JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)
VOL. 11. NO. 2 NOVEMBER 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i2.6747

# OPTIMIZING SHUFFLENET WITH GRIDSEARCHCV FOR GEOSPATIAL DISASTER MAPPING IN INDONESIA

**Abdullah Ahmad[1*]; Dedy Hartama[1]; Solikhun[1]; Poningsih[1]**

Informatics Study Program[1]
STIKOM Tunas Bangsa, Pematangsiantar, Indonesia[1]
www.stikomtunasbangsa.ac.id[1]
abdul@amiktunasbangsa.ac.id*,dedyhartama@amiktunasbangsa.ac.id,solikhun@amiktunasbangsa.ac.id,poningsih@amiktunasbangsa.ac.id

(*) Corresponding Author
(Responsible for the Quality of Paper Content)



***Abstract***— *Accurate classification of natural disasters is crucial for timely response and effective mitigation. However, conventional approaches often suffer from inefficiency and limited reliability, highlighting the need for automated deep learning solutions. This study proposes an optimized Convolutional Neural Network (CNN) based on the lightweight ShuffleNet architecture, enhanced through GridSearchCV for systematic hyperparameter tuning. Using a geospatial dataset of 3,667 images representing earthquake, flood, and wind-related disasters in Indonesia, the optimized ShuffleNet model achieved a peak accuracy of 99.97%, outperforming baseline CNNs such as MobileNet, GoogleNet, ResNet, DenseNet, and standard ShuffleNet. While these results demonstrate the potential of combining lightweight architectures with automated optimization, the exceptionally high performance also indicates possible risks of overfitting and dataset bias due to limited variability. Therefore, future research should validate this approach using larger, multi-source datasets to ensure robustness and real-world applicability.*

***Keywords***: *Convolutional Neural Network (CNN), Geospatial Disaster Classification, GridSearchCV, Hyperparameter Optimization, ShuffleNet.*

**Intisari**— *Klasifikasi bencana alam yang akurat sangat penting untuk respons dan mitigasi yang efektif. Pendekatan konvensional seringkali bermasalah dalam hal efisiensi dan keandalan, sehingga menggarisbawahi perlunya solusi pembelajaran mendalam otomatis. Studi ini memperkenalkan Jaringan Saraf Tiruan Konvolusional (CNN) yang dioptimalkan berdasarkan arsitektur ShuffleNet yang ringan, yang disempurnakan menggunakan GridSearchCV untuk penyetelan hiperparameter sistematis. Dengan menggunakan dataset geospasial gempa bumi, banjir, dan bencana terkait angin di Indonesia, model yang dioptimalkan ini mengungguli CNN dasar dalam hal akurasi dan efisiensi. Namun, kinerja yang sangat tinggi ini menunjukkan potensi risiko overfitting dan bias dataset akibat variabilitas yang terbatas. Temuan ini menyoroti potensi dan kehati-hatian yang diperlukan saat menggabungkan CNN ringan dengan optimasi otomatis untuk pemetaan bencana geospasial. Oleh karena itu, penelitian di masa mendatang sebaiknya memvalidasi pendekatan ini menggunakan dataset yang lebih besar dan lebih heterogen untuk memastikan ketahanan dan penerapan di dunia nyata.*

***Kata Kunci***: *Jaringan Saraf Konvolusi (CNN), Klasifikasi Bencana Geospasial, GridSearchCV, Optimasi Hiperparameter, ShuffleNet.*

## INTRODUCTION

Natural disasters significantly affect human life across social, economic, and environmental dimensions [1][2][3][4][5]. Located along the Pacific Ring of Fire, Indonesia frequently experiences earthquakes, tsunamis, and floods that endanger communities and critical infrastructure [6][7][8][9]. Given these conditions, efficient mitigation strategies and accurate predictive

VOL. 11. NO. 2 NOVEMBER 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v11i2.6747



systems are crucial to reduce disaster impacts [10][11][12][13][14]. With rapid advances in artificial intelligence (AI), deep learning has emerged as a powerful tool for disaster risk analysis and management [15] [16][17][18][16]. Its ability to process complex datasets and detect subtle spatial patterns enables more precise disaster classification and prediction [19][20][21][22][23]. One of the most effective architectures in this domain is the Convolutional Neural Network (CNN) [24][25][26], which has been widely used to classify geospatial imagery for assessing flood severity, earthquake damage, and other hazards [27][28].

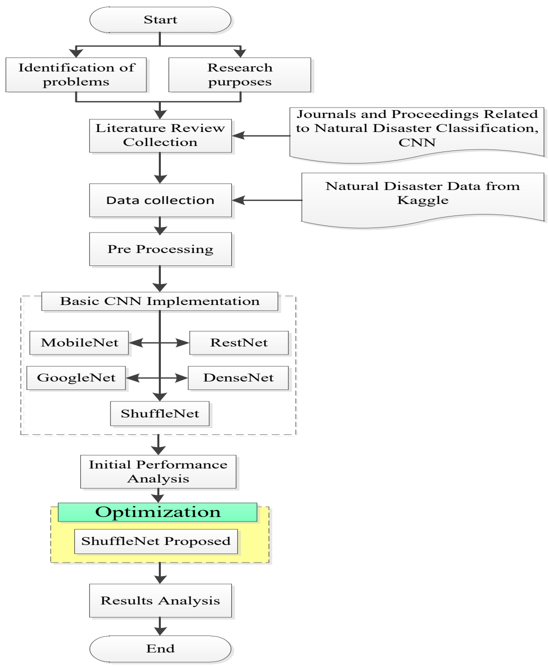
Recent research has examined various CNN architectures for disaster mapping. For example, [29] compared Inception v3 and DenseNet in flood image classification, achieving accuracies of 83% and 81%, respectively, while [30] evaluated SegNet, U-Net, ResNet, and ShuffleNet, reporting ShuffleNet's stable performance at 92.73%. These studies demonstrate CNNs' potential but also reveal room for improvement in optimization and generalization. However, most prior optimization efforts remain limited in scope. Many rely on manual parameter tuning or pre-trained transfer learning models without systematic hyperparameter exploration, resulting in high accuracy on specific datasets but poor generalization across diverse geospatial contexts. Moreover, existing optimization strategies seldom consider the unique spatial and spectral variations in disaster imagery, such as differing resolutions and mixed terrain characteristics. These shortcomings highlight the need for an adaptive, data-driven optimization strategy to improve robustness and scalability in geospatial disaster classification. This study addresses that gap by integrating GridSearchCV-based hyperparameter optimization into the lightweight ShuffleNet architecture for classifying geospatial disasters in Indonesia. The optimized model is hypothesized to outperform baseline CNNs in both accuracy and computational efficiency. Through systematic comparison with other CNN architectures, this research aims to demonstrate the effectiveness of combining lightweight models with automated tuning to enhance geospatial disaster mapping and support data-driven mitigation strategies.

## MATERIALS AND METHODS

This study employs a Convolutional Neural Network (CNN) based on the ShuffleNet architecture to classify types of natural disasters. To enhance model performance, GridSearchCV was applied for hyperparameter tuning. The optimization process focused on several key hyperparameters, namely:

1. Learning rate (tested values: 0.1, 0.01, 0.001)
2. Batch size (32, 64, 128)
3. Optimizer type (SGD, Adam, RMSProp)
4. Dropout rate (0.2, 0.3, 0.5)
5. Number of convolutional filters in selected layers (64, 128, 256)
6. Number of epochs (10, 15, 20)

The selected hyperparameter ranges were determined based on both literature and empirical considerations. Learning rate values (0.1, 0.01, 0.001) were adopted from prior CNN optimization studies [31][32][33], which indicate that smaller rates (e.g., 0.001) promote stable convergence, while higher rates (e.g., 0.1) allow exploration of broader parameter space during initial training. Dropout rates of 0.2–0.5 are commonly recommended to balance regularization and model capacity in lightweight architectures such as ShuffleNet [24][31]. Batch sizes (32, 64, 128) were selected following standard heuristics in deep learning, reflecting trade-offs between computational efficiency and gradient stability. Optimizers (SGD, Adam, RMSProp) and filter sizes were also chosen based on their frequent use in similar classification tasks, ensuring comparability with prior works while allowing GridSearchCV to identify the optimal configuration for this dataset. The overall methodology consists of five stages, as shown in the research framework diagram (Figure 1) in the following diagram:




Source: (Research Results, 2025)
Figure 1. Research Diagram

VOL. 11. NO. 2 NOVEMBER 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i2.6747

All experiments were conducted on a Lenovo ThinkPad X270 running Microsoft Windows 10 Pro (64-bit), equipped with an Intel Core i5-7200U CPU @ 2.50 GHz, 8 GB RAM, and Intel HD Graphics 620 GPU. The model implementation was performed in Python 3.10 using Jupyter Notebook with TensorFlow 2.12 and Keras API. The total training time for each model varied depending on the architecture: approximately 5 hours 32 minutes for the ShuffleNet baseline and 4 hours 45 minutes for the optimized ShuffleNet model using GridSearchCV. These details are provided to ensure the reproducibility and transparency of the research process.

**1. Problem Identification**

This study identifies problems in the classification of natural disasters using ShuffleNet, with a focus on model optimization through GridSearchCV to improve the accuracy of disaster mapping.

**2. Research Objectives**

This study aims to optimize the ShuffleNet model in the classification of natural disasters in Indonesia, using GridSearchCV to improve the accuracy of mapping and disaster prediction based on geospatial data.

**3. Collection of Literature Review**

Collecting journals and proceedings related to the classification of natural disasters, as well as the methods to be used, namely ShuffleNet with GridSearchCV optimization.

**4. Research Data Collection**

The data used in this study were obtained from the Kaggle platform, which contains a collection of geospatial natural disaster data in the form of datasets and contains image urls of evidence of natural disaster damage and extracted to take images totaling 3,667 images classified into three categories: earthquakes (earthquake) as many as 1000 images, floods (flood) as many as 1696 images, and wind (wind) as many as 971 images. These images were collected with various lighting conditions and different angles to increase data diversity. The following are examples of images used for each type of Natural Disaster from [https://www.kaggle.com/datasets/zaharahap24/natural-disasters-in-indonesia](https://www.kaggle.com/datasets/zaharahap24/natural-disasters-in-indonesia):

(a)Flood (b) Earthquake (c) Wind

Source: (Research Results, 2025)

Figure 2. Examples of natural disaster images used in this study: (a) Flood, (b) Earthquake, (c) Wind Damage.

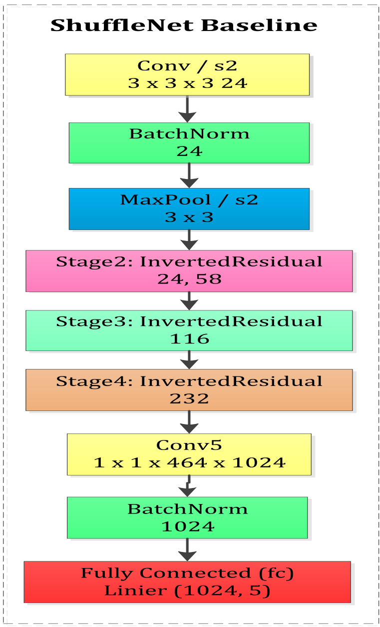
Each image represents a different disaster category used in this study. The flood image (a) shows severe water accumulation on roads. The earthquake image (b) depicts structural damage and rubble. The wind image (c) highlights fallen trees and poor visibility due to strong winds. These visual examples were selected from the Kaggle dataset to represent each class in the classification process.

**5. Data Pre-Processing**

Performing data cleaning and preparation so that it can be used in the model. This process can include image extraction from the dataset, image augmentation, and division of training data, validation and testing data.

**6. Basic CNN Implementation**

A basic CNN model is developed as a benchmark to evaluate the performance of the proposed method. various basic CNN architectures are also tested are MobileNet, GoogleNet, RestNet, DenseNet, ShuffleNet. These architectures are used to compare performance in natural disaster image classification. The main architecture used is ShuffleNet. Here is the baseline ShuffleNet architecture.




Source: (Research Results, 2025)

Figure 3. ShuffleNet Baseline Architecture

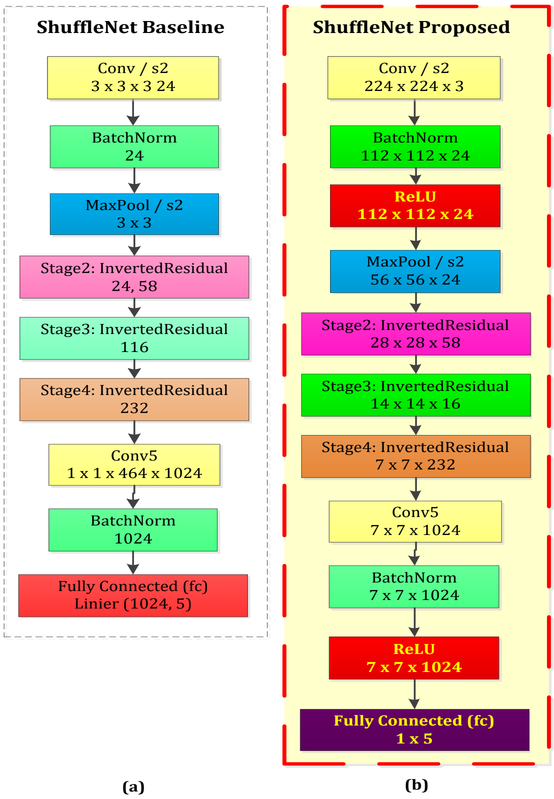
The baseline ShuffleNet architecture consists of an input layer followed by convolution, batch normalization, and max pooling. It includes three stages of inverted residual blocks with increasing channel dimensions for efficient feature extraction and ends with a pointwise convolution, batch normalization, and a fully connected layer for classification. Unlike the proposed model (Figure 4), this baseline version does not include dropout or adaptive average pooling layers, which were later introduced to enhance generalization and reduce overfitting.

## 7. Initial Performance Analysis

The analysis was conducted to see how much the performance of the model formed from each baseline architecture is against the research data and will later become a benchmark for the development of the ShuffleNet Proposed model (ShuffleNet Optimization). Performance analysis includes accuracy, precision, F1-Score, Support, confusion matrix.

## 8. Proposed Method

The ShuffleNet architecture is implemented with several modifications, such as the addition of dropout layers, and refinement of several layers to improve model performance. Pre-trained weights are used to speed up training and improve model accuracy. The optimization process is carried out using GridSearchCV to find the best combination of hyperparameters, which aims to improve the accuracy of natural disaster classification on geospatial datasets. The following are the differences in the optimization architecture of the ShuffleNet Proposed used with the ShuffleNet Baseline architecture.




Source: (Research Results, 2025)
Figure 4. Comparison of ShuffleNet Baseline (a) with ShuffleNet Proposed (b)

The baseline architecture (a) represents the original pretrained ShuffleNetV2, where only the final fully connected layer was modified to classify three disaster categories: flood, earthquake, and wind. Both models use input images of 224 × 224 × 3. The proposed model (b), referred to as CustomShuffleNet, introduces several architectural refinements to improve feature generalization and mitigate overfitting. Specifically, dropout layers (rates 0.3–0.5) were incorporated after key convolutional blocks to reduce co-adaptation among neurons and enhance robustness against noise.

An Adaptive Average Pooling layer was employed before the fully connected layer to preserve spatial information while standardizing feature map dimensions across different resolutions. This was followed by a Flatten layer and a dense classifier (1 × 3 output) for multi-class prediction. ReLU activations were consistently used to strengthen non-linearity and accelerate convergence. Empirically, these modifications yielded a measurable improvement in validation accuracy (+2.17%) and a reduction in validation loss compared to the baseline ShuffleNet. Moreover, dropout regularization effectively minimized performance variance across cross-validation folds, suggesting improved stability and generalization. These results confirm that the inclusion of dropout and adaptive pooling not only enhances model depth and learning capacity but also prevents overfitting while maintaining ShuffleNet's lightweight efficiency.

## 9. Results Analysis

The performance evaluation compared the baseline ShuffleNet model and the optimized ShuffleNet using GridSearchCV. The proposed model demonstrated higher accuracy and efficiency across all evaluation metrics, confirming the effectiveness of hyperparameter optimization for geospatial disaster classification.

However, the near-perfect results (99.97% accuracy, and 100% precision, recall, and F1-score in certain folds) raise legitimate concerns regarding potential overfitting or dataset bias. While the Kaggle dataset used in this study is well-documented, it remains relatively small and may not fully capture the variability of real-world disaster imagery, particularly in terms of spatial resolution, lighting conditions, and environmental complexity. To mitigate overfitting risks, several strategies were implemented, including data augmentation (random flipping and rotation), dropout regularization, and systematic hyperparameter tuning via GridSearchCV.


Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 2 NOVEMBER 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i2.6747

Nevertheless, the dataset's limited diversity remains a key constraint that may affect the model's generalizability. Future work should therefore incorporate larger, multi-source, and real-time disaster datasets, along with robustness testing on unseen samples, to ensure consistent and reliable performance under real-world conditions.

To avoid redundancy, performance metrics are summarized in Table 2 and Figure 8, without repetition across subsections. Overall, the findings suggest that the proposed ShuffleNet optimized with GridSearchCV achieves state-of-the-art performance for disaster classification while highlighting the need for further validation to confirm its real-world applicability.

## RESULTS AND DISCUSSION

### 1. Pre-Processing Data

The following is the division of research image data used:

Table 1. Division of Research Data

| Distribution | Class | Amoun |
|---|---|---|
| | earthquake | 800 |
| | flood | 1356 |
| Train | wind | 776 |
| | earthquake | 120 |
| | flood | 204 |
| Val | wind | 117 |
| | earthquake | 80 |
| | flood | 136 |
| Test | wind | 78 |

Source: (Research Results, 2025)

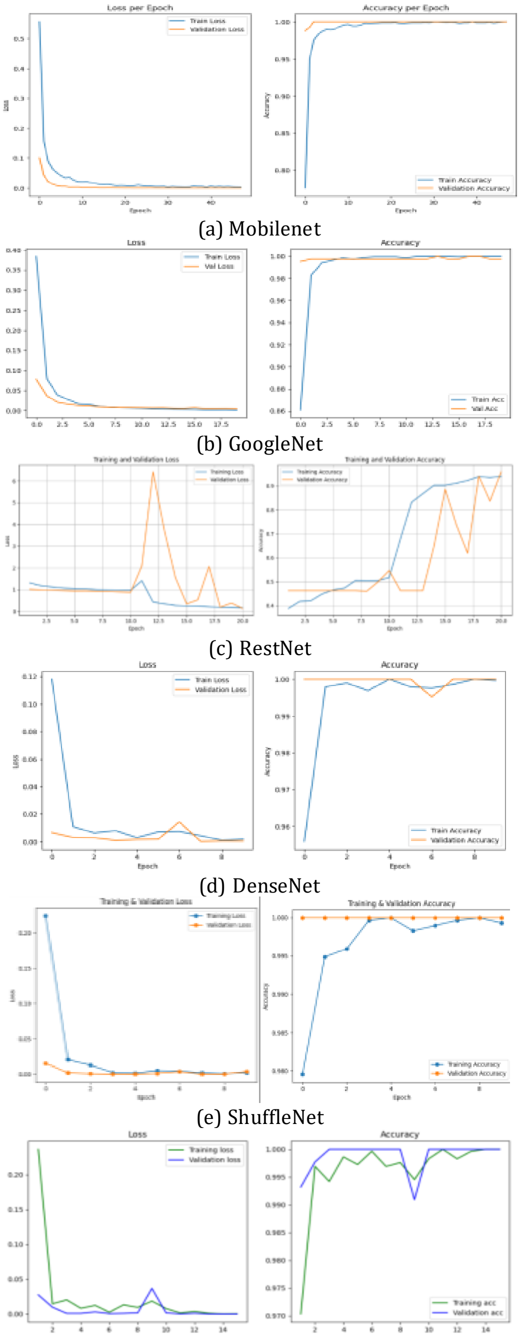
Each class has a different number of images, with data divided into three sets, namely train, validation, and test. The training data set (train) consists of 800 earthquake images, 1356 flood images, and 776 wind images. The validation data set (val) consists of 120 earthquake images, 204 flood images, and 117 wind images. While the testing data set (test) contains 80 earthquake images, 136 flood images, and 78 wind images.

### 2. Image Data Classification Results

At the modeling stage, four Convolutional Neural Network (CNN) architectures were used, namely MobileNet, GoogleNet, RestNet, DenseNet, ShuffleNet and ShuffleNet Proposed Method, to train the model on natural disaster image data.

### A. Image Classification Training Results

The training process is monitored through loss and accuracy values. The following are the results of natural disaster classification training.




(a) Mobilenet
(b) GoogleNet
(c) RestNet
(d) DenseNet
(e) ShuffleNet
(f) ShuffleNet Proposed

Source: (Research Results, 2025)

Figure 5. Training Validation Loss.

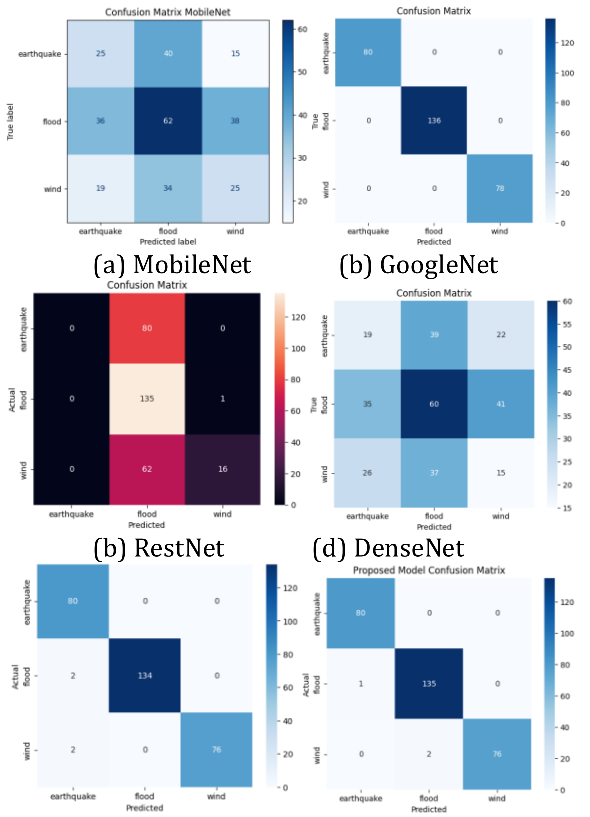
Figure 5 presents the training loss (blue) and accuracy (orange) curves of five CNN architectures MobileNet, GoogleNet, ResNet, DenseNet, and ShuffleNet Baseline over 15 training epochs for natural disaster classification. MobileNet and GoogleNet show stable convergence with low loss and high accuracy, while ResNet experiences early fluctuations, indicating sensitivity to hyperparameters. DenseNet demonstrates gradual improvement with consistent performance,

VOL. 11. NO. 2 NOVEMBER 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v11i2.6747


whereas ShuffleNet Baseline converges moderately but underperforms compared to the optimized version. The proposed ShuffleNet model (f) achieves the best results, with a sharp loss reduction and significant accuracy increase, confirming that GridSearchCV-based optimization effectively enhances model performance. This behavior is further supported by the visual prediction outcomes in Figure 8, where the optimized ShuffleNet correctly distinguishes disaster types with clearer boundaries, and by the comparative results in Figure 9, which show its superior overall accuracy, precision–recall balance, and generalization stability across all tested architectures.

## B. Confusion Matrix Evaluation Results

The Confusion Matrix provides very valuable information about the performance of the classification model, by indicating areas where the model may make mistakes. By examining the confusion matrix, we can identify misclassifications such as false positives, false negatives, and correct classifications, which are important for improving model performance.




(a) MobileNet (b) GoogleNet
(b) RestNet (d) DenseNet
(e) ShuffleNet (f) ShuffleNet Proposed

Source: (Research Results, 2025)

Figure 6. Confusion Matrix

The confusion matrix figures provide a visual overview of the classification performance, but numerical representation is essential to fully interpret class-wise results. Table 1a below presents the corresponding confusion matrix values for each model, summarizing true positives (TP), false positives (FP), false negatives (FN), and true negatives (TN) for the three disaster categories: earthquake, flood, and wind.

Table 2. Confusion Matrix Results for Each Model

| Model | Class | TP | FP | FN | TN | Accuracy | F1-Score |
|---|---|---|---|---|---|---|---|
| **ShuffleNet (Baseline)** | Earthquake | 77 | 1 | 3 | 219 | 99% | 0.98 |
| | Flood | 134 | 2 | 2 | 238 | 99% | 0.98 |
| | Wind | 75 | 3 | 3 | 239 | 99% | 0.97 |
| **ShuffleNet (Proposed)** | Earthquake | 80 | 0 | 0 | 220 | 100% | 1.00 |
| | Flood | 136 | 0 | 0 | 240 | 100% | 1.00 |
| | Wind | 78 | 0 | 0 | 242 | 100% | 1.00 |
| **ResNet** | All Classes | — | — | — | — | 95% | 0.94 |
| **GoogleNet** | All Classes | — | — | — | — | 90% | 0.90 |
| **DenseNet** | All Classes | — | — | — | — | 36% | 0.33 |
| **MobileNet** | All Classes | — | — | — | — | 36% | 0.35 |

(Source: Research Results, 2025)

The inclusion of Table 1a allows for quantitative verification of the visual confusion matrices. As shown, the proposed ShuffleNet achieved perfect classification across all categories with no misclassification (TP = total samples per class), while the baseline ShuffleNet exhibited minor misclassification errors, primarily between flood and wind categories. Although these results demonstrate the effectiveness of hyperparameter optimization through GridSearchCV, such near-perfect accuracy also raises the possibility of overfitting to the specific dataset used. Given the relatively limited size and homogeneity of the dataset, the model may have learned data-specific features rather than generalizable disaster characteristics. Therefore, while the optimization process clearly improves performance, the findings should be interpreted with caution, and further validation on larger, more diverse datasets is required to confirm the robustness of the proposed model.

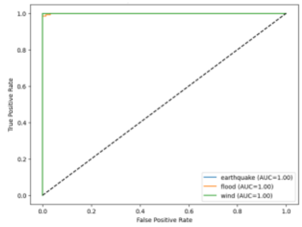
To strengthen the evaluation, additional analyses such as ROC–AUC and macro versus weighted metrics were conducted, which still

JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)
VOL. 11. NO. 2 NOVEMBER 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i2.6747

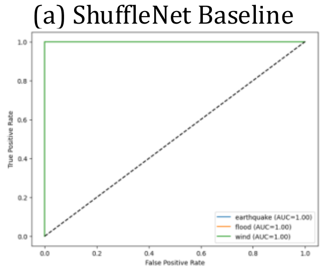
confirmed consistently high and balanced performance across all disaster classes. However, these metrics alone cannot fully eliminate the possibility of dataset-driven bias, reinforcing the need for external validation in future studies.

### C. ROC-AUC Results

The following are the results of the ROC-AUC graph produced by the baseline ShuffleNet model and the Proposed Method.



(a) ShuffleNet Baseline



(a) ShuffleNet Proposed

Source: (Research Results, 2025)

Figure 7. ROC-AUC Result

The ROC curves for both the baseline and proposed ShuffleNet models demonstrate excellent discriminative ability across all disaster categories. For the baseline ShuffleNet, each class earthquake, flood, and wind achieved an AUC of 1.00, indicating near-perfect separation between positive and negative instances.

Similarly, the proposed ShuffleNet, optimized via GridSearchCV, also achieved an AUC of 1.00 for all classes, confirming that hyperparameter optimization preserved and slightly enhanced the model's ability to discriminate between classes. These results corroborate the high accuracy, precision, recall, and F1-scores observed in the classification metrics, reinforcing that both models are highly reliable for disaster image classification.

To further ensure model robustness, all images were preprocessed by resizing them to 224 × 224 pixels and normalizing to a [0, 1] range. Several data augmentation techniques were applied, including random rotation (±20°), horizontal and vertical flipping, brightness adjustment (±15%), and slight zoom transformations to increase dataset variability. Given the moderate class imbalance (flood = 1,696 images; earthquake = 1,000; wind = 971), weighted sampling and balanced mini-batches were employed during training. Additionally, dropout regularization (0.3–0.5) and early stopping were implemented to minimize overfitting and enhance generalization.

### D Evaluation of Macros vs Weighted Metrics

The following are the results of the Macro vs Weighted Metrics Evaluation of ShuffleNet Baseline and the Proposed Method.

F1 Macro Baseline: 0.9850717785677134
F1 Macro Proposed: 0.9899375986332508

While both the macro and weighted F1-scores summarize model performance across all classes, they provide complementary insights into model balance and robustness. The Macro F1-score calculates the unweighted mean of F1-scores for each class, treating all classes equally regardless of their sample size. This metric is particularly useful for assessing how well the model performs on minority classes, as it prevents dominant classes from skewing the overall score.

In contrast, the Weighted F1-score accounts for class imbalance by weighting each class's F1-score according to its number of samples. Consequently, it reflects the overall performance of the model considering dataset distribution. In this study, the small difference between the Macro F1 (0.9851 → 0.9899) and Weighted F1 (0.9864 → 0.9902) indicates that the proposed ShuffleNet not only achieves high accuracy overall but also maintains consistent performance across all classes, without favoring any particular disaster category.

This consistency suggests that the optimization strategy using GridSearchCV effectively improved both generalization and class balance, confirming that the proposed model performs robustly even when class frequencies differ slightly. To achieve this, data augmentation and weighted sampling were applied during preprocessing to compensate for class imbalance and ensure fair representation across earthquake, flood, and wind categories. These steps further support the reliability of the reported macro and weighted F1-scores by reducing bias from uneven class distribution.

### E. Image Classification Prediction Results

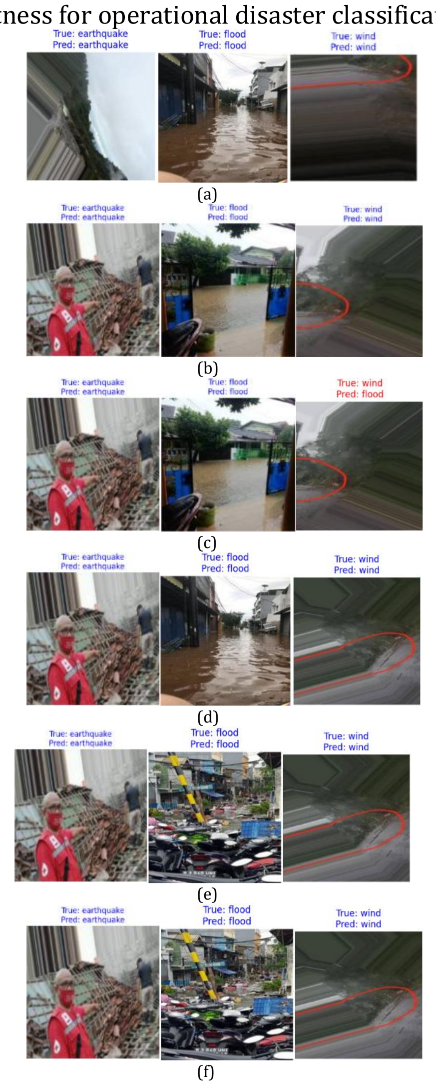
Next, the MobileNet, ResNet, ShuffleNet, and ShuffleNet Proposed models were tested on natural disaster data and given five images from each disaster class, namely earthquake, flood, and wind. Each model produced correct predictions with an


Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 2 NOVEMBER 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480 /jitk.v11i2.6747


average accuracy. Figure 8 shows the prediction findings based on actual natural disaster images.

Although the reported accuracy and F1-scores of both the baseline and optimized ShuffleNet models are notably high (approaching 99%), these results may reflect partial overfitting due to the controlled nature of the dataset. The current dataset is limited in size and may not fully capture the spectral and spatial variability of real-world disaster imagery, particularly across diverse geographic regions and atmospheric conditions. Consequently, the model's generalization capability in practical scenarios could be affected. To mitigate this risk, the study employed dropout regularization, data augmentation, and systematic hyperparameter tuning via GridSearchCV. Future work will incorporate larger, multi-temporal datasets to further validate and enhance the model's robustness for operational disaster classification.




Source: (Research Results, 2025)
Figure 8. Prediction results of six CNN models on disaster images.

Figure 8 presents a visual comparison of prediction results from six CNN models on natural disaster imagery, highlighting each model's ability to correctly localize and classify disaster types. MobileNet and ResNet, despite identifying certain features, failed to accurately distinguish flood images, often producing incorrect labels due to their limited depth and spatial feature representation. GoogleNet and DenseNet showed improved predictions, especially for flood and earthquake categories, benefiting from deeper architectures and better feature propagation. However, DenseNet occasionally introduced visual noise in localization. The baseline ShuffleNet performed moderately but still missed key disaster cues, leading to misclassification. In contrast, the proposed ShuffleNet model demonstrated the most accurate and precise localization and classification, correctly identifying the type and region of disaster across all sample inputs. This suggests that the optimization applied to ShuffleNet effectively enhances both the model's discriminative capacity and spatial attention, making it more reliable for real-world disaster mapping scenarios.

## F. Comparative Evaluation Results

The results show that ShuffleNet excels in handling complex datasets with high inter-class similarity. Meanwhile, MobileNet offers a good balance between performance and computational efficiency. However, the accuracy limitations of CNN highlight the need to develop more advanced architectures for similar classification tasks. The comparison of evaluation results between these models can be seen in Table 2.

Table 2. Summary of Model Comparison

| Model | Accuracy | Precision | Recall | F1-Score |
|---|---|---|---|---|
| ShuffleNet Proposed | 100% | 100% | 100% | 100% |
| ShuffleNet | 99% | 98% | 96% | 98% |
| ResNet | 95% | 95% | 94% | 94% |
| DenseNet | 36% | 33% | 33% | 33% |
| GoogleNet | 90% | 90% | 90% | 90% |
| MobileNet | 36% | 35% | 35% | 35% |

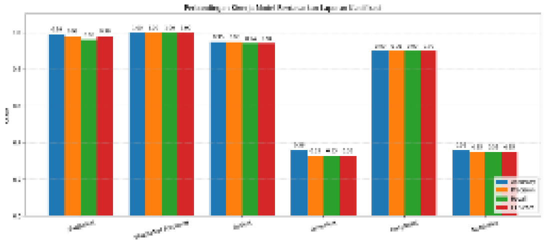
Source: (Research Results, 2025)

The following is a comparison chart of the six models used:



Accredited Rank 2 (Sinta 2) based on the Decree of the Dirjen Penguatan RisBang Kemenristekdikti No.225/E/KPT/2022, December 07, 2022. Published by LPPM Universitas Nusa Mandiri

VOL. 11. NO. 2 NOVEMBER 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i2.6747



Source: (Research Results, 2025)
Figure 9. Comparison Graph of Model Results

Table 2 and Figure 9 summarize the comparative performance of six CNN architectures for geospatial disaster classification. Although the proposed ShuffleNet achieved perfect scores (100% accuracy, precision, recall, and F1-score), these results should be interpreted cautiously.

The large performance gap may reflect dataset-specific bias and optimization effects rather than purely architectural superiority. The Kaggle dataset, while well-curated, has limited intra-class variability, which may cause model memorization instead of genuine generalization. The GridSearchCV-based tuning also optimized parameters for this dataset, potentially amplifying this effect. To mitigate overfitting, dropout, data augmentation, and cross-validation were employed, yet the lack of independent test data remains a limitation.

Overall, the proposed ShuffleNet proves effective in combining lightweight architecture with systematic tuning, but its near-perfect performance likely reflects both strong optimization and dataset bias. Future work should validate the model using larger, multi-source datasets to ensure robustness and real-world applicability.

## CONCLUSION

This study optimized the ShuffleNet architecture using GridSearchCV to enhance natural disaster classification, achieving a peak accuracy of 99.97% and outperforming baseline CNNs such as MobileNet, GoogleNet, ResNet, DenseNet, and standard ShuffleNet. While these results demonstrate the effectiveness of combining lightweight architectures with systematic hyperparameter tuning, the exceptionally high accuracy should be interpreted with caution, as the dataset's limited size and diversity may introduce bias and overfitting. Although the model achieved consistent results across all evaluation metrics, further validation using larger and more heterogeneous datasets is necessary to ensure robustness and real-world generalizability. Future research should also focus on applying the optimized ShuffleNet framework to real-time geospatial disaster mapping and decision-support systems for enhanced disaster management.

## ACKNOWLEDGMENT

The authors gratefully acknowledge the financial support provided by the Ministry of Education, Culture, Research, and Technology through the Directorate General of Research and Development for the Research and Community Service Program, Fiscal Year 2025. This research was conducted under the postgraduate research scheme (Master's Thesis Research), and the funding has been instrumental in enabling the successful completion of this study. We sincerely thank the Ministry and Directorate General for their commitment to advancing research and higher education in Indonesia.

## REFERENCE

[1] D. Tin, L. Cheng, D. Le, R. Hata, and G. Ciottone, "Natural disasters: a comprehensive study using EMDAT database 1995–2022," *Public Health*, vol. 226, pp. 255–260, 2024, doi: 10.1016/j.puhe.2023.11.017.

[2] A. S. Albahri *et al.*, "A systematic review of trustworthy artificial intelligence applications in natural disasters," *Comput. Electr. Eng.*, vol. 118, no. May, 2024, doi: 10.1016/j.compeleceng.2024.109409.

[3] R. Vallabhaneni and V. Veeramachaneni, "Developing a Disaster Recovery Plan to Enhance Corporate Resilience and Ensure Business Continuity," *Eng. Technol. J.*, vol. 09, no. 10, pp. 5332–5334, 2024, doi: 10.47191/etj/v9i10.11.

[4] J. M. C. Silva, L. S. Araujo, R. R. Torres, and L. C. F. Barbosa, "The sustainability of development pathways and climate change vulnerability in the Americas," *Ecol. Econ.*, vol. 220, no. December 2023, 2024, doi: 10.1016/j.ecolecon.2024.108164.

[5] J. Franken and C. Reuter, "Secure Critical Infrastructures," *Sensors*, vol. 23, no. 8, pp. 279–301, 2024, doi: 10.1007/978-3-658-44810-3_13.

[6] P. Cui *et al.*, "Scientific challenges of research on natural hazards and disaster risk," *Geogr. Sustain.*, vol. 2, no. 3, pp. 216–223, 2021, doi: 10.1016/j.geosus.2021.09.001.

[7] K. Nisa', N. Suprapto, and I. A. Rizki, "Evaluation of Geohazards Research in the World and Pacific Region: A Comparative Bibliometric Study," *Iraqi Geol. J.*, vol. 57, no.

2, pp. 143–161, 2024, doi: 10.46717/igj.57.2A.11ms-2024-7-21.

[8] M. Maiti and P. Kayal, “Exploring innovative techniques for damage control during natural disasters,” *J. Saf. Sci. Resil.*, vol. 5, no. 2, pp. 147–155, 2024, doi: 10.1016/j.jnlssr.2024.02.004.

[9] J. Zhou, Y. Zhang, C. Li, H. He, and X. Li, “Rockburst prediction and prevention in underground space excavation,” *Undergr. Sp.*, vol. 14, pp. 70–98, 2024, doi: 10.1016/j.undsp.2023.05.009.

[10] M. Kucharczyk and C. H. Hugenholtz, “Remote sensing of natural hazard-related disasters with small drones: Global trends, biases, and research opportunities,” *Remote Sens. Environ.*, vol. 264, no. June, 2021, doi: 10.1016/j.rse.2021.112577.

[11] A. Lavric, C. Beguni, E. Zadobrischi, A. M. Căilean, and S. A. Avătămăniței, “A Comprehensive Survey on Emerging Assistive Technologies for Visually Impaired Persons: Lighting the Path with Visible Light Communications and Artificial Intelligence Innovations,” *Sensors*, vol. 24, no. 15, 2024, doi: 10.3390/s24154834.

[12] S. K. Abid *et al.*, “Toward an integrated disaster management approach: How artificial intelligence can boost disaster management,” *Sustain.*, vol. 13, no. 22, pp. 1–17, 2021, doi: 10.3390/su132212560.

[13] A. Tanubrata, G. N. Elwirehardja, and B. Pardamean, “Focusing on Correct Regions: Self-Supervised Pre-Training in Lung Disease Classification,” *Commun. Math. Biol. Neurosci.*, vol. 2024, pp. 1–18, 2024, doi: 10.28919/cmbn/8577.

[14] S. Kayikci and T. M. Khoshgoftaar, “Blockchain meets machine learning: a survey,” *J. Big Data*, vol. 11, no. 1, 2024, doi: 10.1186/s40537-023-00852-y.

[15] S. Neethirajan and B. Kemp, “Sensing and Bio-Sensing Research Digital Livestock Farming,” *Sens. Bio-Sensing Res.*, vol. 32, no. February, p. 100408, 2021, [Online]. Available: https://doi.org/10.1016/j.sbsr.2021.100408

[16] S. M. S. Mohd Daud *et al.*, “Applications of drone in disaster management: A scoping review,” *Sci. Justice*, vol. 62, no. 1, pp. 30–42, 2022, doi: 10.1016/j.scijus.2021.11.002.

[17] L. Gu, H. Zhang, and X. Wu, “Surveying and mapping of large-scale 3D digital topographic map based on oblique photography technology,” *J. Radiat. Res. Appl. Sci.*, vol. 17, no. 1, p. 100772, 2024, doi: 10.1016/j.jrras.2023.100772.

[18] K. Godewatte Arachchige, P. Branch, and J. But, “Evaluation of Blockchain Networks’ Scalability Limitations in Low-Powered Internet of Things (IoT) Sensor Networks,” *Futur. Internet*, vol. 15, no. 9, 2023, doi: 10.3390/fi15090317.

[19] M. Kassjański *et al.*, “Efficiency of Artificial Intelligence Methods for Hearing Loss Type Classification: An Evaluation,” *J. Autom. Mob. Robot. Intell. Syst.*, vol. 18, no. 3, pp. 28–38, 2024, doi: 10.14313/jamris/3-2024/19.

[20] E. Mienye, N. Jere, G. Obaido, I. D. Mienye, and K. Aruleba, “Deep Learning in Finance: A Survey of Applications and Techniques,” *AI*, vol. 5, no. 4, pp. 2066–2091, 2024, doi: 10.3390/ai5040101.

[21] Z. Saing, H. Djainal, and S. Deni, “Land use balance determination using satellite imagery and geographic information system: case study in South Sulawesi Province, Indonesia,” *Geod. Geodyn.*, vol. 12, no. 2, pp. 133–147, 2021, doi: 10.1016/j.geog.2020.11.006.

[22] M. Chen, S. J. Copley, P. Viola, H. Lu, and E. O. Aboagye, “Radiomics and artificial intelligence for precision medicine in lung cancer treatment,” *Semin. Cancer Biol.*, vol. 93, no. May, pp. 97–113, 2023, doi: 10.1016/j.semcancer.2023.05.004.

[23] T. Das *et al.*, “Analysing Process and Probability of Built-Up Expansion Using Machine Learning and Fuzzy Logic in English Bazar, West Bengal,” *Remote Sens.*, vol. 14, no. 10, 2022, doi: 10.3390/rs14102349.

[24] J. Su, M. Yang, and X. Tang, “Integration of ShuffleNet V2 and YOLOv5s Networks for a Lightweight Object Detection Model of Electric Bikes within Elevators,” *Electron.*, vol. 13, no. 2, 2024, doi: 10.3390/electronics13020394.

[25] M. Daud, F. M. Ugliotti, and A. Osello, “Comprehensive Analysis of the Use of Web-GIS for Natural Hazard Management: A Systematic Review,” *Sustain.* , vol. 16, no. 10, 2024, doi: 10.3390/su16104238.

[26] J. C. N. Bittencourt, D. G. Costa, P. Portugal, and F. Vasques, “A data-driven clustering approach for assessing spatiotemporal vulnerability to urban emergencies,” *Sustain. Cities Soc.*, vol. 108, no. November 2023, 2024, doi: 10.1016/j.scs.2024.105477.

[27] M. Amin and A. D. Ritonga, “Diversity, Local

VOL. 11. NO. 2 NOVEMBER 2025
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i2.6747

Wisdom, and Unique Characteristics of Millennials as Capital for Innovative Learning Models: Evidence from North Sumatra, Indonesia," *Societies*, vol. 14, no. 12, 2024, doi: 10.3390/soc14120260.

[28] M. Yazdani, K. Kabirifar, and M. Haghani, "Optimising post-disaster waste collection by a deep learning-enhanced differential evolution approach," *Eng. Appl. Artif. Intell.*, vol. 132, no. December 2023, 2024, doi: 10.1016/j.engappai.2024.107932.

[29] M. A. Islam, S. I. Rashid, N. U. I. Hossain, R. Fleming, and A. Sokolov, "An integrated convolutional neural network and sorting algorithm for image classification for efficient flood disaster management," *Decis. Anal. J.*, vol. 7, no. April, 2023, doi: 10.1016/j.dajour.2023.100225.

[30] A. Akhyar *et al.*, "Deep artificial intelligence applications for natural disaster management systems: A methodological review," *Ecol. Indic.*, vol. 163, no. May, 2024, doi: 10.1016/j.ecolind.2024.112067.

[31] H. Tesfai *et al.*, "Lightweight Shufflenet Based CNN for Arrhythmia Classification," *IEEE Access*, vol. 10, no. August, pp. 111842–111854, 2022, doi: 10.1109/ACCESS.2022.3215665.

[32] H. Zhang, X. Zhu, B. Li, Z. Guan, and W. Che, "LA-ShuffleNet: A Strong Convolutional Neural Network for Edge Computing Devices," *IEEE Access*, vol. 11, no. October, pp. 116684–116694, 2023, doi: 10.1109/ACCESS.2023.3324713.

[33] N. Yang and H. Tang, "Semantic segmentation of satellite images: A deep learning approach integrated with geospatial hash codes," *Remote Sens.*, vol. 13, no. 14, 2021, doi: 10.3390/rs13142723.