Korea Computer Industry Association

BIC 2024
The 8th International Conference on Big data, IoT, and Cloud Computing

Comparative Performance of Deep Learning Models in Detecting Invasive Ductal Carcinoma

Nguyen Nang Hung Van[1], Phuc Hao Do[2,*], Tran Duc Le[3,*], Ngo Van Uc[4], Truong Duy Dinh[5], Van Dai Pham[6]

[1]University of Science and Technology, University of Da Nang, Vietnam
[2]The Bonch-Bruevich Saint Petersburg State University of Telecommunications
[3]University of Wisconsin-Stout, USA
[4]Dong A University, Vietnam
[5]Posts and Telecommunications Institute of Technology, Vietnam
[6]Swinburne Vietnam, FPT University, Vietnam
do.hf@sut.ru, let@uwstout.edu, nguyenvan@dut.udn.vn, ngovanuc.1508@gmail.com, duydt@ptit.edu.vn, daipv11@fe.edu.vn
* Corresponding author

Abstract. Breast cancer detection is a critical area of medical research where early and accurate diagnosis can significantly improve patient outcomes. This study conducts a comprehensive comparative analysis of various deep-learning models for detecting Invasive Ductal Carcinoma (IDC), the most prevalent subtype of breast cancer. We evaluate the performance of Convolutional Neural Networks (CNNs), MobileNet, Transformer, and EfficientNet models using a comprehensive dataset of IDC images. The models are assessed based on key metrics, including accuracy, F1-score, precision, recall, and training time. Our results indicate significant differences in performance metrics, with EfficientNet achieving the highest accuracy and demonstrating exceptional reliability and precision. MobileNet also performs strongly, providing a balance of high accuracy and efficiency. These findings offer valuable insights into the suitability of different deep learning architectures for breast cancer detection, guiding future research and clinical applications toward more effective diagnostic tools.

Keywords: Breast Cancer Detection, Invasive Ductal Carcinoma (IDC), CNNs, MobileNet, Transformer, EfficientNet

1 Introduction

Breast cancer remains one of the most common and fatal diseases affecting women worldwide [1]. According to the World Health Organization, breast cancer accounts for approximately 25% of all cancer cases and 15% of all cancer deaths among women globally [2]. The incidence of breast cancer has been steadily rising, with many cases diagnosed at advanced stages due to limited access to early detection and diagnostic facilities. Early-stage diagnosis significantly improves the chances of successful treatment and survival rates.

Artificial intelligence (AI) and deep learning (DL) advancements have revolutionized medical diagnostics [3]. This paper aims to evaluate the effectiveness of several popular deep-learning models in the context of breast cancer detection, specifically focusing on the detection of Invasive Ductal Carcinoma (IDC), the most common subtype of breast cancer. We perform a comparative analysis of Convolutional Neural Networks (CNNs), MobileNet, Transformer, and EfficientNet models using a comprehensive dataset of IDC images. The performance of these models is assessed based on critical metrics, including accuracy, F1-score, precision, and recall. Our study contributes to ongoing research in medical diagnostics by providing insights into the strengths and weaknesses of different deep-learning architectures in detecting breast cancer.

2 Related works

In recent years, machine learning (ML) and deep learning (DL) techniques have emerged as promising tools for enhancing the accuracy and efficiency of breast cancer diagnosis. Research has explored various ML approaches, ranging from traditional algorithms like decision trees, support vector machines (SVM), and neural networks to advanced DL models. Comparative studies, such as those by Fatima et al. [4] and Naji et al. [5], have highlighted the efficacy of ensemble methods and emphasized the importance of feature selection and data preprocessing in enhancing predictive accuracy.

Specifically, DL techniques have garnered attention for their adeptness in handling complex and high-dimensional breast cancer data. Studies by Tiwari et al. [6] and Yala et al. [7] have shown that CNNs excel in accuracy and computational efficiency compared to traditional methods like SVM and k-nearest neighbors (KNN). Furthermore, fusion-based approaches and ensemble methods, as proposed by Siddiqui et al. [8] and Saleh et al. [9], have improved prediction accuracy by integrating DL with traditional ML techniques.

In summary, advancements in ML and DL have significantly advanced breast cancer prediction capabilities. While traditional ML algorithms remain effective, DL models such as CNNs and recurrent neural networks (RNNs) demonstrate superior performance in capturing intricate patterns within medical imaging data. Fusion-based methods further augment accuracy by harnessing the complementary strengths of diverse models, promising advancements in early detection, and personalized treatment strategies for breast cancer patients.

3 Methodology

In this study, we employ the IDC dataset [10] and follow a comprehensive implementation process. First, the data undergoes rigorous preprocessing to ensure quality and suitability for model training, including normalization, augmentation, and partitioning into training and testing sets. Next, we implement and fine-tune several state-of-the-art deep learning models – CNN, MobileNet, Transformer, and EfficientNet – to perform binary classification tasks that distinguish between cancerous and non-cancerous tissues.

3.1 Dataset

Figure 1 shows sample images categorized as "*No breast cancer [IDC (-)]*" and "*Breast cancer – [IDC (+)]*."

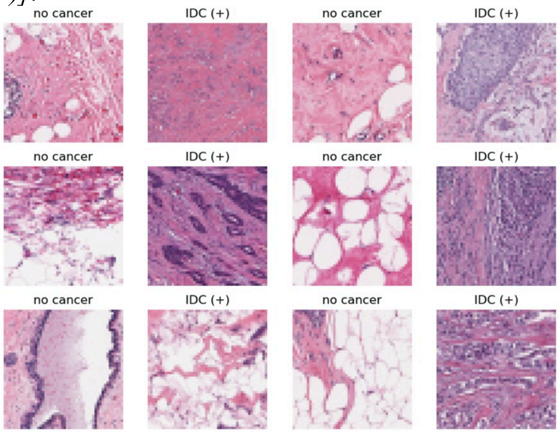

Fig. 1. Sample images "*No breast cancer [IDC (-)]*" and "*Breast cancer – [IDC (+)]*."

The dataset comprises 162 whole-mount slide images of breast cancer specimens scanned at 40x magnification. These images were divided into 277,524 patches, each 50x50 pixels in size, with 198,738 IDC-negative patches and 78,786 IDC-positive patches. Each patch is labeled with a format that includes the patient ID, the x and y coordinates of the patch's origin, and its class label (0 for non-IDC and 1 for IDC). IDC-negative patches typically exhibit a pink coloration, while IDC-positive patches generally appear darker. This darker hue in positive samples indicates stronger cell development, a characteristic commonly observed in cancerous tissues. The dataset was divided into a training set (80%) and a test set (20%) to ensure robust model training and evaluation.

3.2 Model architectures

In this study, we evaluate the performance of four state-of-the-art deep learning models – CNN, MobileNet, EfficientNet, and Transformer – in detecting Invasive Ductal Carcinoma from the IDC dataset $\mathcal{D} = \{(\mathbf{x}_i, y_i)\}_{i=1}^{N}$ Here, $\mathbf{x}_i \in \mathbb{R}^{H\times W\times C}$ represents the i-th input image with height H, width W, and C channels, and $y_i \in \{0,1\}$ is the binary label indicating the presence (1) or absence (0) of IDC.

CNN model: The CNN architecture is designed to automatically and adaptively learn spatial hierarchies of features through backpropagation by using multiple building blocks, such as convolution layers, pooling layers, and fully connected layers. The convolution layers apply filters to the input to create feature maps, the pooling layers reduce the dimensionality of these feature maps, and the fully connected layers generate the final classification output. This architecture is well-suited for image recognition tasks because it can effectively capture spatial features. This architecture processes the input image $\mathbf{x}_i$ through several layers to learn spatial hierarchies of features. The key components include:

- **Convolutional Layers:** Apply filters to extract features.
- **Pooling Layers:** Reduce dimensionality by downsampling.
- **Fully Connected Layers:** Flatten the feature maps and classify the input.

The output is a probability $\hat{y}$ representing the lielihood of IDC: $\hat{y} = \sigma(W_{out}a^{(L)} + b_{out})$ where $a^{(L)}$ is the activation from the final layer, and σ is the sigmoid function.

MobileNet: MobileNet employs depthwise separable convolutions to significantly reduce the number of parameters compared to traditional CNNs, making it efficient and lightweight. This architecture is particularly advantageous for applications requiring low computational power and memory usage, without substantially compromising accuracy. MobileNet's efficiency makes it suitable for mobile and embedded vision application deployment. This model enhances efficiency by using depthwise separable convolutions, which split the convolution into depthwise and pointwise operations: $Z = \sigma((X * K_d) * K_p + b)$, where K_d and K_p are the depthwise and pointwise filters. This reduces the number of parameters and computation.

EfficientNet: EfficientNet scales the depth, width, and resolution of the network in a balanced manner using a compound scaling method, which improves both accuracy and efficiency. This architecture uses mobile inverted bottleneck convolution (MBConv) blocks, which help capture fine-grained features while maintaining computational efficiency. EfficientNet is known for achieving state-of-the-art performance on various image classification benchmarks. This model uses compound scaling to uniformly scale network dimensions (depth, width, resolution). It employs mobile inverted bottleneck convolution (MBConv) blocks, enhancing feature extraction while maintaining efficiency: $Z = \sigma((X * K_{dw}) * K_{pw} + b)$, where K_{dw} and K_{pw} are the filters used in the MBConv blocks.

Transformer: The Transformer model, initially designed for natural language processing, has been adapted for image analysis tasks. It processes the input image as a sequence of patches, embeds them, and uses multi-head self-attention mechanisms to capture global dependencies. This approach allows the model to focus on different image parts and understand complex patterns. The Transformer architecture is highly flexible and can capture long-range dependencies within the image data. This model processes images as sequences of patches. Each image $\mathbf{x}_i$ is divided into patches, embedded, and processed through multi-head self-attention mechanisms to capture global dependencies: $Z = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V$, where where Q, K, and V are the query, key, and value matrices.

4 Experimental Results

Analysis of training accuracy across multiple epochs: Figure 2 illustrates the training accuracy progression of four distinct deep learning models over 100 epochs. The CNN model shows a gradual increase in training accuracy, starting at around 0.78

and reaching approximately 0.91 by the 100th epoch. This slower progression can be attributed to the CNN's conventional architecture, which relies on sequential layers of convolutions and pooling, requiring more epochs to converge compared to the other models. Conversely, MobileNet demonstrates rapid early-stage accuracy enhancement, achieving around 0.98 by the 10th epoch and stabilizing at 0.99 after that. This swift convergence can be explained by MobileNet's efficient architecture, which uses depthwise separable convolutions to reduce the number of parameters and computational complexity, allowing it to learn and generalize from the training data quickly.

EfficientNet showcases exceptional initial accuracy, surpassing 0.90 within the first ten epochs and maintaining near-perfect levels of about 0.99 from the 20th epoch onward. EfficientNet's compound scaling method, which balances depth, width, and resolution, likely contributes to its ability to capture detailed features quickly and efficiently. The Transformer model, starting with a lower initial accuracy of around 0.75, shows steady improvement, reaching about 0.96 by the 100th epoch. While it does not match the peak accuracies of MobileNet and EfficientNet, its steady improvement suggests that the self-attention mechanism effectively captures long-range dependencies in the data. The slower start could be due to the model's complexity and the need for more epochs to fine-tune its numerous parameters effectively.

In summary, MobileNet and EfficientNet quickly achieve and maintain high accuracy, while CNN and Transformer models require more epochs to approach similar performance levels. These differences highlight the trade-offs between model complexity, training time, and accuracy, providing valuable insights for selecting appropriate models for breast cancer detection tasks.

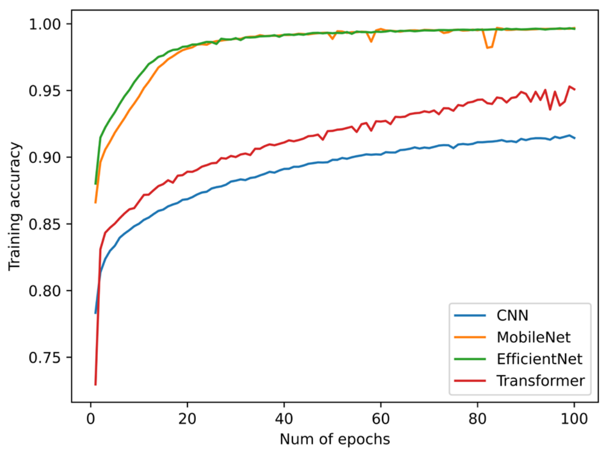

Fig. 2. The training accuracy across multiple epochs between CNN, MobileNet, EfficientNet, and Transformer model

Analysis of metrics (F1-score, Precision, Recall) between different models: Figure 3 presents a comparative analysis of the models' F1-score, Precision, and Recall metrics. Performance metrics range from 0.75 to 0.925 on the y-axis, with each model listed on the x-axis.

EfficientNet, the top-performing model across all metrics, demonstrates its robustness and reliability in breast cancer detection. Its high precision (0.93) indicates it makes very few false positive errors, while its high recall (0.91) ensures it successfully identifies most cancer cases. This balance is reflected in its high F1-score (0.92), making EfficientNet a highly dependable choice for clinical use where accurate diagnosis is paramount.

MobileNet, a close second, also shows strong performance with high precision (0.92) and recall (0.89), suggesting it effectively detects cancer cases while minimizing false alarms. Although demonstrating balanced performance, CNN lags behind EfficientNet and MobileNet, highlighting the benefits of more advanced architectures.

The Transformer model, with its higher recall (0.88) but lower precision (0.77) and F1-score (0.80), indicates it is better at identifying cancer cases but at the cost of higher false positives. This trade-off might be acceptable in scenarios where missing a cancer case is far more detrimental than having a false positive. However, it highlights the need for further optimization to improve precision.

Overall, these findings emphasize the importance of choosing models like EfficientNet and MobileNet for clinical settings. The goal is to maximize detection rates and accuracy, thereby providing reliable and precise breast cancer detection.

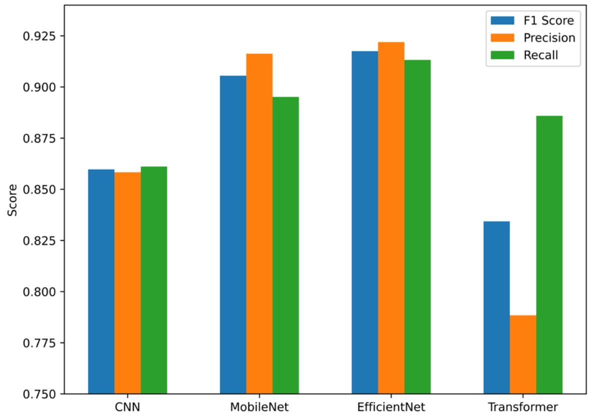

Fig. 3. The metrics (F1-score, Precision, Recall) between CNN, MobileNet, EfficientNet and Transformer model

Analysis of accuracy test data between different models: Figure 4 presents the test accuracy results of four deep learning models. Each model's test accuracy, ranging from 0.80 to 0.94, is depicted on the y-axis, with models listed on the x-axis. CNN's moderate accuracy (0.85) reflects its conventional architecture, which, while effective, may not capture complex patterns as efficiently as more advanced models. MobileNet's higher accuracy (0.90) demonstrates its ability to efficiently process data with its depthwise separable convolutions, making it a strong candidate for real-world applications requiring both accuracy and computational efficiency. EfficientNet's top performance (0.92) can be attributed to its balanced scaling of network dimensions, which allows it to capture intricate features with high precision. This makes EfficientNet particularly suitable for clinical settings where precise detection of cancerous tissues is critical. Transformer's lower accuracy (0.83) indicates that while it has potential, it may require further optimization for this specific task. Its performance suggests that the model's complexity and focus on global dependencies may not be as advantageous for this application compared to other models.

In summary, EfficientNet stands out as the most accurate and reliable model for breast cancer detection, and MobileNet also performs strongly. These findings provide valuable insights into selecting the most appropriate deep-learning models for clinical applications, emphasizing the need for accuracy and efficiency in medical diagnostics.

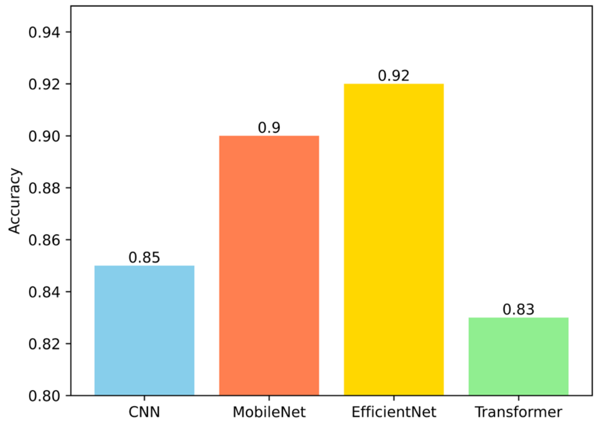

Fig. 4. The accuracy test data with different models (CNN, MobileNet, EfficientNet, and Transformer model)

Analysis of model training time between different models: The analysis revealed significant differences among the four models. The CNN model demonstrated the shortest training time, clocking in at 926.32 seconds, attributable to its simpler architecture compared to the others. MobileNet, however, required a considerably longer training duration of 4320 seconds. This extended training time can be explained by its deeper architecture and the complexity of its depthwise separable convolutions, which are optimized for mobile devices.

EfficientNet, renowned for its superior classification performance, exhibited the longest training time at 17459.63 seconds. This prolonged training period is due to its intricate architecture, which scales across multiple dimensions, balancing depth, width, and resolution to enhance performance. With a training time of 3857 seconds, the Transformer model, while longer than CNN's, showcases its potential through innovative attention mechanisms that effectively capture extensive data dependencies.

These results underscore the trade-offs between model complexity and training time. While CNN offers quick training due to its straightforward design, models like MobileNet and EfficientNet, with their more complex architectures, require longer training times but deliver enhanced performance and precision. The Transformer's moderate training time highlights its capability to handle sophisticated attention mechanisms, positioning it as a promising option despite the need for further optimization.

5 Conclusion

In conclusion, this comparative study underscores the critical importance of selecting appropriate deep-learning models for breast cancer detection, highlighting their distinct strengths and weaknesses. CNNs and MobileNet exhibit efficient performance with shorter training times, making them suitable for applications with constrained computational resources. However, while their performance metrics are commendable, they may not match the accuracy and precision of more advanced models such as Transformer and EfficientNet.

Transformer models, leveraging innovative attention mechanisms, show potential in capturing intricate data dependencies but require longer training durations. EfficientNet stands out with superior overall performance metrics, demonstrating robustness in accurately identifying Invasive Ductal Carcinoma (IDC). Its balanced scaling of network dimensions enhances its capability to capture detailed features with high precision, making it particularly suitable for clinical settings where accurate diagnosis is paramount.

These findings provide valuable insights into the application of deep learning models in medical diagnostics. They emphasize the need to balance model complexity, training time, and accuracy to achieve reliable and precise breast cancer detection. Our future research should continue to optimize these models, considering both computational efficiency and diagnostic performance.

References

1. Akram, M., Iqbal, M., Daniyal, M., Khan, A.U.: Awareness and current knowledge of breast cancer. Biological research 50, 1-23 (2017)
2. Sung, H., Ferlay, J., Siegel, R.L., Laversanne, M., Soerjomataram, I., Jemal, A., Bray, F.: Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA: a cancer journal for clinicians 71, 209-249 (2021)
3. Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H., Wang, Y.: Artificial intelligence in healthcare: past, present and future. Stroke and vascular neurology 2, (2017)
4. Fatima, N., Liu, L., Hong, S., Ahmed, H.: Prediction of breast cancer, comparative review of machine learning techniques, and their analysis. IEEE Access 8, 150360-150376 (2020)
5. Naji, M.A., El Filali, S., Aarika, K., Benlahmar, E.H., Abdelouhahid, R.A., Debauche, O.: Machine learning algorithms for breast cancer prediction and diagnosis. Procedia Computer Science 191, 487-492 (2021)
6. Tiwari, M., Bharuka, R., Shah, P., Lokare, R.: Breast cancer prediction using deep learning and machine learning techniques. Available at SSRN 3558786 (2020)
7. Yala, A., Lehman, C., Schuster, T., Portnoi, T., Barzilay, R.: A deep learning mammography-based model for improved breast cancer risk prediction. Radiology 292, 60-66 (2019)
8. Siddiqui, S.Y., Naseer, I., Khan, M.A., Mushtaq, M.F., Naqvi, R.A., Hussain, D., Haider, A.: Intelligent breast cancer prediction empowered with fusion and deep learning. Computers, Materials and Continua 67, 1033-1049 (2021)
9. Saleh, H., Abd-El Ghany, S.F., Alyami, H., Alosaimi, W.: Predicting breast cancer based on optimized deep learning approach. Computational Intelligence and Neuroscience 2022, 1820777 (2022)
10. Janowczyk, A., Madabhushi, A.: Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases. Journal of pathology informatics 7, 29 (2016)