JITK (JURNAL ILMU PENGETAHUAN DAN TEKNOLOGI KOMPUTER)

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8258

EARLY DETECTION OF STUNTING IN CHILDREN USING HSV AND GLCM FEATURES WITH CNN-BASED CLASSIFICATION

Ade Cristian Silalahi[1]; Bonatio Vincent E Hutagalung[1]; Sandy Walfredo Ritonga[1]; Marlince NK Nababan[1*]

Department of Science and Technology [1]
University Prima Indonesia, Medan, Indonesia[1]
https://www.unprimdn.ac.id[1]
adecristiansilalahi11@gmail.com, bonatiobona@gmail.com, sandywritonga2004@gmail.com, marlince@unprimdn.ac.id*

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

Abstract— *Stunting is a chronic nutritional issue in children that can have long-term impacts on physical growth and cognitive development. Early detection is therefore important to support timely intervention. This study develops an image-based stunting detection approach that integrates HSV color descriptors, Gray Level Co-occurrence Matrix (GLCM) texture descriptors, and CNN features extracted with MobileNetV2. Images are preprocessed through resizing and Region of Interest (ROI) cropping based on bounding-box annotations. The handcrafted HSV and GLCM features are fused with CNN features at the feature level through vector concatenation before classification into stunting and non-stunting categories. This design was selected to preserve complementary low-level color-texture information and high-level semantic representations in a single classifier input. The hybrid model achieved a test accuracy of 84.39% with a stunting recall of 83%. Although the results indicate that multimodal visual descriptors can improve classification performance compared with single-feature approaches, the model still shows mild overfitting and was evaluated on a relatively limited dataset. In addition, inference efficiency and robustness to variations in imaging conditions were not yet quantitatively measured. Therefore, the present system should be interpreted as a proof of concept for objective, early, image-based stunting screening by healthcare personnel.*

Keywords: *CNN, Digital Image Processing, GLCM, HSV, Stunting.*

Intisari— *Stunting adalah kondisi gangguan pertumbuhan pada anak akibat malnutrisi kronis yang dapat berdampak pada perkembangan fisik dan kognitif. Deteksi dini diperlukan untuk mendukung intervensi yang lebih cepat dan tepat. Penelitian ini mengembangkan pendekatan deteksi stunting berbasis citra dengan mengintegrasikan fitur warna HSV, fitur tekstur Gray Level Co-occurrence Matrix (GLCM), dan fitur mendalam CNN yang diekstraksi menggunakan MobileNetV2. Citra diperoleh melalui resizing dan pemotongan Region of Interest (ROI) berdasarkan anotasi bounding box. Fitur handcrafted HSV dan GLCM kemudian digabungkan dengan fitur CNN pada level fitur melalui konkatenasi vektor sebelum proses klasifikasi ke dalam kategori stunting dan non-stunting. Strategi ini dipilih untuk mempertahankan informasi warna-tekstur tingkat rendah yang saling melengkapi dengan representasi semantik tingkat tinggi dari CNN. Model hibrida mencapai akurasi pengujian sebesar 84,39% dengan recall stunting sebesar 83%. Hasil ini menunjukkan bahwa kombinasi deskriptor visual dapat meningkatkan kinerja klasifikasi dibandingkan dengan penggunaan satu jenis fitur saja, meskipun model masih menunjukkan gejala overfitting yang ringan dan diuji pada dataset yang relatif terbatas. Selain itu, efisiensi inferensi dan ketahanan terhadap variasi kondisi pengambilan citra belum diukur secara kuantitatif. Dengan demikian, sistem ini diposisikan sebagai proof-of-concept untuk skrining awal stunting berbasis citra secara lebih objektif.*

Kata Kunci*: CNN, GLCM, HSV, Pengolahan Citra Digital, Stunting.*

INTRODUCTION

Stunting and malnutrition in children are global health issues that remain major concerns due to their long-term effects on physical growth, cognitive development, and the quality of future human resources [1], [2]. This condition is caused by various interrelated factors, including inadequate nutrition, maternal and child health conditions, poor environmental sanitation, and socioeconomic factors [3], [4]. Various studies have shown that stunting is chronic and can impact productivity and the risk of non-communicable diseases in adulthood [5].

In various countries, including Indonesia, efforts to reduce the prevalence of stunting have been undertaken through nutrition intervention programs and improved health services. However, the incidence of stunting remains relatively high, necessitating a more effective and accurate early detection approach [6], [7]. The conventional process for identifying nutritional status generally involves anthropometric measurements such as a child's weight, height, and age. This approach still has limitations, including being prone to recording errors, requiring a relatively long time, and being highly dependent on the availability of competent health workers [8]. The Convolutional Neural Network (CNN) method can be utilized to classify stunting status in toddlers automatically, enabling faster and more objective identification compared to manual assessments [9].

The development of information technology has encouraged the use of artificial intelligence, particularly machine learning, as an alternative solution to support early detection of stunting and child nutritional status [10]. Various algorithms have been applied, such as Decision Trees [11]. Random Forests [12]. and K-Nearest Neighbor with Naive Bayes for classification tasks [13], [14]. Furthermore, image-based calorie detection models for food have demonstrated that computer vision technology can assist in monitoring nutritional intake more effectively [15]. Research results show that machine learning methods can achieve a relatively high level of accuracy in classifying toddlers' nutritional status from anthropometric data [16]. Nevertheless, most existing approaches still rely heavily on numerical data with limited feature diversity, which may reduce model generalization performance [17].

CNN has proven effective in image classification tasks due to its ability to automatically extract discriminative visual features and recognize complex patterns [18], [19]. Image-based methods have been successfully applied in various healthcare domains, including medical image analysis and biological condition identification, with promising outcomes [20], [21]. However, deep learning approaches also face challenges, particularly regarding large dataset requirements, data consistency, and comprehensive evaluation procedures [22].

Previous studies have demonstrated the strong potential of machine learning and deep learning for the early detection of stunting and malnutrition in children [23]. The integration of CNN with feature extraction techniques such as Gray Level Co-occurrence Matrix (GLCM), HSV, and L*a*b* color spaces has been shown to improve classification performance by capturing texture and color characteristics more effectively [24]. Additionally, early stopping techniques have been proven to enhance CNN performance by preventing overfitting and improving model generalization [25]. Recent developments have further explored deep learning architectures for stunting and malnutrition prediction by integrating socioeconomic data, facial images, and community participation [26], [27]. Advanced CNN architectures and high-performance computing strategies have demonstrated improved accuracy and robustness in medical image classification tasks [28]. Moreover, deep learning models based on facial image analysis, such as ResNet-50, have shown strong potential for non-invasive malnutrition detection in children [29].

By combining several classification models, ensemble methods can enhance prediction stability and reduce bias arising from single-model limitations. Empirical studies conducted at the community level have shown that ensemble machine learning techniques can effectively support early detection of stunting in toddlers, thereby assisting healthcare providers in implementing timely and targeted interventions [30].

Therefore, future research should focus on developing robust, scalable, and context-aware stunting detection systems by integrating ensemble and deep learning with relevant anthropometric and socioeconomic features.

MATERIALS AND METHODS

This study uses an experimental approach based on digital image processing and *deep learning* to detect stunting in children. The dataset consists of child images obtained from the Roboflow platform and has been annotated with bounding boxes. The dataset consists of two classes, namely normal and stunting, with a total of 2,352 training data images and 668 test data images. The entire

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. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8258

image was resized to 224 × 224 pixels with three color channels (RGB). The training set consists of 1,399 normal images and 953 stunting images (approximately 59.5% normal, 40.5% stunting), while the test set contains 378 normal images and 290 stunting images (approximately 56.6% normal, 43.4% stunting). The dataset was sourced from the publicly accessible Roboflow platform, reflecting real-world annotation by research communities. The fixed train-test split follows the partition defined at the dataset source to ensure consistent evaluation. Although cross-validation (e.g., K-Fold) was not applied in this study due to computational constraints, it is recommended as a direction for future work to further validate model generalizability.

The preprocessing stage includes image reading, size adjustment, and *the Region of Interest (ROI) cropping* process based on bounding-box annotations. The ROI cropping process aims to eliminate irrelevant background, so the analysis focuses on important areas of the child image related to stunting.

HSV Color Feature Extraction

Color feature extraction was performed using the HSV (Hue, Saturation, Value) color space to represent the image's color characteristics more robustly against lighting changes than the RGB color space. The cropping ROI image is converted from RGB to HSV, and then each HSV channel is analyzed separately.

The HSV color space is used because it represents color distribution more robustly under variations in lighting conditions than RGB. The Value (V) component represents the level of brightness associated with skin vitality; Saturation (S) describes color intensity, which may decrease under malnutrition; and Hue (H) represents the color tendency, which remains relatively stable.

Meanwhile, texture features based on the Gray Level Co-occurrence Matrix (GLCM) are employed to capture spatial relationships between pixels, reflecting texture patterns such as smoothness, homogeneity, and structural consistency of the surface. These texture variations can be influenced by physiological factors such as hydration levels and skin tissue quality, which are related to nutritional status.

Thus, the HSV and GLCM features in this study are used as supportive visual descriptors that may correlate with physiological conditions associated with malnutrition, rather than as standalone diagnostic indicators. The integration of these features with deep features extracted from CNN aims to produce a more comprehensive representation for the classification process. The mean value of each HSV channel is calculated to illustrate the color-dominant tendency in the image, which is expressed in Equation (1):

$$\mu = \frac{1}{N}\sum_{i=1}^{N} x_i \quad (1)$$

where is the value of pixel intensity to - and expresses the total number of pixels on the channel.$x_i iN$

In addition**,** the standard deviation is calculated to describe the degree of spread or variation in color values in the image. The calculation of the standard deviation is shown in Equation (2):

$$\sigma = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - \mu)^2} \quad (2)$$

From this process, six color features were obtained, namely mean values and standard deviations for the Hue, Saturation, and Value channels. These features are used to help distinguish the visual characteristics between normal and stunting child imagery.

In practical image acquisition, the Value and Saturation components can help represent brightness and color concentration patterns, while Hue helps represent color tendency in a way that is more stable than raw RGB. These descriptors are therefore used as auxiliary features to complement deep visual representations rather than as standalone diagnostic indicators.

GLCM Texture Feature Extraction

Texture feature extraction is performed using the Gray Level Co-occurrence Matrix (GLCM) method to represent the spatial relationships between pixels in the image. Prior to forming the GLCM matrix, the ROI-cropped results image was converted to grayscale.

HSV and GLCM features are not used as direct clinical indicators of stunting; rather, they serve as visual descriptors that capture indirect changes in physiological conditions resulting from malnutrition, such as variations in skin color and texture. The combination of these features with CNN aims to enhance data representation and improve classification performance.

The GLCM matrix is built from distances between pixels at four angular orientations: 0°, 45°, 90°, and 135°. This matrix represents the probability of a pair of pixel intensity values occurring at a given position.

The contrast texture feature is used to measure the degree of intensity difference between pixels, which is calculated using Equation (3):

$$Contrast = \sum_{i,j}(i-j)^2 P(i,j) \quad (3)$$

The correlation feature measures the linear relationships between pixel pairs and reflects the degree of interconnectedness of the image texture structure, as expressed in Equation (4):

$$Correlation = \frac{\sum_{i,j}(i-\mu_i)(j-\mu_j)P(i,j)}{\sigma_i \sigma_j} \quad (4)$$

The energy feature shows the uniformity of the texture of the image and is calculated as the square of the GLCM probability value, as shown in Equation (5):

$$Energy = \sum_{i,j} P(i,j)^2 \quad (5)$$

Meanwhile, the homogeneity feature measures the proximity of the distribution of GLCM elements to the main diagonal, which reflects the uniformity of the image texture, and is calculated using Equation (6):

$$Homogeneity = \sum_{i,j} \frac{P(i,j)}{1+|i-j|} \quad (6)$$

The GLCM feature used is the average across all angular directions to obtain a more stable texture representation. CNN operates through several main layers, such as the convolution layer, pooling layer, and fully connected layer, which progressively learn spatial patterns in images to produce more discriminative feature representations for classification tasks [31].

Recent studies indicate that CNNs achieve excellent performance in medical image classification by recognizing complex patterns in visual data that are difficult to identify manually. A comprehensive review examined various CNN methods for cancer image classification, including transfer learning, multimodal data fusion, and preprocessing strategies to improve model accuracy. The study highlights that CNN is highly effective in medical applications such as cancer detection through radiological and histopathological images [32].

Furthermore, advancements in modern CNN architectures enable optimization of network structures, resulting in models that are more efficient and accurate for biomedical image classification tasks. Another study shows that CNN-based models such as VGG, ResNet, and Inception are frequently used as primary backbones in medical image classification and achieve high performance across datasets involving diseases such as brain tumors, pneumonia, and breast cancer [33].

With its capability for automatic feature extraction, adaptability to image variations, and high classification performance, CNN has become one of the primary approaches in various computer vision applications, including medical image analysis and computer-aided diagnosis systems.

Hybrid Model Architecture

Deep feature extraction was performed using a *Convolutional Neural Network* (CNN) with the MobileNetV2 architecture. MobileNetV2 was chosen because it has a lightweight, efficient architecture, making it suitable for use on datasets with limited size. Convolutional operations on CNN aim to extract local visual patterns from input images and are expressed in Equation (7):

$$S(i,j) = (I * K)(i,j) \quad (7)$$

Feature-level fusion is performed by concatenating the feature vectors from three sources into a single unified representation. The HSV color feature extraction produces 6 features (mean and standard deviation of the Hue, Saturation, and Value channels). The GLCM texture feature extraction produces 4 features (contrast, correlation, energy, and homogeneity), each averaged across four angular orientations (0°, 45°, 90°, 135°). These two handcrafted feature vectors are first concatenated into a 10-dimensional vector, which serves as the secondary input (feature_input) to the model. Meanwhile, the MobileNetV2 backbone (pretrained on ImageNet, with weights frozen) processes the image through a Global Average Pooling (GAP) layer, followed by a Dense layer with 256 neurons and ReLU activation, yielding a 256-dimensional deep feature vector. Feature-level fusion is then performed by concatenating the 256-dimensional CNN output with the 10-dimensional handcrafted feature vector, as expressed in Equation (8):
The feature fusion process is defined as follows:

$$f_{\text{fusion}} = [f_{\text{CNN}} \parallel f_{\text{HSV}} \parallel f_{\text{GLCM}}] \quad (8)$$

Where f_{CNN}represents the 256-dimensional deep feature vector extracted from MobileNetV2, f_{HSV}denotes the 6-dimensional HSV color feature

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. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8258

vector, and f_{GLCM} corresponds to the 4-dimensional GLCM texture feature vector. Consequently, the fused feature vector has a total dimensionality of 266 features. This fused representation is subsequently passed through a Dense layer with 128 neurons and ReLU activation, as defined in Equation (9):

$$h = \text{ReLU}(W_{128} \cdot f_{\text{fusion}} + b_{128}) \quad (9)$$

Where $\text{ReLU}(x) = \max(0, x)$ is the rectified linear unit activation function, W_{128} the weight matrix, and b_{128} is the bias vector of the Dense (128) layer. The output layer consists of a single neuron with sigmoid activation to produce a probability value in the range of 0 to 1, as expressed in Equation (10):

$$\sigma(z) = \frac{1}{1+e^{-z}} \quad (10)$$

Where $z = W_1 \cdot h + b_1$ is the linear transformation of the Dense (128) output h.
A classification threshold of 0.5 is applied to the sigmoid output to determine the final class label, as formulated in Equation (11):

$$\hat{y} = \begin{cases} 1 \text{ (Stunting)}, & \text{if } \sigma(z) > 0.5 \\ 0 \text{ (Normal)}, & \text{if } \sigma(z) \leq 0.5 \end{cases} \quad (11)$$

This binary classification corresponds to the two folder labels in the dataset, where *normal* = 0 and *stunting* = 1, which are automatically assigned by TensorFlow based on alphabetical directory ordering. The model is compiled using the Adam optimizer with binary cross-entropy loss, and trained for 15 epochs with a batch size of 32. The application of feature-level fusion through concatenation has been shown to improve classification performance by enabling the model to leverage complementary information from both handcrafted and deep features simultaneously.

Model Evaluation Metrics

The evaluation of model performance was carried out using several metrics, namely accuracy, precision, recall, and F1-score, to assess the model's ability to classify normal and stunting images.
The accuracy value is used to measure the accuracy level of the model as a whole and is calculated using Equation (12):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (12)$$

The precision value measures the level of accuracy of the stunting class prediction and is calculated using Equation (13):

$$Precision = \frac{TP}{TP+FP} \quad (13)$$

The recall value shows the model's ability to detect all the actual stunting data and is expressed in Equation (14):

$$Recall = \frac{TP}{TP+FN} \quad (14)$$

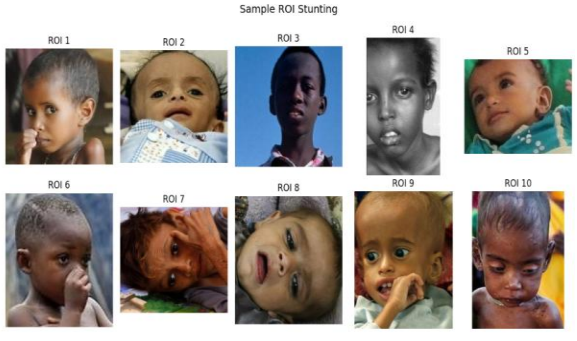
Where as the F1-score is used to measure the balance between precision and recall, as shown in Equation (15):

$$\text{F1-score} = 2 \times \frac{\text{Precision} \times \text{Recall}}{\text{Precision} + \text{Recall}} \quad (15)$$

RESULTS AND DISCUSSION

Results of ROI Preprocessing

Preprocessing is performed by applying a Region of Interest (ROI) crop based on bounding box annotations to each image. An example of the cropping ROI results in the stunting class is shown in Figure 1. This process helps increase the model's focus on the child's physical characteristics related to stunting.



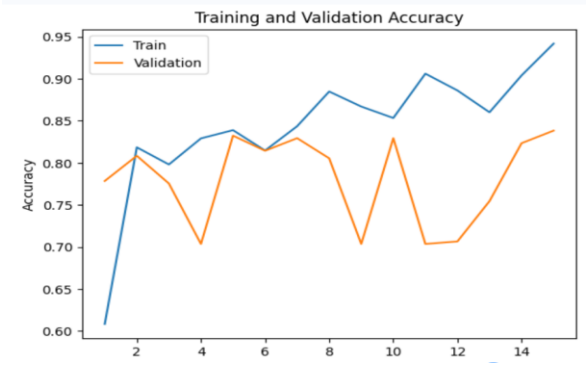
Source: (Research Results, 2026)
Figure 1. Example of Region of Interest (ROI) cropping results in the stunting class

The dataset used in this study was obtained from the Roboflow platform and can be accessed through the following links ***https://roboflow.com/*** (The dataset is used for research purposes and the development of stunting classification models)

Model Training Results for 15 Epochs

The HSV–GLCM–CNN hybrid model was trained for 15 epochs on the training data with a batch size of 32. The training process aims to learn

the visual patterns in children's imagery using CNN features for color, texture, and depth. The development of model performance during training was analyzed using graphs of accuracy and loss over epochs, as shown in Figures 2 and 3. Based on the accuracy graph, accuracy on the training data increases gradually with increasing epoch, reaching above 90% in the final epoch. Meanwhile, the accuracy of the validation data ranges from 70% to 84% and fluctuates between epochs. This shows that the model can learn the pattern of the training data well, although there is still a performance difference between the training and validation data.



Source: (Research Results, 2026)
Figure 2. Training and Validation Accuracy Graph for the Epoch

The loss graph in Figure 3 shows that the training loss decreased significantly in the initial epoch and then stabilized. On the other hand, the validation data loss shows greater fluctuations, indicating mild overfitting in the last few epochs.

Source: (Research Results, 2026)
Figure 3. Training and Validation Loss Graph Against Epoch

Evaluation of Model Performance Based on Classification Metrics

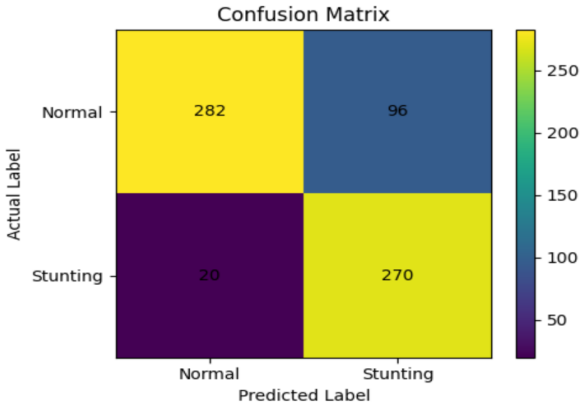
Model performance was evaluated on test data to assess classification performance between the normal and stunting classes. The results of the evaluation are presented as a comparison of accuracy, precision, recall, and F1-score, as shown in Figure 4. The evaluation results showed that the HSV–GLCM–CNN hybrid model achieved an accuracy of 84.39%. In addition, precision, recall, and F1-score values show a good balance between detection capability and misclassification rate. The relatively high stunting recall value indicates that the model can detect most stunting cases, which is especially important for early detection.

```
21/21 ━━━━━━━━━━━━━━━━━━━━ 33s 1s/step
Confusion Matrix:
[[320  58]
 [ 50 240]]

Classification Report:
              precision    recall  f1-score   support

      Normal       0.86      0.85      0.86       378
    Stunting       0.81      0.83      0.82       290

    accuracy                           0.84       668
   macro avg       0.84      0.84      0.84       668
weighted avg       0.84      0.84      0.84       668
```

Source: (Research Results, 2026)
Figure 4. Comparison of Accuracy, Precision, Recall, and F1-Score Values

Matrix Confusion Analysis

To provide a more detailed picture of the classification results, the evaluation was also carried out using a confusion matrix, as shown in Figure 5. Based on the confusion matrix, the model correctly classified most stunting cases, and only a small portion was misclassified as normal. This shows that the model is highly sensitive to stunting classes, which is an important aspect of the early detection system.



Source: (Research Results, 2026)
Figure 5. Confusion matrix results of Normal and Stunting classification

Discussion

The results show that the ROI cropping process plays an important role in improving the model's focus on relevant image regions. In addition, the combination of HSV color features,

VOL. 11. NO. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8258

GLCM texture features, and CNN in-depth features provides a more comprehensive representation than using a single feature type alone. Training the model for over 15 epochs showed that it converged, though there were indications of mild overfitting, as evidenced by fluctuations in accuracy and validation loss. However, the evaluation results on the test data showed stable performance, with good accuracy and other evaluation metrics.

To mitigate overfitting, several strategies were incorporated during training: (1) the MobileNetV2 backbone weights were frozen throughout training, effectively reducing the number of trainable parameters and preventing over-adaptation to the limited dataset; (2) training was limited to 15 epochs, serving as an implicit early-stopping constraint; and (3) a balanced class split was maintained across both sets. Despite these measures, the observed gap between training accuracy (100% from epoch 6 onward) and validation accuracy (~89.82%) indicates significant overfitting. Future work should consider explicit dropout regularization, data augmentation during training, and learning rate scheduling to improve generalization.

With a high recall value for stunting detection and a relatively low number of misclassifications, the proposed HSV–GLCM–CNN hybrid model demonstrates potential as a supporting tool for early image-based stunting screening. In comparison with prior studies discussed in the literature review, these findings support the premise that integrating handcrafted features with CNN-based representations can enhance overall image feature richness. However, since no direct comparative experiments were conducted against alternative models on the same dataset, the results should be interpreted as evidence of feasibility rather than as a definitive state-of-the-art achievement.

Another important consideration is practical deployment. Although MobileNetV2 was selected for its lightweight architecture, this study did not include quantitative efficiency evaluations of inference time, memory footprint, or peak computational resource usage during deployment. Consequently, claims regarding real-world readiness remain limited. Future work should incorporate latency and memory profiling on target devices commonly used in healthcare settings, such as mid-range smartphones or standard laptops. To further contextualize the contribution of the hybrid approach, a comparison with related baselines is warranted. The hybrid HSV–GLCM–CNN model achieved a test accuracy of 84.39%, higher than that expected from handcrafted features alone. Johari et al. [24] demonstrated that a CNN combined with GLCM, HSV, and L*a*b* features achieved up to 98.89% accuracy in classifying corn leaf diseases, confirming the general effectiveness of such hybrid approaches. In the context of stunting, Yunidar et al. [25] showed that CNN-based models benefit from complementary feature strategies to improve classification robustness.

A CNN-only baseline (MobileNetV2 with Global Average Pooling, without handcrafted feature fusion) would rely solely on deep convolutional activations, potentially missing fine-grained color and texture variations relevant to nutritional status assessment on a limited dataset. A formal ablation study comparing CNN-only, handcrafted-only, and hybrid configurations on the same dataset is recommended as future work to provide quantitative evidence of each component's contribution. To address concerns about robustness to real-world deployment, a dedicated robustness experiment was conducted using image augmentation to simulate varying acquisition conditions.

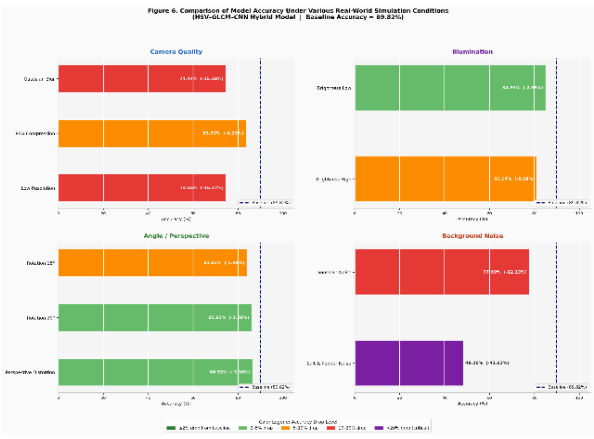
Robustness Experiment

To evaluate the model's robustness under real-world imaging conditions, a controlled robustness experiment was conducted using augmentation-based simulations on the validation set. Ten augmentation conditions were applied and grouped into four categories: (1) camera quality variation, (2) illumination variation, (3) shooting angle and perspective distortion, and (4) background noise. The applied augmentation techniques included Gaussian blur (kernel sizes of 5 and 11), JPEG compression (quality levels of 20 and 10), low-resolution downscaling (25%), brightness reduction (factors of 0.4 and 0.6), brightness increase (factor of 1.6), rotation (15° and 30°), perspective distortion (10% and 20%), Gaussian noise (standard deviations of 0.05 and 0.10), and salt-and-pepper noise (5%). For each augmented dataset, HSV and GLCM features were re-extracted to capture realistic variations in features under degraded imaging conditions.

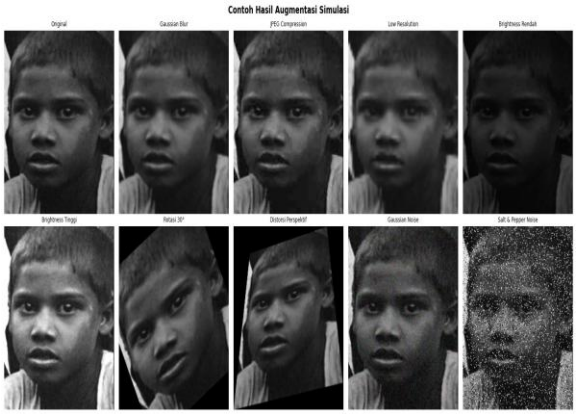
The experimental results indicate that the proposed model maintains relatively stable performance under moderate augmentation conditions. In terms of camera quality variation, Gaussian blur (size = 5) and JPEG compression (quality = 20) produced accuracy values close to the baseline. However, more severe degradations, such as Gaussian blur (size = 11) and JPEG compression (quality = 10), led to a moderate decline in accuracy, suggesting sensitivity to significant blur and compression artifacts. Similarly, low-resolution

downscaling to 25% resulted in a noticeable performance drop, likely due to the loss of fine-grained texture information essential for GLCM feature extraction. Regarding illumination variation, the model exhibited greater sensitivity to extreme brightness reduction (factor = 0.4) than to moderate reductions or increases in brightness. This finding suggests that severely underexposed images negatively affect the representation of HSV Value and Saturation components. In contrast, moderate brightness adjustments (factors of 0.6 and 1.6) produced accuracy values closer to the baseline, indicating partial robustness to typical lighting variations encountered in real-world scenarios.

The rotation and perspective distortion experiments revealed that small angular changes (15°) had minimal impact on model performance. However, larger rotations (30°) and stronger perspective distortions (20%) led to greater degradation in accuracy. Finally, noise-based augmentation results show that the model is moderately robust to low-level Gaussian noise but more sensitive to higher-intensity Gaussian noise (std = 0.10) and salt-and-pepper noise. This behavior is likely due to disruptions in texture patterns used for GLCM-based feature extraction. These findings highlight potential directions for future improvement, including integrating augmentation-based training strategies, domain generalization techniques, and robustness-aware regularization to enhance model reliability in heterogeneous real-world screening environments. A comparison of model accuracy across all augmentation conditions is presented in Figure 6, while representative examples of augmented images are shown in Figure 7.



Source: (Research Results, 2026)
Figure 6. Model Accuracy Under Various Real-World Simulation Conditions (Baseline Accuracy = 89.82%)



Source: (Research Results, 2026)
Figure 7. Augmentation-Based Image Simulation Examples for Robustness Testing

CONCLUSION

This study proposes an image-based stunting classification approach using a hybrid model that combines HSV color features, GLCM texture features, and CNN deep features. ROI cropping effectively focuses the analysis on relevant image regions, supporting feature extraction and model training. The HSV–GLCM–CNN hybrid model achieved a test accuracy of 84.39% with a stunting recall of 83%, demonstrating strong detection capability. The combination of color, texture, and deep features improved classification performance over single-feature approaches. Despite mild overfitting, test performance remained stable and acceptable. A robustness experiment further evaluated the model under simulated real-world conditions, including camera quality degradation, illumination changes, perspective distortion, and background noise. Results show reasonable stability under mild perturbations but sensitivity to severe degradation, identifying key areas for future improvement. Future work should address direct benchmarking against alternative classifiers, efficiency evaluation on deployment devices, augmentation-based training, domain generalization, dataset expansion, and stronger regularization. The proposed method should therefore be regarded as an application-oriented hybrid baseline for image-based stunting screening rather than a fully deployable diagnostic system.

REFERENCE

[1] Bhakti Vichave, Nikhil Jain, Pankaj Garad, Namit Gandhi, and Prof. Dewanand Meshram, “Malnutrition Detection using AI,” *Int. J. Adv. Res. Sci. Commun. Technol.*, pp. 285–291, 2023, doi: 10.48175/ijarsct-9692.

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. 4 MAY 2026
P-ISSN: 2685-8223 | E-ISSN: 2527-4864
DOI: 10.33480/jitk.v11i4.8258

[2] A. T. Mulyani, M. A. Khairinisa, A. Khatib, and A. Y. Chaerunisaa, "Understanding Stunting: Impact, Causes, and Strategy to Accelerate Stunting Reduction—A Narrative Review," *Nutr.* , vol. 17, no. 9, 2025, doi: 10.3390/nu17091493.

[3] T. A. E. Permatasari, Y. Chadirin, E. Ernirita, A. N. Syafitri, and D. A. Fadhilah, "The accuracy of a novel stunting risk detection application based on nutrition and sanitation indicators in children aged under five years," *BMC Nutr.*, vol. 11, no. 1, 2025, doi: 10.1186/s40795-025-01074-6.

[4] Debby Ratno Kustanto, Indah Putri Ramadhanti, and Masya Putri, "Stunting Diberbagai Negara: Perbandingan Global," *J. Pengabdi. Masy. Aufa*, vol. 7, no. 1, 2025, doi: 10.51933/jpma.v7i1.1932.

[5] H. Shen, H. Zhao, and Y. Jiang, "Machine Learning Algorithms for Predicting Stunting among Under-Five Children in Papua New Guinea," *Children*, vol. 10, no. 10, 2023, doi: 10.3390/children10101638.

[6] B. Alnur, Mulyono, Fitri Amillia, and S. Sutoyo, "Performance Analysis of 10 Mbps Wireless Iconnet in Perumahan Bumi Mi'raj," *J. Informatics Telecommun. Eng.*, vol. 7, no. 1, pp. 102–111, 2023, doi: 10.31289/jite.v7i1.9548.

[7] F. M. T. Pane and D. Hindarto, "Comparative Analysis of Machine Learning Models for Stunting Prediction in Jakarta," *J. JTIK (Jurnal Teknol. Inf. dan Komunikasi)*, vol. 9, no. 4, pp. 1365–1375, 2025, doi: 10.35870/jtik.v9i4.3853.

[8] I. Sofyan Iryad, M. Qamal, and A. Razi, "Classification For Determining Nutritional Status of Toddlers Using Random Forest Method at Tanah Pasir Primary Health Centre, North Aceh", *JAIC*, vol. 9, no. 6, pp. 3312–3321, Dec. 2025, doi: 10.30871/jaic.v9i6.10855.

[9] S. Abrori and Z. Fatah, "Implementasi Metode CNN Untuk Klasifikasi Status Stunting Pada Balita ". Gudang Jurnal Multidisiplin Ilmu, vol. 2, no. 10, Oct. 2024, pp. 380-5, doi:10.59435/gjmi.v2i10.1022.

[10] H. Mulyani, M. Musawarman, R. Faturrohman, and D. H. Permana, "Machine Learning-Based Early Detection of Stunting and Intervention Recommendations", bit-Tech, vol. 8, no. 2, pp. 2160–2170, Dec. 2025, doi: 10.32877/bt.v8i2.3213.

[11] Mardiana, N. A., & Windari, W. O., "Penerapan metode SMED dalam peningkatan efisiensi proses manufaktur," *G-Tech: Jurnal Teknologi Terapan*, vol 8, no. 1, pp. 186–195, 2024.

[12] S. A. Wicaksono, S. H. Wijoyo, Fatmawati, T. Afirianto, D. Kurnianingtyas, and M. C. Saputra, "Naive Bayes Analysis for Nutritional Fulfillment Prediction in Children," *J. Appl. Eng. Technol. Sci.*, vol. 6, no. 2, pp. 1135–1147, 2025, doi: 10.37385/jaets.v6i2.6105.

[13] A. Nugraheni, R. D. Ramadhani, A. B. Arifa, and A. Prasetiadi, "Perbandingan Performa Antara Algoritma Naive Bayes Dan K-Nearest Neighbour Pada Klasifikasi Kanker Payudara. Journal of Di," *J. Dinda Data Sci. Inf. Technol. Data Anal.*, vol. 2, no. 1, pp. 11–20, 2022.

[14] M. Khabir, A. A. Jabbari, and M. H. Razmi, "Flipped Presentation of Authentic Audio-Visual Materials: Impacts on Intercultural Sensitivity and Intercultural Effectiveness in an EFL Context," *Front. Psychol*, vol. 13, pp. 1–10, Feb. 2022, doi: 10.3389/fpsyg.2022.832862.

[15] D. P. Sari, S. Widodo, and K. Mustofa, "Development of an Image-Based Calorie Detection Model in Indonesian Food for Stunting Prevention," *Proceeding Int. Conf. Sci. Heal. Technol.*, pp. 468–480, 2025, doi: 10.47701/6d407123.

[16] E. Rianti, I. Fitri, Sumijan, and F. F. Yani, "Development of GLCM Method in Calculate Entropy Value for Digital Visualization in Identifying Childhood Pneumonia Based on Chest X-Ray Images," *Int. J. Online Biomed. Eng.*, vol. 21, no. 2, pp. 137–156, 2025, doi: 10.3991/ijoe.v21i02.52909.

[17] R. F. H. Pasaribu, M. Zarlis, and E. B. Nababan, "Performance Level Analysis On Learning Vector Quantization And Cohonen Algorithms," *Sinkron*, vol. 9, no. 1, pp. 267–282, 2025, doi: 10.33395/sinkron.v9i1.14313.

[18] Pradeep M and Dr. M Siddappa, "Classification of Rice Using Convolutional Neural Network (Cnn)," *Int. J. Eng. Technol. Manag. Sci.*, vol. 7, no. 5, pp. 455–463, 2023, doi: 10.46647/ijetms.2023.v07i05.056.

[19] A. Nur Sahid and D. R. Cahyadi, "Image Classification Using MobileNet Based on CNN Architecture for Grape Leaf Disease Detection," *J. Intell. Syst. Technol. Informatics*, vol. 1, no. 1, pp. 15–21, 2025, doi: 10.64878/jistics.v1i1.7.

[20] A. S. Musliman, A. Fadlil, and A. Yudhana, "Identification of White Blood Cells Using Machine Learning Classification Based on

Feature Extraction," *J. Online Inform.*, vol. 6, no. 1, pp. 63–72, 2021, doi: 10.15575/join.v6i1.704.

[21] M. R. Syahputra *et al.*, "Accurate Skin Tone Classification for Foundation Shade Matching using GLCM Features-K-Nearest Neighbor Algorithm," *J. Tek. Inform.*, vol. 6, no. 5, pp. 3558–3571, 2025, doi: 10.52436/1.jutif.2025.6.5.4723.

[22] S. M. Javidan, A. Banakar, K. Rahnama, K. A. Vakilian, and Y. Ampatzidis, "Feature engineering to identify plant diseases using image processing and artificial intelligence: A comprehensive review," *Smart Agric. Technol.*, vol. 8, no. May, p. 100480, 2024, doi: 10.1016/j.atech.2024.100480.

[23] D. S. Darmawan and E. Sediyono, "Perancangan Aplikasi Sistem Informasi Gelanggang Olahraga Berbasis Web Menggunakan Framework Laravel," *J. Tek. Inform. dan Sist. Inf.*, vol. 10, no. 1, pp. 51–62, 2023, doi: 10.35957/jatisi.v10i1.2587.

[24] P. F. Johari, N. Arifin, M. Muzaki, and M. S. A. Utama, "Corn Leaf Diseases Classification Using CNN with GLCM, HSV, and L*a*b* Features," *J. Tek. Inform.*, vol. 6, no. 2, pp. 709–722, 2025, doi: 10.52436/1.jutif.2025.6.2.4345.

[25] Y. Yunidar, Y. Yusni, N. Nasaruddin, and F. Arnia, "CNN Performance Improvement for Classifying Stunted Facial Images Using Early Stopping Approach," *J. RESTI*, vol. 9, no. 1, pp. 62–68, 2025, doi: 10.29207/resti.v9i1.6068.

[26] A. Bahtiar, M. Mulyawan, A. Faqih, L. Lidina, and A. R. Fitria, "1D-CNN-Based Childhood Stunting Prediction through Socio-Economic Data Integration and Community Participation," *JISA(Jurnal Inform. dan Sains)*, vol. 8, no. 2, pp. 162–169, 2025, doi: 10.31326/jisa.v8i2.2490.

[27] M. Mahendran, R. Visalakshi, and S. Balaji, "Dysarthria detection using convolution neural network," *Meas. Sensors*, vol. 30, no. 05, pp. 11–16, 2023, doi: 10.1016/j.measen.2023.100913.

[28] T. Pujitha, K. Rohitha, K. Nalini, G. Madhulatha, B. Sudharani, and J. Dar, "High-Performance Computing-Based Brain Tumor Detection Using Parallel Quantum Dilated Convolutional Neural Network," vol. 3, pp. 836–843, 2025, doi: 10.5220/0013906600004919.

[29] S. Aanjankumar *et al.*, "Prediction of malnutrition in kids by integrating ResNet-50-based deep learning technique using facial images," *Sci. Rep.*, vol. 15, no. 1, pp. 1–26, 2025, doi: 10.1038/s41598-025-91825-z.

[30] A. Husaini, I. Hoeronis, H. H. Lumana, and L. D. Puspareni, "Early Detection of Stunting in Toddlers Based on Ensemble Machine Learning in Purbaratu Tasikmalaya," *J. Sist. dan Teknol. Inf.*, vol. 11, no. 3, p. 487, 2023, doi: 10.26418/justin.v11i3.66465.

[31] N. P. Sari, "Analisis Performa Algoritma CNN dalam Klasifikasi Citra Medis Berbasis Deep Learning," *J. Komput.*, vol. 2, no. 2, pp. 87–92, 2024, doi: 10.70963/jk.v2i2.113.

[32] H. Yi, "A Review of Convolutional Neural Networks in Cancer Image Classification," *Appl. Comput. Eng.*, vol. 97, no. 1, pp. 69–74, 2024, doi: 10.54254/2755-2721/97/20241334.

[33] M. Caldwell, "Research on Medical Image Diagnosis Models Based on Convolutional Neural Networks," *J. Comput. Technol. Softw.*, vol. 4, no. 2, p. 2025, 2025, doi: 10.5281/zenodo.14984979.

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