Applied Information System and Management (AISM)
Volume 9, (1) 2026, p. 117–124
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v9i1.50124

Application of EfficientNet Transfer Learning with Incremental Fine-Tuning for Road Damage Detection

Riki Winanjaya[1]*, Abdi Rahim Damanik[2], Anton Abdulbasah Kamil[3]

***Abstract*—Image-based road damage detection is an essential component of intelligent infrastructure monitoring systems. However, conventional transfer learning often fails to adapt pre-trained models to domain-specific characteristics such as fine-crack textures, illumination variations, and perspective distortions. This study proposes an EfficientNet-based road damage classification model that leverages incremental fine-tuning and multi-stage data augmentation to enhance feature adaptation and model robustness. The experiments were conducted using the Road Damage Detection dataset from Kaggle, comprising 1,400 labeled images across several road damage classes. The dataset was partitioned into 80:10:10 splits for training, validation, and testing, with stratification. The proposed approach gradually unfreezes EfficientNet layers through a structured incremental fine-tuning schedule while applying staged augmentation to expand data diversity. Experimental results show that the baseline EfficientNet transfer learning model achieved 78.26% accuracy, whereas the proposed model improved performance to 97.10% accuracy, with 97.60% macro precision, 97.20% macro recall, and 97.30% macro F1-score. The results demonstrate that incremental fine-tuning effectively enhances feature adaptation to road damage textures, while multi-stage augmentation improves model robustness. These findings indicate that the proposed approach provides an effective strategy for improving deep-learning-based road damage detection systems in real-world infrastructure monitoring applications.**

***Index Terms*—Deep learning, EfficientNet, image classification, incremental fine-tuning, multi-stage data augmentation, road damage detection, transfer learning.**

I. Introduction

The condition of road infrastructure is a vital element in determining the quality of mobility and transportation safety. Road surface deterioration, including longitudinal, transverse, and alligator cracking, as well as potholes, can increase accident risks and hinder logistics distribution efficiency if not promptly addressed [1], [2], [3], [4]. Currently, identifying road damage relies on manual inspections by field personnel, which are time-consuming, expensive, and highly dependent on the inspector's expertise. These challenges necessitate a modern approach capable of automatically, efficiently, and accurately detecting road damage [5], [6], [7], [8].

The development of deep learning-based computer vision has enabled automated detection of road damage through image analysis. Traditional CNN architectures such as VGG [9], [10], [11], ResNet [12], [13], and MobileNet have been used in various previous studies. However, their results are limited by computational efficiency and the models' generalizability across diverse environmental conditions. Road damage images are often captured under varying lighting conditions, varying camera angles, and in the presence of visual noise, necessitating a more adaptive and robust model for feature extraction [14], [15], [16].

A previous study by [17] introduced the Indian Footpath Damage Segmentation Dataset, the first dataset specifically dedicated to detecting sidewalk damage in India using a deep learning approach. Its advantages include high-resolution imagery, thorough pixel-level annotation using LabelMe, and severity labels (low, medium, and high) determined in collaboration with civil engineers. The dataset is accompanied by a standardized data collection protocol, complete metadata, and advanced model baselines, including U-Net with EfficientNet-B3 and ResNeXt50, which demonstrate strong performance on segmentation and severity classification tasks. However, its weaknesses include limited geographic coverage restricted to Pune, making it less representative of infrastructure variations across other cities; the absence of 3D data or additional sensors such as LiDAR that could improve damage depth analysis; and the dataset's focus solely on visual damage without considering other risk factors such as lighting, pedestrian density, or surface friction coefficient. Nevertheless, this dataset remains an important contribution to pedestrian safety research and computer-vision-based infrastructure monitoring.

Received: 19 January 2026; Revised: 22 February 2026; Accepted: 16 March 2026.
*Corresponding author
[1]Riki Winanjaya, STIKOM Tunas Bangsa, Indonesia (e-mail: riki@amiltunasbangsa.ac.id).
[2]Abdi Rahim Damanik, STIKOM Tunas Bangsa, Indonesia (e-mail: abdirahimdmk@stikomtunasbangsa.ac.id).
[3]Anton Abdulbasah Kamil, Istanbul Gelisim University, Turkiye (e-mail: akamil@gelisim.edu.tr).

In contrast, the research by [18] examines the adoption of the Internet of Things (IoT) in Indonesian hospitals and the factors that influence it, including technological readiness, management support, perceived benefits, and implementation barriers, to provide an overview of the digital transformation landscape in the health sector. Its strengths lie in its focus on the Indonesian context, the use of a systematic analytical framework, and the identification of enabling and inhibiting factors, making it a useful reference for policymakers. However, its weaknesses include a limited sample that does not represent all hospital types, potential bias due to some data collected through surveys or interviews, and insufficient discussion of technical aspects of IoT infrastructure, which may affect the interpretation of the results.

EfficientNet has emerged as a modern CNN architecture, offering a combination of high accuracy and computational efficiency through a compound scaling mechanism [19], [20], [21]. EfficientNet's lightweight yet precise capabilities offer significant potential for application to road-damage detection problems. However, EfficientNet's optimal performance is largely determined by the transfer learning and fine-tuning strategies used. In the context of heterogeneous road-damage datasets, such as those from Kaggle, a more structured, gradual training approach is required to adapt the model to domain characteristics [1], [22], [23].

One relevant approach is incremental fine-tuning, a technique for gradually unfreezing EfficientNet layers to enrich feature representation without compromising model stability [24], [25], [26]. This approach is crucial to avoid catastrophic forgetting and overfitting, which frequently occur when all model layers are unfrozen simultaneously. By utilizing this strategy, this research explores how this gradual feature adjustment process can improve image-based road damage detection capabilities [21], [27], [28].

The road damage detection dataset available on the Kaggle platform provides a comprehensive data source for this research. Featuring diverse road damage types across various environments, the dataset is suitable for evaluating transfer learning and incremental fine-tuning techniques. Accordingly, the study employs EfficientNet with a stepwise approach to achieve improved accuracy and efficiency in automated road condition monitoring. This approach is expected to make a significant contribution to the development of a road maintenance system based on modern technology oriented toward smart infrastructure [29], [30], [31].

II. RESEARCH METHOD

This section describes the methodological approach used in this research to build and evaluate an EfficientNet-based road damage detection model using transfer learning and incremental fine-tuning. The research method is organized into three main subsections: the research dataset, a comparison of the baseline model with the proposed model enhanced through multi-stage data augmentation, and the overall research design, which provides an overview of the experimental flow from start to finish.

A. Dataset (Data Collection)

The dataset constitutes the core component of this research, as data quality and diversity substantially affect model performance in accurately identifying road damage. The research data were collected from the road damage detection dataset available on the Kaggle platform. This dataset was chosen because it contains a wide variety of road damage types and has been widely used in computer vision studies, making it a reliable evaluation reference. The images in the dataset were captured under various environmental conditions, including varying lighting, weather, camera angles, and road surface quality. This diversity presents unique challenges but also improves the model's ability to generalize to real-world data.

The experiments were conducted using the EfficientNet architecture with transfer learning. In this study, EfficientNet-B0 was used as the backbone network due to its balance between computational efficiency and classification performance. All input images were resized to 224 × 224 pixels and normalized to the range [0, 1] before being fed into the model.

Fig. 1. Sample research dataset.

Figure 1 shows that the dataset contains road images classified into four damage categories: longitudinal cracks, transverse cracks, alligator cracks, and potholes. The Kaggle dataset used comprised 1,400 images. Each image was labeled with a damage type, allowing the dataset to be used for classification tasks. To ensure optimal training, the dataset was split into training, validation, and test sets at 80%, 10%, and 10%, respectively. This partitioning was carried out in a controlled manner to maintain balanced class distributions in each subset, preventing the model from becoming biased toward any particular class.

Before training, all images underwent preprocessing. The images were resized to the dimensions appropriate for the EfficientNet architecture, normalized to a specific pixel range, and filtered to remove corrupted or unprocessable data. This step ensured that all data were in a uniform format and ready for use in the model training phase. Preprocessing also improves

Applied Information System and Management (AISM)
Volume 9, (1) 2026, p. 117–124
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v9i1.50124

the stability of the training process, enabling the model to focus on learning visual patterns in road damage without interference from non-essential variations in the input.

B. Comparison of the Baseline Model and the Proposed Model

The baseline model uses EfficientNet with a standard transfer learning approach, where most network layers are frozen, and only the classifier portion is retrained based on the number of classes in the road damage dataset. This approach provides good initial stability, but is limited in recognizing domain-specific visual patterns such as cracks and complex road surface textures. As only a few layers are updated, the baseline tends to produce feature representations that are less adaptive to the dataset's characteristics and show limited performance improvement after a few training epochs.

Table 1.
Comparison of Baseline Model and Proposed Model

Comparative Aspects	EfficientNet Baseline (Standard Transfer Learning)	Proposed EfficientNet Model (Incremental Fine-Tuning + Multi-Stage Augmentation)
Training strategy	Only the classifier is trained; most EfficientNet layers are frozen	Layers are gradually unlocked; progressive fine-tuning from the classifier to the deep layers
Adaptation to road damage dataset	Low; dominant features remain from ImageNet, thus poorly capturing crack patterns	High; incremental fine-tuning increases sensitivity to road texture patterns
Data augmentation	Basic augmentations (rotation, flipping)	Multi-stage augmentation (basic → intermediate → complex) covers variations in light, perspective, and noise.
Training stability	Stable initially, but quickly plateaus	Stable across stages, more consistent accuracy improvements
Risk of overfitting	Higher, especially when the dataset is limited	Lower; multi-stage augmentation expands the diversity of training data
Generalization ability	Moderate; performance decreases with different image conditions	The high model is more adaptive to real-world conditions
Training complexity	Low; faster processing	Higher requires learning rate tuning and fine-tuning stages
Final performance	Accuracy and F1-score increase, but are limited	Accuracy, precision, recall, and F1-score significantly improve across classes.
Main advantages	Fast to train and stable in the early stages	Deeper feature adaptation and more accurate classification results
Main disadvantages	Less responsive to domain-specific features	Requires longer training time and more careful hyperparameter management

Unlike the baseline, the proposed model uses EfficientNet's incremental fine-tuning to improve its adaptability to variations in road damage. Fine-tuning is performed in stages, starting with classifier training and then introducing feature-extraction layers trained with a lower learning rate. This stepwise approach allows for finer-grained parameter adjustments, enabling the model to more accurately capture the visual details of damage without losing the underlying knowledge from the ImageNet pre-training. Furthermore, applying multi-stage data augmentation provides a broader range of inputs than the baseline, making the proposed model more resilient to variations in lighting, perspective, and environmental conditions.

Overall, both models use the same architecture, but their different training strategies yield significantly different performance. The baseline model offers stability but lacks the flexibility to learn domain-specific features. In contrast, the proposed model demonstrates better generalization and significant accuracy improvements through the combination of incremental fine-tuning and multi-stage augmentation.

C. Research Design

The research design was developed to provide a comprehensive overview of the experimental stages. The research process began with dataset collection from Kaggle and verification of data quality and completeness. Once the data were acquired, preprocessing and augmentation steps were applied to prepare them for training. In the modeling stage, two model groups (a baseline model and a proposed model) were built and configured according to the experimental requirements. The baseline model served as a performance benchmark, while the proposed model was developed via transfer learning and incremental fine-tuning.

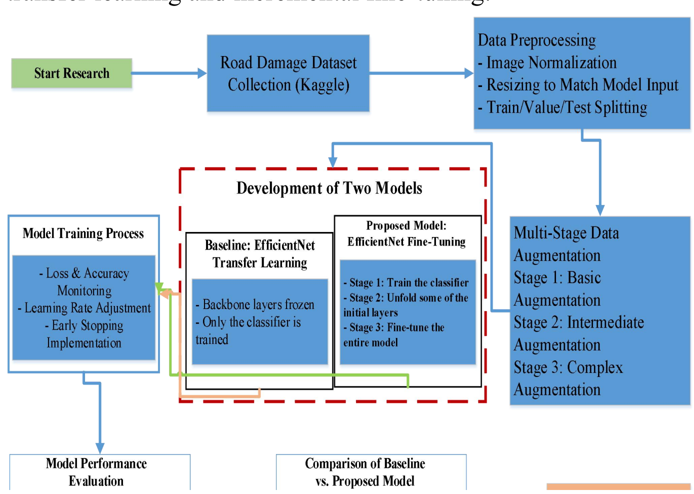


Fig. 2. Research design.

Figure 2 outlines the research design, which begins with the collection of the road damage detection dataset from Kaggle. The preprocessing phase involves image normalization, image resizing, and partitioning into training, validation, and testing subsets. The dataset was then enriched through multi-stage data augmentation, ranging from basic to complex transformations, to simulate variations in real-world conditions. In the modeling phase, two approaches were developed: the baseline EfficientNet trained with standard transfer learning, and the proposed EfficientNet model trained with incremental fine-tuning via gradual unfreezing of model layers. Both models underwent complete training, with loss and accuracy monitored and an early-stopping mechanism to maintain training stability. Following the training phase, model performance was evaluated using multiple metrics, including accuracy, precision, recall, F1-score, confusion matrix analysis, and training curves. A comparative analysis of both models was conducted to assess the effectiveness of incremental fine-tuning and multi-stage augmentation strategies in improving road damage recognition. The study concludes with key findings and recommendations for future research directions.

The next phase involves model training. In the baseline model, training was conducted using a standard transfer learning strategy with basic augmentation. In contrast, the proposed model was trained using a two-stage approach: initial transfer learning and incremental fine-tuning. Each training stage was evaluated using validation data to ensure stability and detect signs of overfitting. After training was completed, each model was tested on a test set not previously used during training. Evaluation was conducted using accuracy, precision, recall, and F1-score metrics to provide a comprehensive overview of model performance. The final stage involved analyzing the results to compare the baseline model's performance with that of the proposed model. The analysis was conducted quantitatively using evaluation metrics and qualitatively by observing the distribution of model predictions across specific damage classes. This analysis determines whether the proposed model provides significant improvements and how incremental fine-tuning and multi-stage augmentation strategies contributed to these improvements. This systematic research design is expected to yield reproducible conclusions that will guide other researchers in developing image-based road-damage detection systems.

III. Results and Discussion

This section presents the research results and an in-depth analysis of the performance of two models: the baseline EfficientNet using the standard transfer learning method and the proposed EfficientNet model with incremental fine-tuning. The evaluation was conducted using common multiclass classification metrics, such as accuracy, macro precision, macro recall, and macro F1-score. The research results were then comprehensively analyzed to assess the contribution and effectiveness of the proposed method for road damage detection. The interpretations provided in this section are based on the conceptual framework established in the introduction, specifically regarding the need for a more adaptive and robust model capable of handling variations in road damage.

A. Model Performance Analysis Based on Key Metrics

The test results show a stark difference in performance between the baseline and proposed models (Table 2). The baseline EfficientNet, which uses only standard transfer learning, recorded an accuracy of 0.7826. Although the precision and recall values were in the moderate range, at 0.8300 and 0.7775, respectively, the overall performance was limited in capturing the full visual patterns of road damage. This fact is reflected in the macro F1-score of 0.7810, indicating that the balance between the model's ability to detect positives and avoid false positives was suboptimal.

Table 2.
Result of Model Performance Evaluation

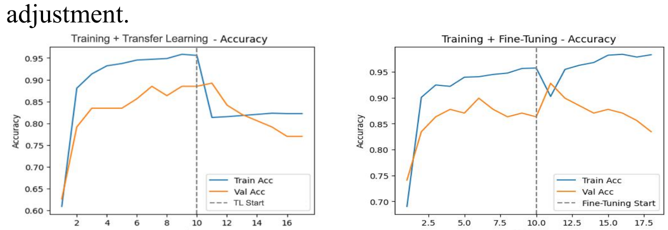
Model	Accuracy	Macro Precision	Macro Recall	Macro F1-score
Baseline: EfficientNet transfer learning	0.7826	0.8300	0.7775	0.7810
Proposed model: EfficientNet fine-tuning	0.9710	0.9760	0.9720	0.9730

In contrast, the proposed incremental fine-tuning model demonstrated significant performance improvements. Accuracy increased to 0.9710, representing an absolute improvement of more than 19% compared to the baseline. This improvement was not limited to accuracy; it was evident across all other metrics as well. Macro precision reached 0.9760, macro recall reached 0.9720, and macro F1-score reached 0.9730. The consistently high performance across all metrics demonstrates that the model can not only accurately predict the class but also consistently detect each damage class without bias toward the majority class.

The improved performance of the proposed model strongly indicates that the incremental fine-tuning strategy is highly effective at adapting EfficientNet's feature representation to road damage patterns, which exhibit finer, more complex textural characteristics than those of common objects in the ImageNet dataset. These results also demonstrate that incremental augmentation significantly increases the diversity of the training data, making the model better able to handle real-world imagery.

B. Training Curve Comparison and Model Stability

Although the training curves are not shown in the table, interpreting the performance metrics provides implicit insight into the model's stability during training. In the EfficientNet baseline, the higher precision than recall suggests the model tends to make more conservative predictions. This phenomenon indicates that the baseline model avoids prediction errors but struggles to detect subtler variations in damage. This imbalance indicates that the feature representation derived from the pre-trained weights is not optimal for the road damage domain without further adjustment.



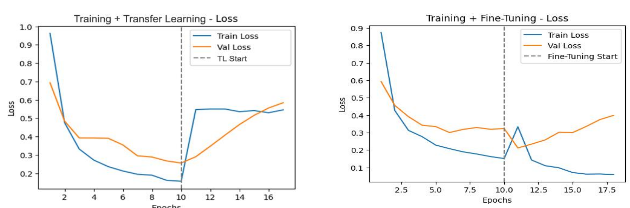
a) Baseline: EfficientNet transfer learning	b) Proposed model: EfficientNet fine-tuning

Fig. 3. Training accuracy curves: a) Baseline: EfficientNet transfer learning and b) Proposed model: EfficientNet fine-tuning.

Applied Information System and Management (AISM)
Volume 9, (1) 2026, p. 117–124
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v9i1.50124

Figure 3 compares the accuracy curves of the baseline EfficientNet model with standard transfer learning and the proposed EfficientNet model trained with an incremental fine-tuning strategy. In the baseline graph (a), training accuracy increases sharply in the first few epochs. However, validation accuracy stagnates and begins to decline after a certain point, particularly after the transfer learning process begins, suggesting overfitting and insufficient adaptability to the characteristics of road-damage data. In contrast, the proposed model graph (b) shows a more stable training pattern; both training and validation accuracy increase consistently until the fine-tuning phase, where they recover after slight fluctuations. After fine-tuning begins, the model shows a significant increase in accuracy compared to the baseline, with a smaller gap between training and validation accuracy, indicating better generalization. Overall, this comparison demonstrates that incremental fine-tuning positively affects learning by enhancing training stability and reducing overfitting, resulting in significantly superior final performance.

In contrast, the proposed model exhibits strong alignment between precision and recall. The difference between the two is very small (0.9760 vs. 0.9720), indicating that the model maintains a balance between prediction accuracy and the ability to detect all existing damage patterns. This consistency demonstrates that the gradual fine-tuning process harmoniously improves the model's internal layers, without causing overfitting or imbalance in the learning process.



a) Baseline: EfficientNet transfer learning b) Proposed model: EfficientNet fine-tuning

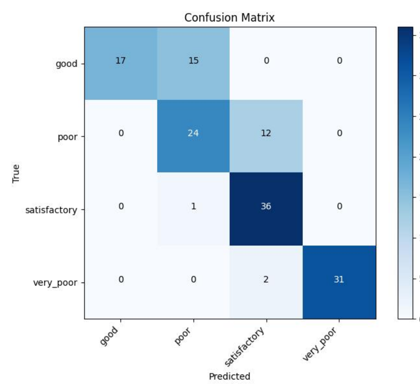
Figure 4. Training loss curves: a) Baseline: EfficientNet transfer learning and b) Proposed model: EfficientNet fine-tuning.

Figure 4 compares the loss curves of the two models, revealing a striking difference in learning dynamics between the baseline EfficientNet and EfficientNet with incremental fine-tuning. In the baseline model (a), both the training and validation losses decline steadily until around the 10th epoch. However, after the transfer learning process begins, the training loss spikes sharply, and the validation loss increases consistently, indicating significant instability and overfitting. The sharp spike in training loss indicates that the model is unable to adapt well when certain layers are retrained. In contrast, the increase in validation loss indicates a loss of generalization.

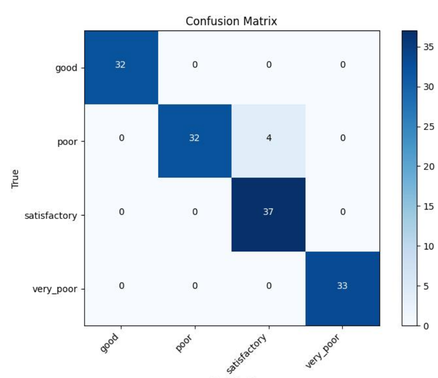
In contrast to the baseline, the proposed model (b) exhibits a much more stable and controlled loss decline. Until the fine-tuning stage, the training loss consistently decreases; despite slight fluctuations during the fine-tuning transition, the model successfully reduces the loss again after additional layers are introduced for training. The validation loss of the proposed model also shows a more balanced trend: it decreases in the initial phase. It increases slightly toward the end of training, while remaining substantially lower than the baseline. Overall, this comparison demonstrates that the incremental fine-tuning strategy produces a more stable learning process, minimizes overfitting, and effectively adapts the model to the characteristics of the road damage data. Based on these results, the incremental fine-tuning strategy not only improves final accuracy but also enhances training dynamics, making the model more stable and generalizable.

C. Confusion Matrix Analysis and Per-Class Performance

Although the confusion matrix is not explicitly included in the table, the pattern of metrics in the results provides a strong indication of predictive behavior across classes. The macro recall of 0.9720 in the proposed model indicates that almost all damage classes are recognized with a very high detection rate. Macro recall reflects the model's ability to identify each damage type equally; this significant improvement indicates that misclassification between visually similar classes, such as longitudinal cracks and transverse cracks, has been drastically reduced.



a) Baseline: EfficientNet transfer learning



b) Proposed model: EfficientNet fine-tuning

Fig. 5. Confusion matrices; a) Baseline: EfficientNet transfer learning and b) Proposed model: EfficientNet fine-tuning.

Based on the confusion matrices Fig. 5, a significant difference is observed between the baseline EfficientNet transfer learning model and the proposed EfficientNet fine-tuning model. In the baseline model, predictions showed several misclassifications, particularly in the good, poor, and satisfactory classes, resulting in an accuracy of around 78%. The most significant errors occurred in the poor class, which was frequently misclassified as satisfactory. Following fine-tuning, the proposed model's performance improved significantly, reaching 97% accuracy and correctly classifying almost all samples. The proposed model demonstrated improvements in precision, recall, and F1-score for almost all classes, and eliminated almost all prediction errors observed in the baseline. Therefore, fine-tuning has been shown to provide substantial improvements in the model's ability to recognize each class more consistently and accurately.

In the baseline model, the macro recall of 0.7775 indicates that it failed to detect some damage categories adequately. This bias is common in standard transfer learning because the feature representations have not fully adapted to the road domain. This indication is also consistent with the baseline F1-score, which is below 0.80, indicating an imbalance between accuracy and sensitivity.

Conceptually, the proposed model's superiority in the confusion matrix can be explained by more specific feature adaptation through incremental fine-tuning. Incremental fine-tuning allows EfficientNet to construct a more refined representation of the damage texture than the baseline, thereby reducing cross-class prediction errors.

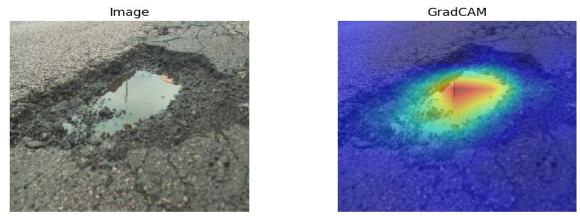
D. Theoretical and Practical Implications

The findings of this study have several important implications. Theoretically, these results reinforce the argument established in the introduction that standard transfer learning is not always adequate for specific domains such as road damage detection, which exhibit highly specific visual patterns. The baseline EfficientNet model proved insufficient for generalizing road-damage features without thorough fine-tuning (Fig. 6). In contrast, incremental fine-tuning provides empirical evidence that gradually unfreezing pre-trained layers can produce more adaptive and capable visual representations.

From a practical perspective, the proposed model's performance, approaching 97% across all metrics, demonstrates its potential for implementation in automated camera-based infrastructure monitoring systems. This high performance also makes the model more reliable under various field conditions, including variations in weather, lighting, and camera angles. These results present opportunities for the model's use in real-time monitoring applications on vehicles, drones, or roadside devices.

The significant performance improvement also confirms that the multi-stage augmentation approach makes a meaningful contribution to building model robustness against visual noise and diverse environmental conditions. This finding is consistent with the literature, which suggests that comprehensive augmentation is particularly effective in domains with high variability.

a) Baseline: EfficientNet transfer learning



b) Proposed model: EfficientNet fine-tuning

Fig. 6. Gradcam results; a) Baseline: EfficientNet transfer learning and b) Proposed model: EfficientNet fine-tuning

E. Summary of Findings

Overall, this study demonstrates that the proposed EfficientNet model with incremental fine-tuning and multi-stage augmentation achieves significant performance improvements over the baseline EfficientNet. The baseline model achieved only 78% accuracy, while the proposed model achieved 97%, representing a nearly 20% improvement. All other metrics, including precision, recall, and F1-score, also improved substantially, with the difference between the baseline and the proposed model exceeding 0.19 points in some metrics. These results confirm the effectiveness of the proposed methodology and support the argument that the road damage detection domain requires an adaptive training strategy, rather than standard transfer learning alone.

This study provides strong evidence that an incremental fine-tuning approach can be a superior method for improving the performance of deep learning models in image-based infrastructure damage analysis tasks, while also providing a relevant scientific foundation for further research in this field.

IV. Conclusion

Based on the results, applying EfficientNet with an incremental fine-tuning strategy and multi-stage data augmentation yields significant performance improvements over the baseline approach that uses only standard transfer learning for the road damage detection task. The baseline model exhibits limited adaptation to the characteristics of the road damage domain, reflected in lower training accuracy and stability, as well as a tendency toward overfitting. In contrast, the proposed model achieved significantly superior performance, with an accuracy of up to 97% and high consistency across all evaluation metrics, including precision, recall, and F1-score. The training accuracy and loss curves demonstrate that incremental fine-tuning results in a more stable learning process, better adjusts EfficientNet's feature representation, and improves the model's generalization to real-world image variations. Therefore, this study provides empirical evidence that the incremental fine-tuning approach is a more appropriate and efficient strategy for developing a deep-learning-based road damage detection system, with strong potential for implementation in automated and real-time infrastructure monitoring applications.

ACKNOWLEDGMENT

The authors would like to express their sincere gratitude to STIKOM Tunas Bangsa for the support and facilities provided during the completion of this research. The institution's support has greatly contributed to the successful implementation and completion of this study. The authors also appreciate all parties who have directly or indirectly contributed to this research, including colleagues and students who provided valuable discussions, insights, and technical assistance throughout the research process. Their contributions and encouragement have been very meaningful in improving the quality of this work.

REFERENCES

[1] S. K. Gooda, "Automatic detection of road cracks using EfficientNet with residual U-Net-based segmentation and YOLOv5-based detection," *Int. J. Recent Innov. Trends Comput. Commun.*, vol. 11, pp. 84–91, 2023, doi: 10.17762/ijritcc.v11i4s.6310.

[2] J. Garcia-Alviz, G. Galindo, J. Arellana, and R. Yie-Pinedo, "Planning road network restoration and relief distribution under heterogeneous road disruptions," *OR Spectr.*, vol. 43, no. 4, pp. 941–981, Dec. 2021, doi: 10.1007/s00291-021-00644-x.

[3] Y. Wang, S. Wang, and X. Hong, "Road extraction using high resolution satellite images based on receptive field and improved Deeplabv3+," *Journal of Physics Conference Series*, vol. 2320, no. 1, Art. no. 012021, Aug. 2022, doi: 10.1088/1742-6596/2320/1/012021.1.

[4] C. Li, "Road recognition method of photovoltaic plant based on improved DeepLabv3+," *Shanghai Jiaotong Daxue Xuebao/Journal Shanghai Jiaotong Univ.*, vol. 58, no. 5, pp. 776–782, 2024, doi: 10.16183/j.cnki.jsjtu.2022.224.

[5] G. Kathiresan, M. Anirudh, M. Nagharjun, and R. Karthik, "Disease detection in rice leaves using transfer learning techniques," *Journal of Physics Conference Series*, vol. 1911, no. 1, Art. no. 012004, May 2021, doi: 10.1088/1742-6596/1911/1/012004.

[6] A. Narula, "Development of CNN-LSTM combinational architecture for COVID-19 detection," *J. Ambient Intell. Humaniz. Comput.*, vol. 14, no. 3, pp. 2645–2656, 2023, doi: 10.1007/s12652-022-04508-2.

[7] Z. Situ, "Real-time sewer defect detection based on YOLO network, transfer learning, and channel pruning algorithm," *J. Civ. Struct. Heal. Monit.*, vol. 14, no. 1, pp. 41–57, 2024, doi: 10.1007/s13349-023-00681-w.

[8] H. Wang, "GAN-STD: small target detection based on generative adversarial network," *J. Real-Time Image Process.*, vol. 21, no. 65, pp. 547–559, 2024, doi: 10.1007/s11554-024-01446-4.

[9] T. A. T. K. Azaharan, "Investigation of VGG-16, ResNet-50 and AlexNet performance for brain tumor detection," *Int. J. online Biomed. Eng.*, vol. 19, no. 8, pp. 97–109, 2023, doi: 10.3991/ijoe.v19i08.38619.

[10] A. D. Goswami, "Electrocardiogram signal classification using VGGNet: a neural network based classification model," *Int. J. Inf. Technol.*, vol. 15, no. 1, pp. 119–128, 2023, doi: 10.1007/s41870-022-01071-z.

[11] F. Yang, "Facial emotion classification based on the modified VGG and MobileNet," *Journal of Physics Conference Series*, vol. 2646, no. 1, Art. no. 012038, Dec. 2023, doi: 10.1088/1742-6596/2646/1/012038.

[12] C. Guo, "Radiographic imaging and diagnosis of spinal bone tumors: AlexNet and ResNet for the classification of tumor malignancy," *J. Bone Oncol.*, vol. 48, no. 3, pp. 342–355, 2024, doi: 10.1016/j.jbo.2024.100629.

[13] X. Yin, "Fault diagnosis of wind turbine gearbox based on multiscale residual features and ECA-stacked ResNet," *IEEE Sens. J.*, vol. 23, no. 7, pp. 7320–7333, 2023, doi: 10.1109/JSEN.2023.3244929.

[14] S. Vasavi, "Classification of buildings from VHR satellite images using ensemble of U-Net and ResNet," *Egypt. J. Remote Sens. Sp. Sci.*, vol. 26, no. 4, pp. 937–953, 2023, doi: 10.1016/j.ejrs.2023.11.008.

[15] F. Muftic, "Deep learning for magnetic resonance imaging brain tumor detection: evaluating ResNet, EfficientNet, and VGG-19," *Int. J. Electr. Comput. Eng.*, vol. 14, no. 6, pp. 6360–6372, 2024, doi: 10.11591/ijece.v14i6.pp6360-6372.

[16] H. Wang, "Contrast-enhanced magnetic resonance image segmentation based on improved U-Net and Inception-ResNet in the diagnosis of spinal metastases," *J. Bone Oncol.*, vol. 42, no. 65, pp. 678–690, 2023, doi: 10.1016/j.jbo.2023.100498.

[17] P. Chakurkar and D. Vora, "Context-aware deep learning based Indian footpath damage segmentation dataset for risk assessment," vol. 23, no. 65, pp. 1–13, 2025, doi: https://doi.org/10.1038/s41597-025-06207-x.

[18] P. Arafin, "Concrete structure damage classification and detection using convolutional neural network," M.S. Thesis, Dept. of Civil Engineering, Lakehead University, Ontario, 2022. Accessed on: Jan. 19, 2026. Available: https://knowledgecommons.lakeheadu.ca/server/api/core/bitstreams/fd53ca74-69c1-48ec-9c6b-4ab2138485f8/content

[19] J. Wang, "Automatic liver segmentation using EfficientNet and attention-based residual U-Net in CT," *J. Digit. Imaging*, vol. 35, no. 6, pp. 1479–1493, 2022, doi: 10.1007/s10278-022-00668-x.

[20] A. Srivastava, "Agile testing efficiency using an integrated approach of fuzzy-based MOORA and AHP," *Int. J. Performability Eng.*, vol. 17, no. 2, pp. 208–215, 2021, doi: 10.23940/ijpe.21.02.p5.208215.

[21] A. Ghosh, "A fine-tuned EfficientNet B1 based deep transfer learning framework for multiple types of brain disorder classification," *Iran. J. Sci. Technol. - Trans. Electr. Eng.*, vol. 48, no. 3, pp. 1279–1299, 2024, doi: 10.1007/s40998-024-00726-w.

[22] Y. Çiçek, "Classifying weed development stages using deep learning methods: Classifying weed development stages with DenseNET, Xception, SqueezeNET, GoogleNET, EfficientNET CNN models using ROI images," *Int. J. Adv. Comput. Sci. Appl.*, vol. 16, no. 2, pp. 619–626, 2025, doi: 10.14569/IJACSA.2025.0160263.

[23] F. E. Botchey, Z. Qin, and K. Hughes-Lartey, "Mobile money fraud prediction—A cross-case analysis on the efficiency of support vector machines, gradient boosted decision trees, and naïve bayes algorithms," Information, vol. 11, no. 8, Art. no. 383, Jul. 2020, doi: 10.3390/info11080383.

[24] M. A. R. Khan, "Car make and model recognition using convolutional neural network: fine-tune AlexNet architecture," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 33, no. 1, pp. 370–379, 2024, doi: 10.11591/ijeecs.v33.i1.pp370-379.

[25] A. Ghosh, "Transfer learning-based deep feature extraction framework using fine-tuned EfficientNet B7 for multiclass brain tumor classification," *Arab. J. Sci. Eng.*, vol. 49, no. 9, pp. 12027–12048, 2024, doi: 10.1007/s13369-023-08607-w.

[26] F. M. J. M. Shamrat et al., "LunGNeT22: A Fine-Tuned model for multiclass classification and prediction of lung disease using x-ray images," *Journal of Personalized Medicine*, vol. 12, no. 5, Art. no. 680, Apr. 2022, doi: 10.3390/jpm12050680.

[27] Q. Hou, R. Xia, J. Zhang, Y. Feng, Z. Zhan, and X. Wang, "Learning visual overlapping image pairs for SfM via CNN fine-tuning with photogrammetric geometry information," *International Journal of Applied Earth Observation and Geoinformation*, vol. 116, Art. no. 103162, Dec. 2022, doi: 10.1016/j.jag.2022.103162.

[28] A. Sekhar, "Brain tumor classification using fine-tuned googlenet features and machine learning algorithms: IoMT enabled CAD system," *IEEE J. Biomed. Heal. Informatics*, vol. 26, no. 3, pp. 983–991, 2022, doi: 10.1109/JBHI.2021.3100758.

[29] T. E. Anju, "Tissue and tumor epithelium classification using fine-tuned deep CNN Models," *Int. J. Adv. Comput. Sci. Appl.*, vol. 13, no. 9, pp. 306–314, 2022, doi: 10.14569/IJACSA.2022.0130936.

[30] T. Tsukamoto, "Comparison of fine-tuned deep convolutional neural networks for the automated classification of lung cancer cytology images with integration of additional classifiers," *Asian Pacific J. Cancer Prev.*, vol. 23, no. 4, pp. 1315–1324, 2022, doi: 10.31557/APJCP.2022.23.4.1315.

[31] N. Becherer, J. Pecarina, S. Nykl, and K. Hopkinson, "Improving optimization of convolutional neural networks through parameter fine-tuning," *Neural Comput. Appl.*, vol. 31, no. 8, pp. 3469–3479, 2019, doi: 10.1007/s00521-017-3285-0.