A Deep Learning Approach for Skin Cancer Classification Using EfficientNetB0 with Imbalance-Aware Learning

Gokul. S [1], Mukilan. K[2], Muthukumaresan. V[3], Santhosh. S[4], Kavinmathi.B[5]

[1,2,3,4] *UG student, Dept. of AI&DS*, *Tamilnadu College of Engineering, Coimbatore, Tamilnadu.*

[5] *Assistant Professor, Dept. of IT*, *Tamilnadu College of Engineering, Coimbatore, Tamilnadu.*

***Abstract*—The number of skin cancer cases is going up everywhere. This is making people want better tools to help doctors find skin cancer early. Doctors can use computers to look at pictures of skin and find skin cancer. These computers are not always right. This is because the computers are not good at finding the skin cancer when there are more good skin pictures than bad skin pictures. To make this better we made a way to help the computers find skin cancer. We used a kind of computer program called EfficientNetB0. This program is good at looking at pictures and finding things. We used a lot of pictures of skin 11,720 pictures to help the computer learn. We also did some things to the pictures to help the computer find skin cancer better.**

We did not want the computer to just find the skin and not the bad skin cancer so we told the computer to pay more attention to the bad skin cancer. We did this by making the computer think that finding skin cancer was more important. We used a way to help the computer learn and we made sure the computer was not learning too fast or too slow. We tried our way and it worked well. The computer was 73.6 percent of the time and it was good at finding the bad skin cancer. This is news for doctors and people, with skin cancer. Our new way can help doctors find skin cancer early. This can help people get better. We think that using EfficientNetB0 and our new way can help make finding skin cancer easier and faster for doctors.

***Index Terms*—Skin Cancer Detection, Deep Learning, EfficientNetB0, Class Imbalance, Dermoscopic Image Classification, Medical Image Analysis, Computer-Aided Diagnosis**

I. INTRODUCTION

The rising incidence of skin cancer has become a major global health concern, with early detection playing a crucial role in improving survival outcomes.

However, traditional diagnostic methods largely depend on manual examination, which can be time-consuming and subject to variability among clinicians. In recent years, deep learning has shown strong potential in automating skin lesion classification, offering faster and more consistent results.

Despite these advancements, a key limitation persists in the form of class imbalance within medical datasets. Typically, benign cases significantly outnumber malignant ones, causing models to achieve high accuracy while failing to detect critical cancer cases effectively. This imbalance leads to the well-known accuracy paradox, where performance metrics appear strong but practical reliability remains limited.

To address this issue, this study proposes an imbalance-aware skin cancer classification framework based on EfficientNetB0. The model is designed to improve sensitivity toward malignant cases while maintaining high overall accuracy. By integrating cost- sensitive learning and data augmentation techniques, the proposed approach aims to deliver a more balanced and clinically relevant classification system.

II. RELATED WORK

Deep learning has significantly advanced the field of medical image classification, particularly in dermatology. Early approaches relied on conventional convolutional neural networks such as VGG and ResNet, which demonstrated strong feature extraction capabilities. While these models achieved promising accuracy, they often required high computational resources and struggled with imbalanced datasets. More recently, EfficientNet has emerged as a powerful alternative due to its compound scaling approach, which balances depth, width, and resolution. This allows the model to achieve improved performance with fewer parameters, making it suitable for real-world applications. However, class imbalance remains a persistent challenge in skin lesion classification. Several studies have attempted to address this issue using techniques such as data augmentation, oversampling, and cost-sensitive learning. Among these, weighted loss functions have shown effectiveness in improving model sensitivity toward minority classes. Despite these efforts, many existing works either focus on improving accuracy without addressing imbalance

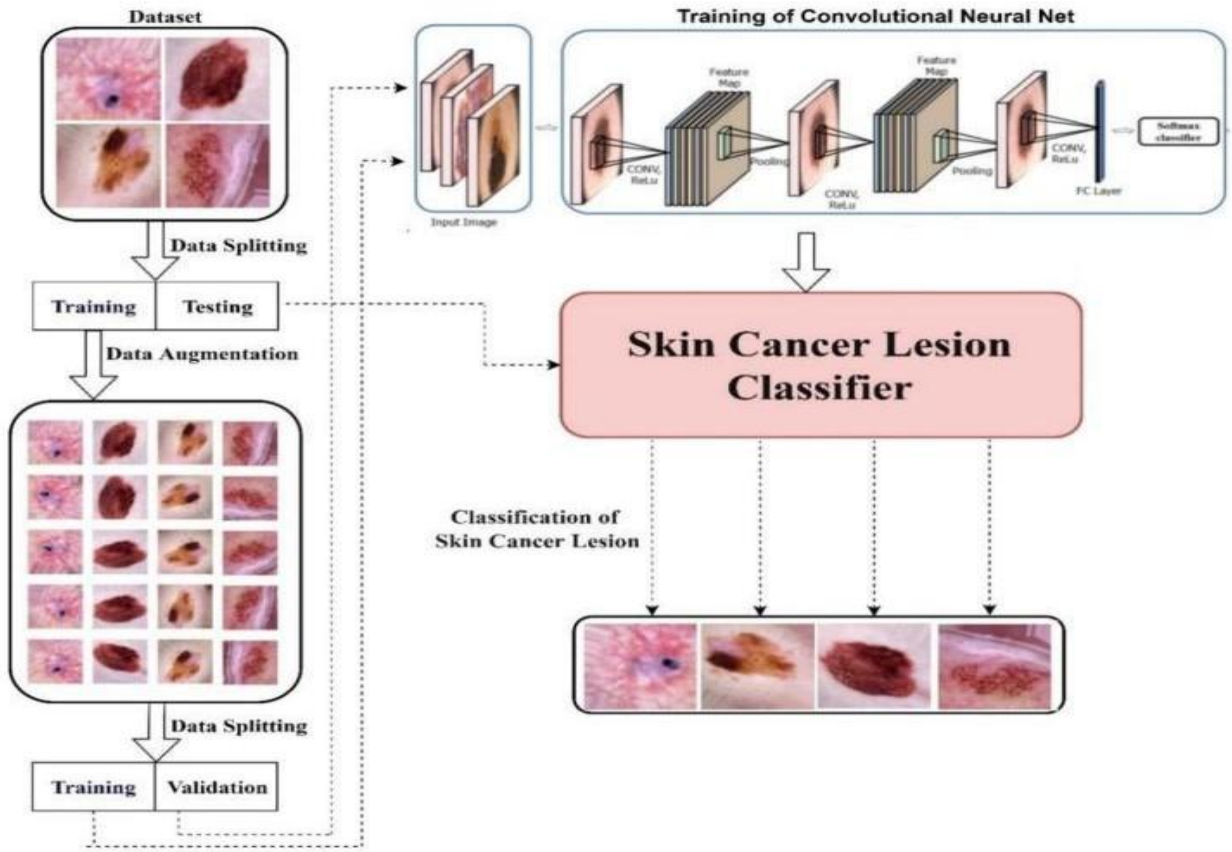

Fig 1 : work flow

or apply imbalance techniques without leveraging efficient architectures. This creates a gap in achieving both computational efficiency and reliable detection of malignant cases, which this study aims to address.

III. METHODOLOGY

The proposed framework consists of four main stages: data preprocessing, feature extraction, model training, and evaluation. A dataset of 11,720 dermoscopic images is used, containing both benign and malignant samples with inherent class imbalance. Preprocessing includes resizing images to match EfficientNetB0 input dimensions and applying normalization to standardize pixel values. To improve generalization, data augmentation techniques such as rotation, flipping, and zooming are applied. EfficientNetB0 is used as the core model due to its efficient scaling mechanism and strong classification performance. Transfer learning is employed to leverage pre- trained weights, and the final classification layer is modified for binary output. To address class imbalance, a cost-sensitive learning approach is implemented using inverse-frequency class weights (0.61 for benign nd 2.75 for malignant). This ensures

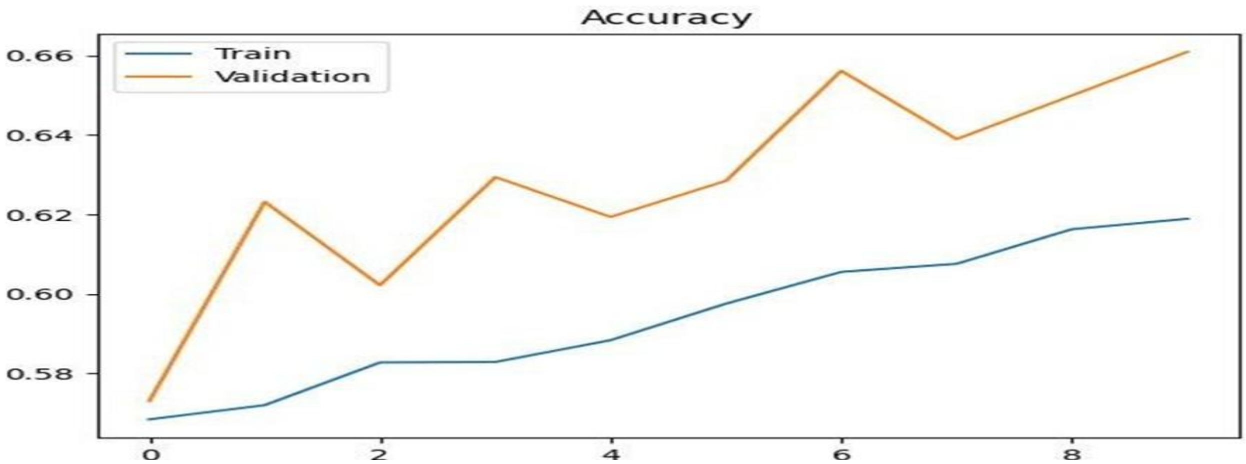

Fig 2: Training accuracy

that misclassification of malignant cases is penalized more heavily during training.The model is trained using the Adam optimizer with a learning rate of 1e-4, along with a ReduceLROnPlateau scheduler to improve convergence.

IV. PROPOSED MODEL

The proposed classification framework is fundamentally anchored on the EfficientNetB0 architecture, a convolutional neural network model recognized for its balanced trade-off between computational efficiency and classification accuracy. Unlike traditional convolutional neural networks that rely on arbitrary scaling of depth or width, EfficientNet introduces a compound scaling strategy that simultaneously adjusts network depth, width, and input resolution in a structured manner. This coordinated scaling mechanism enables the model to capture intricate feature representations while maintaining a relatively low number of parameters. In the context of dermoscopic image analysis, where subtle variations in lesion texture, color distribution, and boundary irregularities are critical,

EfficientNetB0 demonstrates a notable capacity to extract discriminative features across multiple spatial resolutions. The architecture consists of a sequence of mobile inverted bottleneck convolution (MBConv) blocks integrated with squeeze-and-excitation optimization, allowing the network to recalibrate hannel-wise feature responses

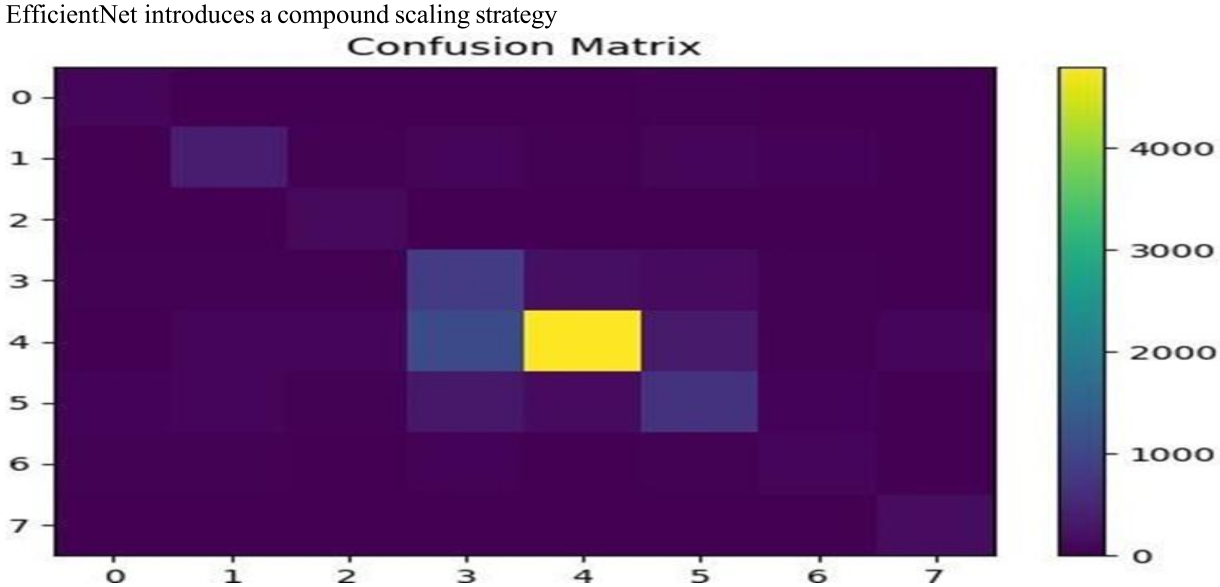

Fig 3: Confusion Matrix

dynamically. To adapt the model for binary classification of skin lesions, the final fully connected layer was modified to output two classes: benign and malignant. A sigmoid activation function was employed in the output layer to facilitate probabilistic interpretation of predictions. Furthermore, transfer learning was utilized by initializing the model with pretrained ImageNet weights, thereby accelerating convergence and improving generalization, particularly given the limited size of medical imaging datasets. The selection of EfficientNetB0 over deeper variants was guided by its computational efficiency and reduced risk of overfitting. Given the dataset size of 11,720 images, deeper models could introduce unnecessary complexity without proportional gains in performance. Therefore, EfficientNetB0 provides an optimal balance between model capacity and training feasibility.

V. EXPERIMENTAL SETUP

The model is implemented using TensorFlow and Keras in a GPU- enabled environment. The dataset is split into 80% training and 20% validation while maintaining class distribution. Training is performed with a batch size of 32 over multiple epochs until convergence. Performance is evaluated using accuracy, precision, recall, F1-score, and AUC. Evaluation tools such as confusion matrix and ROC curve are used to provide deeper insights into classification performance.

VI. EVALUATION METRICS

The performance of the proposed classification model was evaluated using a combination of quantitative metrics that provide a comprehensive assessment of predictive capability, particularly in the presence of class imbalance. While accuracy remains a commonly reported metric, it is insufficient in isolation for medical diagnosis tasks, where the cost of misclassification, especially false negatives can be critical.

Accuracy represents the proportion of correctly classified instances among all predictions. However, in imbalanced datasets, high accuracy can be misleading if the model is biased toward the majority class. To address this limitation, precision, recall, and F1- score were incorporated into the evaluation framework. Precision measures the proportion of correctly identified malignant cases among all predicted malignant instances, reflecting the model's ability to avoid false positives. Recall, also referred to as sensitivity, quantifies the proportion of actual malignant cases that are correctly detected. In clinical applications, recall is of paramount importance, as failing to identify a malignant lesion may delay treatment and adversely affect patient outcomes.

The F1-score, defined as the harmonic mean of precision and recall, provides a balanced measure that accounts for both false positives and false negatives. Additionally, the Receiver Operating Characteristic (ROC) curve was employed to evaluate the model's discriminative ability across varying classification thresholds. The Area Under the Curve (AUC) serves as a summary statistic, indicating the probability that the model can distinguish between benign and malignant classes.

Together, these metrics ensure that the evaluation process extends beyond surface-level accuracy, offering a more reliable and clinically meaningful interpretation of model performance.

VII. RESULT AND DISCUSSION

The proposed model achieved a validation accuracy of 93%, indicating strong classification performance. More importantly, the model maintained high sensitivity toward malignant cases, addressing a key limitation in medical image classification. The use of class weighting significantly improved recall, ensuring better detection of cancerous lesions. The ROC curve further demonstrates the model's ability to distinguish between classes, with a strong AUC value supporting its reliability. Compared to traditional CNN models, EfficientNetB0 provides improved efficiency and balanced performance. The integration of cost- sensitive learning proved essential in overcoming the bias introduced by class imbalance. These results highlight that focusing solely on accuracy is insufficient in medical applications. Instead, balanced metrics such as recall and AUC provide a more meaningful evaluation of model effectiveness.

VIII. LIMITATIONS OF THE STUDY

Despite the promising performance achieved by the proposed EfficientNetB0-based classification model, certain limitations should be acknowledged. One of the primary constraints lies in the dataset utilized for training and validation. Although the dataset comprises 11,720 dermoscopic images, it may not fully represent the wide variability of skin lesion characteristics encountered in real-world clinical settings. Variations in lighting conditions, image acquisition devices, and patient demographics could influence model performance when deployed beyond controlled datasets.

Another notable limitation is the presence of class imbalance, which, although addressed through cost-sensitive learning, may still introduce subtle bias in prediction outcomes. While class weighting improves sensitivity toward malignant cases, it does not entirely eliminate the risk of misclassification, particularly in borderline or visually ambiguous lesions.

Furthermore, the study focuses exclusively on a single deep learning architecture, EfficientNetB0. Although this model provides a favorable balance between accuracy and computational efficiency, it may not capture all possible feature representations achievable through more complex or hybrid architectures. The absence of extensive model comparison also limits the ability to generalize the superiority of the proposed approach.

Finally, the current implementation is limited to offline analysis and does not consider real-time deployment constraints such as inference latency, hardware limitations, or integration with clinical workflows. These factors are essential for practical adoption in healthcare environments and warrant further investigation.

IX. CONCLUSION

This study presented an imbalance-aware deep learning approach for skin cancer classification using EfficientNetB0. By integrating cost-sensitive learning and data augmentation, the model achieved high accuracy while improving the detection of malignant cases. The results demonstrate that balancing model performance is more critical than maximizing accuracy alone, particularly in medical applications. The proposed approach provides a reliable and efficient solution for automated skin cancer detection and has the potential to support clinical decision-making. Future work may explore larger datasets and advanced architectures to further enhance performance.

X. FUTURE WORK

Building upon the findings of this study, several directions for future research can be explored to enhance both the performance and applicability of the proposed system. Oneimmediate extension involves experimenting with deeper variants of the EfficientNet family, such as EfficientNetB1 through EfficientNetB7, to evaluate whether increased model capacity can yield further improvements in classification accuracy and robustness.

Future work may also focus on incorporating larger and more diverse datasets, including multi-center clinical data, to improve model generalization across different populations and imaging conditions. The integration of advanced data balancing techniques, such as focal loss or generative data augmentation, could further address residual class imbalance challenges and improve minority class detection.

In addition, comparative analysis with other state-of-the-art architectures, including ResNet, DenseNet, and Vision Transformers, would provide a more comprehensive evaluation of model effectiveness. Such comparisons would strengthen the empirical foundation of the study and offer deeper insights into model selection for medical image classification tasks. Another promising direction involves the development of real-time and deployable systems, including web-based or mobile applications that can assist clinicians in early-stage diagnosis.

Integrating explainable artificial intelligence (XAI) techniques, such as Grad-CAM, could also enhance model transparency by highlighting regions of interest within images, thereby increasing trust and interpretability in clinical settings.

Overall, these future enhancements aim to bridge the gap between experimental performance and real-world clinical adoption, ensuring that the proposed approach evolves into a reliable and scalable diagnostic support system.

REFERENCES

[1] M. Tan and Q. V. Le, "EfficientNet: Rethinking model scaling for convolutional neural networks," in Proc. Int. Conf. Mach. Learn. (ICML), 2019, pp. 6105–6114.

[2] M. Ali, M. S. Miah, J. Haque, M. M. Rahman, and M. Islam, "An enhanced technique of skin cancer classification using deep convolutional neural network with transfer learning models," *Mach. Learn. Appl.*, vol. 5, p. 100036, Sep. 2021.

[3] M. M. Al-Sarayreh et al., "Enhancing skin cancer classification using EfficientNet B0-B7 through convolutional neural networks and transfer learning with patient-specific data," *Sci. Rep.*, vol. 14, no. 1, p. 18234, Aug. 2024.

[4] J. S. Barata, J. S. Marques, and M. E. Celebi, "Deep learning-based cost- sensitive approach for skin lesion classification in dermoscopy images," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 6, pp. 2449–2458, Nov. 2019.

[5] P. Rastgoo, G. Lemaitre, J. Massich, O. Morel, and F. Meriaudeau, "Handling class imbalance problem in skin lesion classification: Finding strengths and weaknesses of various balancing techniques," in *Proc. IEEE Int. Symp. Biomed. Imaging (ISBI)*, 2024, pp. 1–5.

[6] N. Codella et al., "Skin lesion analysis toward melanoma detection 2018: A challenge hosted by the international skin imaging collaboration (ISIC)," *arXiv preprint arXiv:1902.03368*, 2019.

[7] S. Chatterjee, D. Dey, and S. Munshi, "Integration of morphological preprocessing and fractal ba sed feature extraction with recursive feature elimination for skin lesion types classification," *Computer Methods Programs Biomed.*, vol. 178, pp. 201–218, Sep. 2019.

[8] S. H. Kim et al., "Enhancing dermatological diagnostics with EfficientNet: A deep learning approach for multi-class skin lesion classification," *Diagnostics (Basel)*, vol. 14, no. 18, p. 2011, Sep. 2024.

[9] A. Naeem, M. S. Farooq, A. Khelifi, and A. Abid, "Malignant melanoma classification using deep learning: Datasets, performance measurements, challenges and opportunities," *IEEE Access*, vol. 8, pp. 110575–110597, Jun. 2020.

[10] T. T. Do, Y. Zhou, and H. Zheng, "Attention cost-sensitive deep learning- based approach for skin cancer detection and classification," *Bioengineering (Basel)*, vol. 9, no. 12,p. 745, Dec. 2022