EDUMATIC
Jurnal Pendidikan Informatika

e-ISSN 2549-7472
Edumatic: Jurnal Pendidikan Informatika
Volume 10, Issue 1, **100-109.**
https://e-journal.hamzanwadi.ac.id/index.php/edumatic
DOI: 10.29408/edumatic.v10i1.33699

An Efficient Two Stage Detection Segmentation Framework for Automated Road Crack Assessment

Alvin Hujaya [1,*], Muhammad Rizky Pribadi [1]

[1] Universitas Multi Data Palembang, Indonesia
[*] Corresponding author: Alvin Hujaya, Universitas Multi Data Palembang, Indonesia
✉ 2226250103@mhs.mdp.ac.id

Received: 24 December 2025 | Revised: 12 January 2026 | Accepted: 12 March 2026 | Published: 16 March 2026

Abstract
Road cracks significantly degrade infrastructure quality and pose a threat to traffic safety. To minimize manual inspection inefficiencies, this study investigates a segmentation model integrating MobileNetV3-Small as a backbone for the U-Net architecture to reduce processing time. The performance of the proposed MobileNetV3-Small-U-Net is benchmarked against a standard U-Net using three public datasets: DeepCrack (537 images), CFD (118 images), and Crack500 (3368 images) sourced from GitHub and Kaggle. This research explores the influence of optimization algorithms on evaluation results across these diverse datasets. Specifically, the study evaluates Adam, RMSprop, and SGD optimizers at an image resolution of 224 x 224 pixels, with a 0.001 learning rate and 0.9 momentum. On-the-fly augmentation techniques, including horizontal flips and brightness adjustments (0.8 to 1.2), were implemented during training. Experimental results demonstrate that MobileNetV3-Small-U-Net enhances computational efficiency by achieving a 9 ms inference time, which is 2 ms faster than the standard U-Net. These findings confirm that a MobileNetV3-Small backbone accelerates inference, despite a slight trade-off in evaluation metrics. Additionally, results reveal that the SGD optimizer is unsuitable for these segmentation tasks due to high error rates and the lack of an adaptive learning rate.

Keywords: infrastructure monitoring; road crack; semantic segmentation; ssd-mobilenetv3; u-net

To cite this article: Hujaya, A., & Pribadi, M. R. (2026). An Efficient Two Stage Detection Segmentation Framework for Automated Road Crack Assessment. *Edumatic: Jurnal Pendidikan Informatika*, *10*(1), 100–109. https://doi.org/10.29408/edumatic.v10i1.33699

INTRODUCTION

Road cracks are a form of structural damage that significantly reduces infrastructure performance and threatens global traffic safety (Cao et al., 2023; López et al., 2026). These cracks allow corrosive elements to penetrate pavement layers, accelerating structural failure and lowering the tensile strength of the material, which makes periodic detection essential (Luo et al., 2023). This structural degradation is strongly linked to severe accidents and economic losses; for instance, China recorded 273,098 accidents in 2021 and invested 1.29 trillion CNY in road maintenance in 2022, highlighting the urgent need for accurate early crack monitoring (Deng et al., 2023; Jia et al., 2026).

Manual inspections are increasingly viewed as ineffective, being subjective, costly, and time-consuming, which has driven the adoption of computer vision and Artificial Intelligence (AI) for automated, high-precision detection (Liu et al., 2026; Xu et al., 2024). These detection technologies improve object localization and situational awareness, supporting diverse applications such as autonomous navigation and industrial quality control (Trigka & Dritsas, 2025). Image segmentation provides pixel-level accuracy essential for high-precision tasks and autonomous navigation. This granularity allows for a comprehensive understanding of complex environments, which is vital for robust infrastructure monitoring (Joy et al., 2026).

According to previous research, the SSD-MobileNet architecture has proven to be more memory-efficient and superior in accuracy compared to YOLOv8 (Dharma et al., 2025). This is further supported by a study conducted by (Abdusalomov et al., 2025), which states that MobileNetV3-Small + SSD features a highly compact file size of 8.7 MB and achieves an inference speed of 62.5 FPS. Consequently, the MobileNet model is highly suitable for deployment in various public facilities due to its small memory footprint, relatively high accuracy, and cost-effective implementation.

While MobileNet's advantages are significant, its classification performance can be further enhanced by utilizing semantic segmentation models such as U-Net (Mayya & Alkayem, 2024). The U-Net architecture is renowned for its robust feature extraction capabilities, even surpassing the performance of Swin-Unet and Deeplabv3 with precision and recall scores of 0.9461 in road crack case studies (He & Lau, 2024). This architecture is capable of precisely automating the quantification of microstructure parameters without manual intervention, while simultaneously enhancing detection robustness in autonomous systems through reliable sensor fusion (Patel et al., 2026). U-Net is widely recognized as a benchmark model for road crack segmentation due to its ability to detect cracks with significant depth and precision.

Integrating these two architectures offers significant potential for infrastructure monitoring by leveraging U-Net’s robust feature extraction and MobileNet’s computational efficiency (Morellos et al., 2024). In UAV-based applications, the MobileNet-U-Net fusion has achieved a precision of 85.1% (Lourenco et al., 2025). This demonstrates the feasibility and effectiveness of such a combination, as evidenced by its strong performance in previous UAV-related studies. Furthermore, optimizer selection plays a crucial role in influencing convergence speed, accuracy, and generalization (Dogo et al., 2022; Kompanets et al., 2025). Optimization algorithms assist neural networks in iteratively minimizing the cost function to identify optimal parameters (Mehmood et al., 2023).

While numerous studies have explored the integration of MobileNet and U-Net across various domains, the application of MobileNetV3 remains significantly underexplored compared to its predecessors. Existing literature often relies on single-dataset evaluations to validate findings; however, this limitation restricts a comprehensive understanding of how these models generalize across diverse road textures and environmental conditions. as in the research discussed by (Yu et al., 2024), which suffers from a limited variety of data. Furthermore, prior research typically involves training models with a single optimizer, raising questions regarding whether peak performance has truly been achieved.

This research aims to investigate the fusion of MobileNetV3-Small and U-Net by utilizing MobileNetV3-Small as the backbone for the U-Net architecture. By leveraging three diverse public datasets DeepCrack, CFD, and Crack500 this study analyzes the impact of optimization algorithms on model performance and investigates whether dataset variations yield significantly divergent results. Additionally, this study adopts the Dice coefficient for evaluation, whereas previous research has primarily relied on Confusion Matrices and MioU. Furthermore, this study will compare the impact of utilizing MobileNet as a backbone against a standard U-Net to evaluate the specific contributions and improvements derived from the MobileNet architecture.

METHOD

This study adopts a systematic quantitative approach to transform raw data into strategic insights through advanced deep learning experiments and architectures. To ensure model validity, the research utilizes three diverse public datasets sourced from GitHub and Kaggle, covering various real-world complexities. This methodological structure facilitates an objective evaluation across these environments.

Specifically, this study utilizes three primary public datasets: DeepCrack, consisting of 537 images; CFD, with 118 images; and Crack500, comprising 3,368 images. These datasets integrate various technical complexities, ranging from variations in crack width to environmental noise, to ensure a robust evaluation of the models under study. The use of diverse data sources enables a more comprehensive analysis of various road surface conditions encountered in real-world scenarios. Consequently, this extensive dataset integration not only enriches the training process but also strengthens the validity of the test results across different scenarios. The balance between the quantity of images and the level of environmental difficulty within these datasets serves as a critical foundation for achieving optimal segmentation accuracy.

The U-Net algorithm was first introduced for segmentation in 2015. Structurally, U-Net comprises an encoder and a decoder; in this study, the architecture is modified by integrating MobileNetV3-Small as the backbone for the encoder section. The model is implemented using the PyTorch framework, with a training duration of 50 epochs and an input image resolution of 224 x 224 pixels. Each dataset is partitioned into 80% training data and 20% validation data. RMSprop, Adam, and SGD (0.9 momentum) optimizers are evaluated with a uniform learning rate of 0.001. The training process employs on-the-fly data augmentation techniques, including horizontal flips with a 0.5 probability and brightness adjustments ranging from 0.8 to 1.2, further complemented by an early stopping mechanism with a patience of 7 epochs.

Performance quality will be assessed using a confusion matrix, encompassing metrics such as accuracy, precision, recall, and F1-score. This matrix serves as the primary basis for evaluating and comparing the performance of the implemented classifiers (Valero-Carreras et al., 2023). In addition to the confusion matrix, model evaluation will incorporate Mean Intersection over Union (MIoU) and the Dice coefficient, which function to accurately measure the degree of overlap between the ground truth and the model's predictions.

RESULTS AND DISCUSSION

Results

This research integrates three primary public datasets to facilitate road damage analysis. The DeepCrack, CrackForestDataset (CFD), and Crack500 datasets are implemented for segmentation modeling using the U-Net architecture with a MobileNetV3-Small backbone. The DeepCrack dataset, serves as a cornerstone for segmentation due to its high-quality annotations. It comprises 537 RGB images with a standardized resolution of 544x384 pixels, each paired with a manually labeled binary mask. A defining feature of this dataset is its extreme variation in crack widths, ranging from 1 to 180 pixels, which allows the model to learn various distress scenarios on both asphalt and concrete surfaces under diverse textures. Example images from the DeepCrack dataset can be seen in Figure 1.

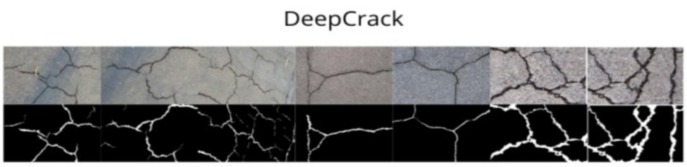

Figure 1. DeepCrack dataset

The CrackForestDataset (CFD) utilizes 118 images captured from various road environments in Beijing, China, to evaluate model performance under heterogeneous urban conditions. These images, resized to 460x320 pixels, intentionally include environmental 'noise' such as shadows, oil stains, and puddles. The inclusion of these elements is crucial for testing the robustness of the U-Net model, as it must learn to distinguish actual road cracks from common visual disturbances frequently encountered in everyday urban scenarios. Sample images from the CFD dataset are presented in Figure 2.

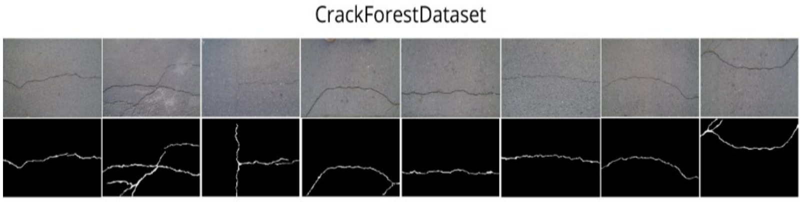

Figure 2. CFD dataset

Crack500 contains 3,368 images captured at a resolution of 2000x1500 pixels. Following cropping and filtering techniques from previous research, these images were partitioned into smaller regions, where only patches containing more than 1,000 pixels were retained. This process resulted in a dataset split of 1,896 training images, 348 validation images, and 1,124 test images. Representative samples from the Crack500 dataset are presented in Figure 3.

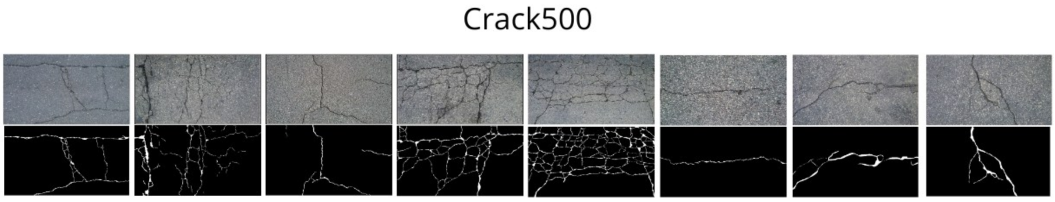

Figure 3. Crack500 Dataset

U-Net training for semantic segmentation utilized a mini-batch size of 8 and early stopping with a 7-epoch patience threshold to prevent overfitting. Inputs were standardized to 224 x 224 pixels to balance computational speed and spatial feature preservation. A static 0.001 learning rate with the Adam optimizer was employed for rapid convergence. These settings were applied uniformly across the DeepCrack, CFD, and Crack500 datasets using an 80:20 training-to-validation split to maximize learning capacity from sparse samples.

Table 1. Training results of the u-net model deepcrack

Optimizer	Accuracy	Precision	Recall	F1-Score	MIoU	Dice
Adam	98.97%	91.96%	77.54%	84.14%	85.78%	84.14%
SGD	98.96%	85.14%	76.46%	80.56%	83.19%	80.56%
RMSprop	98.94%	88.33%	82.46%	85.29%	86.63%	85.29%

Table 2. Training results of the u-net model cfd

Optimizer	Accuracy	Precision	Recall	F1-Score	MIoU	Dice
Adam	98.94%	68.98%	52.85%	59.85%	70.82%	59.85%
SGD	98.49%	82.00%	01.56%	03.06%	50.02%	03.06%
RMSprop	98.95%	71.36%	62.18%	66.45%	74.35%	66.45%

Experimental results on the DeepCrack dataset demonstrate that RMSprop outperforms both Adam and SGD in terms of MIoU and Dice metrics. This superiority is primarily attributed to its higher recall value, particularly in identifying fine crack pixels. While all optimizers maintained high accuracy levels, this was largely driven by the dominance of background pixels within the images. As shown in Table 1, RMSprop is identified as the most optimal algorithm for the segmentation tasks on this dataset, owing to its ability to maximize the overlap between the prediction results and the ground truth. These findings indicate that the use of different optimization algorithms can lead to varying model evaluation outcomes.

Based on the training results of the U-Net model on the CFD dataset in Table 2, RMSprop emerged as the superior optimizer, achieving an MIoU of 74.35% and a Dice score of 66.45%. Although all optimizers recorded accuracy levels above 98%, these figures are misleading as they primarily reflect the dominance of background pixels within the dataset rather than the model's ability to specifically detect cracks. The SGD optimizer experienced a critical failure with an extremely low Recall of 1.56% despite reaching a precision of 82.00%, whereas RMSprop proved far more robust in identifying crack features amidst visual disturbances like shadows, achieving a Recall of 62.18%.

The failure of SGD, resulting in Dice and F1-scores of only 3.06%, was driven by the "accuracy paradox," where the model fell into a pattern of predicting only the background to minimize loss rapidly. Lacking an adaptive learning rate like RMSprop, SGD struggled to escape flat local minima, failing to learn the sparse crack details within the dataset. Mathematically, this extreme low Recall dragged down the harmonic mean scores (Dice and F1), while the mIoU of 50.02% confirms that the model's actual performance on the crack class was near zero, with the score being buoyed almost entirely by the background class. These results demonstrate that the use of SGD is highly unsuitable for this dataset. Due to the absence of an adaptive learning rate within its algorithmic structure, SGD faces significant challenges in escaping flat local minima.

Table 3. Training results of the u-net model crack500

Optimizer	Accuracy	Precision	Recall	F1-Score	MIoU	Dice
Adam	97.44%	79.66%	69.52%	74.25%	78.19%	74.25%
SGD	97.27%	79.66%	68.48%	73.65%	77.73%	73.65%
RMSprop	97.51%	81.91%	70.33%	75.68%	79.14%	75.68%

Results on Crack-500 in Table 3 again highlight RMSprop's superiority, achieving 79.14% MIoU and 75.68% Dice. Unlike the CFD anomalies, all optimizers showed consistent stability with accuracy levels near 97%, and SGD delivered competitive results. These results indicate that the higher precision and recall values achieved by RMSprop confirm that adaptive learning rate algorithms are more effective at segmenting complex crack features across various and diverse dataset scales.

Across all datasets, RMSprop consistently achieved the highest MIoU and Dice scores. DeepCrack yielded the best performance due to clear crack features, while CFD was the most challenging, causing a critical failure for SGD due to visual noise. Ultimately, while all tests exceeded 97% accuracy, robust segmentation depends on adaptive optimizers like RMSprop to maintain sensitivity across diverse difficulty levels.

Table 4. Training results of the mobilenetv3-small-u-net model deepcrack

Optimizer	Accuracy	Precision	Recall	F1-Score	MIoU	Dice
Adam	99.08%	88.34%	86.68%	87.50%	77.78%	87.50%
SGD	98.42%	80.26%	73.40%	76.68%	62.18%	76.68%
RMSprop	98.93%	89.42%	83.03%	86.02%	76.48%	86.02%

Consistent with previous U-Net implementations, the MobileNetV3-Small-U-Net utilizes a batch size of 8 and an early stopping mechanism with a patience of 7 epochs. The input resolution is set to 224 x 224 pixels, with a uniform learning rate of 0.001 applied across all optimizers. Furthermore, the dataset is partitioned into an 80:20 ratio, allocating 80% of the data for training and the remaining 20% for validation purposes.

There are notable changes in the results compared to previous findings, where the Adam optimizer emerged as the best-performing algorithm for the MobileNetV3-Small-U-Net on this dataset. Nevertheless, a significant 8% decrease in MIoU was observed relative to previous outcomes, suggesting that the model is less effective in predicting the overlap with the ground truth area. Furthermore, these results indicate that algorithms featuring an adaptive learning rate outperform SGD, which lacks such a mechanism. These findings also suggest that the choice of an appropriate optimization algorithm is contingent not only on the dataset but also on the specific model architecture being utilized.

Table 5. Training results of the mobilenetv3-small-u-net model cfd

Optimizer	Accuracy	Precision	Recall	F1-Score	MIoU	Dice
Adam	98.80%	66.42%	44.62%	53.38%	36.40%	53.38%
SGD	97.89%	0%	0%	0%	0%	0%
RMSprop	98.72%	72.54%	44.94%	55.50%	38.41%	55.50%

According to these results, RMSProp remains the most effective optimization algorithm for this dataset, consistent with prior findings. Nevertheless, there are significant declines in key metrics, specifically a 35% reduction in MIoU and a 17% decrease in Recall. This indicates that the model struggles to accurately predict the spatial overlap with the ground truth. Furthermore, SGD exhibited a substantial regression, suggesting that this optimization algorithm is unsuitable for the characteristics of this particular dataset. These results provide deeper evidence regarding the influence of optimization algorithms on model evaluation and dataset performance. Furthermore, they highlight the impact of adaptive learning rates in enabling the model to effectively learn from its previous iterations and errors.

Table 6. Training results of the mobilenetv3-small-u-net model crack500

Optimizer	Accuracy	Precision	Recall	F1-Score	MIoU	Dice
Adam	97.44%	78.62%	73.69%	76.07%	61.38%	76.07%
SGD	97.18%	77.97%	69.46%	73.47%	58.06%	73.47%
RMSprop	97.36%	79.74%	69.87%	74.48%	59.33%	74.48%

Based on the results in Table 6, the Adam optimizer emerged as the most effective algorithm for this dataset. This finding contradicts previous outcomes, which identified RMSprop as the superior optimizer. Nevertheless, the results presented in this table exhibit substantial improvements in accuracy, precision, recall, F1-score, and the Dice coefficient. However, a persistent deficiency in MIoU remains, consistent with the trends observed in Table 4 and Table 5, specifically showing a significant 16% decline. This suggests that the integration of MobileNet and U-Net may lead to a diminished capacity for the model to accurately predict the spatial overlap with the ground truth.

Based on the aggregate results, it is evident that the highest performance was achieved using the DeepCrack dataset with the Adam optimizer, reinforcing previous findings that DeepCrack is the most compatible dataset for this model. These outcomes suggest that the architecture is highly effective at segmenting cracks with significant depth compared to other datasets. Conversely, the results also indicate that the model's performance remains limited when faced with substantial environmental noise, such as the shadows, oil stains, and puddles prevalent in the CFD.

Subsequent experiments will focus on the inference speed and the time required for the model to generate predictions, alongside a comparative analysis of its accuracy in detecting road cracks. The images for this testing phase are randomly sourced from Google; this approach aims to ensure that the evaluation utilizes uncurated imagery that has not been pre-processed or specifically tailored for a dataset.

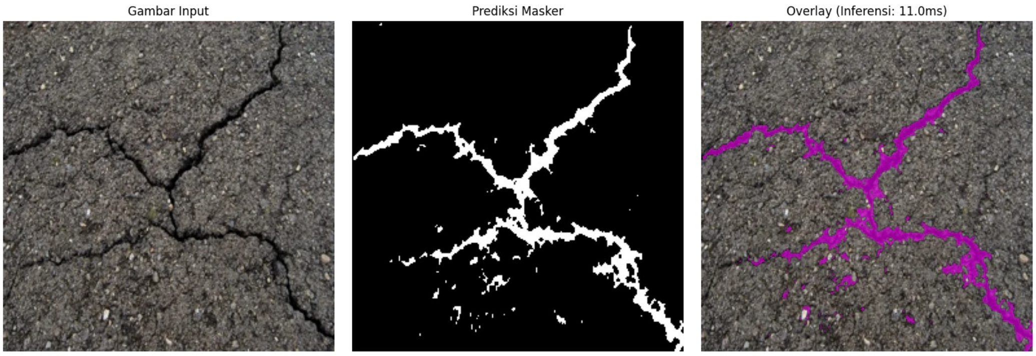

Figure 4. Experiments utilizing the standard u-net

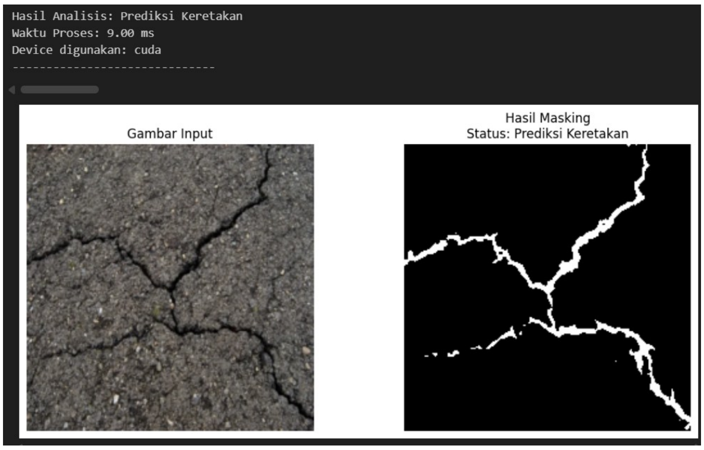

Figure 5. Experiments utilizing the mobilenetv3-small-u-net architecture

Figures 4 and 5 illustrate that the processing time for the standard U-Net is 11 ms, whereas the MobileNetV3-Small-U-Net achieves an inference time of only 9 ms. This demonstrates that integrating MobileNet as the backbone (encoder) enhances model speed, with a recorded difference of 2 ms in this study. Furthermore, the generated segmentation masks reveal that the MobileNetV3-Small-U-Net produces significantly cleaner results compared to the standard U-Net. This proves that by utilizing MobileNet, the model can more precisely detect cracks without being distracted by background artifacts.

Discussion

The experimental results demonstrate that the integration of MobileNetV3-Small as the encoder backbone in the U-Net architecture introduces a trade-off between computational efficiency and segmentation accuracy. While the proposed MobileNetV3-Small-U-Net model achieved faster inference (9 ms) compared with the standard U-Net (11 ms), the segmentation overlaps metrics such as MIoU decreased across several datasets. This indicates that the lightweight architecture improves computational efficiency but slightly reduces the model's ability to capture fine spatial details of crack boundaries. This phenomenon is consistent with the nature of lightweight convolutional networks, where parameter reduction often limits the representational capacity required for highly detailed segmentation tasks.

A deeper examination of the results across the three datasets reveals that dataset characteristics play a crucial role in determining model performance. The DeepCrack dataset consistently produced the highest segmentation accuracy, particularly when optimized using Adam. This can be attributed to the dataset's relatively clear crack patterns and high-quality annotations, which provide strong supervision during training. In contrast, the CFD dataset presented significant challenges due to the presence of complex background noise such as shadows, oil stains, and water puddles. These visual disturbances increase the difficulty of distinguishing crack pixels from non-crack features, leading to lower recall and MIoU values. Similar findings were reported by Liu et al. (2026), who noted that environmental noise remains one of the primary limitations in automated crack segmentation systems.

The CFD presents significant challenges for segmentation models due to substantial environmental noise such as shadows, oil stains, and puddles that visually resemble crack patterns. These disturbances make it difficult for the model to distinguish cracks from background artifacts, leading to lower recall and MIoU values. This finding indicates that segmentation models remain highly sensitive to background complexity, and environmental noise continues to be a major challenge in automated road crack detection.

Adaptive optimizers mitigate this issue by dynamically adjusting learning rates based on gradient statistics during training. RMSprop, for instance, maintains a running average of squared gradients to normalize parameter updates, enabling the model to continue learning even when gradients become sparse. This capability allows the optimizer to maintain sensitivity to small crack features despite severe class imbalance. The results of this study therefore reinforce previous research by Dogo et al. (2022), which demonstrated that adaptive optimizers significantly improve convergence stability in convolutional neural networks for image-based classification and segmentation tasks.

In addition to accuracy metrics, inference speed represents another critical factor for practical deployment in real-world infrastructure monitoring systems. The MobileNetV3-Small-U-Net architecture demonstrated a measurable improvement in processing efficiency by reducing inference time by approximately 2 ms compared to the standard U-Net. Although this improvement may appear modest in isolated experiments, the impact becomes substantial in large-scale monitoring systems where thousands of images must be processed continuously. Lightweight architectures such as MobileNet are particularly advantageous for deployment in edge computing environments, including mobile inspection systems, drones, and embedded monitoring devices where computational resources are limited.

Another noteworthy observation from this study is the variation in optimizer effectiveness across different model architectures. While RMSprop consistently produced the best results for the standard U-Net architecture, Adam performed better in the MobileNetV3-Small-U-Net configuration. This finding suggests that optimizer performance is not solely dependent on dataset characteristics but also influenced by the structural design of the neural network architecture. Consequently, selecting an appropriate optimizer should be considered an integral component of model design rather than a secondary training parameter.

This study identifies several limitations concerning the datasets and noise sensitivity. The evaluation currently relies on public datasets sourced from GitHub and Kaggle, which may not fully represent real-world outcomes for real-time applications due to the varying road conditions across different countries. Furthermore, the dataset size used in this research remains limited, highlighting the need for a larger volume of data in future studies. Regarding noise, as evidenced in Tables 2 and 5, the model still struggles with environmental interference, necessitating further investigation. Consequently, future research should prioritize the collection of city-specific road data to validate the model's performance in diverse real-time environments. Moreover, future studies could incorporate preprocessing filters such as Gaussian or Median filters to reduce image noise and enhance the model's accuracy in crack detection.

CONCLUSION

This study proposes a lightweight road crack segmentation framework by integrating MobileNetV3-Small as the encoder backbone within the U-Net architecture to improve computational efficiency while maintaining segmentation performance. Experimental evaluations on three public datasets (DeepCrack, CFD, and Crack500) demonstrate that the proposed MobileNetV3-Small-U-Net achieves faster inference (9 ms), which is 2 ms faster than the standard U-Net, indicating its suitability for near real-time infrastructure monitoring applications. However, the integration of MobileNetV3-Small introduces a slight reduction in segmentation overlap performance, particularly in MIoU, suggesting a trade-off between model efficiency and spatial segmentation accuracy. The results further reveal that dataset characteristics and optimizer selection significantly influence model performance, where adaptive optimizers such as Adam and RMSprop consistently outperform SGD, especially in handling class imbalance and complex crack patterns. Overall, the proposed framework offers a practical balance between accuracy and computational efficiency for automated road crack detection, although future research should focus on improving robustness against environmental noise and incorporating more diverse datasets to enhance generalization in real-world scenarios.

REFERENCES

Abdusalomov, A., Umirzakova, S., Kutlimuratov, A., Mirzaev, D., Dauletov, A., Botirov, T., Zakirova, M., Mukhiddinov, M., & Cho, Y. I. (2025). Lightweight UAV-Based System for Early Fire-Risk Identification in Wild Forests. *Fire*, *8*(8). https://doi.org/10.3390/fire8080288

Cao, H., Gao, Y., Cai, W., Xu, Z., & Li, L. (2023). Segmentation Detection Method for Complex Road Cracks Collected by UAV Based on HC-Unet++. *Drones*, *7*(3). https://doi.org/10.3390/drones7030189

Deng, L., Zhang, A., Guo, J., & Liu, Y. (2023). An Integrated Method for Road Crack Segmentation and Surface Feature Quantification under Complex Backgrounds. *Remote Sensing*, *15*(6). https://doi.org/10.3390/rs15061530

Dharma, A. S., Pardosi, C. N. S., & Silaen, Z. P. (2025). Comparative Performance of Yolov8 and Ssd-mobilenet Algorithms for Road Damage Detection in Mobile Applications. *Sinkron*, *9*(3), 1159–1169. https://doi.org/10.33395/sinkron.v9i3.15008

Dogo, E. M., Afolabi, O. J., & Twala, B. (2022). On the Relative Impact of Optimizers on Convolutional Neural Networks with Varying Depth and Width for Image Classification. *Applied Sciences (Switzerland)*, *12*(23). https://doi.org/10.3390/app122311976

He, M., & Lau, T. L. (2024). CrackHAM: A Novel Automatic Crack Detection Network Based on U-Net for Asphalt Pavement. *IEEE Access*, *12*, 12655–12666. https://doi.org/10.1109/ACCESS.2024.3353729

Jia, J., Zhang, M., Zhang, J., & Jing, C. (2026). Cracking Characteristics of Asphalt Pavement Under Thermal Stresses. *Materials*, *19*(4), 771. https://doi.org/10.3390/ma19040771

Joy, J., Ram, B. G., & Sun, X. (2026). Real-time deep learning segmentation for agricultural applications: Insights for precision weed management. *Smart Agricultural Technology*, *13*, 101894. https://doi.org/10.1016/j.atech.2026.101894

Kompanets, A., Duits, R., Pai, G., Leonetti, D., & Snijder, H. H. (Bert). (2025). Loss function inversion for improved crack segmentation in steel bridges using a CNN framework. *Automation in Construction*, *170*, 105896. https://doi.org/10.1016/j.autcon.2024.105896

Liu, Y., Du, X., & Li, W. (2026). DCP-TransUNet: An Approach for Crack Segmentation on Roads. *Sensors*, *26*(3). https://doi.org/10.3390/s26031071

López, P., Zubasti, P., García, J., & Molina, J. M. (2026). A UAV-Based Framework for Visual Detection and Geospatial Mapping of Real Road Surface Defects. *Drones*, *10*(2), 119. https://doi.org/10.3390/drones10020119

Lourenco, M., Caetano, D., Mora, A., Oliveira, L. B., & Oliveira, H. (2025). Accurate Rural Road Network With Improved Dataset for Strengthen Wildfire Emergency Response. *IEEE Access*, *13*, 146983–147001. https://doi.org/10.1109/ACCESS.2025.3600192

Luo, J., Lin, H., Wei, X., & Wang, Y. (2023). Adaptive Canny and Semantic Segmentation Networks Based on Feature Fusion for Road Crack Detection. *IEEE Access*, *11*, 51740–51753. https://doi.org/10.1109/ACCESS.2023.3279888

Mayya, A. M., & Alkayem, N. F. (2024). Enhance the Concrete Crack Classification Based on a Novel Multi-Stage YOLOV10-ViT Framework. *Sensors*, *24*(24). https://doi.org/10.3390/s24248095

Mehmood, F., Ahmad, S., & Whangbo, T. K. (2023). An Efficient Optimization Technique for Training Deep Neural Networks. *Mathematics*, *11*(6). https://doi.org/10.3390/math11061360

Morellos, A., Dolaptsis, K., Tziotzios, G., Pantazi, X. E., Kateris, D., Berruto, R., & Bochtis, D. (2024). An IoT Transfer Learning-Based Service for the Health Status Monitoring of Grapevines. *Applied Sciences (Switzerland)*, *14*(3). https://doi.org/10.3390/app14031049

Patel, R., Rajput, P. A., Prabhu, G., Sharma, P., & Karthik, G. M. (2026). Deep learning-driven automated microstructural quantification in tungsten heavy alloys using U-net segmentation and image processing. *International Journal of Refractory Metals and Hard Materials*, *138*, 107719. https://doi.org/https://doi.org/10.1016/j.ijrmhm.2026.107719

Trigka, M., & Dritsas, E. (2025). A Comprehensive Survey of Machine Learning Techniques and Models for Object Detection. *Sensors*, *25*(1). https://doi.org/10.3390/s25010214

Valero-Carreras, D., Alcaraz, J., & Landete, M. (2023). Comparing two SVM models through different metrics based on the confusion matrix. *Computers and Operations Research*, *152*, 106131. https://doi.org/10.1016/j.cor.2022.106131

Xu, Y., Xia, Y., Zhao, Q., Yang, K., & Li, Q. (2024). A Road Crack Segmentation Method Based on Transformer and Multi-Scale Feature Fusion. *Electronics (Switzerland)*, *13*(12), 1–21. https://doi.org/10.3390/electronics13122257

Yu, J., Qian, S., & Chen, C. (2024). Lightweight Crack Automatic Detection Algorithm Based on TF-MobileNet. *Applied Sciences (Switzerland)*, *14*(19). https://doi.org/10.3390/app14199004