CLaGAtt: A Hybrid CNN-LSTM-GRU-Attention Model for Stunting Classification Based on Anthropometric Sequences

Sofiansyah Fadli[1,*], Ahmad Tantoni[2], Novia Arista[3], M. Khairul Anam[4], Muhammad Bambang Firdaus[5]

[1,2]*STMIK Lombok, Jl. Basuki Rahmat, Lombok Tengah and 83511, Indonesia*

[3]*Universitas Nahdatul Ulama Nusa Tenggara Barat, Jl. Pendidikan No. 06, Mataram and 83125, Indonesia*

[4]*Universitas Samudra, Jl. Prof. Dr. Syarief Thayeb, Langsa and 24416, Indonesia*

[5]*Institut Teknologi Sepuluh Nopember, Jl. Raya ITS, Surabaya and 60111, Indonesia*

(Received: February 9, 2026; Revised: April 15, 2026; Accepted: July 15, 2026; Available online: August 9, 2026)

Abstract

Stunting is a chronic nutritional problem that requires accurate early identification because it affects child growth, cognitive development, and long-term human capital. This study adapts the CLaGAtt model, a hybrid CNN–LSTM–GRU–Attention architecture, for stunting classification using anthropometric sequence data. Rather than proposing a new deep learning architecture, the main contribution of this study lies in adapting the existing CLaGAtt framework through an integrated preprocessing pipeline, sequence construction strategy, class balancing using SMOTE, and an evaluation protocol specifically designed for stunting prediction. The preprocessing pipeline included irrelevant-column removal, data transformation, label encoding, standard scaling, class balancing using SMOTE, and sequence generation with a time step of five and a step of one. Three train–test split scenarios were evaluated, namely 90:10, 80:20, and 70:30. Experimental results showed that the 90:10 split produced the best performance, with 91.42% accuracy, 91.50% precision, 91.50% recall, and 91.43% F1-score. The 80:20 and 70:30 scenarios achieved 87.14% and 85.71% accuracy, respectively, indicating that larger training proportions improved model generalization in the available dataset. These findings suggest that the adapted CLaGAtt framework can effectively integrate convolutional feature extraction, sequential learning, and temporal attention for stunting classification from structured anthropometric data. Future work should validate the model on external datasets and integrate regional visualization to support priority intervention mapping.

Keywords: Stunting, CNN, LSTM, GRU, Attention

1. Introduction

Stunting is a form of chronic malnutrition in early life that affects the quality of human capital. Its impact is not limited to physical growth, but also extends to cognitive ability, educational attainment, productivity in adulthood, and long-term health risks [1], [2], [3], [4]. In Indonesia, reducing stunting has become a priority agenda, but challenges remain due to uneven prevalence across regions and the fact that some areas still exhibit high-risk pockets [5], [6], [7], [8]. These conditions indicate the need for accurate classification models that can support early identification of stunting risk. Although regional hotspot mapping and intervention-priority visualization are important for public health planning, these components are not implemented in the present experiment and are positioned as future development directions.

Monitoring of stunting in primary healthcare services generally relies on anthropometric data such as age, gender, weight, and height or length. Such data is often available through Posyandu, Puskesmas, or nutrition reporting systems; however, the quality of record-keeping, consistency of formats, and the utilization of longitudinal data remain challenges. Descriptive analysis can aid reporting, but is not always sufficient to support rapid and consistent risk identification. Consequently, machine learning and deep learning models are increasingly being used to detect stunting status or risk from anthropometric data [9], [10], [11], [12], [13], [14], [15]. Previous research has shown mixed results.

***Corresponding author: Sofiansyah Fadli (sofiansyah182@gmail.com)**
DOI: https://doi.org/10.47738/jads.v7i3.1454

An LSTM model with SMOTE was reported to achieve an accuracy of 85.79%, whilst a Streamlit-based dashboard approach using a BiLSTM model and SMOTE reported very high performance on certain datasets [13], [16]. Another study used Deep Neural Networks for the classification of stunting in infants with an accuracy of around 83.85%, whilst LSTM optimization using a genetic algorithm achieved an accuracy of around 91.10% [10], [17]. Research based on SVM, Random Forest, Decision Tree, and hybrid approaches also demonstrates strong potential in the classification and risk mapping of stunting [5], [12], [18]. Nevertheless, the very high performance observed in some studies should be interpreted with caution, as external validation, class distribution, and cross-regional generalization have not always been discussed in depth.

This study presents an adaptation of the CLaGAtt architecture for stunting classification using structured anthropometric data transformed into fixed-length sequential representations. Unlike the original CLaGAtt design, which was developed for general sequential learning tasks, the proposed adaptation is specifically tailored to healthcare classification by integrating sequence construction from anthropometric records, class balancing using SMOTE, and a hybrid CNN–LSTM–GRU–Attention architecture. CNN is employed to capture local feature interactions among anthropometric variables, while the parallel LSTM and GRU branches learn complementary sequential dependencies from the constructed sequences. The attention mechanism further emphasizes the most informative sequence representations before classification. Therefore, the scope of this study is limited to evaluating the classification performance of the adapted CLaGAtt model. Hotspot mapping, regional intervention prioritization, and dashboard development are not part of the current experimental implementation, but may be integrated in future work using the model outputs as decision-support inputs.

2. Literature Review

Stunting classification has become an important research topic in health informatics because early identification can support preventive intervention and child growth monitoring. Several previous studies have applied computational methods to detect stunting risk using anthropometric and health-related variables. Husaini et al. [2] developed an ensemble machine learning approach for early detection of stunting in toddlers, showing that machine learning can assist health workers in identifying potential stunting cases more systematically. Kelvin et al. [4] also explored stunting classification using a big data approach with Support Vector Regression, indicating that computational models can be applied to health datasets collected from clinical or hospital environments.

Machine learning-based stunting detection has also been studied in mobile and decision-support contexts. Sabilillah et al. [9] compared several machine learning algorithms for stunting detection in the Centing mobile application, demonstrating that classification models can be integrated into digital health applications. Mulyani et al. [12] further emphasized the role of machine learning not only for early detection but also for intervention recommendation. These studies suggest that stunting prediction models should not only focus on classification accuracy, but also consider their practical use as screening and decision-support tools.

Several studies have used conventional machine learning algorithms for stunting prediction. Reza and Rohman [11] implemented Random Forest with random search optimization to improve stunting prediction performance. Aziz et al. [19] applied Learning Vector Quantization for classifying families at risk of stunting, showing that vector-based learning methods can also be used for structured stunting data. Meanwhile, Shen et al. [14] investigated machine learning algorithms for predicting stunting among under-five children in Papua New Guinea, indicating that stunting prediction is not only relevant in Indonesia but also in broader public health contexts. However, conventional machine learning models generally treat input data as static tabular features and may not fully capture sequential patterns in anthropometric records.

Deep learning approaches have been introduced to address more complex representation learning in stunting data. Lestari et al. [20] proposed a Deep Neural Network model for multiclass stunting prediction, showing that multilayer neural networks can learn nonlinear relationships among health-related attributes. However, deep learning models require careful preprocessing, sufficient data, and appropriate validation strategies to avoid overfitting. Therefore, the use of deep learning for stunting classification should be accompanied by evaluation metrics beyond accuracy, such as precision, recall, F1-score, and confusion matrix analysis.

Class imbalance is another important issue in health classification tasks. In many medical datasets, minority classes may represent clinically important cases, including children at risk of stunting. If the imbalance is not handled properly, the model may produce biased predictions toward the majority class. Several studies in machine learning and health-related classification have applied oversampling strategies such as SMOTE to improve minority-class recognition [21], [22]. Although these references are not all specific to stunting, they provide methodological support for applying class balancing in predictive modeling, especially when the goal is to improve the model's sensitivity to underrepresented classes.

Hybrid deep learning architectures have also shown potential in sequential and structured data classification. CNN-based components are commonly used to extract local feature patterns, while recurrent layers such as LSTM and GRU are used to model sequential dependencies. Previous research has shown that CNN-LSTM-based models can improve representation learning in prediction tasks [23], [24]. GRU-based models have also been used because they provide a simpler recurrent structure with fewer parameters than LSTM, making them computationally efficient for sequence modeling [25]. In addition, attention mechanisms have been widely applied to improve model focus on the most informative features or time steps [26]. These methodological developments support the design of a hybrid architecture that combines CNN, LSTM, GRU, and attention mechanisms.

Based on the reviewed studies, several research gaps can be identified. First, many stunting classification studies still rely on conventional machine learning models that focus on tabular classification without explicitly modeling sequential representations. Second, although deep learning has been applied in stunting prediction, the integration of local feature extraction, recurrent learning, and attention-based weighting remains limited. Third, some studies report high accuracy, but their generalization capability m ay still depend on dataset characteristics, validation strategy, and class distribution. Therefore, this study proposes the CLaGAtt model as a hybrid CNN-LSTM-GRU-Attention architecture for stunting classification based on anthropometric sequences.

Although previous studies have demonstrated the effectiveness of machine learning and deep learning for stunting prediction, several methodological limitations remain. Most existing studies process anthropometric variables as independent tabular observations without explicitly modeling sequential relationships among measurements. Furthermore, only a limited number of studies investigate hybrid architectures that simultaneously combine convolutional feature extraction, recurrent sequence modeling, and attention mechanisms within a single framework. Existing studies also rarely provide sufficient methodological details regarding sequence generation, class balancing procedures, and training configuration, making reproducibility difficult. These limitations motivate the present study to adapt the CLaGAtt architecture specifically for anthropometric sequence classification while providing a more transparent experimental pipeline. To position this study more clearly, table 1 summarizes several related studies that are relevant to stunting classification and methodological development.

Table 1. Summary of Related Studies on Stunting Classification

Study	Method	Research Focus	Relevance to This Study
[2]	Ensemble Machine Learning	Early detection of stunting in toddlers	Supports the use of machine learning for stunting screening
[4]	Support Vector Regression	Big data-based stunting classification	Shows the use of computational modeling in clinical stunting data
[5]	Hybrid Machine Learning	Classification, prediction, and clustering of stunting prevalence	Supports hybrid approaches for stunting analysis
[9]	Machine Learning Comparison	Stunting detection for mobile application	Shows the potential integration of ML into digital health systems
[20]	Deep Neural Network	Multiclass stunting prediction	Provides deep learning baseline for stunting classification
[11]	Random Forest + Random Search	Optimized stunting prediction	Demonstrates the role of optimization in improving model performance
[19]	Learning Vector Quantization	Classification of families at risk of stunting	Provides a conventional classification baseline

[12]	Machine Learning	Early detection and intervention recommendation	Connects prediction results with practical intervention support
[14]	Machine Learning Algorithms	Prediction of stunting among under-five children	Shows cross-context relevance of ML-based stunting prediction
This Study	CLaGAtt	CNN-LSTM-GRU-Attention for anthropometric sequences	Integrates local feature extraction, sequential learning, and attention

Overall, previous studies confirm that machine learning and deep learning methods are relevant for stunting classification. However, existing approaches still vary in terms of dataset structure, preprocessing method, model architecture, validation protocol, and performance interpretation. The proposed CLaGAtt model contributes by combining CNN for local feature extraction, LSTM and GRU for sequential representation learning, and attention for emphasizing informative features. This hybrid design is expected to provide a stronger representation of anthropometric sequences and support more reliable stunting classification.

3. Methodology

This study employs a computational experimental approach to evaluate the proposed CLaGAtt model on stunting classification data. The overall research workflow is illustrated in figure 1. In general, the process begins with data preparation and preprocessing, followed by feature extraction using Convolutional Neural Network (CNN), sequence modeling using parallel Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), attention mechanism processing, and final classification through the output layer. As illustrated in figure 1, this workflow was designed to capture both local feature patterns and sequential dependencies within the stunting dataset.

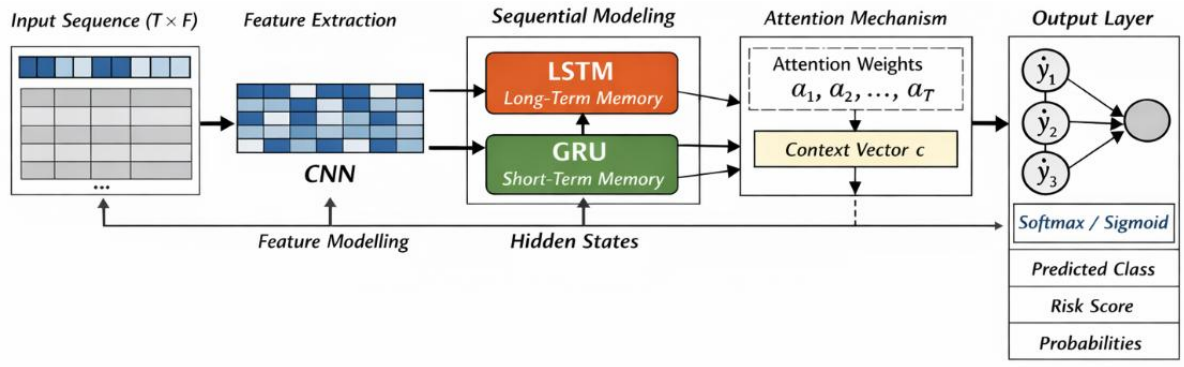

Figure 1. Experimental Workflow of the CLaGAtt Model

3.1. Dataset and Preprocessing

The dataset used in this study consists of structured anthropometric records for binary stunting classification. The data contain demographic and anthropometric variables commonly used in child growth assessment, including age, sex, body weight, body height or length, and additional nutritional attributes. The target variable consists of two classes representing stunting and non-stunting status based on anthropometric assessment. The original dataset was first subjected to data quality assessment before preprocessing. Duplicate records, incomplete observations, and irrelevant attributes were identified and removed to ensure data consistency and reliability.

Prior to model development, several preprocessing steps were performed, including irrelevant-column removal, data transformation, label encoding, and feature standardization using StandardScaler. These steps were conducted to improve data quality, reduce inconsistencies among variables, and ensure comparable feature scales during model optimization [23], [27]. Because the original dataset exhibited class imbalance, the Synthetic Minority Oversampling Technique (SMOTE) was applied to the training data to improve the representation of the minority class while preventing the evaluation results from being biased toward the majority class. The validation and testing datasets were retained for performance evaluation without additional resampling. SMOTE has been widely used in imbalanced classification tasks to improve minority-class recognition and reduce prediction bias toward the majority class.

After preprocessing, the structured anthropometric records were transformed into fixed-length sequential representations using a sliding-window strategy with a time step of five and a step size of one. This transformation

enables the proposed CLaGAtt architecture to capture local feature interactions through the CNN layer and sequential dependencies through the parallel LSTM and GRU layers before the attention mechanism identifies the most informative sequence representations.

To avoid temporal inconsistency and potential data leakage, the balancing procedure was performed after sequence construction and data partitioning. The generated sequences were first divided into training, validation, and testing sets. SMOTE was then applied only to the training sequences, while the validation and testing sequences were kept unchanged. Applying SMOTE before sequence construction may introduce artificial temporal patterns because synthetic tabular records could be combined into sequences that do not reflect realistic anthropometric progression. Therefore, oversampling was performed at the sequence-level training data to preserve temporal consistency in the validation and testing sets and to ensure unbiased model evaluation [28], [29], [21], [30]. The distribution of the processed dataset across the three experimental split scenarios is presented in table 2.

Table 2. Dataset Distribution After Sequence Construction and SMOTE

Split	Train 0	Train 1	Val. 0	Val. 1	Test 0	Test 1
90:10	314	312	18	17	18	17
80:20	279	277	35	35	35	35
70:30	244	243	52	52	53	52

3.2. Sequence Construction

After preprocessing, the structured anthropometric records were transformed into fixed-length sequential representations using a sliding-window strategy. A time step of five and a step size of one were used to construct the input sequences. The time step of five was selected to provide sufficient contextual information for the recurrent layers while maintaining a compact input representation suitable for the available dataset size. A shorter sequence may provide limited contextual patterns, whereas a longer sequence may reduce the number of generated samples and increase the risk of overfitting.

To avoid temporal inconsistency and potential data leakage, sequence construction was performed before data balancing. The generated sequences were first divided into training, validation, and testing sets. SMOTE was then applied only to the training sequences, while the validation and testing sequences were kept unchanged. This procedure was used to prevent synthetic samples from influencing the evaluation data and to preserve temporal consistency.

A potential leakage risk may still occur if measurements from the same child appear in both training and testing sets. Therefore, future experiments should apply subject-level splitting when child identifiers are available. In this study, alternative sequence lengths were not evaluated; thus, the selected time step should be interpreted as an initial experimental configuration rather than an optimized sequence length. Future work should conduct sensitivity analysis using different sequence lengths, such as three, five, seven, and ten time steps.

3.3. CLaGAtt Architecture

The proposed CLaGAtt architecture combines Convolutional Neural Network (CNN), Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Attention mechanisms into a hybrid deep learning framework. The architecture begins with a sequence-based input layer consisting of stunting-related health data. The input data are first processed using a Conv1D layer to extract local feature representations and important spatial patterns from the dataset [31].

The extracted features are then forwarded into two parallel recurrent branches, namely LSTM and GRU. The LSTM branch is used to capture long-term sequential dependencies [24], [32], while the GRU branch provides a more computationally efficient recurrent mechanism for learning sequential information [25]. The outputs from both recurrent layers are subsequently combined using a concatenation layer.

An Attention mechanism is then applied to highlight the most relevant sequential features generated by the recurrent layers [26]. This mechanism enables the model to focus more effectively on informative patterns that contribute significantly to stunting classification. The attention output is subsequently processed using global average pooling,

 ISSN 2723-6471

followed by a dense layer and dropout regularization before generating the final classification result through the output layer. The mathematical formulation of the proposed CLaGAtt architecture is described as follows. The input sequence is represented as:

$$X = \{X_1, X_2, \ldots, X_T\}, x_t \in \mathbb{R}^F \quad (1)$$

X denotes one input sequence, *T* represents the number of time steps, and *F* represents the number of features at each time step. In this study, each sequence consists of anthropometric features arranged into a fixed-length window. Therefore, x_t represents the feature vector at time step *t*.

The CNN layer performs feature extraction as follows:

$$Z = \sigma(Conv1D(X; W_c, b_c)) \quad (2)$$

W_c denotes the convolution kernel weights, b_c is the bias term, and σ(.) represents the activation function, such as ReLU. This layer is used to capture local interactions among anthropometric features within the constructed sequence. The output *Z={z1, z2,...,zT}* represents the extracted feature sequence produced by the convolutional layer. The extracted CNN features are processed in parallel using LSTM and GRU:

$$h_t^L = LSTM(z_t, h_{t-1}^L)$$
$$h_t^G = GRU(z_t, h_{t-1}^G) \quad (3)$$

h_t^L and h_t^G denote the hidden representations generated by the LSTM and GRU branches at time step t, respectively. The previous hidden states h_{t-1}^L and h_{t-1}^G allow both recurrent layers to learn sequential dependencies from earlier time steps. The LSTM branch is intended to capture longer-term dependencies, while the GRU branch provides a more compact recurrent representation. The outputs of both recurrent layers are concatenated to combine sequential representations:

$$h_t = [h_t^L; h_t^G] \quad (4)$$

[;] denotes the concatenation operation. This operation combines the complementary information learned by the LSTM and GRU branches into a single representation h_t for each time step. The temporal attention mechanism calculates the relevance score for each time step:

$$e_t = v^T \tanh(W_a h_t + b_a)$$
$$\alpha_t = \frac{\exp(e_t)}{\sum_{k=1}^{T} \exp(e_t)} \quad (5)$$

$W_a, h_t,$ and v are trainable attention parameters. The value e_t represents the relevance score of the hidden representation at time step *t*. The softmax function converts these scores into normalized attention weights α_t, where higher values indicate that the corresponding time step contributes more strongly to the classification decision.

In this study, the attention mechanism was implemented as additive temporal attention after the concatenation of the LSTM and GRU outputs. Additive attention was selected because it learns a trainable scoring function to measure the relevance of each time-step representation. The relevance score is computed using a feed-forward transformation followed by a non-linear activation, as shown in Equation (5). Therefore, the attention mechanism used in this model is not multiplicative attention or self-attention, but an additive attention mechanism applied to the recurrent sequence representations. The attention layer assigns different weights to the sequential representations generated by the recurrent branches, allowing the model to emphasize more informative time steps before global average pooling and final classification. Thus, the attention mechanism functions as a feature-weighting component rather than as a separate classifier. The context vector is computed as:

$$c = \sum_{t=1}^{T} \alpha_t h_t \quad (6)$$

c represents the weighted summary of all sequential representations. Each hidden representation h_t is multiplied by its corresponding attention weight α_t, so the final context vector gives greater influence to more informative time steps. The final binary classification output is generated using the sigmoid activation function:

$$\hat{y} = \sigma(W_y\, c + b_y) \quad (7)$$

W_y and b_y denote the output layer weight and bias, while $\hat{y}$ represents the predicted probability of the positive class. Because this study uses binary stunting classification, the sigmoid function maps the output value into a probability range between 0 and 1. The binary cross-entropy loss function used in this study is defined as:

$$\mathcal{L} = -[y \log(\hat{y}) + (1 - y)\log(1 - \hat{y})] \quad (8)$$

y denotes the actual class label and $\hat{y}$ denotes the predicted probability. This loss function penalizes incorrect predictions by increasing the loss when the predicted probability differs from the actual label. Therefore, minimizing this loss during training helps the model improve its binary classification performance. A summary of the proposed CLaGAtt architecture is presented in table 3.

Table 3. Summary of the proposed CLaGAtt architecture

Layer	Output shape	Parameters
Input	5 x 14	0
Conv1D	5 x 64	2,752
Batch normalization	5 x 64	256
Max pooling	2 x 64	0
Dropout	2 x 64	0
LSTM	2 x 64	33,024
GRU	2 x 64	24,960
Concatenate	2 x 128	0
Attention	2 x 128	0
Global average pooling	128	0
Dense + dropout	64	8,256
Output dense	1	65

Based on table 3, the proposed architecture integrates multiple deep learning components into a unified framework. The LSTM and GRU layers contribute the largest number of trainable parameters because they are responsible for learning sequential dependencies from the input data. Meanwhile, the attention mechanism enhances the model's ability to emphasize important sequential representations that contribute to the final classification decision.

The proposed CLaGAtt model should therefore be interpreted as an adaptation of the hybrid CNN–LSTM–GRU–Attention architecture rather than as an entirely new deep learning model. The main novelty of this work lies in adapting this architecture to anthropometric sequence-based stunting classification through an integrated preprocessing pipeline, sequence construction strategy, class balancing procedure, and evaluation protocol specifically designed for the stunting prediction task.

3.4.Training Configuration

To improve reproducibility, all experimental scenarios were trained using the same configuration. The proposed CLaGAtt model was optimized using the Adam optimizer with a learning rate of 0.001. Binary cross-entropy was used as the loss function because the task was formulated as binary stunting classification. The model was trained using a batch size of 32 for a maximum of 100 epochs. Early stopping was applied by monitoring validation loss, with a patience value of 10 epochs, and the best model weights were restored for final evaluation. A fixed random seed was used during data splitting, model initialization, and training to reduce experimental variability and ensure more consistent results across runs.

3.5. Evaluation Setting

The evaluation process was conducted to measure the performance of the proposed CLaGAtt model in classifying stunting data. Several evaluation metrics were used, including accuracy, precision, recall, and F1-score. Accuracy measures the overall proportion of correct predictions, while precision and recall evaluate the balance between false positive and false negative predictions. The F1-score was used as a combined metric to assess classification

performance more comprehensively. In addition to these metrics, confusion matrix analysis was employed to observe the distribution of prediction errors across classes. The evaluation results were then analyzed to determine the effectiveness of the proposed hybrid architecture in supporting stunting classification.

The evaluation was conducted using three train–test split scenarios, namely 90:10, 80:20, and 70:30. These scenarios were selected to observe how different proportions of training and testing data influence the performance of the adapted CLaGAtt model. However, this validation strategy should be interpreted as an internal split-based evaluation rather than a full robustness assessment. Cross-validation was not applied in the present experiment; therefore, the reported results may still be influenced by data partitioning. Future studies should employ stratified k-fold cross-validation or external validation datasets to provide stronger evidence of model robustness and generalization.

4. Results and Discussion

4.1. Overall Performance

Table 4 presents the performance comparison of the proposed CLaGAtt model across three data split scenarios, namely 90:10, 80:20, and 70:30. The evaluation was conducted using accuracy, precision, recall, and F1-score. Based on the results, the 90:10 split achieved the highest observed performance, with an accuracy of 91.42%, precision of 91.50%, recall of 91.50%, and F1-score of 91.43%. This result is likely influenced by the larger amount of training data available in the 90:10 scenario, which allows the model to learn more patterns from the dataset.

The 80:20 split scenario achieved an accuracy of 87.14%, precision of 87.92%, recall of 87.14%, and F1-score of 87.08%. Meanwhile, the 70:30 split scenario produced the lowest observed performance, with an accuracy of 85.71%, precision of 86.06%, recall of 85.67%, and F1-score of 85.67%. As shown in table 4, the decrease in performance across the different data split scenarios suggests that reducing the amount of training data may affect the model's ability to learn sequential patterns effectively.

Table 4. Performance comparison across split scenarios

Split	Accuracy	Precision	Recall	F1-score
90:10	91.42%	91.50%	91.50%	91.43%
80:20	87.14%	87.92%	87.14%	87.08%
70:30	85.71%	86.06%	85.67%	85.67%

The performance difference between the 90:10 and 70:30 scenarios reached 5.71 percentage points in accuracy and 5.76 percentage points in F1-score. However, this finding should not be interpreted as evidence of stronger generalization because the 90:10 scenario uses a smaller testing set and may not fully represent broader data variation. Therefore, the 90:10 result indicates better internal split-based performance, while further validation using cross-validation or external datasets is required to assess generalization more rigorously.

Figure 2 illustrates the comparison of classification metrics across the three split scenarios. The bar chart shows that the 90:10 split consistently achieved the highest values for accuracy, precision, recall, and F1-score. The metric values gradually decreased in the 80:20 and 70:30 scenarios, reflecting the influence of reduced training data on model performance within the internal split-based evaluation.

It should also be noted that this study does not include statistical significance testing or confidence interval estimation. Consequently, the observed performance differences among the three train–test split scenarios may be influenced by sampling variability associated with different data partitions. The differences shown in figure 2 should be interpreted as descriptive comparisons and not as evidence of statistically significant differences.

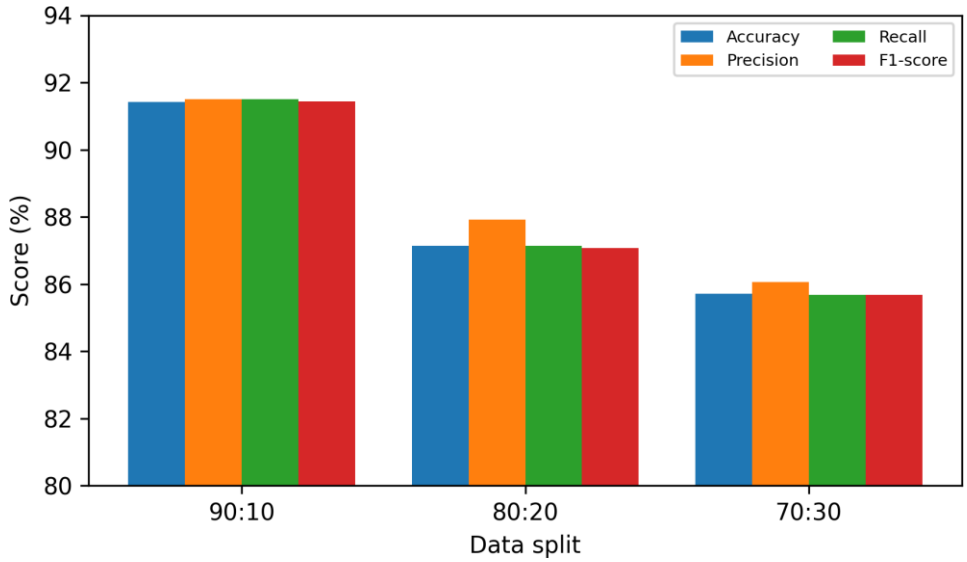

Figure 2. Comparison of classification metrics across data splits

The visualization also demonstrates that the differences among the four evaluation metrics are relatively small within each split scenario. This indicates that the proposed CLaGAtt model produces balanced classification performance without significant discrepancies between precision and recall. Such balanced performance suggests that the model is capable of minimizing both false positive and false negative predictions effectively.

4.2. Configuration Matrix Analysis

The confusion matrix analysis shows that the 90:10 scenario produced 32 correct predictions out of 35 test samples, comprising 16 correct predictions for class 0 and 16 correct predictions for class 1. Misclassifications consisted of one class 0 sample predicted as class 1 and two class 1 samples predicted as class 0. In the 80:20 scenario, the model correctly classified 61 out of 70 test samples, while the 70:30 scenario correctly classified 90 out of 105 test samples. However, the confusion matrix also indicates that the testing set is relatively small, particularly in the 90:10 scenario, which contains only 35 test samples. Consequently, the reported performance metrics for this split should be interpreted with caution, as a small number of misclassifications can noticeably influence the calculated accuracy, precision, recall, and F1-score. Therefore, the results are best regarded as an initial internal validation rather than conclusive evidence of the model's robustness. A summary of the confusion matrix is presented in table 5.

In the context of stunting detection, false negative errors require particular attention as they may result in at-risk children not being prioritized for follow-up. Therefore, recall is an important metric alongside accuracy. As shown in table 5, the 90:10 data split achieved the highest aggregate recall, whereas the 70:30 split exhibited a lower recall for Class 1, suggesting that additional training data or adjustment of the classification threshold may further improve the model's performance.

Table 5. Confusion matrix summary

Split	True 0 Pred. 0	True 0 Pred. 1	True 1 Pred. 0	True 1 Pred. 1
90:10	16	1	2	16
80:20	33	2	7	28
70:30	48	5	10	42

Based on table 5, the 90:10 split scenario produced the most balanced classification result, with relatively low misclassification in both classes. The 80:20 scenario showed an increase in false negative errors, where seven class 1 samples were incorrectly classified as class 0. Meanwhile, the 70:30 scenario produced the highest number of false negative cases, with ten class 1 samples misclassified as class 0. This indicates that reducing the proportion of training data may affect the model’s ability to detect class 1 patterns. Therefore, the 90:10 split can be considered the most effective scenario for the proposed CLaGAtt model in this experiment.

4.3. Discussion

The findings suggest that the integrated CNN, LSTM, GRU, and attention architecture is relevant for stunting classification based on anthropometric sequences. CNN is designed to capture local feature interactions, LSTM and GRU are used to model sequential representations, and the attention mechanism is intended to emphasize informative time-step representations. However, this study did not conduct an ablation analysis by removing individual components from the architecture. Therefore, the specific contribution of each module to the final performance cannot be determined conclusively from the current experiment.

When compared with previous studies, the best result of 91.42% provides a general indication that the proposed approach performs competitively within the existing literature. For example, this result is higher than the DNN-based study reporting 83.85% accuracy and is close to the LSTM model optimized using a genetic algorithm, which reported approximately 91.10% accuracy [12]. However, this comparison should not be interpreted as a direct superiority claim because previous studies used different datasets, validation designs, sample sizes, class distributions, and preprocessing procedures. Several studies reported very high accuracy, for example above 98%, but these results also need to be interpreted carefully in relation to their validation design, data size, class distribution, and potential overfitting risk [11], [14], [15]. Therefore, the current result should be viewed as preliminary internal evidence rather than definitive proof that the proposed architecture outperforms previous methods.

A limitation of this study is that the evaluation is still based on internal data and has not yet undergone external validation across regions or healthcare facilities. Furthermore, the functions of mapping priority area hotspots and the GUI planned in the initial proposal have not been the primary focus of this experiment's results. Future development should test the model on external data, incorporate contextual variables such as sanitation, parental education, and socio-economic factors, and integrate the prediction results with a dashboard for priority intervention areas.

4.4. Class-Level Interpretation

The class-level evaluation provides deeper insight into the classification behavior of the proposed CLaGAtt model beyond overall accuracy. In the 90:10 split scenario, class 0 achieved a precision of 0.889, recall of 0.941, and F1-score of 0.914, while class 1 achieved a precision of 0.941, recall of 0.889, and F1-score of 0.914. These results indicate that the model produced balanced performance between both classes in the best-performing split scenario. The 80:20 split scenario still showed relatively strong performance; however, the recall value for class 1 decreased to 0.800, indicating an increase in false negative predictions. A similar pattern was observed in the 70:30 scenario, where the recall for class 1 decreased to 0.807.

These findings suggest that evaluating stunting classification using only accuracy is insufficient. From a public health perspective, false negative predictions are particularly critical because children who are actually at risk of stunting may not receive appropriate early intervention. The results presented in table 6 suggest that future implementation should incorporate threshold optimization or cost-sensitive evaluation strategies to improve recall for the at-risk class, which is particularly important in early screening settings.

Table 6. Class-level performance summary

Split	Class	Precision	Recall	F1-score	Support
90:10	0	0.889	0.941	0.914	17
90:10	1	0.941	0.889	0.914	18
80:20	0	0.825	0.943	0.880	35
80:20	1	0.933	0.800	0.862	35
70:30	0	0.828	0.906	0.865	53
70:30	1	0.896	0.807	0.849	52

Based on table 6, the model consistently achieved higher precision values for class 1 compared to class 0 across all split scenarios. However, the recall values for class 1 gradually decreased as the proportion of testing data increased. This trend indicates that the model became less effective in identifying all positive stunting cases when the amount of training data was reduced. Nevertheless, the F1-score values remained relatively stable across all scenarios,

demonstrating that the proposed CLaGAtt architecture maintained balanced classification capability despite variations in data split configurations.

Figure 3 presents the confusion matrix comparison for each split scenario. The visualization illustrates the distribution of correct and incorrect predictions across classes in the 90:10, 80:20, and 70:30 scenarios. In the 90:10 scenario, the confusion matrix shows that most samples were correctly classified, with only a small number of misclassifications between classes. The 80:20 and 70:30 scenarios exhibited a gradual increase in false negative predictions, particularly for class 1, which is consistent with the decrease in recall values reported in figure 3.

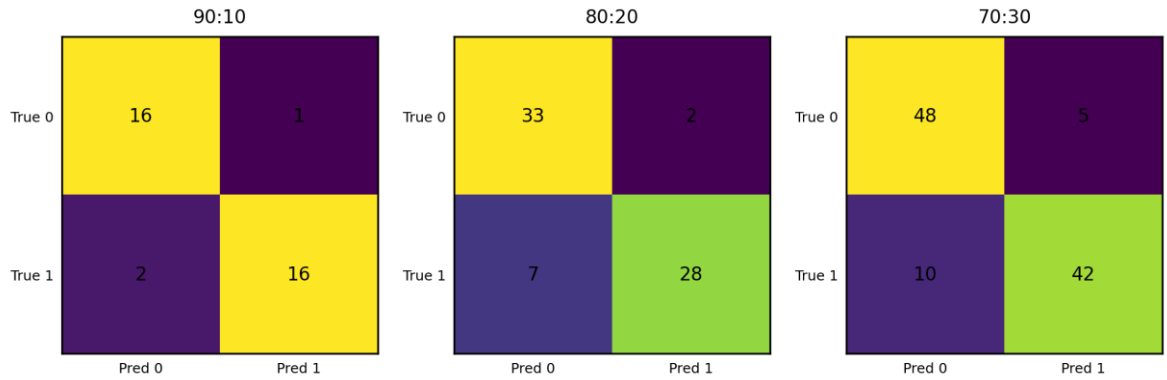

Figure 3. Confusion matrix comparison for each split scenario

4.5. Practical Implications

For implementation, several safeguards are needed. The input data must be standardized, measurement errors should be checked, and missing values should be handled transparently. The system should also store a clear log of predictions and provide interpretable summaries for users. These steps are important because stunting-related decisions may affect intervention allocation and public health planning. However, the present study does not include an explainability analysis to identify which anthropometric variables or temporal patterns contribute most strongly to the model predictions. Therefore, the proposed model should be regarded as a decision-support tool rather than a replacement for clinical judgment. Future work should integrate Explainable Artificial Intelligence (XAI) techniques, such as feature attribution or attention visualization, to improve model transparency, support interpretation by health professionals, and increase trust in practical healthcare applications.

4.6. Limitation and Future Work

This study has several limitations. The evaluation relies on internal train–test split scenarios rather than external validation across independent regions or healthcare facilities. Consequently, the reported performance should be interpreted as an initial experimental result rather than evidence of broad generalization. In addition, the validation strategy is limited to three train–test split scenarios without cross-validation. Although this design provides an initial comparison of model performance under different training and testing proportions, it does not fully assess the robustness of the proposed model across multiple random data partitions. Furthermore, the dataset is limited to structured anthropometric features, whereas stunting risk is also influenced by contextual factors such as sanitation, maternal education, household economic status, infection history, and dietary patterns.

Although the proposed CLaGAtt model shows promising performance for stunting classification, the current evidence is not yet sufficient to claim definitive performance improvement over existing approaches. The present experiment did not conduct statistical significance testing, confidence interval estimation, extensive baseline comparison, or ablation analysis. Consequently, the observed performance differences among the evaluated split scenarios may be influenced by sampling variability and cannot be interpreted as statistically significant differences. Therefore, the reported results should be interpreted as preliminary evidence of feasibility rather than conclusive proof of model superiority.

Future work should evaluate the proposed model using more rigorous validation protocols, such as stratified k-fold cross-validation, repeated holdout validation, leave-region-out validation, or prospective testing on newly collected

data. External validation using datasets from different regions or healthcare facilities is also necessary to assess the model's generalizability. Statistical significance testing and confidence interval estimation should also be incorporated to determine whether observed performance differences are statistically reliable across multiple experimental runs. Fair baseline comparisons using conventional machine learning models, single deep learning models, and reduced variants of CLaGAtt should also be conducted to assess the specific contribution of each architectural component. In addition, ablation analysis is needed by removing or modifying the CNN, LSTM, GRU, and attention components to identify their individual effects on classification performance.

Because this model is intended for healthcare-related screening support, leakage control and explainability should also be prioritized. Future studies should ensure that data splitting, sequence construction, and SMOTE are performed without information leakage between training and testing data. Explainable AI techniques should be incorporated to improve model interpretability and enable health workers to better understand which variables or temporal patterns contribute most significantly to the prediction. Finally, the hotspot mapping module should be developed and validated to ensure that aggregated prediction results can provide meaningful regional intervention priorities.

4.7. Comparison with Previous Studies

The proposed model was compared with several recent studies to provide a general context for its performance rather than to establish a direct superiority claim. Because previous studies employed different datasets, data characteristics, class distributions, validation protocols, and experimental settings, the reported performance values should not be interpreted as results obtained under identical benchmarking conditions. Therefore, the comparison presented in this study is intended to position the proposed approach within the existing literature rather than to provide a strictly controlled performance ranking. Table 7 shows that CLaGAtt achieves better performance than the DNN model reported by Lestari et al. and is close to the LSTM model optimized with genetic algorithm. The result is lower than studies reporting extremely high accuracy, such as BiLSTM+SMOTE and Random Forest with grid search optimization. However, very high accuracy should be interpreted carefully because differences in data source, label construction, class balance, and validation protocol can strongly influence the reported scores.

Compared with single sequential models, CLaGAtt combines local feature extraction and recurrent modeling within a unified architecture. However, differences in reported accuracy across studies should be interpreted with caution because they are influenced by variations in dataset characteristics, sample size, class distribution, preprocessing strategies, validation procedures, and evaluation protocols. Consequently, the comparison presented in table 7 should be regarded as a qualitative comparison of methodological approaches rather than definitive evidence that the proposed model outperforms all previous methods. A fairer comparison would require standardized benchmark datasets and identical experimental settings across all evaluated models.

Table 7. Comparison with related stunting classification studies

Study	Method	Reported result	Main note
[16]	LSTM + SMOTE	85.79% accuracy	Anthropometry data
[13]	BiLSTM + SMOTE	99.43% accuracy	Dashboard and visualization
[17]	DNN	83.85% accuracy	Regularized deep model
[10]	LSTM + GA	91.10% accuracy	Learning-rate optimization
[18]	RF + grid search	99.59% accuracy	Conventional ML baseline
This study	CLaGAtt	91.42% accuracy	CNN-LSTM-GRU-Attention

From the perspective of applied health informatics, the model does not need to be interpreted only as a competition for the highest accuracy. The more important contribution is whether the model can support early screening, provide consistent risk estimation, and become part of a larger workflow for monitoring child growth. Therefore, the next stage should combine model evaluation with usability testing, data governance assessment, and validation with health program stakeholders.

4.8. Reproducibility Consideration

Reproducibility is a critical issue for deep learning studies in health data because small differences in preprocessing can produce different evaluation results. In this experiment, the most influential steps are label encoding, feature standardization, class balancing using SMOTE, and sequence generation. Each step should be documented in the final experimental repository so that future researchers can reproduce the same train-validation-test partitions and verify whether the observed performance is stable.

Another important issue is data leakage. Because the dataset is transformed into sequences, records that originate from the same child or from temporally adjacent measurements should be handled carefully when splitting data. If related samples appear in both training and testing sets, the model may appear to generalize well even though it only learns repeated patterns from the same subject. For that reason, future experiments should use subject-level or region-level splitting when the data structure allows it.

The current report shows promising results but should be considered an internal validation stage. A stronger manuscript would benefit from additional information such as dataset size before and after SMOTE, the exact list of input variables, random seed, optimizer setting, number of epochs, batch size, and early stopping criteria. Providing these details will improve transparency and help reviewers assess whether the proposed model can be fairly compared with previous stunting classification studies. For practical deployment, reproducibility also relates to system maintenance. A deployed screening model must be updated when measurement standards, population characteristics, or data-entry practices change. Therefore, the next version of this work should include a model monitoring plan, periodic recalibration, and a mechanism to flag uncertain predictions for manual review by health workers.

For practical deployment, reproducibility also relates to system maintenance. A deployed screening model must be updated when measurement standards, population characteristics, or data-entry practices change. Therefore, future development should include a model monitoring plan, periodic recalibration, and a mechanism to flag uncertain predictions for manual review by health workers. To facilitate reproducibility, future versions of this work should provide the complete preprocessing workflow, sequence construction algorithm, model hyperparameters, random seed initialization, and training scripts. These implementation details are essential to ensure that the reported performance can be independently verified and fairly compared with future studies.

5. Conclusion

This study adapted the CLaGAtt model for stunting classification based on anthropometric sequences by combining Conv1D, LSTM, GRU, and attention mechanisms. The experimental pipeline included the removal of irrelevant columns, data transformation, label encoding, standardization, sequence formation with a time step of 5 and a step size of 1, and class balancing using SMOTE. The experimental results show that the 90:10 scenario achieved the highest observed performance, with an accuracy of 91.42%, precision of 91.50%, recall of 91.50%, and F1-score of 91.43%. The 80:20 and 70:30 scenarios also yielded performance above 85%, but showed a decline compared to the 90:10 scenario.

These findings suggest that the adapted CLaGAtt model has potential as an initial computational approach for anthropometric data-based stunting classification, particularly when sufficient training data are available and class imbalance is properly handled. The main contribution of this article is to provide preliminary empirical evidence that the integration of local feature extraction, sequential learning, and attention mechanisms is feasible for stunting classification on structured anthropometric data. However, because this study did not compare the proposed architecture with simpler baseline models trained on the same dataset, the results should not be interpreted as conclusive evidence of performance improvement over existing methods.

Further research is needed to validate the model on external datasets, evaluate thresholds to reduce false negatives, and conduct direct baseline comparisons using simpler machine learning and deep learning models on the same dataset. Future work should also develop a priority area mapping dashboard so that prediction results can support more targeted nutritional interventions. However, priority area mapping and dashboard development were not implemented in the present experiment and should be considered future development directions rather than part of the current study scope.

6. Declarations

6.1. Author Contributions

Conceptualization: S.F., and M.K.A.; Methodology: S.F.; Software: A.T.; Validation: S.F., M.K.A., and N.A.; Formal Analysis: S.F., A.T., and N.A.; Investigation: S.F.; Resources: M.B.F.; Data Curation: S.F.; Writing Original Draft Preparation: M.K.A., M.B.F., and N.A.; Writing Review and Editing: M.B.F., N.A., and A.T.; Visualization: A.T.; All authors have read and agreed to the published version of the manuscript.

6.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

6.3. Funding

This research was funded by the Ministry of Higher Education, Science, and Technology of the Republic of Indonesia (Kementerian Pendidikan Tinggi, Sains, dan Teknologi/Kemendiktisaintek) under the Fundamental Regular Research Scheme.

6.4. Institutional Review Board Statement

Not applicable.

6.5. Informed Consent Statement

Not applicable.

6.6. Declaration of Competing Interest

The authors declare that they have no known competing financial interests.

References

[1] A. Saleh, S. Syahrul, V. Hadju, I. Andriani, and I. Restika, "Role of Maternal in Preventing Stunting: a Systematic Review," *Gac. Sanit.*, vol. 35, no. July, pp. S576–S582, 2021, doi: 10.1016/j.gaceta.2021.10.087.

[2] 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," *JustIN*, vol. 11, no. 3, pp. 487-495, 2023, doi: 10.26418/justin.v11i3.66465.

[3] Supranoto, L. Sasmito, S. Indriastuti, H. Prayitno, and M. Sawir, "From parallel to partnership governance: Strengthening institutional synergy for stunting reduction in decentralized Indonesia," *Soc. Sci. Humanit. Open*, vol. 12, no. 1, pp. 1–12, 2025, doi: 10.1016/j.ssaho.2025.102051.

[4] K. Kelvin, R. A. F. Adriansyah, C. Juliandy, F. M. Sinaga, F. Liko, and A. Angkasa, "Classification of Big Data Stunting Using Support Vector Regression Method at Stella Maris Medan Maternity Hospital," *Indones. J. Artif. Intell. Data Min.*, vol. 7, no. 2, pp. 497-509, 2024, doi: 10.24014/ijaidm.v7i2.31112.

[5] N. Hasdyna, R. K. Dinata, Rahmi, and T. I. Fajri, "Hybrid Machine Learning for Stunting Prevalence: A Novel Comprehensive Approach to Its Classification, Prediction, and Clustering Optimization in Aceh, Indonesia," *Informatics*, vol. 11, no. 4, pp. 1–32, 2024, doi: 10.3390/informatics11040089.

[6] V. Hadju and M. Maddeppungeng, "Stunting prevalence and its relationship to birth length of 18–23 months old infants in Indonesia," *Enferm. Clin.*, vol. 30, no. S4, pp. 205–209, 2020, doi: 10.1016/j.enfcli.2019.10.069.

[7] D. L. Lestari, Y. H. Gusmira, M. A'arif, and A. Y. Amelia, "Free Nutritious Meal Policy As A Solution To Overcoming The Stunting Problem In Indonesia," *INNOVATIVE: J. Soc. Sci. Res.*, vol. 4, no. 4, pp. 10021–10031, 2024, doi: 10.31004/innovative.v4i4.14373.

[8] Masdalena, "Factors Affecting Stunting in Indonesia: A Study PRISMA (Preferred Reporting Items for Systematic Review & Meta-Analysis)," *J. Pharm. Sci.*, vol. 7, no. 1, pp. 159–174, 2024, doi: 10.36490/journal-jps.com.

[9] F. T. Sabilillah, C. A. Sari, R. B. Abiyyi, and A. Susanto, "Comparison Of Machine Learning Algorithms On Stunting Detection For 'Centing' Mobile Application To Prevent Stunting," *SINKRON*, vol. 8, no. 4, pp. 2360–2368, 2024, doi: 10.33395/sinkron.v8i4.13967.

[10] R. M. Fikri, A. Purbasari, A. Zulianto, F. R. Rinawan, and A. I. Susanti, "LSTM Parameter Optimization with Genetic Algorithm for Stunting Prediction," *PIKSEL*, vol. 13, no. 1, pp. 133–144, 2025, doi: 10.33558/piksel.v13i1.10761.

[11] A. A. R. Reza and M. S. Rohman, "Prediction Stunting Analysis Using Random Forest Algorithm and Random Search Optimization," *J. Inform. Telecommun. Eng.*, vol. 7, no. 2, pp. 534–544, 2024, doi: 10.31289/jite.v7i2.10628.

[12] 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, 2025, doi: 10.32877/bt.v8i2.3213.

[13] T. A. Zuraiyah, N. Widanti, Yamato, A. Chairunnas, K. Mauludin, and B. A. Setha, "Classification and Visualization Model of Stunting Zone Distribution Using Artificial Intelligence and Streamlit Approaches," *JOIV: Int. J. Inform. Vis.*, vol. 9, no. 5, pp. 2040–2047, 2025, doi: 10.62527/joiv.9.5.3174.

[14] 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, pp. 1-18, 2023, doi: 10.3390/children10101638.

[15] C. Hayat and I. A. Soenandi, "Hybrid Architecture Model of Genetic Algorithm and Learning Vector Quantization Neural Network for Early Identification of Ear, Nose, and Throat Diseases," *J. Inf. Syst. Eng. Bus. Intell.*, vol. 10, no. 1, pp. 1–12, 2024, doi: 10.20473/jisebi.10.1.1-12.

[16] F. M. Amin, D. Candra, and R. Novitasari, "Identification of Stunting Disease using Anthropometry Data and Long Short-Term Memory (LSTM) Model," *Comput. Eng. Appl.*, vol. 11, no. 1, pp. 25–36, 2022, doi: 10.18495/comengapp.v11i1.395.

[17] W. S. Lestari, Caroline, and Y. M. Saragih, "Deep Learning Approach for Stunting Classification in Toddlers," in *Proc. 2024 2nd Int. Conf. Technol. Innov. Appl. (ICTIIA)*, IEEE, vol. 2024, no. December, pp. 1–5, 2024, doi: 10.1109/ICTIIA61827.2024.10761797.

[18] M. I. Elim and E. Utami, "Performance Comparison of Child Stunting Prediction Support Vector Machine vs Random Forest with Grid Search Optimization," *JUTIF*, vol. 6, no. 5, pp. 5305–5319, 2025, doi: 10.52436/1.jutif.2025.6.5.5285.

[19] A. Aziz, F. Insani, J. Jasril, and F. Syafria, " Implementation of the Learning Vector Quantization (LVQ) Method for the Classification of Families at Risk of Stunting," *BITS*, vol. 5, no. 1, pp. 12–20, 2023, doi: 10.47065/bits.v5i1.3478.

[20] W. S. Lestari, Y. M. Saragih, and C. Caroline, "Multiclass Classification for Stunting Prediction Using Deep Neural Networks," *JITK*, vol. 10, no. 2, pp. 386–393, 2024, doi: 10.33480/jitk.v10i2.5636.

[21] I. G. B. A. Budaya and I. K. P. Suniantara, "Comparison of Sentiment Analysis Algorithms with SMOTE Oversampling and TF-IDF Implementation on Google Reviews for Public Health Centers," *MALCOM: Indones. J. Mach. Learn. Comput. Sci.*, vol. 4, no. 3, pp. 1077–1086, 2024, doi: 10.57152/malcom.v4i3.1459.

[22] G. Husain, D. Nasef, R. Jose, J. Mayer, M. Bekbolatova, T. Devine, and M. Toma, "SMOTE vs. SMOTEENN: A Study on the Performance of Resampling Algorithms for Addressing Class Imbalance in Regression Models," *Algorithms*, vol. 18, no. 1, pp. 1–16, 2025, doi: 10.3390/a18010037.

[23] Y. A. Singgalen, "A Hybrid CNN-LSTM Model with SMOTE for Enhanced Sentiment Analysis of Hotel Reviews," *BITS*, vol. 6, no. 3, pp. 1363–1373, 2024, doi: 10.47065/bits.v6i3.6301.

[24] K. Niu, G. Lu, X. Peng, Y. Zhou, J. Zeng, and K. Zhang, "CNN autoencoders and LSTM-based reduced order model for student dropout prediction," *Neural Comput. Appl.*, vol. 35, no. 30, pp. 22341–22357, 2023, doi: 10.1007/s00521-023-08894-2.

[25] M. Khodayar, A. Farajzadeh Bavil, and M. E. Khodayar, "Deep Attention GRU-GRBM with Dropout for Fault Location in Power Distribution Networks," in *2024 IEEE Texas Power and Energy Conference (TPEC 2024)*, IEEE, vol. 2024, no. March, pp. 1–6, 2024, doi: 10.1109/TPEC60005.2024.10472203.

[26] W. Qu, L. Guo, J. Cui, and X. Jin, "Multimodal Attention-Based Instruction-Following Part-Level Affordance Grounding," *Appl. Sci.*, vol. 14, no. 11, pp. 1–23, 2024, doi: 10.3390/app14114696.

[27] N. Matondang and N. Surantha, "Effects of oversampling SMOTE in the classification of hypertensive dataset," *Adv. Sci. Technol. Eng. Syst.*, vol. 5, no. 4, pp. 432–437, 2020, doi: 10.25046/AJ050451.

[28] M. M. R. Mubarak, Y. H. Chrisnanto, and P. N. Sabrina, "Implementation of Random Forest Using Smote and Smoteenn in Customer Churn Classification in E-Commerce," *Enrichment: J. Multidiscip. Res. Dev.*, vol. 1, no. 8, pp. 463–477, 2023, doi: 10.55324/enrichment.v1i8.69.

[29] R. J. Dhanal and V. R. Ghorpade, "Aspect-based sentiment-analysis using topic modelling and machine-learning," *Int. J. Electr. Comput. Eng.*, vol. 14, no. 6, pp. 6689–6698, 2024, doi: 10.11591/ijece.v14i6.pp6689-6698.

[30] S. Chopannejad, A. Roshanpoor, and F. Sadoughi, "Attention-assisted hybrid CNN-BILSTM-BiGRU model with SMOTE–Tomek method to detect cardiac arrhythmia based on 12-lead electrocardiogram signals," *Digit. Health*, vol. 10, no. March, pp. 1–20, 2024, doi: 10.1177/20552076241234624.

[31] D. Hatta Fudholi, R. N. Abida Nayoan, A. Fathan Hidayatullah, and D. Brahma Arianto, "Hybrid CNN-BiLSTM Model for Drug Named Entity Recognition," *J. Eng. Sci. Technol.*, vol. 17, no. 1, pp. 730–744, 2022.

[32] S. Imron, E. I. Setiawan, Joan Santoso, and Mauridhi Hery Purnomo, "Aspect Based Sentiment Analysis Marketplace Product Reviews Using BERT, LSTM, and CNN," *J. RESTI*, vol. 7, no. 3, pp. 586–591, 2023, doi: 10.29207/resti.v7i3.4751.