Volume XI Issue 3 Year 2026 | Page 742-753 | ISSN: 2527-9866
Received: 17-06-2026 | *Revised: 17-07-2026* | *Accepted: 30-07-2026*

Optimization of MobileNet Architecture with Ghost Module for Dental Enamel Caries Classification

M. Zaky Naufal Farisky[1], Yohannes[2]
[1,2] Multi Data Palembang University, Palembang, South Sumatra, Indonesia, 30113
e-mail: mzakynaufalfarisky_2226250122@mhs.mdp.ac.id[1], yohannesmasterous@mdp.ac.id[2]
*Correspondence: mzakynaufalfarisky_2226250122@mhs.mdp.ac.id

Abstract: *Dental and Oral diseases, particularly dental caries, represent a global health issue that requires early identification at the enamel stage to prevent further demineralization and damage. The utilization of artificial intelligence technology through Convolutional Neural Networks (CNN) has been widely applied for medical image analysis. however, complex conventional models often incur high computational loads. Therefore, this study aims to implement and evaluate MobileNet variants (V1, V2, V3, and V4) optimized using the Ghost Module to classify dental caries images. The integration of the Ghost Module aims to mitigate feature redundancy and enhance feature representation without compromising image extraction quality. A dataset of 2,000 clinical dental images from the public "Caries-Spectra", categorized as advanced enamel caries, early-stage enamel caries, and no enamel caries, was curated and expanded to 12,000 images using preprocessing and augmentation. The dataset splitting was executed with a final learning proportion of 80% training data, 10% Validation data, and 10% testing data. The image preprocessing utilized Histogram Equalization, CLAHE, and Adaptive Thresholding methods. The overall hybrid architecture was then evaluated based on performance metrics (such as Accuracy, Precision, Recall, and F1-Score). The experimental results demonstrate that the integration of the Ghost Module consistently enhances the classification accuracy across all MobileNet variants. The proposed Hybrid MobileNetV1 with Ghost Module achieved the highest performance, securing an overall accuracy of 96.83%, with well-balanced precision and recall across all enamel caries stages.*

Keywords: *Dental Enamel Caries, MobileNet, Ghost Module, Image Classification*

1. Introduction

Dental and oral diseases remain one of the major global health challenges, significantly affecting quality of life and imposing substantial economic burdens worldwide. According to the Global Oral Health Status Report published by the World Health Organization (WHO) in 2022, nearly half of the world's population, or approximately 3.5 billion people, suffer from oral diseases. Among these conditions, dental caries in permanent teeth is one of the most prevalent diseases globally, affecting more than 2 billion people [1]. This study specifically focuses on enamel caries, which represents the early stage of demineralization and damage to the outer enamel layer of the tooth. Early detection at the enamel stage is crucial because structural damage at this phase is the fundamental stage preceding cavitation. Therefore, timely preventive interventions can reduce disease progression and preserve tooth structure [2].

The advancement of artificial intelligence technology has contributed significantly to the transformation of medical image analysis through the application of Deep Learning algorithms, particularly Convolutional Neural Networks (CNNs). In dentistry, intraoral images enable digital analysis of carious lesions to provide additional diagnostic information for medical practitioners [3]. Various CNN architectures, such as ResNet-50, EfficientNet-B4, and MobileNet, have been widely utilized to process these images. However, complex architectures such as ResNet-50 possess a large number of parameters and layers, resulting in high computational costs [4]. Consequently, lightweight architectures have become increasingly relevant to address computational resource limitations, especially when diagnostic systems are deployed on clinical devices with standard or low specifications, or operate independently on mobile devices [5]. Previous studies have also demonstrated that lightweight architectures can effectively detect dental caries while maintaining a significantly smaller model size [6].

MobileNet was selected as the backbone CNN architecture in this study due to its streamlined feature extraction mechanism. This architecture employs depthwise separable convolution to process spatial features effectively [7]. However, despite this architectural advantage, standard MobileNet variants still generate substantial feature redundancy during pointwise convolution operations. The reliability of this architecture is reflected in its continuous evolution, beginning with the introduction of the inverted residual with linear bottleneck structure in MobileNetV2 [8]. MobileNetV2 has also demonstrated excellent performance in various real-world applications, including real-time object detection systems [9]. This evolution continued with the integration of automated optimization techniques in MobileNetV3 [10], and eventually led to the introduction of the Universal Inverted Bottleneck in the latest generation, MobileNetV4 [11].

To address the issue of feature redundancy, this study proposes the integration of the Ghost Module into the MobileNet architecture. As a fundamental optimization component, the Ghost Module limits the generation of redundant feature maps by applying cheap linear operations to intrinsic features [12]. This optimization strategy addresses the limitations of conventional lightweight architectures [13], [14], which often perform excessive channel expansion [15], while also optimizing the pointwise convolution operation in MobileNet, which typically consumes the majority of computational resources [16]. The advantages of integrating Ghost Module into various backbone architectures have been empirically validated. Its application in multi-branch models has been shown to reduce model size by more than 90% while maintaining or even improving classification accuracy [17]. In the medical imaging domain, Ghost Module integration has successfully reduced model parameters substantially while preserving feature extraction quality [18]. Furthermore, studies applying Ghost Module to intraoral dental images have demonstrated its effectiveness in reducing computational costs, achieving reductions of 44.19% in model parameters and 29.01% in computational complexity while maintaining robust feature extraction capabilities [19]. Previous studies have also shown that integrating Ghost Module into lightweight architectures consistently outperforms baseline models in terms of classification accuracy [20]. Therefore, this study focuses on integrating Ghost Module into the MobileNet architecture for dental caries image classification, aiming to evaluate the performance of an optimized lightweight CNN model for intraoral image-based diagnostic systems.

2. Literature Review

This study refers to the research conducted by Ashi [21] on enamel caries detection. In that study, Ashi developed an automated deep learning-based classification model designed to support dental public health by utilizing a lightweight architecture. The proposed model employed an attention-guided feature fusion approach that combined Modified MobileNetV2 and Modified EfficientNetB0 to enhance feature extraction capabilities. Using a public dataset consisting of 2,000 clinical dental images covering advanced enamel caries, early-stage enamel caries, and no enamel caries categories, the model achieved a highest accuracy of 96.92%.

While a recent study [21] demonstrated exceptional classification accuracy (96.92%) using an attention-guided feature fusion approach, their methodology relies on the parallel execution of a dual-backbone ensemble (Modified MobileNetV2 combined with Modified EfficientNetB0). The integration of two separate deep learning architectures invariably introduces significant structural complexity and redundancy during the initial feature extraction phase. The primary novelty of this present study lies in establishing a streamlined, single-stream alternative that achieves comparable diagnostic accuracy. Instead of employing a multi-branch ensemble, this study structurally optimizes a standalone MobileNet backbone through the integration of the Ghost Module. The Ghost Module acts as a regularizer that systematically mitigates the generation of redundant feature maps inherent in standard pointwise convolutions. This approach empirically proves that an optimized single-backbone network can effectively differentiate enamel caries stages (achieving 96.83% accuracy) without necessitating the architectural overhead required by existing fusion-based frameworks.

3. Methods

This study was conducted through experiments on four lightweight deep learning architectures, namely MobileNetV1, MobileNetV2, MobileNetV3, and MobileNetV4, for the classification of intraoral images from the "Caries-Spectra: A Dataset of Enamel Caries" dataset. The experiments were divided into 14 testing scenarios. These scenarios consisted of seven experiments using the standard models (MobileNetV1, MobileNetV2, MobileNetV3-Small, MobileNetV3-Large, MobileNetV4-Small, MobileNetV4-Medium, and MobileNetV4-Large) and seven additional experiments using models optimized with the Ghost Module.

Table 1. Experimental Scenario

	Scenarios	Architecture	Optimization
1.	Scenario 1	MobileNetV1	None
2.	Scenario 2	MobileNetV2	None
3.	Scenario 3	MobileNetV3-Small	None
4.	Scenario 4	MobileNetV3-Large	None
5.	Scenario 5	MobileNetV4-Small	None
6.	Scenario 6	MobileNetV4-Medium	None
7.	Scenario 7	MobileNetV4-Large	None
8.	Scenario 8	MobileNetV1	Ghost Module
9.	Scenario 9	MobileNetV2	Ghost Module
10.	Scenario 10	MobileNetV3-Small	Ghost Module
11.	Scenario 11	MobileNetV3-Large	Ghost Module
12.	Scenario 12	MobileNetV4-Small	Ghost Module
13.	Scenario 13	MobileNetV4-Medium	Ghost Module
14.	Scenario 14	MobileNetV4-Large	Ghost Module

The systematic steps of the general testing scenario in this study are illustrated in Figure 1.

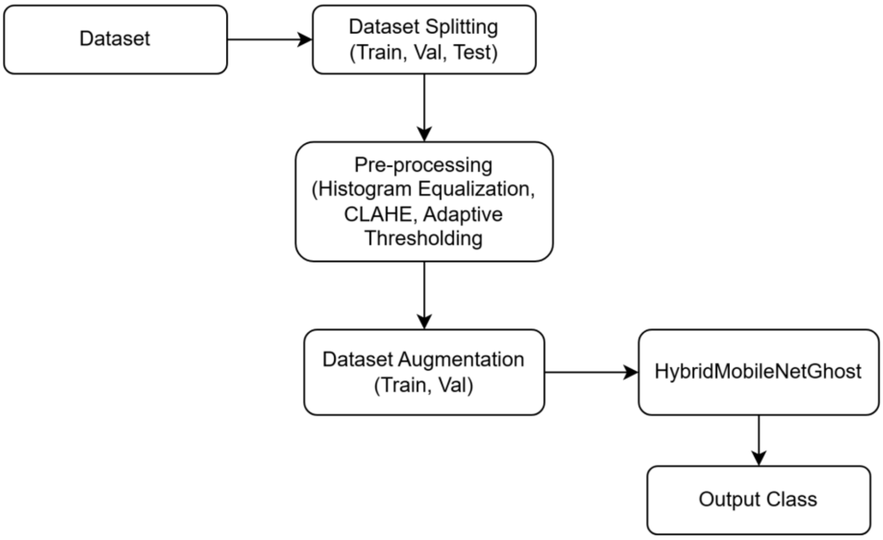


Figure 1. General Experimental Scenarios

A.Data Description

The dataset used in this study is the "Caries-Spectra: A Dataset of Enamel Caries" intraoral dental image dataset collected from medical clinics in Rajshahi and Dhaka, Bangladesh [22]. The dataset consists of 2,000 low-resolution images that are categorized into three classes: AdvanceEnamel_Caries (800 images), EarlyStageEnamel_Caries (800 images), and NoEnamel_Caries (400 images). Examples of images from the dataset are presented in Figure 2.

(a) **(b)** **(c)**

Figure 2. (a) NoEnamel_Caries, **(b)** EarlyStageEnamel_Caries, **(c)** AdvanceEnamel_Caries

B.Dataset Splitting and Pre-processing

From the 2,000 original images, the dataset was first divided into three subsets. A total of 1,200 images were allocated exclusively as testing data without any augmentation. Meanwhile, the remaining 800 images were used for the learning process, consisting of 640 images for training and 160 images for validation. This dataset partition followed the baseline setting used in the study [21].

To enhance the visibility of enamel demineralization textures, the training and validation images underwent several pre-processing steps, including Histogram Equalization, CLAHE (Contrast Limited Adaptive Histogram Equalization), and Adaptive Thresholding. Subsequently, the pre-processed images were converted into RGB format to comply with the input specifications of the MobileNet architecture.

C.Data Augmentation

Data augmentation was applied only to the training and validation datasets to improve the model's generalization ability and reduce the risk of overfitting. The augmentation techniques employed in this study included Random Rotation, Horizontal Flipping, and Vertical Flipping.

Through this process, the number of training images increased from 640 to 9,600 images, while the number of validation images increased from 160 to 1,200 images. Meanwhile, the testing dataset remained unchanged at 1,200 original images without augmentation. Consequently, the final dataset consisted of 12,000 images, comprising 80% training data (9,600 images), 10% validation data (1,200 images), and 10% testing data (1,200 original images without augmentation).

D.MobileNet Architecture

MobileNet is a lightweight Convolutional Neural Network (CNN) architecture specifically designed to address the challenges of high computational cost and memory requirements, making it highly suitable for deployment on resource-constrained devices. The evolution of this architecture began with MobileNetV1, which introduced Depthwise Separable Convolution to replace standard convolutional layers [7]. Subsequently, MobileNetV2 was developed based on the inverted residual structure with a linear bottleneck, eliminating non-linear activation functions in narrow layers to prevent the loss of crucial information [8].

The next generation, MobileNetV3, combines manual design with automated optimization through Hardware-aware Network Architecture Search (NAS), supported by the h-swish activation function and the Squeeze-and-Excitation (SE) module to achieve a balance between performance and latency [10]. Finally, MobileNetV4 was designed as a universal architecture that introduces the Universal Inverted Bottleneck block to operate efficiently across various mobile hardware platforms [11]. In this study, all four MobileNet variants are employed as backbone networks to extract fundamental feature maps from enamel caries images.

E.Ghost Module

Ghost Module was introduced by Han [23] as an optimization component to improve the efficiency of neural networks by exploiting the redundancy that frequently occurs in the generation of feature maps. Instead of producing all features using computationally expensive standard convolutions, this module divides the process into two main stages [23]. The first stage is the Primary Convolution, where standard convolution is applied to generate a small number of intrinsic feature maps while maintaining a low computational cost.

The second stage involves Cheap Operations, in which the intrinsic feature maps are processed using a series of simple linear transformations to generate additional features known as ghost feature maps. Finally, the module concatenates the original intrinsic feature maps with the generated ghost feature maps. Through this approach, Ghost Module produces the same number of output feature maps as standard convolution while significantly reducing the total number of parameters and floating-point operations (FLOPs) without substantially compromising the quality of feature representation [23].

F.Hybrid MobileNetGhost

The proposed system integrates MobileNet acting as a feature extractor backbone. The extracted feature maps are passed into the Ghost Module, which applies a Primary Convolution to extract intrinsic features and utilizes a Cheap Operation to generate ghost features with minimal computational cost [23]. The maps are concatenated, reduced via Adaptive Average Pool 2D, and fed into the Classification Head.

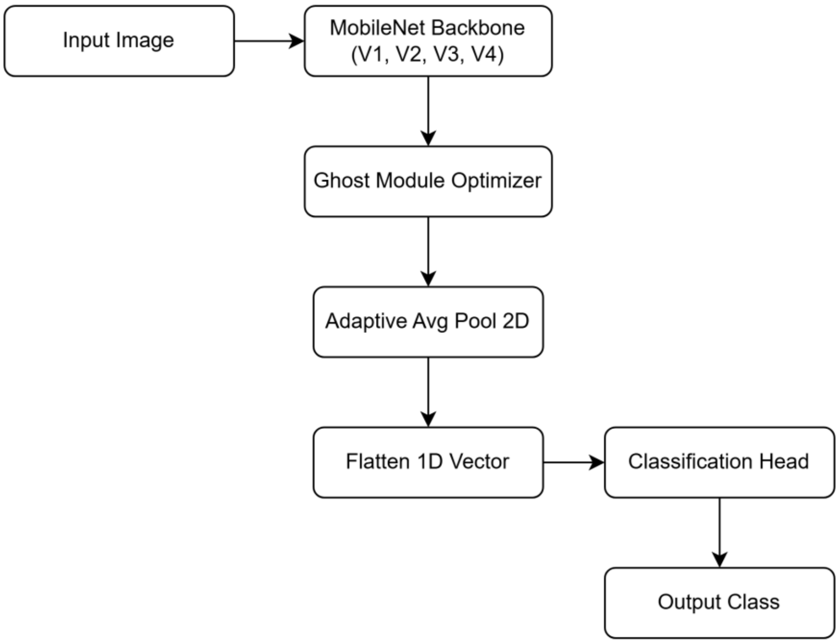


Figure 3. HybridMobileNetGhost Scenarios

G.Model Training

The training process of the HybridMobileNetGhost architecture was performed using the Cross-Entropy Loss function and the Adam optimizer with a learning rate of 0.0001. The training procedure was conducted iteratively for a total of 30 epochs.

H.Model Evaluation

In this stage, the previously created system will be tested on the test data. After the testing phase, the test results are calculated to determine the success rate of the method used with the Confusion Matrix to calculate precision, recall, accuracy, and F1-score values, as shown in Equations (1), (2), (3), and (4).

$$Precision = \frac{TP}{TP + FP} \qquad (1)$$

$$Recall = \frac{TP}{TP + FN} \qquad (2)$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \qquad (3)$$

$$F1\,Score = \frac{2\ x\ Precision\ x\ Recall}{Precision\ +\ Recall} \qquad (4)$$

4. Results and Discussion

The performance results of the 14 experimental scenarios comparing standard MobileNet and MobileNet architectures with Ghost Module are presented in Tables 2 and 3.

Table 2. Precision, Recall, and F1-Score for Each Class Across Experimental Scenarios

	Scenarios	Class	Precision	Recall	F1-Score
1	Scenario 1	AdvanceEnamel_Caries	0.9604	0.9604	0.9604
		EarlyStageEnamel_Caries	0.9464	0.9563	0.9513
		NoEnamel_Caries	0.9702	0.9500	0.9600
2	Scenario 2	AdvanceEnamel_Caries	0.9289	0.9521	0.9403
		EarlyStageEnamel_Caries	0.9332	0.9313	0.9322
		NoEnamel_Caries	0.9476	0.9042	0.9254
3	Scenario 3	AdvanceEnamel_Caries	0.9562	0.9542	0.9552
		EarlyStageEnamel_Caries	0.9572	0.9313	0.9440
		NoEnamel_Caries	0.9252	0.9792	0.9514
4	Scenario 4	AdvanceEnamel_Caries	0.9596	0.9396	0.9495
		EarlyStageEnamel_Caries	0.9261	0.9667	0.9460
		NoEnamel_Caries	0.9694	0.9250	0.9467
5	Scenario 5	AdvanceEnamel_Caries	0.9505	0.9604	0.9554
		EarlyStageEnamel_Caries	0.9466	0.9229	0.9346
		NoEnamel_Caries	0.9393	0.9667	0.9528
6	Scenario 6	AdvanceEnamel_Caries	0.9739	0.9313	0.9521
		EarlyStageEnamel_Caries	0.9050	0.9729	0.9378
		NoEnamel_Caries	0.9911	0.9292	0.9591
7	Scenario 7	AdvanceEnamel_Caries	0.9061	0.9646	0.9344
		EarlyStageEnamel_Caries	0.9414	0.9042	0.9224
		NoEnamel_Caries	0.9868	0.9375	0.9615
8	Scenario 8	AdvanceEnamel_Caries	0.9706	0.9646	0.9676
		EarlyStageEnamel_Caries	0.9593	0.9812	0.9701
		NoEnamel_Caries	0.9828	0.9500	0.9661
9	Scenario 9	AdvanceEnamel_Caries	0.9468	0.9646	0.9556
		EarlyStageEnamel_Caries	0.9561	0.9521	0.9541
		NoEnamel_Caries	0.9700	0.9417	0.9556
10	Scenario 10	AdvanceEnamel_Caries	0.9722	0.9458	0.9588
		EarlyStageEnamel_Caries	0.9450	0.9667	0.9557
		NoEnamel_Caries	0.9545	0.9625	0.9585
11	Scenario 11	AdvanceEnamel_Caries	0.9825	0.9375	0.9595
		EarlyStageEnamel_Caries	0.9214	0.9771	0.9484
		NoEnamel_Caries	0.9657	0.9375	0.9514
12	Scenario 12	AdvanceEnamel_Caries	0.9624	0.9604	0.9614
		EarlyStageEnamel_Caries	0.9481	0.9521	0.9501
		NoEnamel_Caries	0.9582	0.9542	0.9562
13	Scenario 13	AdvanceEnamel_Caries	0.9473	0.9729	0.9599
		EarlyStageEnamel_Caries	0.9559	0.9479	0.9519
		NoEnamel_Caries	0.9870	0.9500	0.9682
14	Scenario 14	AdvanceEnamel_Caries	0.9529	0.9688	0.9607
		EarlyStageEnamel_Caries	0.9603	0.9583	0.9593
		NoEnamel_Caries	0.9828	0.9542	0.9683

Table 2 presents the model performance in distinguishing each enamel caries category based on Precision, Recall, and F1-Score metrics. Further analysis focuses on the best-performing model, namely Scenario 8 (MobileNetV1 + Ghost Module), which demonstrates balanced performance across all testing classes. For the EarlyStageEnamel_Caries class, the model achieved a Recall of 0.9812 and an F1-Score of 0.9701. The high Recall value indicates that the model has an excellent ability to detect early-stage enamel caries, thereby minimizing the risk of missed detections.

In addition, the model exhibited outstanding performance in the NoEnamel_Caries class, achieving the highest Precision value of 0.9828, which indicates a low probability of healthy teeth being misclassified as caries. For the AdvanceEnamel_Caries class, the model achieved an F1-Score of 0.9676, demonstrating its effectiveness in identifying severe enamel lesions. These results confirm that the integration of the Ghost Module enhances the model's ability to classify different levels of enamel caries severity accurately and consistently.

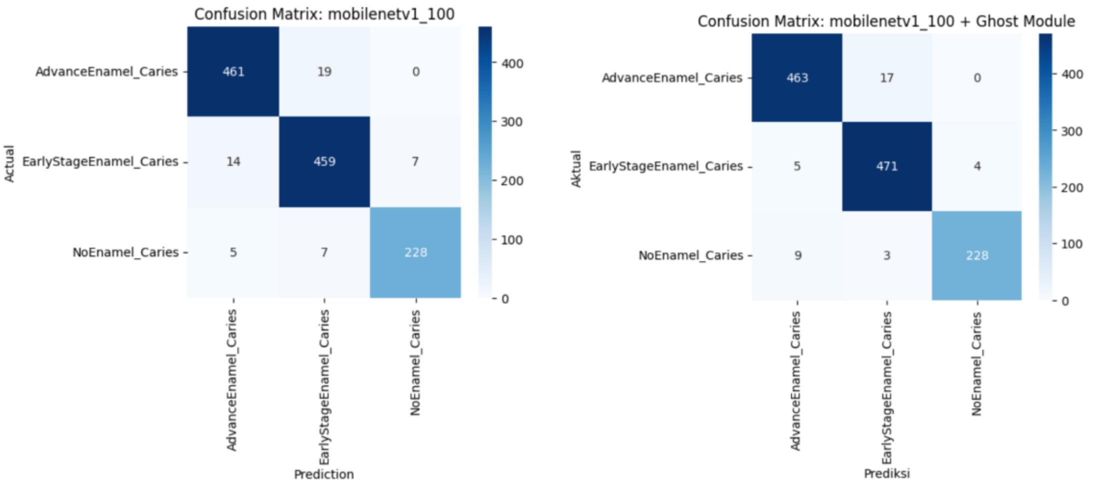
Table 3. Overall Accuracy Comparison of MobileNet Variants with and without Ghost Modules

	Scenarios	Model	Overall Accuracy
1	Scenario 1	MobileNetV1	0.9567
2	Scenario 2	MobileNetV2	0.9342
3	Scenario 3	MobileNetV3-Small	0.9500
4	Scenario 4	MobileNetV3-Large	0.9475
5	Scenario 5	MobileNetV4-Small	0.9467
6	Scenario 6	MobileNetV4-Medium	0.9475
7	Scenario 7	MobileNetV4-Large	0.9350
8	Scenario 8	MobileNetV1 + Ghost Module	**0.9683**
9	Scenario 9	MobileNetV2 + Ghost Module	0.9550
10	Scenario 10	MobileNetV3-Small + Ghost Module	0.9575
11	Scenario 11	MobileNetV3-Large + Ghost Module	0.9533
12	Scenario 12	MobileNetV4-Small + Ghost Module	0.9558
13	Scenario 13	MobileNetV4-Medium + Ghost Module	0.9583
14	Scenario 14	MobileNetV4-Large + Ghost Module	**0.9617**

The performance disparity among the MobileNet variants can be attributed to the correlation between architectural complexity and the characteristics of the intraoral dataset. MobileNetV4 variants incorporate highly complex Universal Inverted Bottleneck (UIB) structures optimized for large-scale, highly variable datasets. When applied to a narrowly focused and localized clinical dataset like enamel caries, such high-capacity architectures are highly susceptible to overfitting, capturing irrelevant clinical noise rather than generalizable lesion features. Conversely, MobileNetV1 possesses a straightforward, linear depthwise separable structure. The integration of the Ghost Module into MobileNetV1 does not merely reduce feature redundancy, it acts as a strict structural regularizer. By restricting the generation of deep feature maps to cheap linear operations, the Ghost Module stabilizes the optimization pathway of the simpler V1 baseline, preventing overfitting and yielding the most robust feature extraction for early-stage and advanced enamel caries. This theoretical dynamic explains why the simpler MobileNetV1 + Ghost Module framework outperforms the more structurally advanced MobileNetV4 variants.

To fully comprehend the performance dynamics presented in Table 3, two architectural phenomena must be objectively analyzed: the baseline superiority of the simpler MobileNetV1 and the extreme regularizing impact of the Ghost Module on complex variants. First, regarding the standard architectures, MobileNetV4 variants are engineered with high-capacity Universal Inverted Bottleneck (UIB) structures optimized for extracting complex, multi-scale semantic geometries from massive datasets [11]. However, the classification of dental enamel caries from localized intraoral images relies heavily on identifying subtle, low-level texture degradations (such as demineralization spots) rather than complex object geometries. Consequently, the massive parameter capacity of MobileNetV4 acts as a detriment. The model overfits by capturing irrelevant clinical noise and illumination variances, explaining why the standard MobileNetV4-Large, which yielded an accuracy of 0.9350, severely underperformed compared to the simpler MobileNetV1 that achieved a higher accuracy of 0.9567. MobileNetV1, utilizing straightforward depthwise separable convolutions without residual bottlenecks [7], inherently possesses a receptive field and capacity strictly aligned with the dataset's low-level feature complexity.

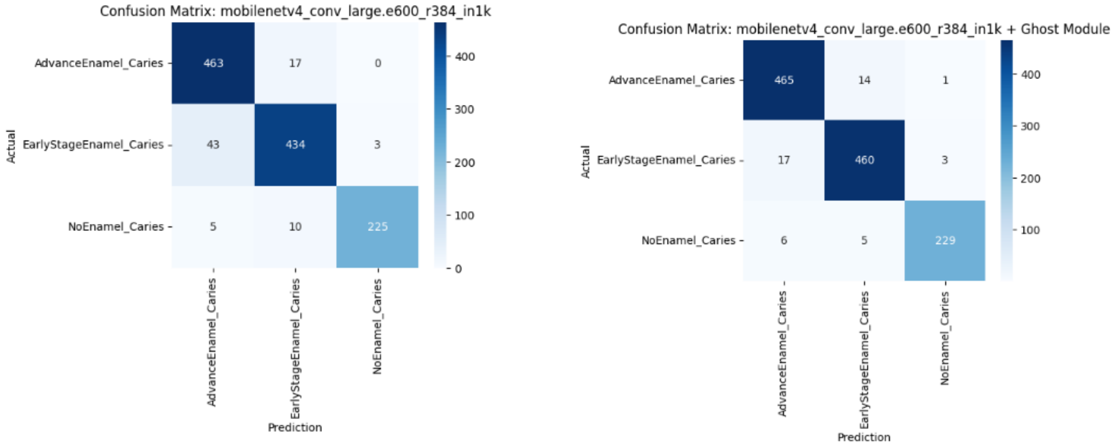
Furthermore, the integration of the Ghost Module fundamentally rescues over-parameterized models while perfecting simpler ones. For MobileNetV4-Large, the Ghost Module structurally restricts the network's dense pointwise operations, forcing the generation of redundant maps through cheap linear transformations [23]. This constraint acts as an aggressive structural regularizer, mitigating the variant's tendency to overfit and resulting in the most significant accuracy surge in the study, increasing from 0.9350 to 0.9617. For MobileNetV1, which already operates at the optimal feature scale, the Ghost Module rigorously filters out redundant channel information without discarding intrinsic texture data. This optimization yields exceptionally crisp feature maps, allowing the Hybrid MobileNetV1 + Ghost Module to achieve the absolute highest overall performance with an accuracy score of 0.9683, particularly excelling in distinguishing the critical classification threshold between Early-Stage and No Enamel Caries.



(a) **(b)**

Figure 4. Confusion Matrix of: **(a)** MobileNetV1, **(b)** MobileNetV1 + Ghost Module

To provide a transparent evaluation of the best-performing architecture, Figure 4 illustrates the confusion matrices for the proposed MobileNetV1 with Ghost Module and its standard baseline. The standard MobileNetV1 accurately classified 461 AdvanceEnamel, 459 EarlyStage, and 228 NoEnamel cases, but exhibited difficulty with class overlap, erroneously predicting 14 EarlyStage lesions as advanced. Upon integrating the Ghost Module, the network's diagnostic precision improved noticeably, most prominently in the EarlyStageEnamel_Caries category, where true positive predictions surged to 471 cases and severe misclassifications dropped to just 5. While the detection of healthy enamel (NoEnamel_Caries) remained consistently stable at 228 correct predictions across both configurations, the Ghost Module effectively sharpened the network's discriminative ability to isolate early-stage demineralization from advanced cavities, solidifying its position as the most clinically robust architecture in this study.



(a) **(b)**

Figure 5. Confusion Matrix of: **(a)** MobileNetV4-Large, **(b)** MobileNetV4-Large + Ghost Module

To provide a clearer understanding of the impact of Ghost Module integration, the Confusion Matrix of the model with the most significant improvement, namely the comparison between the standard MobileNetV4-Large and MobileNetV4-Large integrated with the Ghost Module, is presented in Figure 5. Based on the testing results, the standard model in Figure 5(a) correctly classified 463 AdvanceEnamel_Caries images, 434 EarlyStageEnamel_Caries images, and 225 NoEnamel_Caries images. However, the model still exhibited a relatively high misclassification rate, particularly for the EarlyStageEnamel_Caries class, where 43 images were incorrectly classified as AdvanceEnamel_Caries.

After the integration of the Ghost Module, the model performance improved across all classes, as shown in Figure 5(b). The number of correctly classified images increased to 465 for AdvanceEnamel_Caries, 460 for EarlyStageEnamel_Caries, and 229 for NoEnamel_Caries. In addition, the misclassification of EarlyStageEnamel_Caries images decreased significantly from 43 to only 17 images. These results demonstrate that the Ghost Module effectively enhances the model's performance, enabling it to distinguish the severity levels of enamel caries more accurately than the standard architecture.

While the proposed HybridMobileNetGhost architecture demonstrates high diagnostic metrics, this study acknowledges certain methodological constraints. First, the performance metrics were derived from fixed dataset splits without repeated experimental runs. However, the uniform accuracy improvement observed across all seven evaluated MobileNet variants strongly indicates a robust structural advantage rather than a single-run anomaly, although future multi-run evaluations are necessary to establish formal statistical significance. Second, the image preprocessing pipeline, which comprises Histogram Equalization, CLAHE, and Adaptive Thresholding, was implemented as a unified stack directly adopted from the established methodology by Ashi [21] to ensure a standardized baseline evaluation for the Caries-Spectra dataset. By maintaining this methodological consistency, the isolated effects of each enhancement technique were not evaluated through an ablation study, as the primary scope of this research remained on the architectural optimization of the Ghost Module. Future research will prioritize rigorous cross-validations and detailed preprocessing ablation studies to further quantify these architectural and morphological contributions [24], [25].

5. Conclusions

This study successfully implemented and evaluated the optimization of MobileNet variants by integrating the Ghost Module for dental enamel caries classification. The experimental results demonstrate that the integration of Ghost Module consistently improves the classification performance of all MobileNet variants compared to their standard architectures. The best performing model was MobileNetV1 integrated with Ghost Module, achieving an accuracy of 0.9683. Meanwhile, the most significant performance improvement was observed in MobileNetV4-Large, whose accuracy increased from 0.9350 to 0.9617 after being optimized with Ghost Module.

In addition to improving accuracy, the confusion matrix analysis indicates that the Ghost Module functions as an effective structural regularizer, significantly reducing classification error rates by mitigating feature redundancy. By achieving highly competitive diagnostic performance through a streamlined, single-stream architecture rather than relying on complex multi-backbone ensembles, the proposed model demonstrates strong potential for practical deployment in intraoral image-based diagnostic systems. These findings also provide opportunities for future research, where the model can be evaluated in real-time clinical environments and further developed using object detection approaches to localize caries regions more accurately.

References

[1] World Health Organization, Global Oral Health Status Report: Towards Universal Health Coverage for Oral Health by 2030. Geneva, Switzerland: World Health Organization, 2022. Available: https://www.who.int/publications/i/item/9789240061484

[2] F. Y. I. Asiri, "Automated Classification of Enamel Caries from Intraoral Images Using Deep Learning Models: A Diagnostic Study," Journal of Clinical Medicine, vol. 14, no. 24, p. 8959, 2025. DOI: https://doi.org/10.3390/jcm14248959

[3] J. Kühnisch, O. Meyer, M. Hesenius, R. Hickel, and V. Gruhn, "Caries Detection on Intraoral Images Using Artificial Intelligence," Journal of Dental Research, vol. 101, no. 2, pp. 158–165, 2022. DOI: https://doi.org/10.1177/00220345211032524

[4] M. Oktiana, M. Fitria, H. Habibie, K. Saddami, Y. Elma, H. Rahayu, S. Janura, R. Putri, and R. Novita, "Performance Analysis of ResNet-50, EfficientNet B4, and MobileNet V2 for Dental Caries Classification," JOIV : International Journal on Informatics Visualization, vol. 10, no. 1, pp. 10–18, 2026. DOI: https://doi.org/10.62527/joiv.10.1.3713

[5] A. Garg, J. Lu, and A. Maji, "Towards Earlier Detection of Oral Diseases on Smartphones Using Oral and Dental RGB Images," arXiv preprint arXiv:2308.15705, 2023. DOI: https://doi.org/10.48550/arXiv.2308.15705

[6] A. F. Boy, A. Akhyar, T. Y. Arif, and S. Syahrial, "Development of an Artificial Intelligence Model Based on MobileNetV3 for Early Detection of Dental Caries Using Smartphone Images: A Preliminary Study," Advances in Science and Technology Research Journal, vol. 19, no. 4, pp. 109–116, 2025. DOI: https://doi.org/10.12913/22998624/200308

[7] Tedyyana, A., and O. Ghazali. "W. Purbo, O.(2024). Enhancing intrusion detection system using rectified linear unit function in pigeon inspired optimization algorithm." *IAES International Journal of Artificial Intelligence (IJ-AI)* 13.2: 1526-1534.

[8] M. Sandler, A. Howard, M. Zhu, A. Zhmoginov, and L.-C. Chen, "MobileNetV2: Inverted Residuals and Linear Bottlenecks," in Proc. IEEE/CVFConference on Computer Vision and Pattern Recognition, pp. 4510–4520, 2018. DOI: https://doi.org/10.1109/CVPR.2018.00474

[9] U. Seidaliyeva, D. Akhmetov, L. Ilipbayeva, and E. T. Matson, "Real-Time and Accurate Drone Detection in a Video with a Static Background," Sensors, vol. 20, no. 14, p. 3856, 2020. DOI: https://doi.org/10.3390/s20143856

[10] A. Howard, M. Sandler, G. Chu, L.-C. Chen, B. Chen, M. Tan, W. Wang, Y. Zhu, R. Pang, V. Vasudevan, Q. V. Le, and H. Adam, "Searching for MobileNetV3," in Proc. IEEE/CVF International Conference on Computer Vision (ICCV), pp. 1314–1324, 2019. DOI: https://doi.org/10.1109/ICCV.2019.00140

[11] D. Qin, C. Leichner, M. Delakis, M. Fornoni, S. Luo, F. Yang, W. Wang, C. Banbury, C. Ye, B. Akin, V. Aggarwal, T. Wang, B. Li, S. Choudhary, M. Zhu, and W. Wang, "MobileNetV4: Universal Models for the Mobile Ecosystem," Computer Vision – ECCV 2024, 2024. DOI: https://doi.org/10.1007/978-3-031-73661-2_5

[12] Y. Tang, K. Han, J. Guo, C. Xu, C. Xu, and Y. Wang, "GhostNetV2: Enhance Cheap Operation with Long-Range Attention," Proceedings of the 36th International Conference on Neural Information Processing Systems (NeurIPS), 2022. DOI: https://dl.acm.org/doi/10.5555/3600270.3600994

[13] Yohannes, M. E. A. Rivan, S. Devella, and Tinaliah, "A novel optimization strategy for CNN models in Palembang Songket motif recognition," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 16, no. 1, 2025, DOI: https://doi.org/10.14569/IJACSA.2025.0160180

[14] Yohannes, M. E. A. Rivan, S. Devella, and Tinaliah, "Ghost-Vanilla feature maps: A Novel Hybrid Architecture for Efficient Fine-Grained Songket Motif Classification," International Journal of Advanced Computer Science and Applications (IJACSA), vol. 17, no. 1, 2026, DOI: https://doi.org/10.14569/IJACSA.2026.0170129

[15] K. Hong, G.-H. Kim, and E. Kim, "GhostNeXt: Rethinking Module Configurations for Efficient Model Design," Applied Sciences, vol. 13, no. 5, p. 3301, 2023. DOI: https://doi.org/10.3390/app13053301

[16] X. Xu, Y. Ding, Z. Lv, Z. Li, and R. Sun, "Optimized Pointwise Convolution Operation by Ghost Blocks," Electronic Research Archive, vol. 31, no. 6, pp. 3187–3199, 2023. DOI: https://doi.org/10.3934/era.2023161

[17] A. Liu, T. Su, H. Sun, and X. Zhang, "Inc-ghostNet: A Lightweight Multi-Branch Ghost Module Fusion Network for Image Classification," Chinese Journal of Electronics, vol. 34, no. 6, pp. 1899–1902, 2025. DOI: https://doi.org/10.23919/CJE.2023.00.402

[18] L. Yang, H. Cai, X. Luo, J. Wu, R. Tang, Y. Chen, and W. Li, "A Lightweight Neural Network for Lung Nodule Detection Based on Improved Ghost Module," Quantitative Imaging in Medicine and Surgery, 2023. DOI: https://doi.org/10.21037/qims-21-1182

[19] J. Hong, S. Hu, Z. Tang, H. Li, Z. Zhao, and J. Yi, "Develop an Esthetic-Zonal Non-Invasive Periodontal Assessment Tool Based on YOLOv8 and Intraoral Images," PeerJ Computer Science, vol. 11, p. e3229, 2025. DOI: https://doi.org/10.7717/peerj-cs.3229

[20] M. A. Mulia, S. N. Abdillah, R. Rachman, and A. Yuniarti, "MobileGhost: Lightweight Ghost Module for Image Classification," 2024 Ninth International Conference on Informatics and Computing (ICIC), 2024. DOI: https://doi.org/10.1109/ICIC64337.2024.10957428

[21] H. Ashi, "From Clinic to Community: An Interpretable Artificial Intelligence Framework for Enamel Caries Detection to Support Public Health Dentistry," European Journal of Dentistry, 2026. DOI: https://doi.org/10.1055/s-0046-1816061

[22] G. M. S. Himel, M. M. Islam, and U. H. Hannan, "Caries-Spectra: A Dataset of Enamel Caries," Mendeley Data, ver. 2, 2023. DOI: https://doi.org/10.17632/9jnf2jvghy.2

[23] K. Han, Y. Wang, Q. Tian, J. Guo, C. Xu, and C. Xu, "GhostNet: More Features From Cheap Operations," 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), pp. 1580–1589, 2020. DOI: https://doi.org/10.1109/CVPR42600.2020.00165

[24] I Gusti Agung Putu Mahendra, A. Julianto, and A. Tedyyana, "Klasifikasi Citra Topeng Bali Berdasarkan Karakteristik Visual Menggunakan Arsitektur EfficientNetV2," Jurnal RESISTOR (Rekayasa Sistem Komputer), vol. 8, no. 3, pp. 122–132, Dec. 2025, doi: https://doi.org/10.31598/jurnalresistor.v8i3.1926.

[25] I. G. A. Putu Mahendra, Muhammad Ikhsan Wibowo, and Zuliar Efendi, "Peningkatan Akurasi Klasifikasi Ikan kepe-kepe (Famili Chaetodontidae) dengan EfficientNetV2 dan Bayesian Hyperparameter Tuning," *Jurnal Sistem Komputer dan Informatika (JSON)*, vol. 7, no. 2, pp. 260–270, Dec. 2025, doi: https://doi.org/10.30865/json.v7i2.9028.