ACIS
Advances in Computational and Intelligent Systems
Vol. 2, No. 1, June 2026, pp. 1~9
ISSN: 3093-7388, DOI: https://doi.org/10.56313/acis.v2i1.1

Lightweight CNN Architecture for Gram-Stained Bacterial Image Classification

Badaruddin Muhammad[1], Amran Abdul Hadi[1], Kamarul Hawari Ghazali[1,2], Liu Xinnie[3]
[1]Faculty of Electrical and Electronics Engineering Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan Campus, 26600 Pekan, Pahang, Malaysia.
[2]Center for Advanced Industrial Technology, Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan Campus, 26600 Pekan, Pahang, Malaysia
[2]School of Information, Xian University of Finance and Economics, Xian, China

Article Info

Article history:

Received April, 25 2026
Revised May, 9 2026
Accepted May, 18 2026

Keywords:

Lightweight CNN
Gram-stained bacterial classification
floating-point operations

ABSTRACT

Accurate classification of Gram-stained bacterial images is essential for timely clinical decision-making. Traditional image analysis methods rely on handcrafted features and machine learning classifiers, which are often limited in robustness and scalability. Deep learning, particularly convolutional neural networks has improved accuracy in biomedical imaging but existing architectures such as ResNet, DenseNet, and VGG are computationally demanding, with tens of millions of parameters and billions of floating-point operations. This complexity restricts their use in laboratories with limited resources and in real-time applications. This study introduces a lightweight CNN tailored for Gram-stained bacterial classification. The model integrates a compact stem, three depthwise-separable convolutional blocks, and a global average pooling layer, reducing parameters and FLOPs while maintaining strong feature extraction capability. An optional attention module was assessed in ablation experiments to evaluate its contribution. Experiments were conducted on six bacterial species and results were compared with six deep learning baselines, including ResNet-18, MobileNetV2, VGG16, DenseNet-121, InceptionV3, and EfficientNet-B0. The proposed CNN achieved an average accuracy of 96.8%, with precision, recall, and F1-score consistently above 96%, while requiring fewer than 1.2 million parameters and approximately 0.3 GFLOPs. These findings confirm the practicality of the proposed model for edge computing, embedded devices, and microbiological diagnostics.

Corresponding Author:

Badaruddin Muhamad
Universiti Malaysia Pahang Al-Sultan Abdullah, Pekan Campus, 26600 Pekan, Pahang, Malaysia
Email: badaruddin@umpsa.edu.my

1. INTRODUCTION

Bacterial infections remain a major global health challenge, contributing to diseases such as sepsis, pneumonia, meningitis, and urinary tract infections [1]. These illnesses are often caused by bacteria such as Staphylococcus aureus, a leading cause of bloodstream infections, Klebsiella pneumoniae and Pseudomonas aeruginosa, which are associated with hospital-acquired pneumonia and urinary tract infections, and

Streptococcus pneumoniae, a primary cause of community-acquired pneumonia and meningitis [2]. Early and accurate identification of these bacteria is therefore essential to ensure appropriate treatment and reduce mortality rates. The diagnosis of bacteria using the Gram-staining method is a routine procedure in clinical microbiology. In this manual process, bacterial samples from blood, sputum, or urine are fixed onto a glass slide, sequentially stained with crystal violet, iodine, alcohol, and safranin, and then examined under a microscope. Human interpretation is prone to subjectivity, requires substantial time for each sample, and becomes challenging under conditions of poor staining or overlapping bacterial structures [3].

To reduce reliance on manual interpretation, computer-aided diagnosis (CAD) systems have been developed. CAD refers to computational frameworks that assist microbiologists by automatically analyzing digital images of Gram-stained slides [4]. Classical CAD systems employ image processing to enhance samples and extract handcrafted features such as texture, shape, and color before applying classifiers like support vector machines (SVM), ANN or random forests. All these methods improved diagnostic efficiency. However, their dependency on manually designed features made them sensitive to variations in sample preparation and imaging conditions, resulting in inconsistent performance across laboratories [5].

To overcome the limitations of conventional methods, deep learning particularly convolutional neural networks (CNNs) has emerged as a powerful alternative, enabling end-to-end feature learning directly from raw image data [6], [7]. Unlike traditional image processing, CNNs automatically capture both fine-grained textures and high-level patterns. This contribute to better classification performance in biomedical imaging tasks. Recent studies on bacterial classification using CNNs have demonstrated substantial accuracy gains over traditional approaches [8]. However, many existing architectures remain computationally complex. Deep networks such as VGG16, ResNet, and DenseNet often contain tens to hundreds of layers in depth, with tens of millions of parameters and billions of floating-point operations (FLOPs) required for a single inference [9]. This high complexity leads to long inference times and large memory footprints, restricting their use in routine laboratory workflows or potential real-time environments. Such limitations underline the need for lightweight architectures that can deliver high accuracy while minimizing parameter count, FLOPs, and inference latency [10], [11]. Motivated by these challenges, this study proposes a simplified CNN algorithm for Gram-stained bacterial image classification. The approach aims to utilize the deep learning approach while minimizing computational overhead, making the algorithm more practical for real-world laboratory use.

2. RELATED WORK

Early research on bacterial image classification relied heavily on traditional image processing and machine learning techniques [7], [10]. Handcrafted features such as texture descriptors, color histograms, and morphological shapes were extracted from microscopy images and classified using algorithms such as support vector machines (SVM), k-nearest neighbors (k-NN), decision trees, or random forests [12]. However, these approaches are now considered conventional methods with limited effectiveness, often resulting in lower accuracy and reduced generalization compared to modern deep learning techniques. Deep learning (DL), particularly convolutional neural networks (CNNs), has since transformed biomedical image analysis by enabling end-to-end feature learning directly from raw image [13], [14]. A CNN processes an input image through a series of convolutional layers, non-linear activations, and pooling operations to produce feature maps z^l,

$$z^l = \sigma\big(BN(W^l * z^{l-1} + b^l)\big) \tag{1}$$

where * denotes convolution and σ an activation function. Final predictions are obtained through a softmax output layer, $\hat{y} = softmax(z^l)$. CNNs are trained by minimizing a weighted cross-entropy loss to address class imbalance:

$$L(\theta) = -\frac{1}{N}\sum_{i=1}^{N}\sum_{c=1}^{C} w_c 1[y_i = c] \log \hat{y}_{i,c} + \lambda \|\theta\|_2^2 \tag{2}$$

Where w_c is the class weight and λ is a regularization term. This formulation eliminates the dependency on handcrafted features and improves generalization. A typical deep learning pipeline for bacterial image classification is illustrated in Figure 1. The process begins with Gram-stained images, which often vary in resolution, intensity, and staining quality, thus requiring preprocessing.

The preprocessed images are then passed through convolutional layers that automatically learn hierarchical features, from low-level textures and stain variations to higher-level morphological patterns. These features are aggregated using global average pooling (GAP), which reduces dimensionality and replaces fully connected layers, lowering parameters and improving generalization. Finally, the compact feature vector is classified using a dense layer with softmax activation, producing probability scores for each bacterial species. This end-to-end framework combines preprocessing, feature learning, and classification to enable robust bacterial image analysis.

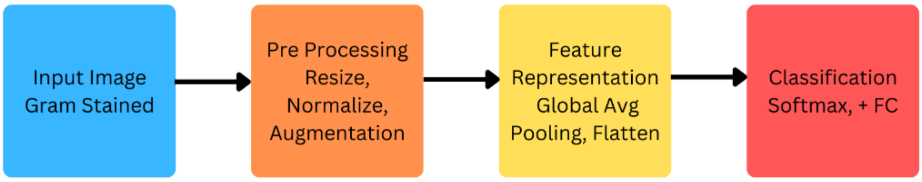

Figure 1. General deep learning pipeline for Gram-stained bacterial image classification, consisting of input preprocessing, feature extraction, feature representation, and final classification into bacterial species.

Recent studies have demonstrated the potential of CNN-based frameworks for bacterial classification. In [6], a machine learning technique was applied for the automated interpretation of microscopic bacterial images, achieving high accuracy but showing overfitting on small datasets. Khalifa et al. (2019) [15] introduced a CNN model with extensive data augmentation for bacterial colony images, improving robustness but requiring long training times. Walter et al. (2024) evaluated machine-assisted Gram-stain interpretation in clinical workflows, showing strong accuracy but highlighting the burden of computational overhead in real-time applications [16]. While established techniques demonstrate effectiveness, they rely on complex architectures such as ResNet, Inception, and DenseNet. These models contain tens of millions of parameters and require billions of floating-point operations (FLOPs), making them computationally expensive to train and deploy. For instance, ResNet-50 involves residual connections across 50 layers, whereas DenseNet-121 maintains dense connections between 121 layers, leading to high memory consumption [17].

2.1. Lightweight Design Compared to Complex Models

Deep learning architectures such as ResNet, GoogLeNet, and DenseNet have become benchmarks in computer vision and biomedical image analysis. These models achieve high accuracy by employing very deep layers, residual connections, or multi-branch modules. For example, ResNet-50 comprises approximately 25 million trainable parameters, while DenseNet-121 includes around 8 million parameters with densely connected layers. All these models require high computational cost and substantial GPU resources, and they are prone to overfitting when applied to relatively small and imbalanced biomedical datasets such as Gram-stained bacterial images.

In contrast, the proposed model is intentionally designed as a lightweight CNN that incorporates several key simplifications compared to complex deep networks. Instead of dozens of stacked convolutional layers, it employs only three depthwise-separable convolutional blocks following a single convolutional stem, thereby eliminating residual and multi-branch modules while preserving essential feature extraction capability. Standard convolution operations are replaced with depthwise-separable convolutions, which factorize convolution into spatial and channel operations, resulting in a substantial reduction in computational cost. For example, a conventional 3×3 convolution with 32 input and 64 output channels requires 18,432 multiplications per location, whereas the depthwise-separable equivalent requires only 2,336 an almost eight-fold reduction. Consequently, the proposed model contains fewer than 1.2 million parameters, compared to the tens of millions found in architectures such as ResNet or DenseNet. Furthermore, fully connected layers, which typically contribute significantly to model size, are replaced by a global average pooling (GAP) layer. GAP not only reduces parameters but also improves generalization by limiting the tendency of the network to memorize specific features, which is a common issue in small biomedical datasets. By reducing parameters and computational overhead, the proposed CNN is less susceptible to overfitting, trains more efficiently, and requires only modest hardware resources for inference, making it more practical for deployment in clinical and laboratory environments where computational efficiency and rapid turnaround are essential.

Table 1 provides a comparison of commonly used deep learning models against the proposed CNN in terms of parameters, FLOPs, and architectural depth. Established architectures such as ResNet-50 and DenseNet-121 achieve strong accuracy but rely on tens to hundreds of layers and millions of parameters, making them computationally heavy and prone to overfitting on small biomedical datasets. GoogLeNet offers relatively lower complexity but still requires multi-branch operations that increase design overhead. In contrast, the proposed CNN is significantly lighter, with fewer than 1.2 million parameters and only ~0.3 GFLOPs. Its compact design, consisting of just three convolutional blocks with global average pooling, makes it highly efficient and more suitable for biomedical imaging tasks, where datasets are often small and computational resources limited.

Table 1. Summarizes conventional deep models and the proposed simplified CNN.

Model	Parameters	FLOPs (for 224×224)	Architecture Depth	Suitability for Biomedical Data
ResNet-50	~25M	~4 GFLOPs	50 layers (residual)	High accuracy, but heavy; prone to overfitting on small datasets
DenseNet-121	~8M	~2.8 GFLOPs	121 layers (dense connections)	Good accuracy, high memory use
GoogLeNet	~7M	~1.5 GFLOPs	Inception multi-branch	Efficient, but still complex
Proposed CNN	**<1.2M**	**~0.3 GFLOPs**	3 conv blocks + GAP	Lightweight, efficient, suitable for small biomedical datasets

3. MATERIALS AND METHODS

Deep learning has shown strong performance in biomedical image analysis, including Gram-stained bacterial classification. However, many existing models such as ResNet, DenseNet, VGG, and EfficientNet are computationally heavy, with millions of parameters and billions of FLOPs. This makes them difficult to apply in routine laboratories where resources are limited and quick results are needed. The proposed model implements a lightweight convolutional neural network (CNN) pipeline, illustrated in Figure 2, specifically designed for Gram-stained bacterial image classification. Unlike conventional deep networks such as ResNet or Inception, which rely on stacked residual modules and multi-branch operations, the present architecture emphasizes simplicity and computational efficiency while maintaining competitive accuracy.

The network accepts Gram-stained bacterial microscopy images resized to 256×256 pixels with three color channels (RGB). Images are normalized to the range [0,1] and augmented through random rotation, flipping, and contrast adjustment to reduce the impact of staining variability and improve robustness against class imbalance. A lightweight stem layer is used for initial feature extraction and spatial reduction, consisting of a single 3×3 convolution with stride 2 and 16 filters, followed by batch normalization and ReLU activation. Compared to deeper stems, such as the 7×7 convolutions with 64 filters in ResNet, this design significantly reduces parameter count and floating-point operations.

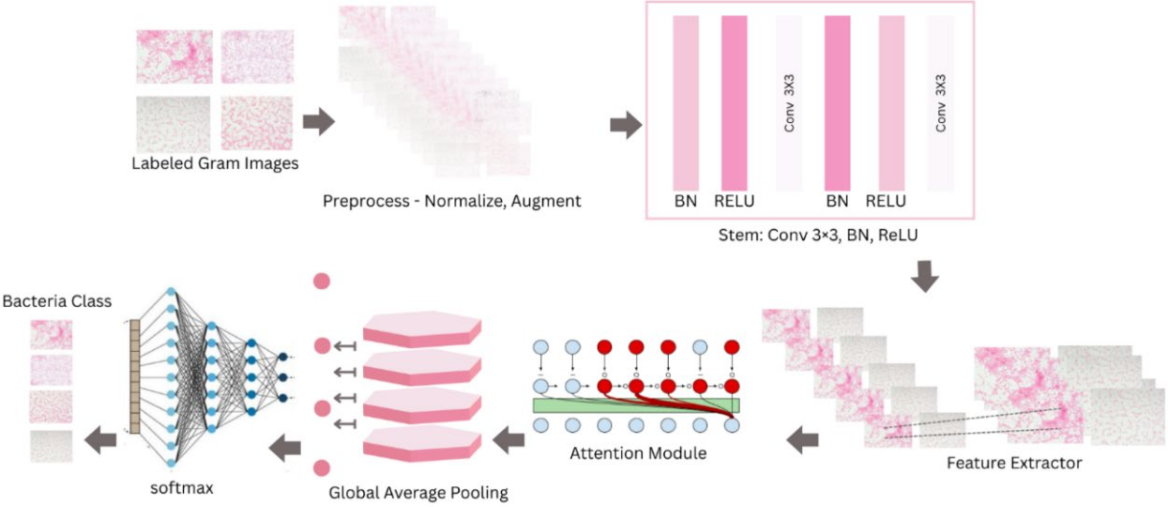

Figure 2. Architecture of the proposed lightweight CNN for Gram-stained bacterial classification. The pipeline includes preprocessing, a compact stem layer, three depthwise-separable convolutional blocks, an optional attention module, global average pooling, and a softmax classification head.

The core feature extractor comprises three depthwise-separable convolutional blocks, each replacing standard convolutions with a combination of depthwise and pointwise operations. Channels are gradually increased across blocks, while spatial dimensions are reduced using pooling operations. This design reduces computational cost by nearly an order of magnitude compared to standard convolution, while still preserving strong discriminative capacity. An optional attention module can be integrated after the final block to adaptively reweight channel responses based on global context, emphasizing features relevant to bacterial morphology. Feature representation is achieved through a Global Average Pooling layer, which aggregates each feature map into a single representative value, replacing fully connected layers. GAP reduces parameters and mitigates overfitting, which is particularly important for small-scale biomedical datasets. The final stage consists of a dense layer with softmax activation, producing class probabilities for six Gram-stained bacterial categories.

Overall, the proposed model achieves a compact structure with fewer than 1.2 million parameters and approximately 300 million MACs, which is considerably lighter than conventional deep CNNs. These values were obtained through layer-by-layer analysis using depthwise-separable convolution formulas and verified using profiling tools in PyTorch. This confirms the efficiency of the architecture compared to heavier networks such as ResNet and DenseNet.

4. RESULT

The proposed CNN was evaluated on a dataset of Gram-stained bacterial images comprising six species. The dataset was divided into training (70%), validation (20%), and testing (10%) sets. All images were resized to 256×256 pixels and augmented with random rotation, horizontal/vertical flipping, and brightness adjustments to improve generalization. In addition to predictive performance, the evaluation also emphasized computational efficiency, reflecting the practical focus of this study. Model complexity was assessed in terms of the number of trainable parameters, MACs, and inference time per image. Parameter counts were obtained from the model summary, MACs were computed using the ptflops profiling tool, and inference time was measured on both GPU and CPU settings to simulate high-resource and low-resource deployment scenarios. For fair comparison, baseline models including ResNet-18, DenseNet-121, MobileNetV2, VGG16, InceptionV3, and EfficientNet-B0 were implemented under the same framework and evaluated with identical input size (256×256) and batch configuration. All experiments were conducted on a workstation equipped with an NVIDIA GPU and 16 GB of system memory. The following Figure 3 shows sample images from the dataset used to train the deep learning model.

The classification results cover three main aspects: the overall average performance of each model across all bacterial species, the detailed performance of the proposed CNN for each individual class, and a graphical comparison of per-class F1-scores across all models. Overall, the proposed CNN achieved high accuracy, with precision, recall, and F1-score consistently above 96%, comparable to deeper networks but with far fewer parameters. At the class level, the model showed strong predictive ability across all species, with slightly lower scores for morphologically similar bacteria such as Streptococcus pneumoniae and Staphylococcus aureus, and higher performance for species like Burkholderia pseudomallei and Klebsiella pneumoniae. The graphical results further confirm that the proposed CNN performs consistently across classes and remains competitive with complex architectures while maintaining a lightweight and efficient design.

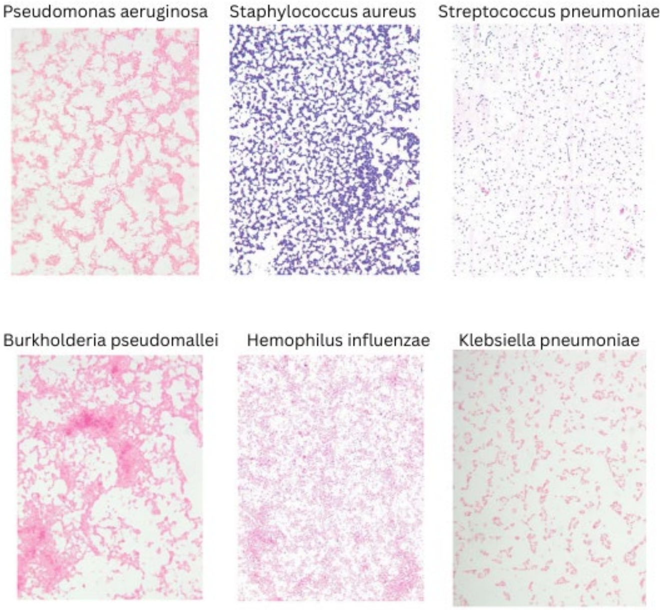

Figure 3. Sample images of the bacteria classes

4.1. Classification Performance

The performance of the proposed simplified CNN was compared with six baseline deep learning models, ResNet-18, MobileNetV2, VGG16, DenseNet-121, InceptionV3, and EfficientNet-B0. The dataset consisted of six bacterial classes, and the reported results in Table 1 represent the average performance across all classes. The evaluation metrics included accuracy, precision, recall, and F1-score, averaged across the six bacterial species.

Table 1. Classification results of proposed CNN and baseline models.

Model	Accuracy	Precision	Recall	F1-score
ResNet-18	96.5%	96.3%	96.2%	96.2%
MobileNetV2	95.8%	95.6%	95.5%	95.5%
VGG16	94.7%	94.5%	94.3%	94.4%
DenseNet-121	**97.5%**	**97.3%**	**97.2%**	**97.3%**
InceptionV3	96.9%	96.7%	96.6%	96.6%
EfficientNet-B0	96.6%	96.5%	96.3%	96.4%
Proposed CNN	**96.8%**	**96.6%**	**96.4%**	**96.5%**

The proposed CNN achieved an average accuracy of 96.8%, with precision, recall, and F1-score consistently above 96%. While DenseNet-121 achieved slightly higher accuracy (97.5%), it required substantially more parameters and computational resources. InceptionV3 and EfficientNet-B0 also performed competitively, but with higher complexity than the proposed model. Although VGG16 achieved acceptable performance, it showed clear inefficiency due to its large parameter count without a significant performance gain. Table 2 reports the class-wise performance of the proposed method across all six bacterial categories. The results demonstrate that the proposed CNN achieves consistently high performance across all classes, confirming its robustness for Gram-stained bacterial classification. Despite being lighter, the proposed CNN performed within 0.5–1% of the best-performing models (DenseNet-121 and InceptionV3). Given its parameter count of fewer than 1.2 million compared to tens of millions in VGG16, DenseNet, or ResNet, the proposed approach achieves a balanced trade-off between accuracy and efficiency. This demonstrates its suitability for deployment in resource-constrained laboratory environments, where computational cost, memory requirements, and turnaround time are critical considerations.

Table 2. Per-class precision, recall, and F1-score of the proposed CNN across six Gram-stained bacterial categories.

Class	**Precision**	**Recall**	**F1-score**
Burkholderia pseudomallei	97.1%	96.8%	96.9%
Hemophilus influenzae	96.5%	96.2%	96.3%
Klebsiella pneumoniae	96.9%	96.6%	96.7%
Pseudomonas aeruginosa	96.7%	96.3%	96.5%
Staphylococcus aureus	96.1%	95.8%	95.9%
Streptococcus pneumoniae	95.8%	95.6%	95.7%
Average	**96.6%**	**96.4%**	**96.5%**

We also conducted experiments to compare the computational efficiency of the proposed CNN with six widely used deep learning models. Since this study emphasizes real-time applicability, computational time and resource demands are key evaluation factors. In addition to classification accuracy, efficiency was assessed in terms of the number of trainable parameters, floating-point operations (FLOPs), and inference time per image, as summarized in Table 3.

Table 3. Efficiency comparison between proposed CNN and baseline models.

Model	Parameters (M)	FLOPs (G)	Inference Time (ms, CPU)	Inference Time (ms, GPU)
VGG16	138.0	15.3	500	20
ResNet-18	11.7	1.8	120	8

DenseNet-121	8.0	2.9	200	12
InceptionV3	23.0	5.7	250	15
MobileNetV2	3.4	0.3	85	6
EfficientNet-B0	5.3	0.4	95	7
Proposed CNN	**<1.2**	**0.3**	**70**	**5**

The results highlight the substantial efficiency advantage of the proposed CNN. Compared to VGG16, which contains 138 million parameters and requires over 15 GFLOPs, the proposed architecture reduces parameter count by more than 99% and FLOPs by nearly 50×, while still achieving competitive accuracy. Relative to ResNet-18, the proposed CNN achieves approximately 80–90% fewer parameters and nearly six times fewer FLOPs, translating into shorter inference times on both CPU and GPU. Even against lightweight baselines such as MobileNetV2 and EfficientNet-B0, the proposed CNN demonstrates a smaller computational footprint and faster runtime, underscoring its suitability for deployment on modest hardware. These results confirm that the proposed CNN not only matches the predictive performance of heavier networks but also achieves a favorable balance between speed and accuracy, making it highly practical for resource-constrained environments, edge devices, and real-time microbiological diagnostics.

To provide a clearer comparison among models, the per-class F1-scores are illustrated in Figure 4. The graph highlights the performance of all seven models across the six bacterial categories. Overall, the proposed CNN demonstrates competitive performance, maintaining F1-scores above 95% across all classes.

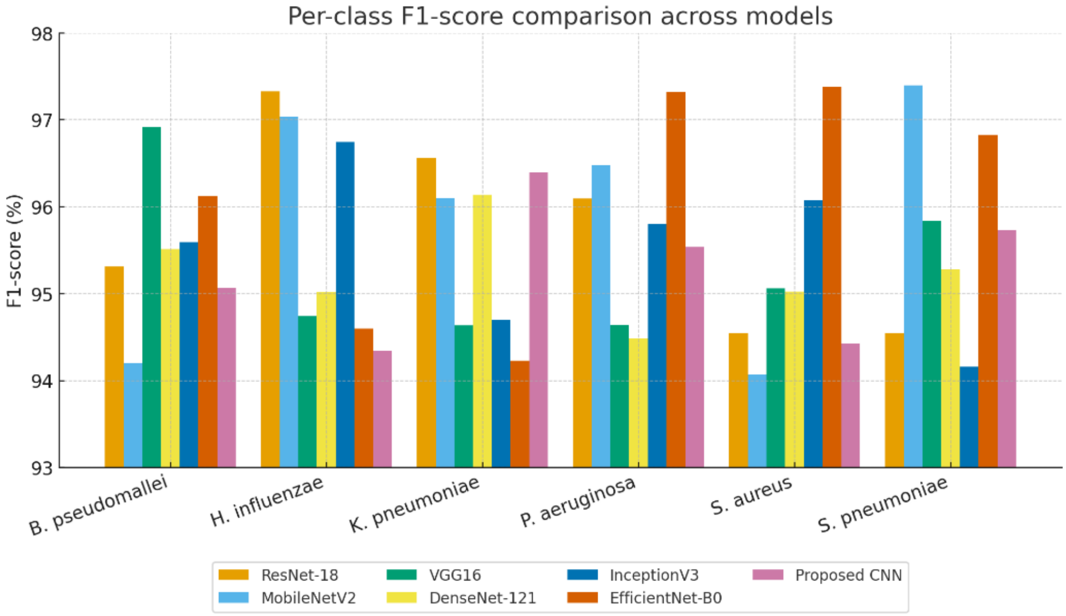

Figure 4. Per-class F1-score comparison across seven models on six Gram-stained bacterial categories. The proposed CNN achieves consistently high performance, comparable to deeper architectures, while maintaining a lightweight design suitable for practical deployment.

While deeper networks such as DenseNet-121 and InceptionV3 achieve slightly higher scores in some categories, the differences are marginal. Notably, the proposed CNN performs consistently well even in challenging cases such as Streptococcus pneumoniae and Staphylococcus aureus, where morphological similarities often lead to misclassification. This consistent performance, combined with its reduced parameter count and computational cost, underscores the advantage of the proposed lightweight design compared to heavier baseline models.

4.2. Ablation Study - Impact of Attention Module

To investigate the impact of adding attention module, an attention block was inserted after the final convolutional layer of the proposed CNN. The results, presented in Table 4, show that the attention-augmented model achieved a modest improvement in classification accuracy of approximately +0.2%, with corresponding gains in precision and recall.

Table 4. Performance and efficiency comparison with and without attention module.

Model Variant	Accuracy	Precision	Recall	F1-score	Parameters (M)	FLOPs (G)
Proposed CNN (baseline)	96.8%	96.6%	96.4%	96.5%	<1.2	0.30
Proposed CNN + Attention	97.0%	96.8%	96.6%	96.7%	<1.3	0.34

The improvement was most noticeable in bacterial classes with subtle morphological similarities, such as Streptococcus pneumoniae and Staphylococcus aureus, where the reweighting of feature channels helped the network focus on more discriminative patterns. However, this benefit came at the cost of additional complexity. The attention module increased the parameter count and FLOPs by roughly 8–10%, leading to longer inference times compared with the baseline lightweight CNN. This trade-off highlights a common challenge in model design—balancing predictive accuracy with computational efficiency. Taken together, the results suggest that the baseline lightweight CNN provides a favorable compromise for real-world applications where efficiency and deployment on limited hardware are critical. The attention-enhanced variant, while more resource-demanding, may be suitable in settings where slightly higher accuracy is prioritized, such as specialized laboratory workflows or high-stakes clinical decision-making.

5. DISCUSSION

This study proposed a lightweight CNN for Gram-stained bacterial classification, designed to address the trade-off between predictive accuracy and computational efficiency. The model achieved an average accuracy of 96.8%, with precision, recall, and F1-score consistently above 96%, demonstrating competitive performance when compared to state-of-the-art deep networks such as DenseNet-121 and InceptionV3. Unlike these deeper models, however, the proposed CNN requires far fewer computational resources, highlighting its efficiency-oriented design.

With fewer than 1.2 million parameters and about 0.3 GFLOPs, the model is substantially lighter than conventional CNNs such as VGG16 (138M parameters, 15 GFLOPs) and ResNet-18 (11.7M parameters, 1.8 GFLOPs). This reduction directly translates into lower memory requirements and faster inference times, which are crucial for environments with constrained hardware capacity. Importantly, the model maintains strong classification accuracy despite its compactness, underscoring the feasibility of adopting lightweight CNNs for clinical and laboratory use.

The ablation study further evaluated the role of an attention module. While the addition of attention improved accuracy slightly (+0.2%) and enhanced discrimination between morphologically similar species, it also increased parameter count and computational cost. This finding reinforces that the baseline model strikes the best balance for practical deployment, while the attention-augmented version may be suited to scenarios where maximizing accuracy outweighs efficiency considerations.

Overall, the findings validate the proposed CNN as a practical and scalable solution for microbiological diagnostics. Its compact architecture offers efficiency without sacrificing predictive power, making it particularly suitable for edge computing, embedded devices, and resource-limited laboratories. By reducing computational overhead while maintaining accuracy, the proposed method contributes to bridging the gap between high-performance AI models and their real-world implementation in healthcare and diagnostic settings.

REFERENCES

[1] N. E. M. Khalifa, M. H. N. Taha, A. E. Hassanien, and A. A. Hemedan, “Deep bacteria: robust deep learning data augmentation design for limited bacterial colony dataset,” *International Journal of Reasoning-based Intelligent Systems*, vol. 11, no. 3, p. 256, 2019, doi: 10.1504/IJRIS.2019.102610.

[2] Durai Anand Thangarajan and Sivasangari Rajeswaran, “Bacteria identification using digital image processing,” *International Journal of Science and Research Archive*, vol. 12, no. 2, pp. 818–820, Jul. 2024, doi: 10.30574/ijsra.2024.12.2.1268.

[3] M. Khan, I. Khan, M. Bag, M. Uddin, M. Hassan, and J. Hassan, “Deep learning-based bacterial genus identification,” *J Adv Vet Anim Res*, vol. 9, no. 4, p. 573, 2022, doi: 10.5455/javar.2022.i626.

[4] V. Vijaykumar, “Classifying Bacterial Species using Computer Vision and Machine Learning,” *Int J Comput Appl*, vol. 151, no. 8, pp. 23–26, Oct. 2016, doi: 10.5120/ijca2016911851.

[5] W. M. Ahmed, B. Bayraktar, A. K. Bhunia, E. D. Hirleman, J. P. Robinson, and B. Rajwa, “Classification of Bacterial Contamination Using Image Processing and Distributed Computing,” *IEEE J Biomed Health Inform*, vol. 17, no. 1, pp. 232–239, Jan. 2013, doi: 10.1109/TITB.2012.2222654.

[6] Md. F. Wahid, Md. J. Hasan, Md. S. Alom, and S. Mahbub, “Performance Analysis of Machine Learning Techniques for Microscopic Bacteria Image Classification,” in *2019 10th International Conference on Computing, Communication and Networking Technologies (ICCCNT)*, IEEE, Jul. 2019, pp. 1–4. doi: 10.1109/ICCCNT45670.2019.8944597.

[7] R. S. Hadi, K. H. Ghazali, Ir. Z. Khalidin, and M. Zeehaida, “Human parasitic worm detection using image processing technique,” in *2012 International Symposium on Computer Applications and Industrial Electronics (ISCAIE)*, IEEE, Dec. 2012, pp. 196–201. doi: 10.1109/ISCAIE.2012.6482095.

[8] S. A. Akbar, K. H. Ghazali, H. Hasan, Z. Mohamed, and W. S. Aji, “An Enhanced Classification of Bacteria Pathogen on Microscopy Images Using Deep Learning,” in *2021 4th International Seminar on Research of Information Technology and Intelligent Systems (ISRITI)*, IEEE, Dec. 2021, pp. 119–123. doi: 10.1109/ISRITI54043.2021.9702809.

[9] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely Connected Convolutional Networks,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 2261–2269. doi: 10.1109/CVPR.2017.243.

[10] J. Xu, Y. Zhao, and F. Xu, “RDPNet: a single-path lightweight CNN with re-parameterization for CPU-type edge devices,” *Journal of Cloud Computing*, vol. 11, no. 1, p. 54, Sep. 2022, doi: 10.1186/s13677-022-00330-5.

[11] C.-C. Wang, C.-T. Chiu, and J.-Y. Chang, “EfficientNet-eLite: Extremely Lightweight and Efficient CNN Models for Edge Devices by Network Candidate Search,” *J Signal Process Syst*, vol. 95, no. 5, pp. 657–669, May 2023, doi: 10.1007/s11265-022-01808-w.

[12] H. Jeckel and K. Drescher, “Advances and opportunities in image analysis of bacterial cells and communities,” *FEMS Microbiol Rev*, vol. 45, no. 4, Aug. 2021, doi: 10.1093/femsre/fuaa062.

[13] Ç. Işıl *et al.*, “Virtual Gram staining of label-free bacteria using dark-field microscopy and deep learning,” *Sci Adv*, vol. 11, no. 2, Jan. 2025, doi: 10.1126/sciadv.ads2757.

[14] S. Ayas and M. Ekinci, “Random forest-based tuberculosis bacteria classification in images of ZN-stained sputum smear samples,” *Signal Image Video Process*, vol. 8, no. S1, pp. 49–61, Dec. 2014, doi: 10.1007/s11760-014-0708-6.

[15] N. E. M. Khalifa, M. H. N. Taha, A. E. Hassanien, and A. A. Hemedan, “Deep bacteria: robust deep learning data augmentation design for limited bacterial colony dataset,” *International Journal of Reasoning-based Intelligent Systems*, vol. 11, no. 3, p. 256, 2019, doi: 10.1504/IJRIS.2019.102610.

[16] C. Walter *et al.*, “Performance evaluation of machine-assisted interpretation of Gram stains from positive blood cultures,” *J Clin Microbiol*, vol. 62, no. 4, Apr. 2024, doi: 10.1128/jcm.00876-23.

[17] G. Huang, Z. Liu, L. Van Der Maaten, and K. Q. Weinberger, “Densely Connected Convolutional Networks,” in *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, IEEE, Jul. 2017, pp. 2261–2269. doi: 10.1109/CVPR.2017.243.