Development of Tracking Features and Recommendation for Mother's Nutrition in the Genting Mobile Application to Prevent Stunting using Vision Transformer (Case Study: PKK Bandung City)

2024 International Conference on Intelligent Cybernetics Technology & Applications (ICICyTA) | 979-8-3315-0649-0/24/$31.00 ©2024 IEEE | DOI: 10.1109/ICICYTA64807.2024.10913053

1st Azka Tsawaab Dhafin
Department of Information Systems
Telkom University
Bandung, Indonesia
azkatdhafin@student.telkomuniversity.ac.id

2nd Rahmat Fauzi
Department of Information Systems
Telkom University
Bandung, Indonesia
rahmatfauzi@telkomuniversity.ac.id

3rd Zalina Fatima Azzahra
Department of Information Systems
Telkom University
Bandung, Indonesia
zalinaza@telkomuniversity.ac.id

***Abstract*—Stunting is a chronic nutritional problem that has a serious impact on child growth and development in Indonesia. It not only affects physical stature but also impairs cognitive development and increases the risk of long-term socio-economic issues. Although several applications have been developed for stunting prevention, many do not provide efficient nutrition tracking features and personalized nutrition recommendations, which are crucial for early prevention of stunting. This research introduces the development of an Android mobile application called "Genting" that leverages the Vision Transformer (ViT) machine learning model for food image classification and Google Generative AI for providing tailored nutrition recommendations. The ViT model achieved an accuracy of 82.12% on the validation dataset after 20 epochs, with a loss value of 0.6725. To ensure model robustness, K-fold cross-validation with 5 folds was employed, resulting in an average accuracy of 95.93 Further testing demonstrated the model's ability to recognize various types of food with the highest probability reaching 99.99%, reflecting high confidence in its predictions. This high accuracy ensures more reliable food classification and improving efficiency in nutrition tracking. Although not publicly released, the Genting application underwent thorough testing in a restricted environment with Posyandu cadres to verify its functionality. This application is designed to provide innovative solutions in nutrition monitoring through advanced machine learning technology, offering more personalized and efficient nutrition recommendations. By doing so, it aims to support early stunting prevention efforts, addressing a significant public health challenge in Indonesia by improving the precision and effectiveness of nutritional interventions at a critical stage in child development.**

Keywords- stunting prevention, vision transformer, nutrition recommendations

I. INTRODUCTION

Stunting is one of the most pressing health issues in Indonesia, especially among children under five. Stunting is characterized by height that is well below average for a child's age, resulting from chronic undernutrition and unhealthy environmental conditions during critical periods of a child's growth. Stunting not only stunts a child's physical growth, but also impacts cognitive development, learning capacity, and productivity in adulthood, which in turn can exacerbate the cycle of poverty [1].

Despite various efforts to reduce stunting, it remains a significant health challenge in Indonesia. Government programs and community interventions have made some progress, but the complexity of the issue stemming from factors like chronic undernutrition, poor maternal health, and inadequate early childhood care requires more targeted and innovative solutions. The persistence of high stunting rates underscores the need for refined strategies, as shown in Fig. 1, indicating concerning levels of stunting among children under five.

Fig. 1. Graph of stunting percentage in Indonesia based on SSGI [2].

Fig. 1. shows that the prevalence of stunting in Indonesia is still at an alarming level, despite some reductions. For example, based on data from the Indonesian Nutrition Status Survey (SSGI) in 2022, the prevalence of stunting in Indonesia fell from 24.4% to 21.6% [2]. Efforts to address stunting in several regions, such as Bandung City, have shown significant results. By 2023, the stunting rate in Bandung City had dropped to 16.3% [3]. However, this figure is still relatively high and has not reached the national target of 14% by 2024 [4].

The causes of stunting are very diverse, ranging from low socioeconomic status, lack of access to health services and clean water, to low maternal education which has an impact on suboptimal child feeding patterns [2]. Lack of knowledge about the importance of balanced nutrition since pregnancy and exclusive breastfeeding during the first six months of a child's life are also significant factors that increase the risk of stunting [5].

There are many existing applications to prevent stunting in Indonesia, such as Simpati, MyBidan, Elsimil, STUNTECH, and estundad, but these applications do not yet have a nutrition tracking and recommendation feature that can consider the mother's specific nutritional preferences and needs for early stunting prevention. This tracking and nutrition recommendation play a crucial role in preventing stunting early on [6]. In this context, mobile apps can offer a promising solution as they can provide more effective health interventions, with advantages such as cost-effectiveness, easy accessibility, and stronger user engagement over traditional methods [7].

To support stunting prevention, the Genting app was developed with a case study in the Bandung City PKK, utilizing machine learning technology. Unlike existing apps, Genting is explicitly created to assist mothers in managing their preconception and maternal nutrition through the provision of personalised nutrition monitoring and recommendations. Its primary goal is to prevent stunting at an early stage. The app utilizes Vision transformer (ViT) to improve efficiency and automation in nutrition monitoring. ViT enables real-time classification of food images, which is then used to display reliable nutrition information from data that has been retrieved from external sources, ensuring users get relevant and accurate information from the captured images. In addition, the Genting app utilizes Google Generative AI to provide nutrition recommendations that are more personalized and in line with user preferences, proving to be more effective than traditional methods [8]. The ViT model was evaluated based on accuracy, loss, label classification ability, and cross-validation, ensuring precise and accurate predictions. This approach is expected to improve the efficiency of nutrition tracking and the effectiveness of early stunting prevention.

II. Literature Review

A. Stunting

Stunting is a child growth disorder with height below two standard deviations from the WHO median, caused by nutritional deficiencies and poor environment [9]. In Indonesia, stunting is triggered by socioeconomic factors, maternal education, and limited access to health. Maternal knowledge about nutrition plays an important role in preventing stunting, as balanced nutrition greatly affects children's growth and development. Joint efforts from the government and the community are needed to overcome this problem [10].

B. Nutrition

Nutrition is essential to help children grow and develop, especially during crucial periods such as pregnancy through to toddlerhood. With a balanced diet, nutrition is key to supporting children's health and growth from an early age [11].

C. Vision Transformer

A Vision Transformer (ViT) is a type of AI model that treats images like puzzles, breaking them down into smaller pieces and understanding how they fit together to identify objects or scenes. ViTs are attention-based models meaning they can focus on specific parts of an image to understand its content. They are commonly used for computer vision tasks such as image classification and object detection. ViTs process images by dividing them into patches which are then processed through encoder transformers, enabling superior performance over CNNs [12]. The model combines patch embedding with positioning to produce high accuracy class predictions.

D. Google Generative AI

Google Generative AI is a powerful tool that can generate new and creative content based on the data it has been trained on. Artificial intelligence can significantly improve the effectiveness of recommendation systems compared to traditional methods by using techniques such as synthetic user profiling and item embedding to increase the diversity and accuracy of recommendations [13].Generative AI, such as GANs, provide more personalized recommendations and overcome the cold start problem [8].

III. Methodology

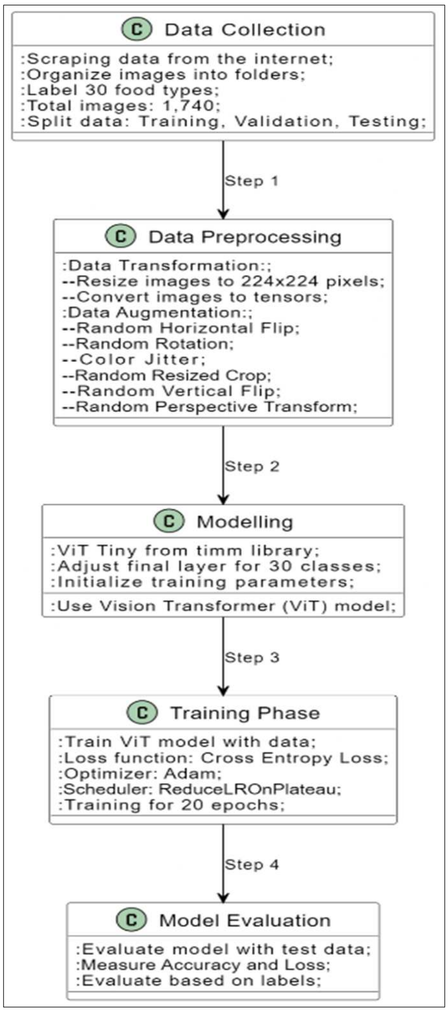

Fig. 2. Flowchart of the methodology used in the study.

Fig. 2 presents a flowchart of the methodology used in this study, starting with data collection, followed by data preprocessing, modeling, training, and finally, model evaluation. Each step builds on the previous one, guiding the process of developing and validating the Vision Transformer (ViT) model for food image classification.

A. Data Collection

The dataset was collected using the scraping method from the internet, containing images of various types of food commonly eaten by Indonesians. After collection, the images were grouped into folders based on food types to facilitate automatic labeling. The dataset consists of 30 food type labels, including avocado, apple, fried chicken, gado-gado, fried fish, pindang fish, grilled salmon, corn, guava, orange, potato, mango, brown rice, white rice, stir-fried long beans, papaya, banana, rendang, sayur asem, sayur sop, watermelon, cassava, fried tofu, omelet, fried tempeh, eggplant balado, stir-fried spinach, stir-fried kale, sweet potato, and fried shrimp. A total of 1,740 images were collected, with 41 images per class used for training, 11 images per class for validation, and 6 images per class for testing.

Although the dataset size is relatively small, it reflects a careful selection process to ensure image quality and relevance. Expanding the dataset with additional images scraped from the internet poses challenges, as many publicly available images lack consistent quality, contain noise, or have unsuitable lighting and angles for effective classification. As a result, the chosen dataset size balances the need for model training with the goal of maintaining high-quality input data. While this limitation may impact the model's ability to generalize to unseen data, it underscores the importance of quality over quantity in image-based machine learning models. Addressing this issue fully would require access to a more curated and diverse dataset, which was beyond the scope of this study.

B. Data Preprocessing

This step involves several transformation and augmentation techniques to prepare the data for use in model training. The main purpose of preprocessing is to increase data diversity and prevent overfitting. Data preprocessing in this research goes through 2 stages, namely data transformation and data augmentation.

1) Data Transformation

The data transformation process is a critical step in preprocessing that aims to prepare the images in the dataset so that they match the input requirements of the Vision transformer (ViT) model. This process includes two steps. The first step involves resizing the images to 224x224 pixels, to ensure uniformity across the dataset for the model to process the images properly. The second step involved converting the images to tensors, which are the underlying data structures used by PyTorch to perform numerical computations.

2) Data Augmentation

Data augmentation is a technique used to artificially expand the size of a dataset by creating modified versions of existing images. This technique helps in increasing the diversity of data used in model training, thus preventing overfitting and improving the generalization ability of the model. Data augmentation is done with various techniques, including: Random Horizontal Flip, Random Rotation, Color Jitter, Random Resized Crop, Random Vertical Flip, and also Random Perspective Transform. These data augmentation techniques are used to improve the performance of machine learning models by increasing the variety of training data, so that the model can be more general and accurate in recognizing objects in various conditions.

3) Modelling

In this research, the Vision transformer (ViT) model is used as the main tool for image classification. ViT is a transformer model pre-trained on the ImageNet dataset, which provides an initial knowledge base from a large dataset, thus speeding up the training process. The ViT model used in this research is ViT Tiny, which is imported using the timm library. Adjustments are made to the final layer to adjust to the number of classes in the research dataset, by replacing the last linear layer so that the model can perform classification according to the research needs. The model was then initialized with training parameter settings such as optimizer and loss function selected for further training.

TABLE I. VIT MODEL ARCHITECTURE USED

Layer	Type	Configuration
0	PatchEmbed	Conv2d(3, 192, kernel_size=(16, 16), stride=(16, 16))
1	LayerNorm	(192,)
2-11	Transformer Blocks	12 Blocks with Multi-Head Attention and MLP
12	LayerNorm	(192,)
13	Linear	in_features=192, out_features=30

Table I shows the architecture of the ViT model used, with details of the configuration at each layer.

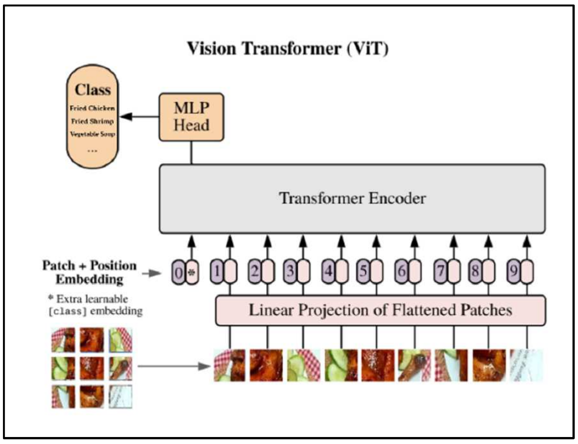

Fig. 3. Illustration of the ViT model architecture used.

Fig. 3. illustrates the structure of the ViT model more visually, showing how the input image is converted into small patches through the Patch Embedding process. These patches are then processed by Transformer Blocks consisting of Multi-Head Attention and Multilayer Perceptron (MLP) mechanisms to capture contextual information locally and globally. Layer Normalization used at multiple layers serves to stabilize and speed up model training by normalizing the output of each layer.

The synergy of these modules enables ViT to process visual data efficiently and effectively, resulting in accurate classification predictions from a given input image. As a result of this combination, ViT adapts well to the image classification task in this study.

C. Training Phase

In the training stage, the Vision transformer (ViT) model is trained using the dataset that has gone through the transformation and augmentation process. The loss function used is Cross Entropy Loss, which is suitable for multi-class classification, while the optimizer used is Adam optimizer, which is famous for its fast convergence. To maintain training efficiency and prevent overfitting, the ReduceLROnPlateau scheduler is used which automatically reduces the learning rate when the evaluation metric shows no improvement. Training was conducted for 20 epochs, where each epoch consisted of a forward pass to generate predictions, loss calculation based on the difference between predictions and original labels, and backpropagation to update model weights. Model performance was monitored using validation data to ensure good generalization ability.

D. Model Evaluation

The evaluation stage aims to test the performance of the Vision Transformer (ViT) model that has been trained using test data. This evaluation process now includes three main aspects: evaluation based on accuracy and loss, evaluation based on labels, and cross-validation. Evaluation based on accuracy and loss uses accuracy and loss metrics to assess model performance. Accuracy measures how well the model's predictions match the actual labels, indicating the model's success in solving the problem, while loss quantifies the error in the model's predictions. The accuracy (1) and loss (2) formulas are as follows:

$$Accuracy\ (\%) = \left(\frac{TP+TN}{TP+TN+FP+FN}\right) \times 100\% \quad (1)$$

$$Loss\ (\%) = \left(\frac{FP+FN}{TP+TN+FP+FN}\right) \times 100\% \quad (2)$$

Description:

- True Positive (TP): The data is predicted correctly by the model and the actual is correct.
- True Negative (TN): The data is predicted wrong by the model and the actual is wrong.
- False Positive (FP): The data is predicted to be wrong but the actual data is correct.
- False Negative (FN): Data predicted to be true but the actual data is false.

In addition to this, model evaluation based on labels is conducted by comparing the true labels and predicted labels to assess the model’s precision in identifying the correct categories. Furthermore, the model’s robustness and reliability were assessed using 5-fold cross-validation. Cross-validation works by splitting the dataset into five equal parts, where the model is trained on four parts and tested on the remaining part, with this process repeated five times. This ensures that the model’s performance is evaluated across different subsets of the data, providing a thorough assessment of its generalizability.

IV. Result and Discussion

A. Model Evaluation Based on Accuracy and Loss

At this stage, the accuracy and loss levels of the model are tested using a scheme with an epoch of 20. This scheme is used to see how accurate and stable a model is in predicting or classifying an object.

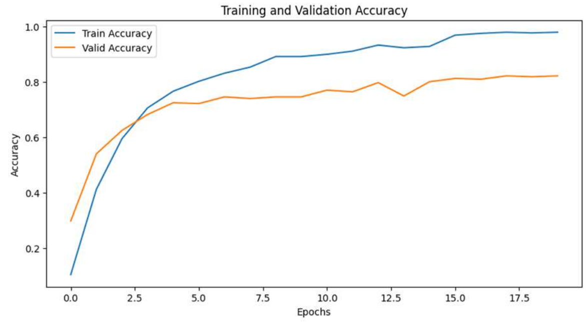

Fig. 4. Training and validation accuracy graph.

Fig. 4. shows the training and validation accuracy graphs of the ViT model for 20 epochs. This graph shows the improvement of the model's accuracy on the training and validation data as the epochs increase. Initially, the training accuracy increases sharply, indicating the model is learning from the data, while the validation accuracy increases more slowly, indicating the model's ability to recognize patterns from new data. Midway through the training, the training accuracy approaches 1.0, while the validation accuracy stabilizes around 0.8, indicating good performance on data not seen before. The difference between these two accuracies indicates a slight overfitting, where the model is better at predicting the training data than the validation data. Overall, this graph provides a visual representation of the model's training effectiveness and its ability to generalize.

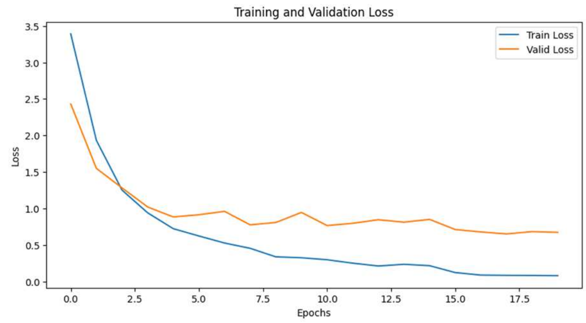

Fig. 5. Training and validation loss graph.

Fig. 5. shows the training and validation loss graphs of the ViT model for 20 epochs. This graph illustrates how the loss values on the training data (train loss) and validation data (valid loss) decrease as the epochs increase. At the beginning of training, the training and validation loss start from a fairly high value, indicating that the model has not learned the pattern from the data well. As the training process progresses, the loss values on the training data decrease significantly, indicating that the model is getting better at predicting the training data. The loss value on the validation data also decreases, but at a slower rate and tends to stabilize after a few epochs.

TABLE II. VIT MODEL TRAINING RESULTS

Epoch	Train Loss	Train Accuracy	Valid Loss	Valid Accuracy
1	3,3907	10,33%	2,4313	29,70%
2	1.9327	41,14%	1,5488	53,94%
3	1,2518	59,43%	1,2823	62,42%
4	0,9407	70,57%	1,0204	68,18%
5	0,7232	76,59%	0,8837	72,42%
6	0,6231	80,16%	0,9139	72,12%
7	0,5267	83,09%	0,9605	74,55%
8	0,4541	85,28%	0,7757	73,94%
9	0,3374	89,11%	0,8078	74,55%
10	0,3248	89,11%	0,9458	74,55%
11	0,2977	89,92%	0,7659	76,97%
12	0,2509	91,06%	0,7967	76,36%
13	0,2119	93,25%	0,8456	79,70%
14	0,2352	92,28%	0,8121	74,85%
15	0,2161	92,76%	0,8502	80,00%
16	0,1223	96,83%	0,7116	81,21%
17	0,0872	97,48%	0,6782	80,91%
18	0,0838	97,89%	0,6513	82,12%
19	0,0822	97,64%	0,6828	81,82%
20	0,0795	97,89%	0,6725	82,12%

In Table II, the training results of the Vision transformer (ViT) model are shown, illustrating the changes in loss and accuracy values over 20 training epochs. This data shows how the model gradually learns from the training data and improves its ability to make predictions on the validation data.

In the first epoch, the model shows a train loss of 3.3907 and training accuracy of 10.33%, indicating poor initial prediction. The validation loss of 2.4313 and accuracy of 29.70% further highlight the model's struggle to generalize early in training.

As training progresses, the train loss decreases and accuracy improves significantly. By the 10th epoch, the train loss drops to 0.3248 with 89.11% accuracy, while the validation loss is 0.9458 with 74.55% accuracy, showing good generalization but slight overfitting.

By the 20th epoch, the model achieves a train loss of 0.0795 with 97.89% accuracy and a validation loss of 0.6725 with 82.12% accuracy. Despite some overfitting, the model's overall performance improves, especially in its generalization capability.

Overall, Table 2 provides a clear picture of the training effectiveness of the ViT model, where the model consistently improves its ability to predict more accurately on both training and validation data. This data also provides important insights into how the model learns and adapts the training parameters to improve its performance throughout the training process.

B. Model Evaluation Based on Labels

The label values will be used to test the Vision transformer model that has been trained and achieved the best accuracy. Testing is done with test data, where the actual label represents the actual value and the predicted label is the result of the model. The test results will be displayed as food images for each category, with actual and predicted labels.

Fig. 6. Model evaluation based on label value.

Fig. 6. shows a sample evaluation result of the Vision transformer model that has been trained with the test data. Each food image in this figure shows the actual label (True) and the predicted label (Predicted) of the model. These images cover various food categories such as tamarind vegetable, corn, brown rice, banana, grilled salmon, watermelon, stir-fried spinach, and apple. The model managed to correctly predict several categories such as corn, brown rice, banana, watermelon, stir-fried spinach, and apple, demonstrating the model's ability to recognize these food types.

C. Model Testing on Mobile Application

This test section describes the results of food image classification on the Genting app. This feature uses Vision transformer (ViT) machine learning technology to recognize and categorize food types taken through a mobile phone camera or selected from a photo gallery. This ViT technology is integrated into the Genting mobile application through Flask and ngrok, thus supporting the application's functionality in recognizing and classifying food images accurately.

TABLE III. FOOD CLASSIFICATION RESULTS BASED ON PROBABILITY

No.	Food Items	Probability
1	Oranges	99,21%
2	Guava	99,45%
3	Corn	99,9%
4	Grilled salmon	83,86%
5	Stir-fried string beans	99,88%
6	Mango	96,65%
7	Tamarind vegetable	99,75%
8	Apple	99,83%
9	Stir-fried spinach	95,41%
10	Brown rice	99,92%
11	Ikan pindang	96,28%
12	Vegetable soup	99,97%
13	Fried shrimp	99,46%
14	Fried chicken	95,18%
15	Omelet	98,07%
16	Fried tempeh	88,12%
17	Stir-fried kale	98,35%
18	Sweet potato	98,63%
19	Watermelon	99,95%
20	Fried tofu	99,90%
21	Cassava	99,68%
22	Vegetable soup	99,97%
23	Fried fish	85,27%
24	Papaya	98,63%
25	Gado-gado	99,99%
26	Potato	93,80%
27	Rendang	99,90%
28	Banana	95,41%
29	Fried tempeh	88,12%

30	White rice	99,81%

Table III shows the results of food image classification by the Vision transformer (ViT) model on the Genting application. The model successfully classified various types of food with the highest probability achieved by the model being 99.99%, reflecting the confidence level of the model in its predictions. During testing, all 30 food labels in the dataset were tested, and the results showed that the model performed well in recognizing each food type with a model accuracy on the validation dataset of 82.12%.

D. Model Evaluation Based on Cross-Validation

The model was evaluated using 5-fold cross-validation to ensure consistent performance across different data subsets. The results from each fold show some variation, with the lowest accuracy recorded at 88.21% in Fold 1, and the highest accuracy of 99.19% in Fold 3. The overall average accuracy across the five folds was 95.93%, indicating that the model demonstrates consistent and reliable performance. Despite slight variations between folds, the model's general performance remains stable, with high accuracy observed throughout the cross-validation process.

E. Tracking and Nutrition Recommendation Feature on Genting Mobile Application

The nutrition tracking feature of the Genting mobile app allows users to log food consumption manually or by taking a picture. The app provides in-app prompts to guide users in taking better photos, ensuring optimal food classification using the Vision Transformer model, which then provides nutritional information such as calories, protein, fat, and carbohydrates. The top five classification results are displayed to help users efficiently monitor daily intake.

The nutrition recommendation feature uses the Gemini 1.5 API to provide tailored food suggestions based on parameters like age, weight, height, activity level, health conditions, meal price, and dietary preferences. Users can also select allergies, specify meal times, and input preferred or disliked foods. The app then generates meal recommendations with servings, nutritional values, ingredients, and alternatives. This feature helps prevent stunting by offering personalized and targeted nutrition recommendations to meet users' needs.

Although not publicly released, the Genting application underwent thorough testing in a restricted environment with Posyandu cadres to verify its functionality. Both features were found to be fully functional, accurately identifying food items and delivering tailored nutritional advice based on the users' needs. This demonstrates the app’s capability to assist in monitoring and managing maternal nutrition effectively.

V. CONCLUSIONS

The nutrition tracking feature using Vision Transformer (ViT) in the Genting mobile app has been successfully implemented, achieving 82.12% accuracy on the validation dataset after 20 epochs with a loss of 0.6725. The model’s highest classification probability reached 99.99% for 30 trained food labels, reflecting high confidence in its predictions. To ensure robustness, 5-fold cross-validation was performed, yielding an average accuracy of 95.93%. While slight overfitting was observed, the model’s strong cross-validation performance suggests good generalization.

Additionally, the app includes a nutrition recommendation feature powered by Google Generative AI, offering personalized recommendations based on user needs.

The Genting app, though not publicly released, was tested with Posyandu cadres in a limited setting to verify its functionality. By automating nutrition tracking and providing personalized recommendations, the app shows potential as an innovative tool for stunting prevention, particularly in assisting mothers in managing their preconception and maternal nutrition.

For future research, it is recommended to integrate the app with nutritionists and dietitians, enabling users to consult and receive personalized health advice based on their needs. This advice would utilize the app’s collected data, such as dietary patterns and health goals. Incorporating expert consultations could enhance the precision of recommendations, improve user engagement, and lead to better health outcomes.

REFERENCES

[1] A. Wigati, F. Yulia, K. Sari, and T. Suwarto, “Pentingnya Edukasi Gizi Seimbang untuk Pencegahan Stunting pada Balita,” 2022.

[2] S. Liza Munira, “Hasil Survei Status Gizi Indonesia (SSGI) 2022,” 2022.

[3] Diskominfo Bandung, “Pemdaprov Jabar Sebut Kota Bandung Sukses Turunkan Stunting Secara Signifikan,” Pemerintah Daerah Provinsi Jawa Barat. Accessed: Aug. 05, 2024. [Online]. Available: https://jabarprov.go.id/berita/pemdaprov-jabar-sebut-kota-bandung-sukses-turunkan-stunting-secara-signifikan-13708

[4] Novrizaldi, “Pemerintah Optimis Target Penurunan Stunting 14 Persen Tercapai di 2024,” KEMENKO PMK. Accessed: Jan. 04, 2024. [Online]. Available: https://www.kemenkopmk.go.id/pemerintah-optimis-target-penurunan-stunting-14-persen-tercapai-di-2024

[5] E. Satriawan, “National Strategy To Accelerate Stunting Reduction 2018-2024,” 2018.

[6] A. A. Permana, A. T. Perdana, N. Handayani, and R. Destriana, “A Stunting Prevention Application ‘nutrimo’ (Nutrition Monitoring),” in *Journal of Physics: Conference Series*, IOP Publishing Ltd, Mar. 2021. doi: 10.1088/1742-6596/1844/1/012023.

[7] N. Stasya and W. Sulistiadi, “The Effectiveness of Mobile Application as Educational Intervention to Prevent Stunting: A Systematic Review,” *The 7th International Conference on Public Health*, p. 170, 2020, doi: 10.26911/the7thicph-FP.02.26.

[8] M. O. Ayemowa, R. Ibrahim, and M. M. Khan, “Analysis of Recommender System Using Generative Artificial Intelligence: A Systematic Literature Review,” *IEEE Access*, 2024, doi: 10.1109/ACCESS.2024.3416962.

[9] World Health Organization, “Stunting in a nutshell.” Accessed: Mar. 15, 2024. [Online]. Available: https://www.who.int/news/item/19-11-2015-stunting-in-a-nutshell

[10] I. Kuswanti and S. Khairani Azzahra, “Hubungan Pengetahuan Ibu tentang Pemenuhan Gizi Seimbang dengan Perilaku Pencegahan Stunting pada Balita,” *Jurnal Kebidanan Indonesia*, vol. 13, no. 1, Jan. 2022, doi: 10.36419/jki.v13i1.560.

[11] F. Mayar, Y. Astuti[2], P. Anak, U. Dini, I. Pendidikanuniversitas, and N. Padang, “Peran Gizi Terhadap Pertumbuhan dan Perkembangan Anak Usia Dini,” 2021.

[12] Z. Fu, “Vision Transformer: Vit and its Derivatives,” May 2022, [Online]. Available: http://arxiv.org/abs/2205.11239

[13] S. Shakir Sumit, “Title: An AI-based solution for the cold start and data sparsity problems in the recommendation systems,” 2023.