SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16129
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Classification of Paddy as Visual Anomaly in Rice Piles Using MobileNetV2-Based Convolutional Neural Network

Barokah Saadah[1)*] Tri Aristi Saputri[2)]
[1) 2)] Computer Science, Faculty of Business and Science Technology, Dharma Wacana University, Indonesia
[1)]barokahsaadah@email.com,[2)]aristy@dharmawacana.ac.id

Submitted :April 20, 2026 | **Accepted** : May 8, 2026 | **Published** : July 5, 2026

Abstract: Rice is a strategic food commodity whose quality is assessed based on the visual appearance of the grains, including the presence of unhusked rice as an undesirable element in piles of milled rice. Manual inspection is subjective, time-consuming, and prone to errors, necessitating a more objective automated approach. To address this issue, this study applies a MobileNetV2-based Convolutional Neural Network with *transfer learning* to classify unhusked grains as visual anomalies in rice piles into two classes: normal rice and anomalous grains. In terms of methodology, the dataset consists of 1,000 self-acquired images stratified into three groups with a 70:15:15 ratio. Image preprocessing was performed via background removal using the *rembg* library and random background simulation with five background color variations. Training was conducted in two phases: Phase 1 (*transfer learning* with a frozen *base model*) and Phase 2 (*fine-tuning* by opening the last 30 layers of the *base model*). The evaluation results on the test data showed an *accuracy* of 90.67%, a macro *precision* of 0.9213, a macro *recall* of 0.9067, and a macro *F1-score* of 0.9058. The *false positive* rate across all tests was 0. Phase 1 was selected as the best model because it produced more stable performance compared to Phase 2. *Grad-CAM* visualizations confirmed that the model focuses its attention on the visual features of the objects, not background patterns. These findings demonstrate that a combination of preprocessing, *transfer learning*, and data augmentation is effective for binary image classification when dealing with limited datasets.

Keywords: Convolutional Neural Network; Grad-CAM; Grains; MobileNetV2; Rice; Transfer learning; Visual Anomalies

INTRODUCTION

Rice is a major food commodity whose quality depends heavily on the visual appearance of the grains, including size, shape, color, uniformity, and chalkiness. In quality standards for milled rice, the presence of unhulled grains or paddy kernels is a factor considered in quality grading; therefore, their appearance in a normal batch of rice is regarded as an undesirable condition (Aznan, Viejo, Pang, & Fuentes, 2021; Ilo, Badjona, Singh, Shenfield, & Zhang, 2025). In conventional practice, rice quality assessment still relies heavily on manual visual inspection, which tends to be subjective, time-consuming, and prone to errors due to fatigue or differences in assessment standards (Aznan et al., 2021; Sokudlor, Laloon, Junsiri, & Sudajan, 2023). Therefore, approaches based on computer vision and *deep learning* are seen as an increasingly important solution to overcome the limitations of manual inspection (He et al., 2023).

Unmilled rice grains appearing among piles of normal rice can be identified as visual anomalies because their presence deviates from the expected visual composition, which should be dominated by milled rice grains. In *precision* agriculture research, visual anomalies have been extensively studied using image-based *deep learning* approaches (Gkountakos, Ioannidis, Demestichas, Vrochidis, & Kompatsiaris, 2024). Convolutional Neural Networks (CNNs) are widely used in image classification due to their ability to automatically learn visual features without relying on manual feature engineering (Gulzar, 2023). Under conditions of limited datasets, *transfer learning* becomes a more effective option because it leverages the knowledge of pre-trained models, thereby reducing the risk of *overfitting* (Bichri, Chergui, & Hain, 2023). Among various CNN architectures, MobileNetV2

*name of corresponding author

was selected because it is known to be lightweight, efficient, and performs well for image classification tasks (Gulzar, 2023).

Previous research on rice image classification has generally focused on variety classification, grain grade categorization, or chalkiness levels using morphological features and conventional *machine learning* (Setiawan, Adi, & Widodo, 2023; Sokudlor et al., 2023). Although some studies have used CNNs for binary classification of rice, the objects studied were chalkiness or general quality, not the presence of grains as visual anomalies (C. Wang et al., 2022; Yusuf, Ruimassa, Tawainella, & Maharani, 2024). Research addressing visual anomalies in agricultural imagery has also not been specifically applied to rice (Mendoza-Bernal, González-Vidal, & Skarmeta, 2024). Given this gap, research that specifically identifies unhusked grains as visual anomalies in rice piles within a binary image classification framework, accompanied by preprocessing strategies to address background dominance and model interpretation via *Grad-CAM*, remains scarce.

This study aims to apply image preprocessing steps and a MobileNetV2-based CNN using a *transfer learning* approach to classify rice grains identified as visual anomalies in a pile of rice into two classes: normal rice and anomalous rice grains. Additionally, this study evaluates the performance of the MobileNetV2 model in classifying normal rice and anomalous grains based on the metrics of *accuracy*, *precision*, *recall*, and *F1-score*, analyzes the model's classification errors, specifically *false positive*s and *false negative*s, and visualizes the model's regions of interest using *Grad-CAM* to identify the visual characteristics that influence the model's predictions.

LITERATURE REVIEW

Rice quality is closely linked to the visual characteristics of the grains, which can be identified through digital images. Image processing to assess the degree of broken grains and the milling level of rice has been studied as a replacement for manual inspection methods (Chen, Li, & Wang, 2022; Suryana, 2023). Studies on morpho-colorimetric parameters also demonstrate the potential of digital images to objectively distinguish rice varieties (Aznan et al., 2021). The presence of undesirable elements in rice piles, including paddy, foreign objects, and abnormal grains, is a major concern in image-based quality control (Aznan et al., 2021; Setiawan et al., 2023).

Various manual feature-based and *machine learning* approaches have been applied to rice samples. The use of HSV color features and GLCM texture features has proven effective for classifying foreign objects in rice with an *accuracy* of 96.83% using SVM (Setiawan et al., 2023). The use of morphological features from binary images with a Random Forest model was also able to achieve an *accuracy* exceeding 99% for milled rice categories (Sokudlor et al., 2023). A comparative study between multi-class SVM and CNN for rice grading tasks also showed highly competitive results, with categorization *accuracy* reaching 98.25% (Thammastitkul & Petsuwan, 2023). Despite their high *accuracy*, these approaches still rely on manual feature engineering, which requires in-depth domain knowledge.

The use of CNNs for rice classification has been applied for various purposes. The VGG16 architecture was used to distinguish between high-quality and low-quality rice in an Android-based system with 99% *accuracy* (Yusuf et al., 2024). Another study applied CNN-based binary classification to rice grain images based on chalkiness and used *Grad-CAM* to identify the model's prediction support regions (C. Wang et al., 2022). Another study also applied a CNN to classify rice into three categories medium, premium, and counterfeit and achieved 95% *accuracy* with perfect *precision* in the counterfeit rice category, underscoring the potential of *deep learning* to support image-based food safety (Bastian et al., 2025). Similar research on rice variety recognition using CNNs has also been conducted with a focus on developing system application interfaces, underscoring the broader relevance of CNNs for rice-related applications (Prabowo & Hadikurniawati, 2023). In a study on rice leaf diseases, a MobileNetV3Large-based CNN with *transfer learning* successfully achieved an *accuracy* of nearly 95% under limited dataset conditions (Faqih, Irsan, & Fathoni, 2024). This study confirms that *transfer learning* is effective for addressing data limitations in image classification tasks (Bichri et al., 2023).

In studies of visual anomalies in agricultural imagery, research using multi-architecture CNNs has shown that data preprocessing and augmentation play a crucial role in improving a model's generalization to new data (Mendoza-Bernal et al., 2024). A comprehensive review of *deep learning*-based anomaly detection in agriculture also confirms that visual data is the primary source for anomaly identification (Gkountakos et al., 2024). The fundamental differences between classification, detection, and counting tasks in grain crop phenotyping have also been discussed in detail, which helps to more clearly position this research within the scope of computer vision (Gong & Fan, 2022). A review of the use of CNNs in computer vision for cereal crop phenotyping also highlights the differences between classification, detection, and segmentation tasks, which are crucial for determining the positioning of this research (Y. Wang & Su, 2022).

Based on the overall review, the CNN-based binary image classification approach using *transfer learning* has a strong theoretical and empirical foundation. However, studies that specifically treat unhusked rice as a visual anomaly in rice piles within a binary classification framework, combined with background preprocessing strategies and *Grad-CAM* interpretation, remain scarce in the existing literature.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026 e-ISSN : 2541-2019
DOI : https://doi.org/10.33395/sinkron.v10i3.16129 p-ISSN : 2541-044X

METHOD

This study is an experimental study in the field of image processing and image classification. The overall flow of the research stages is systematically presented in Figure 1.

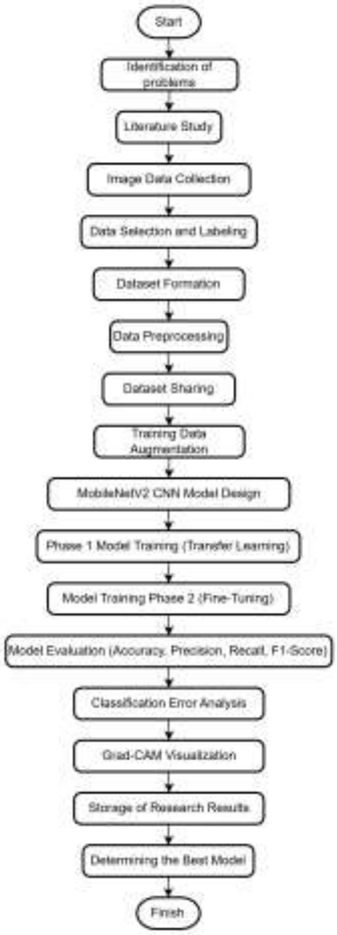

Fig. 1 Research Phases

As shown in Figure 1, the research was conducted through a series of sequential stages, beginning with problem identification, literature review, data collection and selection, preprocessing, dataset splitting, data augmentation, model design and training, and concluding with evaluation and selection of the best model. A detailed explanation of each stage is provided below.

Dataset

The dataset used consists of 1,000 images acquired independently using a smartphone camera, divided equally into two classes: 500 images of normal rice and 500 images of abnormal paddy. The normal rice images represent piles of milled rice without any unhusked grains, while the anomalous unhusked grain images represent piles of rice containing one or more unhusked grains.

Data Preprocessing

The preprocessing stage is performed in two sequential steps on the entire dataset prior to data splitting. First, background removal is applied using the *rembg* library to separate the rice and paddy objects from the uniform dark background. Second, random background simulation is performed by replacing the background of the separated image with one of five color variations: white (255, 255, 255), cream (255, 253, 208), beige (210, 180, 140), light gray (192, 192, 192), and dark gray (105, 105, 105), selected deterministically using a fixed seed. Control of the background and visual conditions of the image is a critical factor because environmental color interference can significantly affect the color representation of the object (Yumono, Yuliana, & Sarbini, 2022).

Dataset Split

The dataset was stratified into training (70%, or 700 images), validation (15%, or 150 images), and test (15%, or 150 images) sets, using a fixed seed of 42 to ensure reproducibility.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16129
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Data Augmentation

Data augmentation is applied only to the training data during the training process without permanently increasing the size of the main dataset. The augmentation techniques used include horizontal flipping, vertical flipping, 90° rotation, and random brightness adjustment, with all operations using a fixed seed. The augmentation strategy has been shown to significantly improve the model's generalization ability under limited dataset conditions (Chun et al., 2022).

Model Architecture

The classification model used is a MobileNetV2-based CNN with pre-trained weights from *ImageNet*. A custom classification layer was added on top of the *base model*, consisting of 2D Global Average Pooling, a 256-unit *Dense layer* with ReLU activation and L2 regularization, 0.5 *Dropout*, a 128-unit *Dense layer* with ReLU activation and L2 regularization, 0.3 *Dropout*, and a 1 unit Dense output layer with *sigmoid* activation for binary classification. The total number of model parameters is 2,618,945 (9.99 MB). Images with a *confidence score* $\geq$ 0.5 are classified as anomalous grains, while those with a score < 0.5 are classified as normal rice.

Table 1 MobileNetV2 Model Architecture

Layer	Output Shape	Parameter
Input Layer	(None, 224, 224, 3)	0
MobileNetV2 (Base)	(None, 7, 7, 1280)	2.257.984
Global Average Pooling 2D	(None, 1280)	0
Dense (256, ReLU) + L2	(None, 256)	327.936
Dropout (0.5)	(None, 256)	0
Dense (128, ReLU) + L2	(None, 128)	32.896
Dropout (0.3)	(None, 128)	0
Dense (1, Sigmoid)	(None, 1)	129
Total Parameter	-	2.618.945 (9,99 MB)

Model Training

The training is conducted in two phases. Phase 1 is a *transfer learning* phase in which all layers of the *base model* are frozen so that only the custom classification layer is trained, using the *Adam* optimizer with a learning rate of 1×10^{-3}, a maximum of 30 epochs, and *early stopping* based on *val_accuracy* with a patience of 8. Phase 2 is the *fine-tuning* phase, where the last 30 layers of the *base model* are unfrozen with a smaller learning rate of 1×10^{-5}. Both phases use *callback*s in the form of ModelCheckpoint, EarlyStopping, and ReduceLROnPlateau. The results of both phases are compared to determine the model with the most optimal performance on the validation and test data.

Evaluation

The model was evaluated using a confusion matrix and four key metrics: *accuracy*, *precision*, *recall*, and *F1-score*, calculated using the macro average approach. The positive class is defined as anomalous paddy; thus, *True positive* (TP) is paddy correctly predicted as paddy, *False negative* (FN) is paddy incorrectly predicted as rice, *False positive* (FP) is rice incorrectly predicted as paddy, and *True negative* (TN) is rice correctly predicted as rice. The formulas used are as follows:

$$Accuracy=\frac{TP+TN}{TP+TN+FP+FN} \quad (1)$$

$$Precision=\frac{TP}{TP+FP} \quad (2)$$

$$Recall=\frac{TP}{TP+FN} \quad (3)$$

$$F1\text{-}score=\frac{2\times Precision\times Recall}{Precision+Recall} \quad (4)$$

In addition to quantitative evaluation, classification error analysis was conducted through visual inspection of the *false negative* images along with their respective *confidence score*s. The *Grad-CAM* technique was applied to the final convolutional layer of MobileNetV2 to visualize the image regions that most influenced the model's classification decisions (Aminudin et al., 2025).

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16129
e-ISSN : 2541-2019
p-ISSN : 2541-044X

RESULT

Dataset and Preprocessing

The research dataset consists of 1,000 images evenly divided into the normal rice and anomalous paddy classes. All original images were acquired against a uniform dark background. After preprocessing, each image has a new background that varies among five color options. The results of the background removal and random background simulation processes are shown in Figure 2.

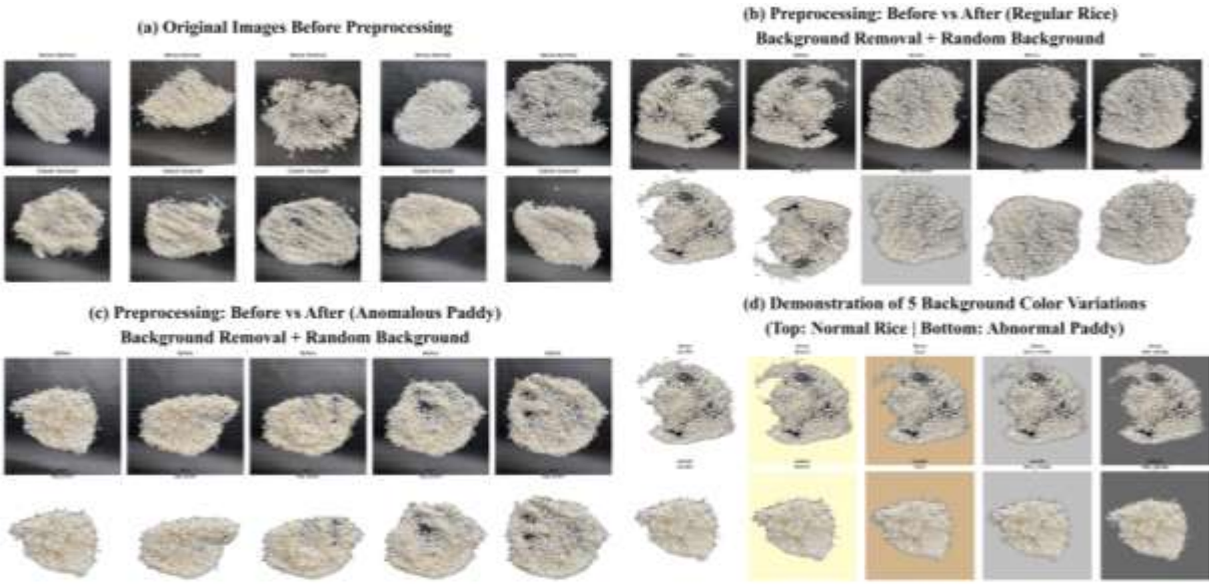

Fig. 2 Dataset Preprocessing: (a) Original Images Before Preprocessing, (b) Normal Rice Preprocessing Results using background removal and random background, (c) Anomalous Paddy Preprocessing Results, and (d) Demonstration of 5 background color variations for both normal rice (top) and abnormal paddy (bottom).

Table 2 Dataset Distribution

Subset	Regular Rice	Abnormal Grain	Total
Training Data (70%)	350	350	700
Validation Data (15%)	75	75	150
Test Data (15%)	75	75	150
Total	500	500	1.000

Model Training

The *accuracy* and loss curves during training are shown in Figure 3, where Figure 3(a) displays the progress in Phase 1 and Figure 3(b) in Phase 2. In Phase 1, training stopped automatically at the 15th epoch because the *early stopping* mechanism was activated, indicating that the model had reached an optimal point. Meanwhile, in Phase 2, or the *fine-tuning* stage, the training process completed more quickly, at the 8th epoch.

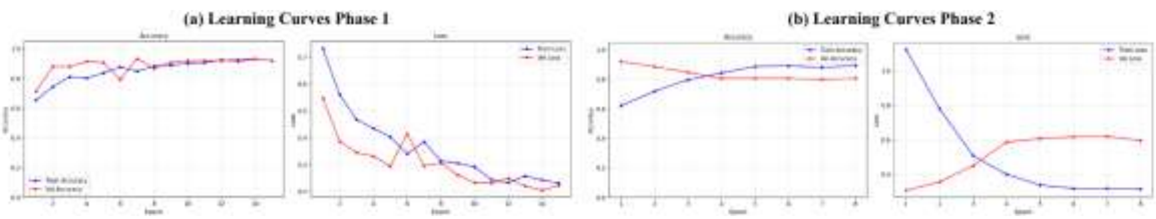

Fig. 3 Learning Curves: (a) Phase 1, (b) Phase 2

Model Performance Evaluation

The confusion matrices for Phase 1 and Phase 2 on the validation and test data are shown in Figure 4.

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16129
e-ISSN : 2541-2019
p-ISSN : 2541-044X

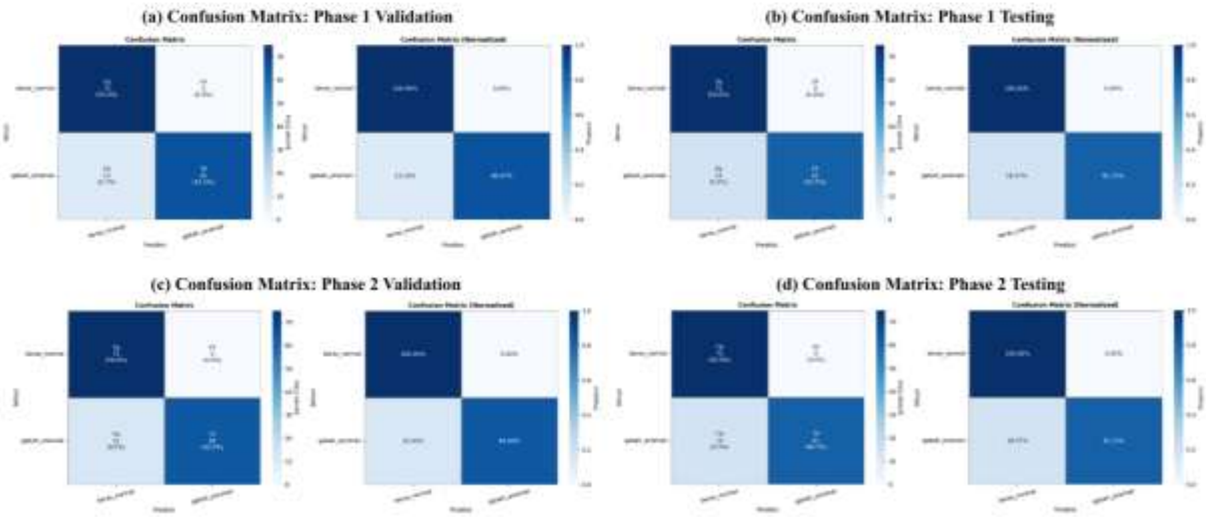

Fig. 4 Confusion Matrix: (a) Phase 1 Validation, (b) Phase 1 Testing, (c) Phase 2 Validation, (d) Phase 2 Testing

The performance of the trained model was then evaluated using the *accuracy*, *precision*, *recall*, and *F1-score* metrics. A summary of the evaluation results for Phase 1 training is presented in Table 3, while the detailed results for Phase 2 after the *fine-tuning* process can be seen in Table 4. Additionally, a comparison of evaluation metrics between Phase 1 and Phase 2 on both the validation and test datasets is presented in Table 5.

Table 3 Phase 1 Metrics Evaluation Results

Metric	Validation Data	Test Data
Accuracy	0,9333 (93,33%)	0,9067 (90,67%)
Precision (macro)	0,9412	0,9213
Recall (macro)	0,9333	0,9067
F1-score (macro)	0,9330	0,9058
True positive (TP)	65	61
True negative (TN)	75	75
False positive (FP)	0	0
False negative (FN)	10	14

Table 4 Phase 2 Metrics Evaluation Results

Metric	Validation Data	Test Data
Accuracy	0,9200 (92,00%)	0,9067 (90,67%)
Precision (macro)	0,9310	0,9213
Recall (macro)	0,9200	0,9067
F1-score (macro)	0,9195	0,9058
True positive (TP)	63	61
True negative (TN)	75	75
False positive (FP)	0	0
False negative (FN)	12	14

Table 5 Phase 1 vs Phase 2 Metrics Comparison

Metric	Phase 1 Validation	Phase 2 Validation	Validation Difference	Phase 1 Testing	Phase 2 Testing	Testing Difference
Accuracy	0,9333	0,9200	-0,0133	0,9067	0,9067	0,0000
Precision (macro)	0,9412	0,9310	-0,0101	0,9213	0,9213	0,0000
Recall (macro)	0,9333	0,9200	-0,0133	0,9067	0,9067	0,0000
F1-score (macro)	0,9330	0,9195	-0,0136	0,9058	0,9058	0,0000

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16129
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Classification Error Analysis

Based on the results of testing on the test data, none of the images produced *false positive*s in either phase. Errors occurred only in the form of *false negative*s, specifically 14 images of abnormal paddy that were incorrectly classified as normal rice in the test data. The *false negative* images and their respective *confidence score*s are shown in Figure 5.

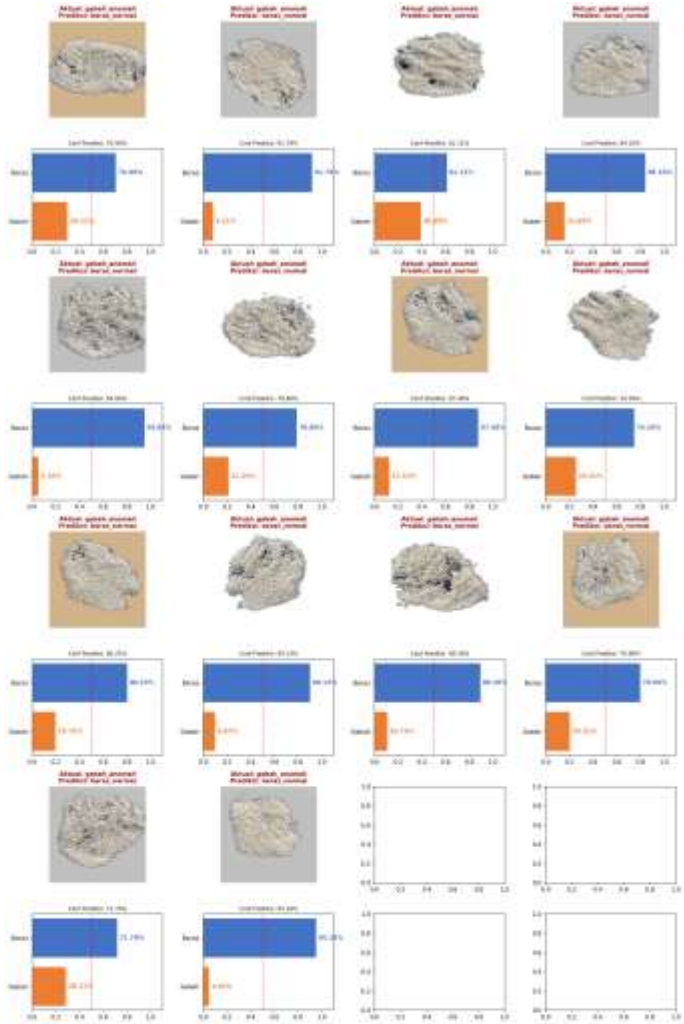

Fig. 5 Visualization of *False negative* Images Using a *Confidence score*

***Grad-CAM* Visualization**

The *Grad-CAM* visualization for a correct prediction is shown in Figure 6(a), while the visualization for a *false negative* is shown in Figure 6(b).

*name of corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16129
e-ISSN : 2541-2019
p-ISSN : 2541-044X

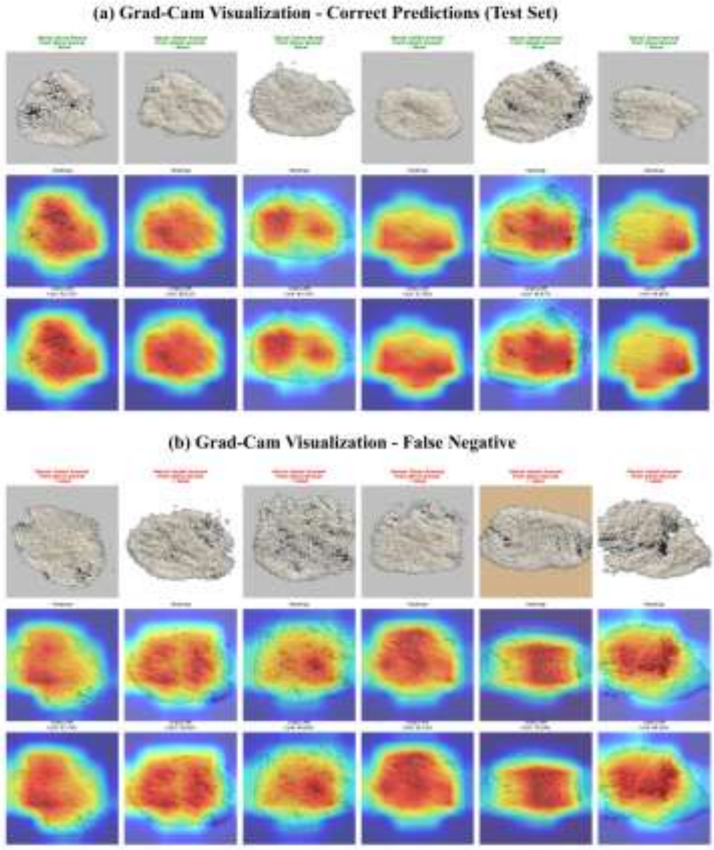

Fig. 6 *Grad-CAM* Visualization: (a) Correct Predictions, (b) *False negative*s

DISCUSSIONS

The MobileNetV2 model trained in this study achieved an *accuracy* of 90.67% on the test dataset, with a macro *precision* of 0.9213, a macro *recall* of 0.9067, and a macro *F1-score* of 0.9058. These results were achieved despite the dataset consisting of only 1,000 self-acquired images, indicating that the combination of *transfer learning*, preprocessing, and data augmentation positively contributes to the model's overall performance.

The most notable finding is that the *false positive* rate consistently remained at 0 across all tests, in both Phase 1 and Phase 2, and on both the validation and test datasets. This is highly positive from a rice quality control perspective, as *false positive*s can lead to the rejection of rice products that are actually of good quality. This indicates that the model has successfully learned the visual features of normal rice very well through the designed training process.

A comparison between Phase 1 and Phase 2 shows that both produced identical performance on the test data; however, Phase 2 yielded slightly lower values on the validation data, with an *accuracy* difference of −0.0133. This indicates mild *overfitting* during *fine-tuning*, meaning the model adapted too closely to the training data, resulting in a slight decline in its generalization ability. This phenomenon is consistent with the findings of Bichri et al. (2023), which state that *fine-tuning* on limited datasets must be performed with great caution because it has the potential to increase the risk of *overfitting*. Based on this, Phase 1 was selected as the best model because it produces more stable and consistent performance.

All classification errors that occurred were *false negative*s, specifically 14 images of abnormal paddy that were incorrectly predicted as normal rice. Based on visual inspection and *confidence score*s ranging from 61.11% to 95.18% for the rice class, the main causes of the errors can be identified as four characteristics: paddy is present in small quantities, resulting in a too small proportion of paddy visual features in the image; the relatively pale color of the paddy, leading to low contrast between classes; the overall appearance of the pile is still dominated by rice grains, so the global visual impression resembles normal rice, and some paddy objects are covered by rice grains, so the distinctive morphological features of paddy are not optimally captured. These conditions indicate that the errors are not solely caused by architectural weaknesses but are influenced by the intrinsic similarity between the two classes under certain visual conditions.

Grad-CAM visualizations directly confirm these findings. In correct predictions, the model focuses red activations on the pile area and specifically on the yellowish or brownish grains of paddy. *Grad-CAM*

*name of corresponding author

SinkrOn
JURNAL & PENELITIAN TEKNIK INFORMATIKA
Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16129
e-ISSN : 2541-2019
p-ISSN : 2541-044X

visualizations of *false negative* samples (Figure 6) reveal three patterns of model failure in recognizing anomalies. First, in a pile dominated by rice, the red activation pattern is centered on the middle area that does not contain grains, while the actual grain objects only receive yellow/green activation. This indicates that the model recognizes the objects but only considers them as supporting features, causing the *confidence score* to fail to exceed the 0.5 threshold. Second, in the case of pale-colored grains, the *heatmap* shows minimal activation differences between the grains and the surrounding rice, indicating the model's limitation in extracting discriminative color features. Third, for partially visible grains (obstructed/incomplete), although the red activation correctly targets the object, the model fails to classify because it cannot fully extract the characteristic morphological features of the grain (long and pointed shape). Overall, this visualization confirms that the errors are not caused by background noise, but rather by the limitations of visual features under certain grain conditions. This pattern confirms that the background removal preprocessing successfully trained the model not to be distracted by background patterns; however, the diversity of visual grain conditions in the training data still needs to be expanded so that the model can recognize grains under more challenging conditions.

Compared to the study by Setiawan et al. (2023), which used manual HSV and GLCM features with an SVM and achieved an *accuracy* of 96.83%, this study achieved a lower *accuracy*. However, the approach used is *end-to-end* without manual feature engineering, so the model relies entirely on automatically learned feature representations. This difference can also be understood from the perspective of task complexity, given that this study uses images of rice piles with a wider variety of visual variations in the grains. Compared to the study by Faqih et al. (2024) which used MobileNetV3Large for rice leaf disease classification and achieved an *accuracy* close to 95%, the results of this study demonstrate the comparable potential of the MobileNetV2 architecture on limited custom datasets within the agricultural domain.

This study has several limitations that must be acknowledged. First, the dataset used is limited to 1,000 self-acquired images with two balanced classes; thus, the diversity of visual conditions of the rice grains in the training data does not cover all possible real-world scenarios that might be encountered in the field, such as rice grains with varying degrees of maturity or under different lighting conditions. Second, this study limits classification to only two classes (normal rice and anomalous grains), so the model is not designed to handle other foreign objects that may be present in the rice pile. Third, the *fine-tuning* process (Phase 2) exhibits mild *overfitting* on the validation data, indicating that the current dataset size is not yet fully optimal to support deeper *fine-tuning*. Fourth, this study does not include testing under real-world conditions outside a controlled environment, so the model's generalization ability regarding images captured under actual field conditions remains unconfirmed.

CONCLUSION

This study successfully applied a MobileNetV2-based Convolutional Neural Network using a *transfer learning* approach to classify unhusked rice as a visual anomaly in rice piles. The preprocessing steps, consisting of background removal and random background simulation with five background color variations, combined with data augmentation on the training data, proved effective in preparing the dataset so that the model could focus more on learning the visual features of rice and unhusked rice objects.

Performance evaluation on the test data showed that the Phase 1 model achieved an *accuracy* of 90.67%, a macro *precision* of 0.9213, a macro *recall* of 0.9067, and a macro *F1-score* of 0.9058, with a consistently zero *false positive* rate across all tests. A comparison between Phase 1 and Phase 2 shows that the *transfer learning* approach with a frozen *base model* is more stable than *fine-tuning* under conditions of limited dataset size; therefore, Phase 1 was selected as the best model.

Classification error analysis reveals that all errors were *false negative*s caused by intrinsic similarities between the two classes, such as pale-colored paddy, paddy present in small quantities, or paddy partially obscured by rice grains. *Grad-CAM* visualizations confirm that in correct predictions, the model focuses its attention on the relevant areas of the paddy grains, whereas in *false negative* cases, the model's activation is concentrated on the center of the pile which contains no paddy or on areas where the paddy is not fully visible, while the areas where paddy actually exists receive only low activation, confirming that errors are more influenced by the limitations of visual features in the image rather than by background noise. These results indicate that increasing the diversity of training data, particularly under conditions where paddy is difficult to recognize, is a crucial step toward improving the model's discriminative ability in future research. Future works should focus on expanding the training dataset to include more diverse visual conditions, such as varying paddy maturity levels, different lighting environments, and a wider range of pile compositions. Additionally, extending the classification scope to detect multiple types of foreign objects in rice piles and exploring real-time deployment on edge devices would further enhance the practical applicability of this system.

REFERENCES

Aminudin, M. A. I., Abdullah, M. N., Mustapha, F., Eng, K. K., Mustapha, M., & Mustapha, A. (2025). Explainable *Deep learning* Framework for Binary Corrosion Image Classification Using *Grad-CAM*.

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 10, Number 3, July 2026
DOI : https://doi.org/10.33395/sinkron.v10i3.16129
e-ISSN : 2541-2019
p-ISSN : 2541-044X

Sensors, *25*(22), 7070. https://doi.org/10.3390/s25227070

Aznan, A., Viejo, C. G., Pang, A., & Fuentes, S. (2021). Computer Vision and *Machine learning* Analysis of Commercial Rice Grains: A Potential Digital Approach for Consumer Perception Studies. *Sensors*, *21*(19), 6354. https://doi.org/10.3390/s21196354

Bastian, A., Priyadi, D., Zaliluddin, D., Mardiana, A., Wahid, A., Rifki, M., & Aziz, M. F. (2025). Penerapan Convolutional Neural Network (CNN) untuk Klasifikasi Kualitas Beras sebagai Strategi Peningkatan Keamanan Pangan di Indonesia. *TEMATIK*, *12*(1), 50–58. https://doi.org/10.38204/tematik.v12i1.2332

Bichri, H., Chergui, A., & Hain, M. (2023). Image Classification with *Transfer learning* Using a Custom Dataset: Comparative Study. *Procedia Computer Science*, *220*, 48–54. https://doi.org/10.1016/j.procs.2023.03.009

Chen, W., Li, W., & Wang, Y. (2022). Evaluation of Rice Degree of Milling Based on Bayesian Optimization and Multi-Scale Residual Model. *Foods*, *11*(22), 3720. https://doi.org/10.3390/foods11223720

Chun, T. H., Hashim, U. R., Ahmad, S., Salahuddin, L., Choon, N. H., & Kanchymalay, K. (2022). Efficacy of the Image Augmentation Method Using CNN *Transfer learning* in Identification of Timber Defect. *International Journal of Advanced Computer Science and Applications*, *13*(5), 107–114. https://doi.org/10.14569/IJACSA.2022.0130514

Faqih, R. R., Irsan, M., & Fathoni, M. F. (2024). Rice Plant Disease Detection System Using *Transfer learning* with MobilenetV3Large. *Sinkron*, *8*(2), 805–812. https://doi.org/10.33395/sinkron.v8i2.13383

Gkountakos, K., Ioannidis, K., Demestichas, K., Vrochidis, S., & Kompatsiaris, I. (2024). A Comprehensive Review of *Deep learning*-Based Anomaly Detection Methods for *Precision* Agriculture. *IEEE Access*, *12*, 197715–197733. https://doi.org/10.1109/ACCESS.2024.3522248

Gong, L., & Fan, S. (2022). A CNN-Based Method for Counting Grains Within a Panicle. *Machines*, *10*(1), 30. https://doi.org/10.3390/machines10010030

Gulzar, Y. (2023). Fruit Image Classification Model Based on MobileNetV2 with Deep *Transfer learning* Technique. *Sustainability*, *15*(3), 1906. https://doi.org/10.3390/su15031906

He, Y., Fan, B., Sun, L., Fan, X., Zhang, J., Li, Y., & Suo, X. (2023). Rapid Appearance Quality of Rice Based on Machine Vision and Convolutional Neural Network Research on Automatic Detection System. *Frontiers in Plant Science*, *14*. https://doi.org/10.3389/fpls.2023.1190591

Ilo, B., Badjona, A., Singh, Y., Shenfield, A., & Zhang, H. (2025). Artificial Intelligence in Rice Quality and Milling: Technologies, Applications, and Future Prospects. *Processes*, *13*(11), 3731. https://doi.org/10.3390/pr13113731

Mendoza-Bernal, J., González-Vidal, A., & Skarmeta, A. F. (2024). A Convolutional Neural Network Approach for Image-Based Anomaly Detection in Smart Agriculture. *Expert Systems with Applications*, *247*(2), 123210. https://doi.org/10.1016/j.eswa.2024.123210

Prabowo, S. T., & Hadikurniawati, W. (2023). Deteksi dan Pengenalan Jenis Beras Menggunakan Metode Convolutional Neural Network. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(1), 163–167. https://doi.org/10.36040/jati.v7i1.6150

Setiawan, A., Adi, K., & Widodo, C. E. (2023). Rice Foreign Object Classification Based on Integrated Color and Textural Feature Using *Machine learning*. *Mathematical Modelling of Engineering Problems*, *10*(2), 572–580. https://doi.org/10.18280/mmep.100226

Sokudlor, N., Laloon, K., Junsiri, C., & Sudajan, S. (2023). Enhancing Milled Rice Qualitative Classification with *Machine learning* Techniques Using Morphological Features of Binary Images. *International Journal of Food Properties*, *26*(2), 2978–2992. https://doi.org/10.1080/10942912.2023.2264533

Suryana, E. A. (2023). Application to Estimate the Rice Polishing Degree Using Image Processing. *Jurnal Pangan*, *32*(1), 33–40. https://doi.org/10.33964/jp.v32i1.585

Thammastitkul, A., & Petsuwan, J. (2023). Thai Hom Mali Rice Grading Using *Machine learning* and *Deep learning* Approaches. *IAES International Journal of Artificial Intelligence (IJ-AI)*, *12*(2), 815–822. https://doi.org/10.11591/ijai.v12.i2.pp815-822

Wang, C., Caragea, D., Narayana, N. K., Hein, N. T., Bheemanahalli, R., Somayanda, I. M., & Jagadish, S. V. K. (2022). *Deep learning* Based High-Throughput Phenotyping of Chalkiness in Rice Exposed to High Night Temperature. *Plant Methods*, *18*(9), 2–23. https://doi.org/10.1186/s13007-022-00839-5

Wang, Y., & Su, W. (2022). Convolutional Neural Networks in Computer Vision for Grain Crop Phenotyping: A Review. *Agronomy*, *12*(11), 1–25. https://doi.org/10.3390/agronomy12112659

Yumono, F., Yuliana, D. E., & Sarbini, R. N. (2022). Histogram Citra Jenis Beras dengan Menyertakan Kertas Putih Untuk Identifikasi Awal Jenis Beras dengan Menggunakan Jaringan Syaraf Tiruan. *Jurnal Informatika dan Rekayasa Perangkat Lunak*, *3*(2), 129–137. https://doi.org/10.33365/jatika.v3i1.1891

Yusuf, M., Ruimassa, R., Tawainella, A. I., & Maharani, D. (2024). Klasifikasi Kualitas Beras Menggunakan Convolutional Neural Network Berbasis Android. *Jurnal Komputer dan Informatika*, *12*(2), 186–192. https://doi.org/10.35508/jicon.v12i2.18004

*name of corresponding author