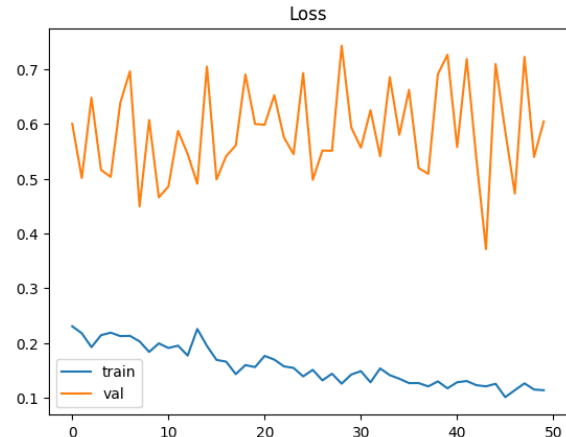


Fig. 9. Training loss (Team-C)

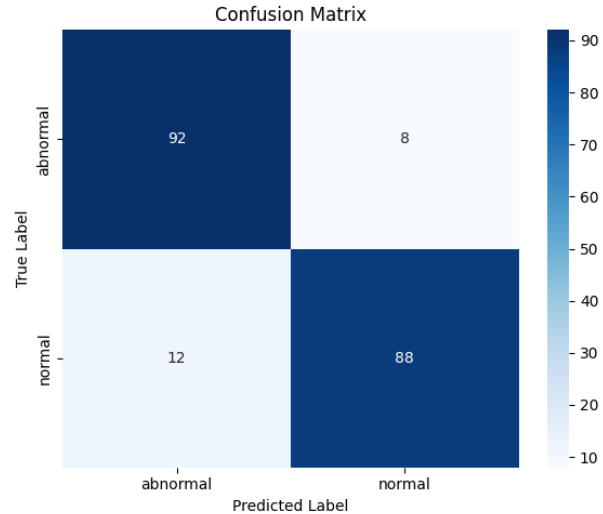


Fig. 7. Confusion matrix using MobileNet (Team-B)

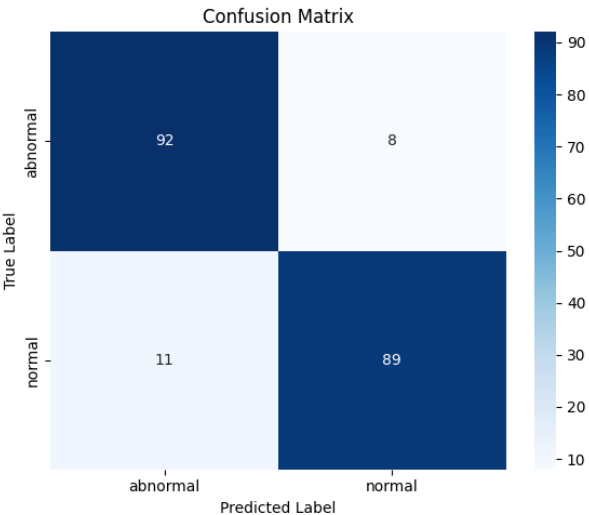


Fig. 10. Confusion matrix (Team-C)

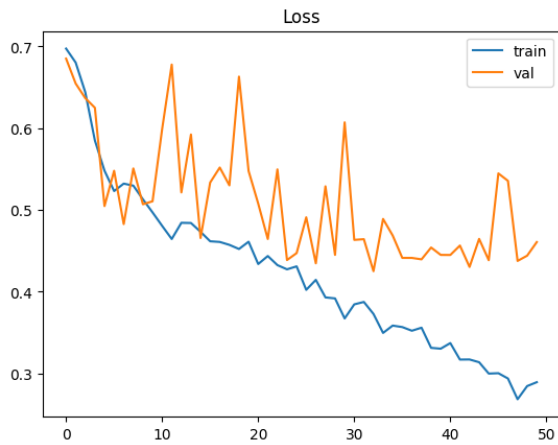


Fig. 11. Training accuracy (Team-D)

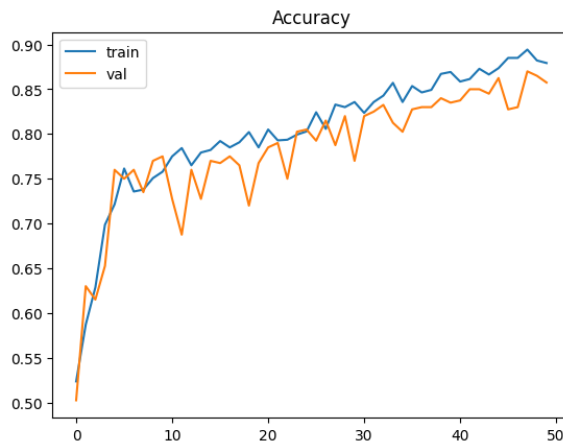


Fig. 12. Training loss (Team-D)

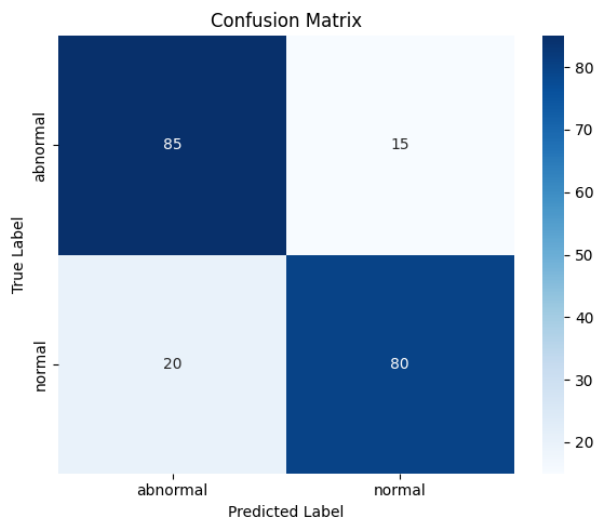


Fig. 13. Confusion matrix (Team-D)

VIII. CONCLUSION

This paper provided a comprehensive and applied introduction to the development of CAD-like systems using convolutional neural networks for mammography image classification. By covering the full pipeline from preprocessing and model design to evaluation and transfer learning. The positive engagement and achieved results underscore the relevance of such training initiatives in advancing the adoption of artificial intelligence in medical imaging.

REFERENCES

- [1] M. L. Giger, "Machine Learning in Medical Imaging," *Journal of the American College of Radiology*, vol. 15, no. 3, 2018.
- [2] R. M. Masoud, R. M. A. Bakir, M. S. Saraya, and S. M. Ayyad, "BREAST-CAD: A Computer-Aided Diagnosis System for Breast Cancer Detection Using Machine Learning," *Technologies*, vol. 13, no. 7, Art. no. 268, 2025, doi: 10.3390/technologies13070268.
- [3] L. Wang, "Mammography with deep learning for breast cancer detection," *Frontiers in Oncology*, vol. 14, 2024, doi: 10.3389/fonc.2024.1281922.
- [4] M. Heath, K. Bowyer, D. Kopans, R. Moore, and P. W. Kegelmeyer, "The digital database for screening mammography," in *Proc. 5th Int. Workshop Digital Mammography (IWDM 2000)*, M. J. Yaffe, Ed. Madison, WI: Medical Physics Publishing, 2001, pp. 212–218, doi: 10.2214/ajr.179.1.1790184.
- [5] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. 3rd Int. Conf. Learning Representations (ICLR)*, San Diego, CA, USA, Sept. 2015, doi: 10.48550/arXiv.1409.1556.
- [6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Computer Vision and Pattern Recognition (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778, doi: 10.1109/CVPR.2016.90.
- [7] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in *Proc. 36th Int. Conf. Machine Learning (ICML)*, Long Beach, CA, USA, May 2019, pp. 6105–6114, doi: 10.48550/arXiv.1905.11946.