

Hybrid Deep Learning Model for Breast Cancer Classification in Low-Middle Income Countries: A MobileNetV2 and Cubic SVM

Kennedy T. Chitiza, Prosper Mbire, Valentine E. Gora, Felix Mazunga, Kudakwashe P. Dzingirai, Shamiso Mabota

Department of Applied Physics and Telecommunications, Midlands State University, Gweru, Zimbabwe

Email: kennychitiza@gmail.com

How to cite this paper: Chitiza, K.T., Mbire, P., Gora, V.E., Mazunga, F., Dzingirai, K.P. and Mabota, S. (2026) Hybrid Deep Learning Model for Breast Cancer Classification in Low-Middle Income Countries: A MobileNetV2 and Cubic SVM. *Advances in Breast Cancer Research*, 15, 55-65.

<https://doi.org/10.4236/abcr.2026.152005>

Received: January 18, 2026

Accepted: March 14, 2026

Published: March 17, 2026

Copyright © 2026 by author(s) and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

Abstract

Breast cancer is one of the major causes of female death in any part of the world, and the burden is highly disproportionately high in low-middle income countries (LMICs) because of the inaccessibility of diagnostic tools. The proposed paper suggests a lightweight hybrid deep learning model that integrates MobileNetV2 to extract features with a cubic Support Vector machine (SVM) to classify the features, specifically to be used in resource constrained settings. The model was trained and tested on the mini-DDSM dataset and obtained 78 percent overall accuracy and 70 percent recall on malignant cases and 98 percent recall on the normal cases. The computational analysis shows that the model is efficient and less than 20 ms per image inference on standard CPU hardware and a small footprint of 14 MB, which is reasonable to implement in regions with limited computational resources like LMICs. The hybrid architecture is especially strong in terms of false positive minimization (84 percent precision in all normal cases) and clinically relevant sensitivity in detecting malignancy. An interface built with Streamlit was created to showcase the practical use in low resource clinical environments.

Keywords

Breast Cancer, Deep Learning, MobileNetV2, Support Vector Machines, Low-Middle Income Countries, Medical Image Classification, Convolutional Neural Networks

1. Introduction

BREAST cancer is a significant health issue in the world, and it is the most fre-

quently diagnosed cancer in women across the world. The data provided by the World Health Organization shows that, in 2022, more than 670,000 breast cancer deaths took place, and 310,720 new cases are expected in 2024 [1]. Early diagnosis has a huge positive impact on the effectiveness of treatment and the survival rate, but the traditional diagnostic tools have certain limitations that are especially acute in the conditions with limited resources. Conventional breast cancer screening techniques mainly use imaging techniques, such as digital mammography, ultrasound and magnetic resonance imaging (MRI). Although mammography is regarded as the gold standard, it has several drawbacks, such as radiation exposure, physical discomfort, and high false positive and false negative rates [2]. Ultrasound is an additional modality that is operator dependent and less efficient in detecting micro-calcium [3]. MRI is highly sensitive but lacks specificity, is quite expensive, and takes time to perform the procedure [4]. With the emergence of artificial intelligence, especially the deep learning methods, medical image analysis has undergone a revolution. Convolutional Neural Networks (CNNs) have proven to have exceptional automated feature detection and pattern recognition in medical images [5]. According to recent research, deep learning models can perform a particular diagnostic task as well as or better than human radiologists do it [6]. Nevertheless, the majority of the state-of-the-art models require a significant number of computational capabilities and big, labeled datasets, which present a deployment challenge in the LMICs. The study provides answers to the urgent need to have accessible breast cancer diagnostics in resource restricted settings in three main ways, it: 1) builds a hybrid MobileNetV2-cubic SVM structure that is optimized to use limited computational resources; 2) evaluates and validates the proposed structure on the mini-DDSM dataset with a specific focus on clinical applicability metrics; and 3) provides an analysis of deployment frameworks in the context of LMIC implementation. The offered model provides a trade-off between diagnostic accuracy and computational efficiency with an overall accuracy of 78 percent and the malignant case recall of 70 percent with a small 14 MB footprint that can be utilized in a mobile or an edge computing device.

2. Related Work

2.1. Deep Learning in Medical Imaging

Deep learning has emerged as a transformative approach in medical image analysis, with CNN architecture demonstrating efficacy in breast cancer detection [7]. Early CNN implementations like LeNet and AlexNet established foundational architectures, while subsequent developments including VGGNet, GoogleNet, and ResNet introduced enhanced depth and feature extraction capabilities [8]. However, these architectures often require substantial computational resources, limiting their applicability in resource-constrained environments.

2.2. Lightweight Architectures for Medical Applications

MobileNet architectures, introduced by Howard *et al.*, revolutionized mobile and

embedded vision applications through depth wise separable convolutions [9]. EfficientNet further advanced model efficiency through compound scaling of network depth, width, and resolution [10]. Both architectures have been adapted for medical imaging tasks, with studies demonstrating their potential for breast cancer classification in limited-resource settings [11]. Recent investigations into hybrid approaches combining deep feature extraction with traditional machine learning classifiers have shown promise. Shen *et al.* demonstrated improved classification performance by integrating CNN features with SVM classifiers [12], while Saber *et al.* achieved 98.96% accuracy on MIAS dataset using transfer learning approaches [13]. However, these studies primarily focused on high-resource environments or utilized computationally intensive architectures.

2.3. Bridging the Resource Gap in LMICs

The disparity in healthcare resources between high-income countries and LMICs necessitates specialized approaches to medical AI deployment [14]. Computational constraints, limited internet connectivity, and infrastructure limitations require models that are both accurate and efficient. Previous work has explored mobile health applications and telemedicine solutions, but few have specifically addressed the intersection of computational efficiency and diagnostic accuracy for breast cancer detection [15].

3. Methodology

3.1. Dataset Description and Preprocessing

The models were developed and tested with the help of the mini-DDSM (Digital Database for Screening Mammography) dataset. It is a publicly available repository of 9682 mammography images of three classes normal, benign, and malignant. To solve the issue of class imbalance in the original data, a balanced training subset was designed by randomly selecting 2700 images in each class, which gave a training set of 8100 images. This balanced subset was the only one that the model was trained and validated on using an 80/10 split (6480 training, 1620 validation).

To assess this, a test set was first made by withholding 10 percent of the original and imbalanced dataset, stratified by class, leaving 968 images. The images were also preprocessed following the same pipeline as the training data; resizing to 224×224 pixels, pixel normalization to $[0, 1]$, and three channels conversion via duplication.

Nevertheless, 361 images of this original test pool were eliminated during the preprocessing phase on one of the following grounds: corruption of original image file, channel conversion process failed, or the metadata was not compatible and could not be loaded. The performance evaluation used valid and high-quality images and after these quality control filters, the final test set consisted of 607 images (200 normal, 203 benign, 204 malignant), where only valid images were used. The final test set was distributed approximately equally in classes so that the per-class metric can be calculated in a fair way. **Table 1** demonstrates detailed characteristics of datasets.

Table 1. Mini-DDSM dataset characteristics.

Feature	Description
Age Range	27 - 91 years
Total Images	9682
Original Resolution	500 × 500 pixels
Processed Resolution	224 × 224 pixels
Image Format	PNG (converted to RGB)
Class Distribution	Normal (2728), Benign (3360), Malignant (3596)
Balanced Subset	2700 images per class

3.2. Hybrid Model Architecture

The proposed hybrid architecture integrates MobileNetV2 for feature extraction with a cubic SVM for classification, as illustrated in **Figure 1**. This design leverages MobileNetV2’s efficiency in feature representation while utilizing SVM’s robustness in high-dimensional classification.

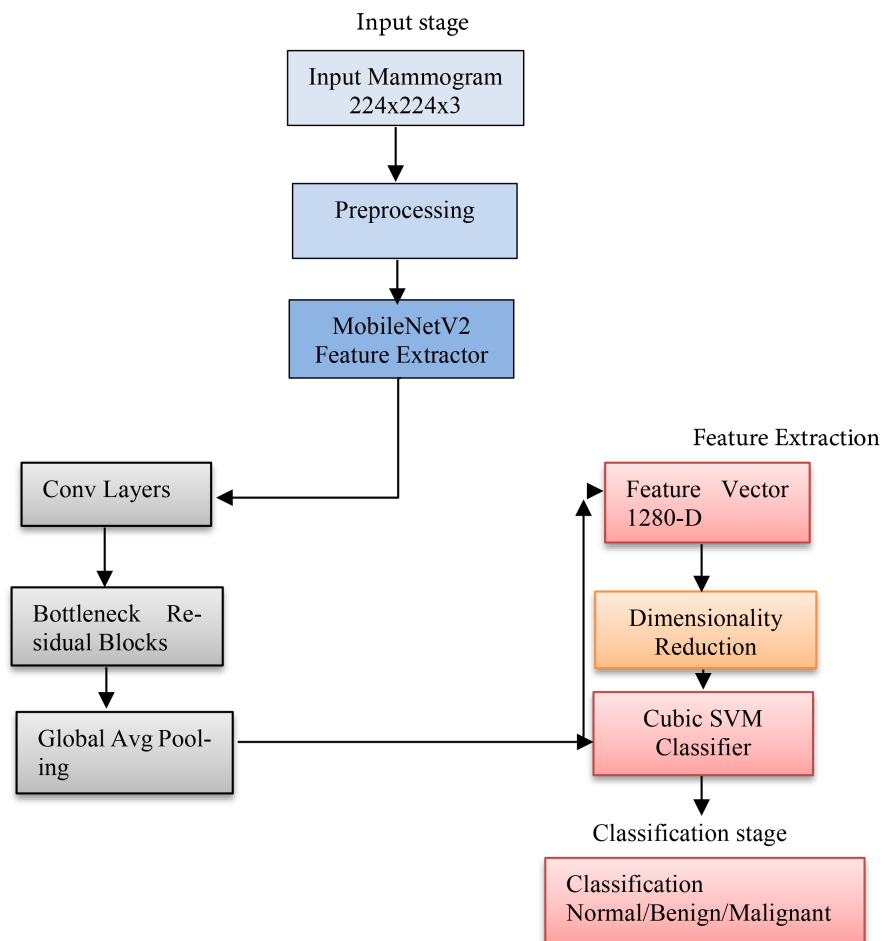


Figure 1. Proposed hybrid MobileNetV2-cubic SVM architecture for breast cancer classification.

3.2.1. MobileNetV2 Feature Extractor

MobileNetV2 uses inverted residual blocks that include linear bottlenecks to ensure the representational power and reduce the cost of computation. Architecture employs depthwise separable convolutions which break down standard convolutions into depthwise and pointwise operations, which drastically reduce the amount of parameters and the computation cost. MobileNetV2 was fine-tuned on ImageNet, after which the top 100 layers are frozen to maintain learned features without changing the pattern of the mammography, but adapt to the specific patterns.

3.2.2. Cubic SVM Classifier

The cubic SVM implements a third-degree polynomial kernel function:

$$(x_i, x_j) = (\gamma x_i^T x_j + r)^d$$

where γ , r , and $d=3$ are kernel parameters. This nonlinear kernel enables effective classification of complex, high-dimensional feature representations extracted by MobileNetV2. MobileNetV2 global average pooling layer feature vectors (1280-dimensional) were diminished using principal component analysis (PCA) prior to SVM training as a remedy to the curse of dimensionality. We retained 200 principal components, which explained about 95 percent of cumulative explained variance of the training data. This dimensionality reduction trade-off allowed retaining features of the model and being computationally efficient.

3.3. Training Configuration

Model training employed categorical cross-entropy loss with Adam optimizer (learning rate = 3×10^{-5}). Early stopping with 15-epoch patience prevented overfitting, while class-weighted loss compensated for initial dataset imbalance. The dataset was partitioned into 80% training, 10% validation, and 10% test sets. Training progressed through 60 epochs with batch size of 32.

3.4. Performance Metrics

Model assessment was done based on overall measures such as accuracy, precision, recall, F1-score, and area under the receiver operating characteristic curve (AUC-ROC). Specific class measures were prioritized to measure clinical utility especially sensitivity on malignant cases and specificity on normal cases. Inference time, model size and memory requirements were all computational metrics.

4. Results and Analysis

4.1. Training Dynamics

Training progression demonstrated effective learning with validation accuracy stabilizing at 77.4% after 50 epochs. **Figure 2** illustrates the convergence patterns for loss and accuracy metrics across training iterations.

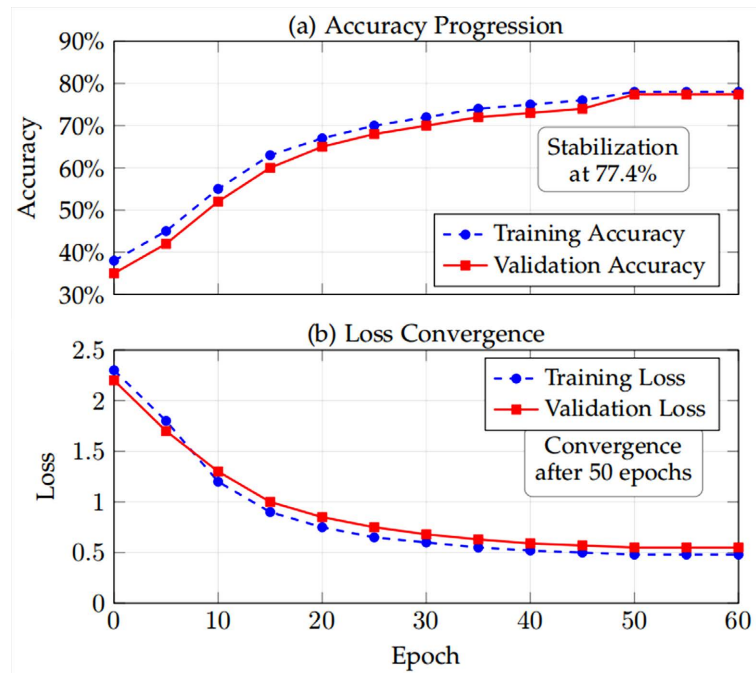


Figure 2. Training progression demonstrated effective learning with validation accuracy stabilizing at 77.4% after 50 epochs. The figure illustrates the convergence patterns for loss and accuracy metrics across training iterations. Loss curves show rapid decrease in initial epochs with stabilization around epoch 50, while accuracy curves demonstrate steady improvement converging to final values of 78% (training) and 77.4% (validation).

4.2. Classification Performance

The hybrid model achieved 78% overall accuracy on the test set with class-specific performance detailed in **Table 2**. Notably, the model demonstrated strong performance for normal cases (98% recall, 84% precision) and clinically relevant sensitivity for malignant detection (70% recall). The confusion matrix in **Table 3** reveals specific misclassification patterns, with benign cases most frequently confused with normal (39 instances) and malignant (21 instances) categories.

Table 2. Classification performance metrics.

Class	Precision	Recall	F1-Score	Support
Normal	0.84	0.98	0.90	200
Benign	0.87	0.72	0.79	203
Malignant	0.74	0.70	0.72	204
Overall	0.82	0.78	0.80	607

Table 3. Confusion matrix analysis.

Actual	Normal	Benign	Malignant
Normal	134	49	17
Benign	39	143	21
Malignant	3	2	199

4.3. Computational Efficiency

The model's lightweight architecture enables efficient deployment in resource-constrained environments. Key computational metrics include:

- Model size: 14 MB (suitable for mobile deployment)
 - Inference time: <20 ms per image on AMD Ryzen 7 5800H CPU
 - Memory requirement: <500 MB RAM during inference
 - Throughput: <50 images per minute on standard hardware
- These characteristics make the model particularly suitable for LMIC deployment where high-end GPU hardware is often unavailable.

4.4. Comparison with Existing Approaches

Table 4 compares the proposed hybrid model with recent breast cancer classification approaches, highlighting the balance between accuracy and efficiency achieved by our method.

Table 4. Comparison with recent breast cancer classification models.

Study	Acc.	Size	Time	Architecture	Dataset
Saber <i>et al.</i> (2023)	98.96%	250 MB	150 ms	InceptionV3	MIAS
Shi <i>et al.</i> (2022)	75.00%	180 MB	120 ms	EfficientNet	CBIS-DDSM
Padelia <i>et al.</i> (2023)	97.14%	210 MB	135 ms	Enh. EfficientNet	MIAS
Ours	78.00%	14 MB	<20 ms	MobileNetV2 -SVM	mini-DDSM

4.5. Clinical Deployment Analysis

A Streamlit-based web application was developed to demonstrate practical deployment in clinical settings. The interface includes user authentication, image upload with preview functionality, and real-time prediction display. The application processes mammography images through the trained model and provides classification results with confidence scores.

4.6. Ablation Study

Justification of Hybrid Architecture. To justify the design decision of the replacement of the standard Softmax classifier of MobileNetV2 with Cubic SVM, an ablation study of both configurations was performed under the same experimental conditions. The hybrid model increased the overall accuracy and malignant recall by 2.8 and 5 percent respectively compared to the Softmax baseline and minimized model size by 4 MB. These are small but clinically significant improvements in screening settings and the additional complexity of the SVM integration is worth the effort especially in resource-limited settings where both model efficiency and diagnostic sensitivity are important.

5. Discussion

5.1. Model Performance Analysis

The hybrid MobileNetV2-cubic SVM architecture provides competitive performance while maintaining computational efficiency, which is crucial for LMIC deployment. The 78% overall accuracy, although lower than some resource-intensive models, represents a favourable trade-off given significantly reduced computational requirements. The high recall for normal cases (98%) is particularly valuable in screening applications, reducing unnecessary follow-ups and healthcare expenditures.

The 70% recall for malignant cases, while clinically significant, indicates room for improvement. Misclassification analysis suggests borderline cases and early-stage malignancies present challenges. Future iterations could incorporate attention mechanisms or ensemble methods to address these limitations.

5.2. Computational Efficiency Implications

The model's compact size (14 MB) and fast inference (<20 ms) enable multiple deployment scenarios relevant to LMICs:

- 1) Mobile deployment on smartphones or tablets used by healthcare workers in remote areas.
- 2) Integration with low-cost diagnostic hardware at primary care facilities.
- 3) Efficient transmission and analysis in bandwidth-limited telemedicine applications.
- 4) Offline operation without continuous internet connectivity.

5.3. Limitations and Future Directions

Several limitations should be acknowledged. Although the mini-DDSM dataset is useful for early development, it may not represent population diversity across geographic and ethnic groups. Future work should include multi-center validation in diverse LMIC populations. While effective, the cubic SVM might not be optimal for MobileNetV2's high-dimensional feature space. Alternative classifiers such as gradient boosting machines or deep neural network heads may improve performance. Additionally, the current model does not integrate clinical context with individual images. Next-generation applications could incorporate patient metadata, multi-view analysis, and temporal comparisons to enhance diagnostic accuracy. Although down sampling mammograms to 224×224 pixels was essential as a computational efficiency measure, it probably decreased the detectability of fine structures, including micro-calcifications and subtle malignant structures. This is a conscious trade-off to allow it to be deployed in low resource settings but could also be the cause of the reported sensitivity of 70 percent on malignant cases. Future directions might include multi-resolution pipeline or patch-based analysis to maintain important detail and deployability.

5.4. Clinical Safety, Risk Mitigation, and Intended Use Case

The fact that the 70% recall (sensitivity) is observed with malignant cases only and does not represent a negligible detection ability, also means that there is a 30% false negative rate that should be carefully analyzed in terms of patient safety. This level of sensitivity is suboptimal when compared to state-of-the-art models that can reach 97 - 98 percent accuracy on smaller, curated datasets such as MIAS (as indicated in **Table 4**). But a direct comparison without taking into consideration the context of deployment can be misleading. The usefulness of the suggested model is not the substitution of gold-standard diagnostics but the consideration of the workflow bottlenecks of the Low-Middle Income Countries (LMICs).

The proposed system would be better applied in pre-screening or a triage position rather than being an initial diagnostic instrument due to the performance profile, namely high recall (98 percent) and moderate recall (70 percent). A radiologist or a clinical officer must look at each mammogram in a normal LMIC screening pipeline. Using this model as a first-pass filter, the system will be able to automatically indicate mammograms that have a high likelihood of being normal with an accuracy of 98%. Such high-confidence-normal cases might be relegated or grouped to be reviewed quicker, which will considerably decrease the load of overworked specialists.

The clinical risk is addressed in this triage situation as follows:

1) Conservative Triage Threshold: Cases that are not categorized as high-confidence normal (*i.e.* any case with a prediction score that is skewed to benign or malignant) would automatically be sent to immediate expert review. This will make sure that though the model is used to clear the normal cases, it does not block any potentially malignant case to get to a human expert.

2) Human-in-the-Loop: The model will not substitute for the clinician, rather it will enhance him or her. The 30% of the malignant cases that the model would misclassify would, in this workflow, be classified in either of two categories: They are wrongly labeled as “Benign” and yet would still be reported to the experts due to not being under the “Normal” category. They are incorrectly categorized under the normal. Although this is the worst-case scenario, the desired workflow presupposes that in areas with very limited resources, the alternative is frequently no screening instead of an ideal screen.

3) Comparison to the Status Quo: Opportunistic screening or no screening is a reality in most resource constrained environments because there are no radiologists. The 70 percent sensitive automated triage tool is not a perfect tool but has the potential to detect most of the malignancies that would not have been detected before it reaches the symptomatic later stages. The missing rate of 30 percent should be balanced with the possibility to reach hitherto unscreened groups.

Moreover, the low rate of false positive on normal cases (high precision) of the model reduces chances of flooding the system with unnecessary follow-ups and biopsies which is of great essence in where the resource of confirmatory testing is scarce.

Thus, the 30% miss rate is unacceptable in stand-alone diagnostic equipment, but it is a possible net positive in the overall health of the population when used as a workload-balancing triage device in a human-in-the-loop system that operates in extreme resource scarcity. Further developments will be made to enhance malignant sensitivity using strategies such as multi-view analysis or ensemble strategy, although the present deployment model is concerned with safety by guaranteeing that all non-normal predictions are subject to expert scrutiny.

6. Conclusions

This paper presents a resource-optimized hybrid MobileNetV2-cubic SVM model for breast cancer classification. The architecture balances diagnostic accuracy (78% overall, 70% malignant recall) with computational efficiency (14 MB, 20 ms inference), addressing key deployment limitations in LMICs.

The model demonstrates strength in identifying normal cases (98% recall), potentially reducing unnecessary referrals within overstretched health systems. While malignant case sensitivity requires further enhancement, current performance represents an important step toward accessible breast cancer diagnostics.

The development strategy, focused on computational efficiency and diagnostic accuracy, provides a blueprint for deploying medical AI in resource-limited environments. As global healthcare systems increasingly adopt AI-aided diagnostics, models like the one presented here will be essential for ensuring equitable access to advanced medical technologies across all resource settings.

Future research directions include multi-modal integration, privacy-preserving model enhancement via federated learning, and clinical validation trials in LMICs to assess real-world efficacy and implementation pathways.

Acknowledgements

The authors gratefully acknowledge Midlands State University for research support and access to computational resources. Special thanks to the open-source medical imaging community for dataset accessibility and tool development.

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

- [1] World Health Organization (2024) Breast Cancer. <https://www.who.int/news-room/fact-sheets/detail/breast-cancer>
- [2] Marinovich, M.L., Lotter, W., Waddell, A. and Houssami, N. (2024) Simulated Arbitration of Discordance between Radiologists and Artificial Intelligence Interpretation of Breast Cancer Screening Mammograms. *Journal of Medical Screening*, **32**, 48-52. <https://doi.org/10.1177/09691413241262960>
- [3] Iacob, R., Iacob, E.R., Stoicescu, E.R., Ghenciu, D.M., Cocolea, D.M., Constantinescu, A., *et al.* (2024) Evaluating the Role of Breast Ultrasound in Early Detection of Breast Cancer in Low- and Middle-Income Countries: A Comprehensive Narrative Review.

- Bioengineering*, **11**, Article 262. <https://doi.org/10.3390/bioengineering11030262>
- [4] Franklin, M., Lewis, S. and Smith, A.L. (2024) Complexities in Supportive Care for People with Metastatic Breast Cancer: A Qualitative Study. *Journal of Cancer Survivorship*, **20**, 198-208. <https://doi.org/10.1007/s11764-024-01646-8>
 - [5] Alzubaidi, L., Zhang, J., Humaidi, A.J., Al-Dujaili, A., Duan, Y., Al-Shamma, O., et al. (2021) Review of Deep Learning: Concepts, CNN Architectures, Challenges, Applications, Future Directions. *Journal of Big Data*, **8**, Article No. 53. <https://doi.org/10.1186/s40537-021-00444-8>
 - [6] Salehi, A.W., Khan, S., Gupta, G., Alabdullah, B.I., Almjally, A., Alsolai, H., et al. (2023) A Study of CNN and Transfer Learning in Medical Imaging: Advantages, Challenges, Future Scope. *Sustainability*, **15**, Article 5930. <https://doi.org/10.3390/su15075930>
 - [7] Younesi, A., Ansari, M., Fazli, M., Ejlali, A., Shafique, M. and Henkel, J. (2024) A Comprehensive Survey of Convolutions in Deep Learning: Applications, Challenges, and Future Trends. *IEEE Access*, **12**, 41180-41218. <https://doi.org/10.1109/access.2024.3376441>
 - [8] Li, Z., Liu, F., Yang, W., Peng, S. and Zhou, J. (2022) A Survey of Convolutional Neural Networks: Analysis, Applications, and Prospects. *IEEE Transactions on Neural Networks and Learning Systems*, **33**, 6999-7019. <https://doi.org/10.1109/tnnls.2021.3084827>
 - [9] Howard, A.G., Zhu, M., Chen, B., Kalenichenko, D., Wang, W., Weyand, T., Andreetto, M. and Adam, H. (2017) MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications. arXiv: 1704.04861.
 - [10] Tan, M. and Le, Q.V. (2019) EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. *Proceedings of the 36th International Conference on Machine Learning*, Long Beach, 9-15 June 2019, 6105-6114.
 - [11] Thapliyal, N., Aeri, M., Kukreja, V. and Sharma, R. (2024) Navigating Landscapes through AI: A Comparative Study of EfficientNet and MobileNetV2 in Image Classification. 2024 *International Conference on Emerging Technologies in Computer Science for Interdisciplinary Applications (ICETCS)*, Bengaluru, 22-23 April 2024, 1-4. <https://doi.org/10.1109/icetcs61022.2024.10543922>
 - [12] Shen, L., Margolies, L.R., Rothstein, J.H., Fluder, E., McBride, R. and Sieh, W. (2019) Deep Learning to Improve Breast Cancer Detection on Screening Mammography. *Scientific Reports*, **9**, Article No. 12495. <https://doi.org/10.1038/s41598-019-48995-4>
 - [13] Saber, A., Sakr, M., Abo-Seida, O.M., Keshk, N. and Chen, H. (2023) A Novel Deep Learning Model for Breast Cancer Diagnosis Using Transfer Learning and Ensemble Classifiers. *Computers in Biology and Medicine*, **162**, Article ID: 107054.
 - [14] Careem, R., Md Johar, M.G. and Khatibi, A. (2024) Deep Neural Networks Optimization for Resource-Constrained Environments: Techniques and Models. *Indonesian Journal of Electrical Engineering and Computer Science*, **33**, 1843-1854. <https://doi.org/10.11591/ijeecs.v33.i3.pp1843-1854>
 - [15] Petrini, D., Dantas, C.R.R., Santos, A.B.M., Barros, R.C. and Torres, R.S. (2023) Two-View Mammography Classification Using Transfer Learning and Deep Convolutional Neural Networks. *Computer Methods and Programs in Biomedicine*, **229**, Article ID: 107296.