Applied Information System and Management (AISM)
Volume 8, (2) 2025, p. 285–296
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v8i2.46680

Leveraging Deep Learning for Cultural Preservation: A Mobile Application for Padang Cuisine

Njoto Benarkah[1*], Vincentius Riandaru Prasetyo[2], Andreas Bayu Prakarsa[3]

***Abstract*—Padang cuisine, originating from West Sumatra, Indonesia, is recognized as one of the most widespread traditional food types due to its prevalence in restaurants across the country. Despite the increasing interest in classifying Indonesian food using artificial intelligence, there have been limited studies that have explicitly focused on classifying Padang dishes using deep learning approaches. This study aimed to develop an intelligent mobile application capable of identifying various Padang dishes from images using transfer learning-based convolutional neural networks (CNNs). Four pre-trained CNN architectures—EfficientNetV2M, MobileNetV2, VGG19, and ResNet152V2—were fine-tuned and evaluated on a dataset of Padang food images. This dataset comprised a total of 1,108 images, categorized into nine distinct Padang dishes, collected from both publicly available repositories and original photographs taken for this study. Among these models, ResNet152V2 achieved the best performance after optimization, with a validation loss of 0.4142 and a test accuracy of 91.33%. The optimized model was converted to TensorFlow Lite and deployed as a mobile application, enabling real-time recognition of Padang dishes. This study presented a deep-learning-based mobile solution for recognising nine traditional Padang dishes with high accuracy, demonstrating the potential of AI-driven applications to support culinary heritage preservation and promote cultural tourism in Indonesia.**

***Index Terms*—Convolutional neural network, food recognition, mobile application, Padang cuisine, transfer learning.**

I. Introduction

Cultural preservation extends beyond the safeguarding of tangible artefacts and monuments to include living traditions. In recognition of this, the United Nations Educational, Scientific and Cultural Organization (UNESCO) introduced the 2003 Convention for the Safeguarding of the Intangible Cultural Heritage (ICH) [1]. ICH comprises the customs, forms of expression, expertise, and abilities that groups and individuals consider fundamental to their cultural identity [2]. While traditionally transmitted through oral traditions, there has been increasing interest in utilizing digital technologies to document, disseminate, and revitalize ICH [3]. Leveraging digital technologies offers novel avenues to enhance access and ensure the longevity of intangible cultural assets in the digital age [4], [5]. Culinary heritage, as a significant component of ICH, often encapsulates a region's history, agricultural practices, and social customs, making its preservation crucial for cultural identity.

Indonesia's vast archipelagic geography and remarkable ethnolinguistic diversity have given rise to over 5,000 documented traditional recipes [6], [7]. Among these, Padang cuisine from West Sumatra is notable for its bold flavors, complex spice blends, and distinctive cooking methods [8], [9]. A hallmark of Padang cuisine is *rendang* [10], a dish that has received global acclaim, notably topping Cable News Network's 'World's 50 Best Foods' list in 2011 [11]. However, survey results from this study reveal that while 94% of respondents recognize *rendang*, awareness of other dishes, such as *telur balado,* drops to 63%, and others remain largely unknown. This skewed recognition risks reducing cultural memory to a single icon and neglecting a more expansive repertoire of equally significant dishes.

Current preservation relies on manual documentation and expert knowledge. While effective, these approaches are not widely accessible and do not offer interactive or real-time identification of dishes. As a result, opportunities to engage a broader audience, including younger generations and tourists, through convenient digital tools remain limited [12], [13].

Padang cuisine presents a significant challenge for food classification due to the subtle visual differences between many of its dishes, making it an ideal case for fine-grained image recognition. Deep learning, particularly Convolutional Neural Networks (CNNs), has revolutionized image classification by automatically learning hierarchical and discriminative features,

Received: 9 June 2025; Revised: 24 July 2025; Accepted: 17 August 2025.
*Corresponding author

[1*]Njoto Benarkah, Department of Informatics Engineering, University of Surabaya, Indonesia (e-mail: benarkah@staff.ubaya.ac.id).
[2]Vincentius Riandaru Prasetyo, Department of Informatics Engineering, University of Surabaya, Indonesia (e-mail: vincent@staff.ubaya.ac.id).
[3]Andreas Bayu Prakarsa, Department of Informatics Engineering, University of Surabaya, Indonesia (e-mail: abayup1498@gmail.com).

making it highly suitable for fine-grained tasks. Existing studies that apply deep learning to food classification have achieved promising results, such as 79.23% accuracy in classifying Turkish cuisine [14] and 85.52% accuracy for broader Indonesian cuisine using Random Forest (RF) [15]. However, these studies typically focus on coarse-grained classification and lack integration into practical, real-time applications.

Crucially, no previous work has applied deep learning to fine-grained classification of Padang dishes and deployed such models into a mobile application for real-time culinary recognition and cultural heritage preservation. This research gap motivates the present study, which aims to develop a CNN-based model capable of accurately classifying Padang dishes and to integrate the model into a mobile app. The application not only identifies dishes from images but also provides information on key ingredients and recipes, thus supporting both culinary education and cultural preservation.

The objectives of this research are:

- To develop a deep learning model designed to classify various Padang dishes based on input images accurately.
- To integrate the model into a mobile application for real-time dish recognition.
- To demonstrate that deep learning can capture the subtle visual cues essential for differentiating similar dishes, thereby providing practical proof of its suitability for preserving culinary heritage.

This work contributes a digital tool for comprehensive documentation, promotes culinary tourism, and offers an educational resource while advancing the application of deep learning in cultural informatics.

The remainder of this paper is organized as follows: Section II discusses related work in food image classification and AI for cultural heritage. Section III details the research method, including dataset collection, model selection, training and optimization procedures, and mobile application development. Section IV presents the experimental results and provides a comprehensive discussion of the model's performance. Finally, Section V concludes the paper and outlines future research directions.

II. Related Work

Recent advancements in digital technologies have significantly contributed to cultural preservation efforts, primarily through mobile applications [16]. However, many of these applications merely present static information such as names, images, and descriptions [14], [17], lacking the integration of intelligent image recognition to identify regional culinary identities. This limitation is especially pronounced for regional cuisines, such as Padang, where nuanced intra-cuisine variations require more sophisticated approaches than simple database lookups.

To address these challenges, deep learning methods, especially CNNs have become the predominant technique for food image classification [18], [19]. Reference [20] developed a high-quality dataset of 1,644 images across 34 Indonesian traditional foods captured under controlled studio lighting, enabling CNN architectures including DenseNet121, ResNet50, InceptionV3, and NasNetMobile, to achieve outstanding results, i.e., DenseNet121 reached 99.4% accuracy, with precision and recall exceeding 0.92, demonstrating the benefits of dataset quality for classifier performance. This illustrates the strength of CNNs for cultural heritage applications when trained on high-quality, uniform datasets. However, a significant weakness was the dataset's lack of variability; models trained under these conditions demonstrated limited generalization to real-world settings, struggling with variations in serving styles and food presentation. Moreover, another study that focuses on broad Indonesian cuisine rather than regional, fine-grained distinctions leaves visual complexities within regional cuisines, such as Padang dishes, insufficiently explored. Our work directly tackles this limitation by targeting detailed classification within unconstrained Padang cuisine images, reflecting realistic use cases. Building on the dataset size and model diversity, a study [21] provided a comprehensive evaluation of 67 deep learning models across 16 architectures. The EfficientNetV2-L model achieved the highest accuracy of 85.44%, demonstrating strong overall classification capabilities on a diverse Indonesian food dataset comprising 24,427 images across 160 categories. This extensive benchmarking is a notable strength, providing valuable insights into the suitability of the architecture for classifying Indonesian food. Rasyidi et al. identified persistent challenges. Specifically, they identified persistent challenges, such as class imbalance, which causes difficulty in learning underrepresented foods, and high visual similarity between dishes, which leads to misclassification, as seen with *bandrek* being confused with *bajigur*. Additionally, the work's reliance on single-label classification limits its ability to handle images containing multiple food items, indicating a need for multi-label or object detection approaches. Our study addresses these issues by focusing on the fine-grained nuances of Padang cuisine and exploring practical deployment scenarios that require robust handling of real-world variability and complex food compositions.

Meanwhile, [22] employed a CNN-based model on nine Indonesian food categories, achieving an evaluation accuracy of 91.11%. Darojat's study highlights the effectiveness of CNNs in feature extraction and classification, but also exposes weaknesses due to the limited diversity of training data, which leads to overfitting and misclassification stemming from lighting differences and visual similarities among foods. Similarly, [23] applied a three-layer CNN model to classify 14 regional Indonesian foods, achieving a lower accuracy of 64.44%. The model’s lower performance was attributed to dataset imbalance and inadequate data preprocessing, resulting in false positives and negatives. These works collectively underscore the challenges in fine-grained regional food classification, particularly when datasets lack diversity or cleanliness, challenges our research mitigates by collecting and utilizing more varied, unconstrained images specific to Padang cuisine.

Beyond Indonesian cuisine, [24] proposed SlowDeepFood,

Applied Information System and Management (AISM)
Volume 8, (2) 2025, p. 285–296
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v8i2.46680

a framework combining semi-automatic dataset creation with transfer learning on EfficientNet for fine-grained regional food classification, achieving 91.91% accuracy on Food-101 and 95.33% on Middle Eastern food classification. The study demonstrated the effectiveness of semi-automatic dataset creation and transfer learning techniques in improving classification performance for diverse, underrepresented regional cuisines. However, despite its versatility, SlowDeepFood's model can struggle to capture the subtle visual distinction critical for fine-grained differentiation with closely related regional dishes, such as those in Padang cuisine. Nogay et al integrated transfer learning on six CNNs with Canny edge detection and data augmentation to classify Turkish cuisine food groups, reaching 79.23% accuracy using MobileNetV2. This dual-task framework, which addresses classification and portion estimation, tackles challenges in dietary assessment for underrepresented cuisines but remains constrained by a relatively small dataset and the difficulty of 2D portion estimation. This work inspires our study's focus on culturally specific Padang cuisine. Beyond deep learning, [15] explored traditional machine learning methods, such as Random Forest with segmentation, for 34 Indonesian foods, achieving an accuracy of 85.52%. This demonstrates the viability of conventional methods for specific tasks. Other technologies, such as augmented reality for promoting Sumatran specialities [17] or web-based culinary recommendation systems [16], also contribute to broader cultural engagement.

Despite the demonstrated successes, a key limitation across the literature is the difficulty of fine-grained food classification under unconstrained, real-world conditions with diverse presentations and multi-item images. Many approaches lack practical deployment in mobile applications, thereby restricting their cultural preservation impact. While current literature confirms the strength of CNNs and transfer learning in food image classification, challenges remain in dataset variability, intra-regional food distinction, class imbalance, and the real-world applicability of these methods. This research directly addresses these gaps by developing a deep learning model tailored for fine-grained Padang cuisine classification using a diverse, unconstrained dataset and deploying it in a mobile application. This work aims to enhance both digital heritage preservation and real-world dietary assessment for culturally significant regional foods.

III. RESEARCH METHOD

This study employed a five-phase methodology: dataset construction, data preprocessing and augmentation, model training with hyperparameter optimisation and evaluation, fine-tuning, and mobile application development, followed by post-deployment usability testing. Figure 1 illustrates the research methodology. Although KDD and CRISP-DM provide general guidance, the workflow of this study is adapted to the specific demands of image-centric tasks. While it follows the conceptual stages of CRISP-DM, such as data preparation, modeling, evaluation, and deployment, it is optimized for convolutional neural network-based food classification.

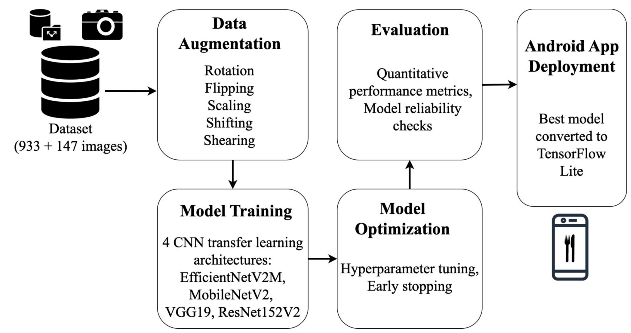


Fig. 1. Overview of the proposed Padang food classification workflow, including dataset construction, augmentation, model training and optimization, evaluation, and mobile deployment using the best-performing CNN model.

The primary goal of this study was to engineer a robust and lightweight deep learning model for a mobile application capable of classifying nine types of Padang dishes from images. The classification task was formulated as a multi-class image classification problem, where a given input image $x \in \mathbb{R}^{224\times224\times3}$ was assigned to a label $y \in \{0, 1, ..., 8\}$, representing one of the nine Padang dish classes. The deep learning model outputs a vector of logits $z = [z_0, z_1, \ldots, z_8]$. The classifier $f_\theta(x)$, parameterized by weight θ, predicted the probability of each class using a softmax function as shown in (1), at the output layer:

$$\hat{y}_i = \frac{e^{z_i}}{\sum_{j=0}^{8} e^{z_j}} \text{, for } i = 0, \ldots, 8 \qquad (1)$$

where z_j are the logits (raw outputs) produced by the concluding fully connected layer. The model's learning process aimed to minimize the sparse categorical cross-entropy loss, as shown in (2):

$$\mathcal{L}(y, \hat{y}) = -\log(\hat{y}_y) \qquad (2)$$

where y denotes the actual class index, and $\hat{y}_y$ indicates the anticipated probability of the appropriate class. The class exhibiting the maximum probability score of $\hat{y}_i$ is subsequently determined as the final predicted label.

A. Dataset Construction

Publicly available datasets from Kaggle are frequently employed for model training and evaluation in machine learning research [25], [26], [27], [28]. In this study, the dataset was constructed by combining two sources: (1) the Padang Cuisine dataset from Kaggle [29] and (2) original photographs of Padang dishes captured by the authors. The Kaggle dataset initially contained 933 images across nine classes, serving as the primary source of data for this analysis. This combined dataset comprises a diverse mix of real-world photographs, promotional materials, and recipe illustrations, contributing to a wide range of visual characteristics, including background complexity, lighting variations, and textual elements. Rather than being treated as noise, this variability is leveraged to enhance model robustness. In real-world applications, especially mobile-based recognition systems, inputs may originate from a wide range of sources, such as printed menus, digital advertisements, or recipe websites. Including such heterogeneous samples helps the model generalize more effectively across different deployment scenarios [30].

The completed dataset consists of 1,108 images segmented into nine Padang dish classes: *Ayam goreng, ayam pop, daging rendang, dendeng batokok, gulai ikan, gulai tambusu, gulai tunjang, telur balado*, and *telur dadar*. Representative images from each class are presented in Fig. 2.

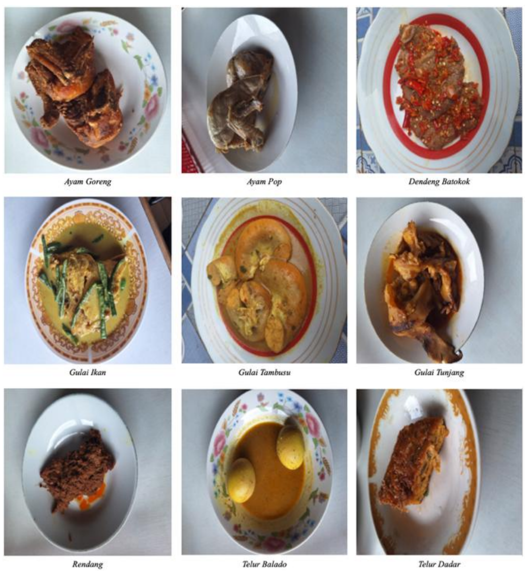


Fig. 2. Representative images of nine Padang dish classes in this study.

The images, which varied in resolution, were randomly partitioned into training (80%), validation (10%), and test (10%) sets. This split ratio was chosen to maximize the number of samples available for model training, to achieve robust feature learning in a relatively small dataset, while reserving independent and statistically meaningful sets for unbiased validation and final testing. With 1,108 total images, the 80:10:10 ratio division ensures that each class retains enough samples, around 12–18 images in the validation and test sets to provide reliable performance estimates, whereas larger validation and test proportions, such as 70:20:10, would reduce training data and potentially degrade model generalization [31].

Table 1.
Class-Wise Distribution of Images Across Training, Validation, and Test Sets for Padang Dish Classification

Padang dishes	Training	Validation	Test
Ayam goreng	100	12	12
Ayam pop	96	18	14
Dendeng batokok	95	12	12
Gulai ikan	96	16	12
Gulai tambusu	105	13	12
Gulai tunjang	99	12	12
Rendang	96	12	12
Telur balado	96	12	12
Telur dadar	96	12	12

Table 1 illustrates the class-wise distribution of images, which confirms a relatively balanced number of images across classes and partitions. This balance is crucial for preventing class bias and ensuring that all categories are equally represented in both training and evaluation. The partitioning was performed programmatically using a randomized allocation according to the specified ratio, ensuring that no image appeared in more than one subset and eliminating manual selection bias.

B. Data Preprocessing and Augmentation

All images were resized to a resolution of 224x224 pixels—a common specification for CNNs—and converted to PNG format. Although the dataset size was relatively modest, the class balance and visual quality were maintained to ensure fair representation for each class.

To enhance variability and reduce overfitting, on-the-fly augmentation was applied during training using the ImageGenerator module:

- Data augmentation for training included rescaling pixels to a range of 1/255, applying random rotations of up to 40 degrees, and introducing translational shifts of 0.2 in width and height, as well as shearing (0.2), zooming (0.2), and horizontal flipping transformations.
- For validation and testing, data only underwent pixel normalization through rescaling by 1/255.

These augmentation strategies increased data variability and improved the model's ability to generalize to unseen images, despite the modest dataset size. The effectiveness of this approach was demonstrated by the high classification performance, with the best model, ResNet152V2, achieving 95.37% validation accuracy and 95.45% test accuracy.

C. Model Architecture and Training

Four pre-trained CNN models were evaluated: EfficientNetV2M, MobileNetV2, VGG19, and ResNet152V2. EfficientNetV2M utilizes a compound scaling technique that involves concurrently adjusting the network's depth, width, and resolution. This approach contributes to superior performance while requiring fewer parameters than conventional models [32]. The model has good performance to be applied to various image classification tasks [33], [34]. MobileNetV2 uses depth-wise separable convolutions and inverted residuals to reduce computational complexity and memory consumption, yet still achieves high accuracy [35], [36]. VGG19 is known for its 19 layers, including 16 convolutional layers, three fully connected layers, and five max pooling layers, and has demonstrated strong performances in various image classification tasks [37], [38]. The ResNet152V2 architecture is a variant of the ResNet model, which incorporates residual blocks to help mitigate the vanishing gradient problem by

Applied Information System and Management (AISM)
Volume 8, (2) 2025, p. 285–296
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v8i2.46680

allowing gradients to flow through shortcut connections. This architecture has deep layers of 152, which are capable to learn complex features from images [39], [40]. The models used in this study have demonstrated strong performance in various image classification tasks. We compared these four models, and the best-performing one was then used as the basis for the mobile application.

Initialized with ImageNet weights, every pre-trained model was then refined on the Padang dish dataset. The input shape for this process was consistently set to (224, 224, 3). To balance feature reuse while still allowing model adaptation to the Padang dish dataset, a portion of the initial layers was frozen during training. The number of initial frozen layers for each model is listed in Table 2.

Table 2.
The Number of Initial Frozen Layers of The Four Transfer Learning Models

Model	Initial frozen layer
EfficientNetV2M	50
MobileNetV2	20
VGG19	4
ResNet152V2	50

The variation was intentional and model-dependent, based on:

1) *Depth and complexity of architecture*
Deeper models, such as ResNet152V2 and EfficientNetV2M, contain more layers, enabling greater abstraction of generic features in the early layers, making them more transferable and thus safer to freeze more extensively.
2) *Sensitive to low-level features*
Simpler models, such as VGG19, benefit from fine-tuning earlier layers, as their shallow depth limits feature abstraction, requiring more adaptation to the new dataset.
3) *Empirical performance*
The number of frozen layers was determined through preliminary experiments to balance generalization and task-specific learning, avoiding underfitting from freezing too many layers or overfitting from unfreezing too many.

A custom classifier block was appended, as illustrated in Fig. 3. It consisted of the following layers in order: a GlobalAveragePooling2D layer, followed by Dropout (rate = 0.3); a Dense layer featuring 256 units, ReLU activation, and L2 regularization; an additional Dropout layer at a 0.3 rate; and finally, a terminal Dense output layer containing 9 units and employing Softmax activation. Model compilation utilized the Adam optimizer, configured with a learning rate of $1e^{-5}$. Sparse categorical cross entropy was designated as the loss function, and accuracy was used as the principal evaluation metric.

The training process involved a batch size of 16 for a maximum of 100 epochs. To optimize training efficiency, an early stopping criterion was implemented, ceasing the process once both validation and training accuracy exceeded 95%, alongside model checkpointing based on validation performance.

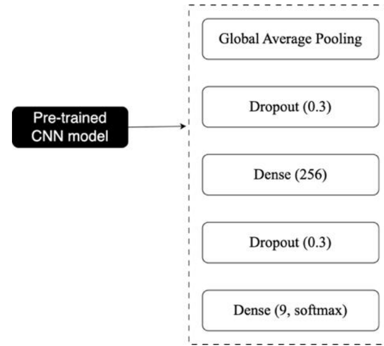


Fig. 3. Custom classifier block.

D. Model Evaluation and Optimization

Key performance indicators, specified in (3)–(6), consisted of accuracy, precision, recall, and F1-score, respectively.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (3)$$

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

$$F1\ Score = \frac{2 \times Precision \times Recall}{Precision + Recall} \quad (6)$$

where TP is true positive, TN is true negative, FP is false positive, and FN is false negative.

The best performing model, ResNet152V2, underwent hyperparameter tuning involving adjustments to dense layer units, dropout ratios, the application of L2 regularization, and additional dropout layers. Subsequently, the enhanced model was adapted to TensorFlow Lite format to facilitate mobile deployment, ensuring that performance remained largely uncompromised.

E. Development of the Mobile Application

The mobile application was developed using Android Studio and Kotlin, integrating the optimized TensorFlow Lite model for on-device inference. The TensorFlow Lite Android support library was used to embed the model, enabling low-latency and memory-efficient predictions.

Users can classify food images by selecting from their gallery or capturing a new image using the device's camera. Before inference, images are automatically resized and normalized to match the model's expected input dimension (224×224 pixels). The mobile application outputs the predicted Padang dish name, the corresponding Padang dish image, a brief description of the dish, a list of ingredients, the recipe, and the top three predicted probabilities.

To enhance the user experience, a confidence threshold mechanism is implemented; if the model's confidence score is below 90%, the mobile application prompts the user to retake or select an alternative image.

IV. RESULTS

This section presents a comparative analysis of four pre-trained CNN models to identify the optimal architecture for classifying Padang cuisine. The discussion critically reflects on model behaviour, supported by performance metrics, visualizations, and insights from prior theoretical research.

A. Model Benchmarking and Training Performance

Table 3 compares the training and validation performance of each model. ResNet152V2 outperformed all other models, recording the maximum validation accuracy of 95.80% and the minimum validation loss of 0.6418. This superior performance reflects its exceptional generalization ability, which is attributed to its architectural depth and residual learning capability. Although VGG19 showed high training accuracy of 98.03%, its elevated validation loss of 0.7765 indicates overfitting, which is consistent with its limited architectural regularization. While computationally efficient, MobileNetV2 lagged in both accuracy and loss, likely due to its reduced representational capacity.

Table 3.
Comparative Performance Analysis of The Four Models

Model	Train Acc	Val Acc	Train Loss	Val Loss
EfficientNetV2M	0.9730	0.9074	0.4998	0.7426
MobileNetV2	0.9545	0.9328	0.5747	0.7064
VGG19	0.9803	0.9074	**0.2935**	0.7765
ResNet152V2	**0.9534**	**0.9580**	0.5694	**0.6418**

Figure 4–7 illustrate the respective training and validation accuracy and loss trends observed for each model. These plots help illustrate the convergence behavior and generalization of each model. As shown in Fig. 4, EfficientNetV2M's learning curve exhibited gradual performance gains, with validation accuracy starting at 12.96% in epoch 1 and surpassing the 85% threshold by epoch 42. The model reached a peak validation accuracy of 92.59% with a loss of 0.7150 at epoch 54. After training on epoch 100, the model maintained a high validation accuracy of 90.74% and a stable loss of 0.7426.

Figure 5 demonstrates that MobileNetV2 showed rapid initial improvement, with validation accuracy climbing from 26.05% at epoch 1 to 63.03% by epoch 12. Its performance continued to grow, reaching a peak validation accuracy of 94.96% at epoch 90. After this point, accuracy plateaued and then slightly decreased to a final value of 93.28% at epoch 100, suggesting stable training near its peak performance.

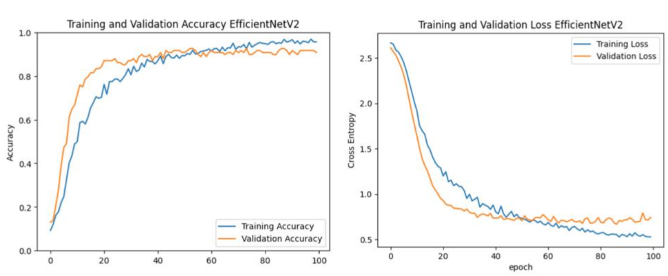


Fig. 4. Training and validation accuracy and loss curves for EfficientNetV2M.

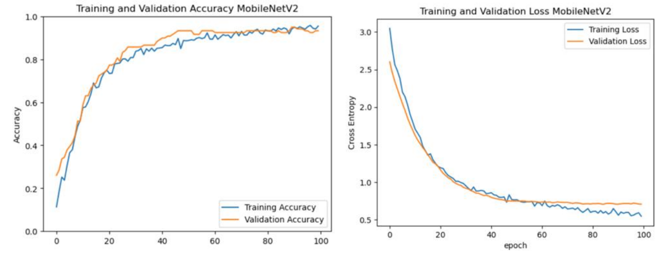


Fig. 5. Training and validation accuracy and loss curves for MobileNetV2.

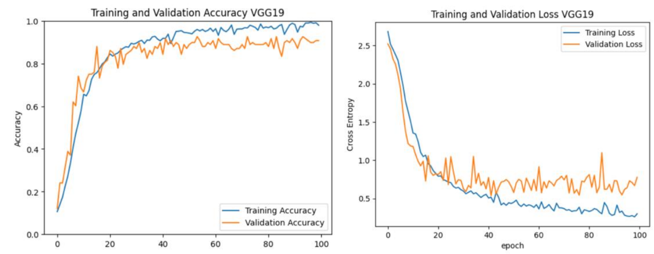


Fig. 6. Training and validation accuracy and loss curves for VGG19.

Figure 6 demonstrates that VGG19 showed fluctuations post-epoch 32, indicating saturation and a risk of overfitting. The validation accuracy of VGG19 peaked at 91.67% at epoch 42, with a low validation loss of 0.5364. However, by epoch 100, the validation accuracy had slightly dropped to 90.74% while the validation loss increased to 0.7765, providing clear quantitative of this saturation.

In contrast, Fig. 7 shows ResNet152V2 displayed smooth convergence, steady loss reduction, and no signs of overfitting. It achieved a good early performance, crossing the 90% validation accuracy threshold at epoch 16 and reaching a peak of 96.64% at epoch 36. Training was stopped early at epoch 38 due to the early stopping criteria, with a final validation accuracy of 95.80% and a low loss of 0.6418. This early stopping affirms the capacity of ResNet152V2 for fine-grained visual tasks and its generalization ability.

The overall findings from this detailed analysis highlighted that ResNet152V2's architectural depth and regularization mechanism contributed to better generalization, outperforming the other models in both performance and training efficiency.

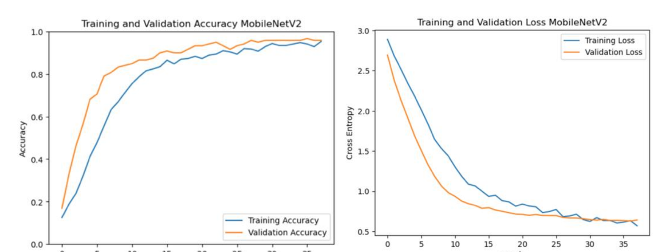


Fig. 7. Training and validation accuracy and loss curves for ResNet152V2.

B. Test Performance and Class-wise Evaluation

Table 4 summarizes test set performance. ResNet152V2 achieved the highest test accuracy of 95.45%, outperforming other CNN architectures, including MobileNetV2 (93.64%), EfficientNetV2M (91.67%), and VGG19 (90.74%). These results confirm that ResNet152V2 is capable of robustly predicting new data and is well-suited for the fine-grained

Applied Information System and Management (AISM)
Volume 8, (2) 2025, p. 285–296
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v8i2.46680

classification task of Padang cuisine.

Table 4.
Test accuracy of CNN models

Model	Test Accuracy
EfficientNetV2M	0.9167
MobileNetV2	0.9364
VGG19	0.9074
ResNet152V2	**0.9545**

To further understand model behavior and misclassification trends, a confusion matrix was generated for the ResNet152V2 model, as shown in Fig. 8. This matrix illustrates the distribution of predictions across true and predicted labels, highlighting that most classes were classified correctly with minimal confusion. Notably, minor misclassifications occurred between visually similar dishes such as *ayam goreng* and *rendang*, and between *gulai ikan* and *gulai tunjang*. These cases suggest areas where additional data augmentation or visual attention techniques could be employed to improve discrimination.

To quantitatively assess the performance specific to each class, Table 5 illustrates the calculated precision, recall, and F1-score metrics for each dish class. Notably, the model attained a perfect F1-score for three out of nine classes: *Ayam pop, dendeng batokok*, and *telur balado*, demonstrating excellent consistency. However, slightly reduced scores for *ayam goreng, rendang*, and *gulai tunjang*, indicate that fine-grained visual differences may lead to occasional misclassification.

Overall, the performance analysis reaffirms the effectiveness of ResNet152V2 for fine-grained food classification. The combination of high-test accuracy, strong class-wise metrics, and confusion matrix insights validates the model's robustness and real-world applicability. Despite overall strong performance, the results suggest opportunities for improving class separability using advanced augmentation or attention mechanisms.

Table 5.
Class-wise Evaluation Metrics for ResNet152V2 Model

Class	Precision	Recall	F1-Score
Ayam Goreng	0.8462	0.9167	0.8000
Ayam Pop	1.0000	1.0000	1.0000
Daging Rendang	0.9167	0.9167	0.9167
Dendeng Batokok	1.0000	1.0000	1.0000
Gulai Ikan	1.0000	0.9167	0.9565
Gulai Tambusu	0.9231	1.0000	0.9600
Gulai Tunjang	0.9167	0.9167	0.9167
Telur Balado	1.0000	1.0000	1.0000
Telur Dadar	1.0000	0.9167	0.9565

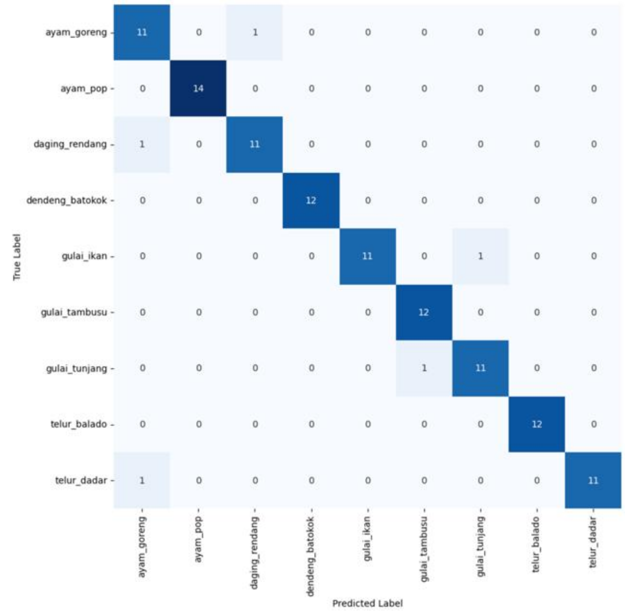


Fig. 8. Confusion matrix for ResNet152V2 model.

C. Hyperparameters Tuning

To enhance generalization and reduce overfitting, the ResNet152V2 model was fine-tuned using KerasTuner with a random search strategy. The tuning process was executed over 5 trials, each running for a maximum of 100 epochs. The hyperparameters optimized included the dropout rate, the number of units in the dense layer, the L2 regularization coefficient, and whether to add additional dropout layers. The optimal values obtained from this process are presented in Table 6.

Table 6.
Optimal Hyperparameters for ResNet152V2 Model

Hyperparameter	Value
Dropout rate (layer 1)	0.2
Dense layer units	832
L2 regularization coefficient	0.00010800795
Additional dropout layer required	Yes
Dropout rate (additional layer)	0.1

This configuration was chosen for its balance of regularization and representational power, reducing overfitting while maintaining classification performance.

D. Performance of the Optimized Model

The optimized model underwent retraining employing the Adam optimizer, configured with a learning rate of $1e^{-5}$. Sparse categorical cross-entropy loss function, and the accuracy as the performance indicators. A maximum of 100 epochs was set for training, alongside the implementation of early stopping, which ceased the process if the validation loss showed no reduction for five consecutive epochs.

Following hyperparameter tuning, the optimized ResNet152V2 configuration sustained high validation accuracy while achieving a reduction in validation loss, indicating improved convergence stability and better generalization control. The final optimized model achieved 99.32% training accuracy, 93.28% validation accuracy, a training loss of 0.1251, and a validation loss of 0.4142, as shown in Table 7 and result of class-wise evaluation metric is shown in Table 8. On the test set, the model reached 91.82% accuracy, confirming robust performance across unseen data.

Table 7.
Performance Metrics for The Optimized ResNet152V2 Model

Metric	Value
Training accuracy	0.9932
Validation accuracy	0.9328
Training loss	0.1251
Validation loss	0.4142
Test accuracy	0.9182

Table 8.
Class-Wise Evaluation Metrics for Optimized ResNet152V2 Model

Class	Precision	Recall	F1-Score
Ayam Goreng	0.7857	0.9167	0.8462
Ayam Pop	0.9333	1.0000	0.9655
Daging Rendang	1.0000	0.8333	0.9091
Dendeng Batokok	1.0000	1.0000	1.0000
Gulai Ikan	1.0000	0.8333	0.9091
Gulai Tambusu	0.8000	1.0000	0.8889
Gulai Tunjang	0.8182	0.7500	0.7826
Telur Balado	1.0000	1.0000	1.0000
Telur Dadar	1.0000	0.9167	0.9565

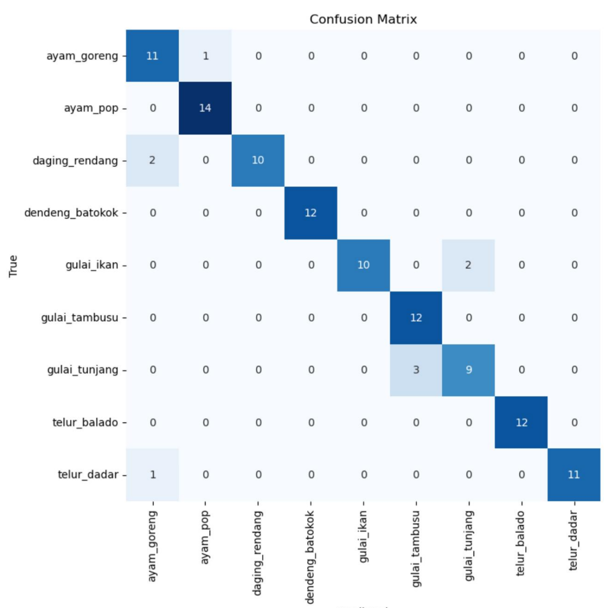


Fig. 9. Confusion matrix for the optimized ResNet152V2 model.

Class-wise evaluation revealed that precision, recall, and F1-scores remained well-balanced across categories, with no single class dominating the error distribution (Fig. 8 and 9). The confusion matrix corroborated this by showing that misclassifications were relatively diffuse rather than concentrated in specific classes. This uniformity is particularly important in fine-grained classification tasks, where class imbalances or systematic bias toward certain categories can compromise real-world utility. In practical deployment scenarios, consistent per-class performance reduces the risk of model blind spots, thereby improving both reliability and trustworthiness.

Overall, the optimization process not only preserved the strong predictive performance established in the pre-optimization phase but also yielded a model with greater stability, improved error distribution, and a performance profile better suited for reliable application in operational settings.

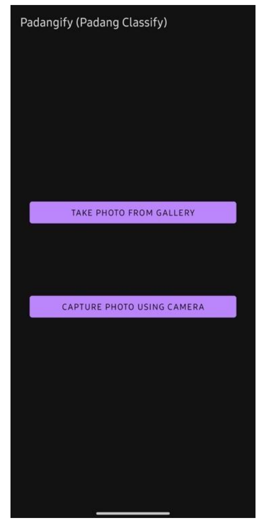


Fig. 10. Home screen of the mobile application.

E. Mobile Application Deployment

To support deployment in real-world scenarios, the optimized ResNet152V2 model was converted to TensorFlow Lite format, ensuring low latency and high accuracy in mobile environments. The mobile application allows users to input food images either by capturing them with the camera or selecting them from the gallery. Upon launch, the mobile application displays a home screen with options to upload or take a photo, as shown in Fig. 10. Once an image is selected, users are taken to a confirmation screen where they can verify the input image before prediction, as shown in Fig. 11.

After confirmation, the image is passed through the integrated TensorFlow Lite model for inference. If the prediction score is 0.9 or higher, the application displays the predicted dish name along with an image, a list of ingredients, and a recipe. If the confidence is below 0.9, a message prompts the user to provide another image. The top-3 prediction outcomes are illustrated in Fig. 12.

The deployment of the ResNet152V2 model using TensorFlow Lite enables robust, near real-time inference suitable for mobile environments. The model demonstrates responsive behavior, with predictions generated within a few seconds after image confirmation. The robustness of the classification was additionally confirmed through trials conducted with a smartphone camera under conventional lighting, which underscores the practical feasibility of this approach for food classification tasks.

The model's integration into the Android application enables automatic Padang cuisine recognition from

Applied Information System and Management (AISM)
Volume 8, (2) 2025, p. 285–296
P-ISSN: 2621-2536; E-ISSN: 2621-2544; DOI: 10.15408/aism.v8i2.46680

user-uploaded images, complete with ingredient lists and recipe steps. The mobile application offers a tangible tool for cultural preservation, culinary exploration, and potential integration into smart tourism or educational platforms.

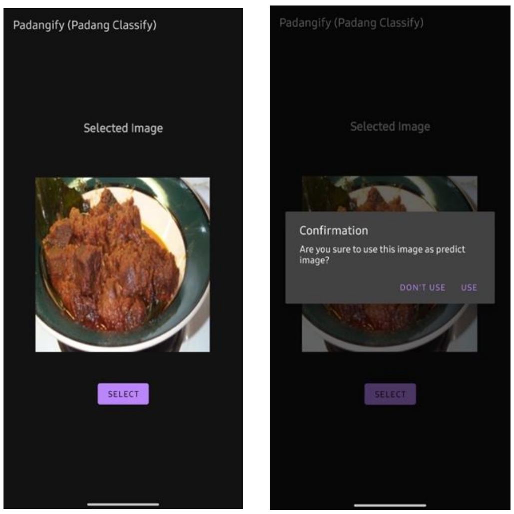


Fig. 11. Image selection (left) and confirmation screen (right).

F. Discussions

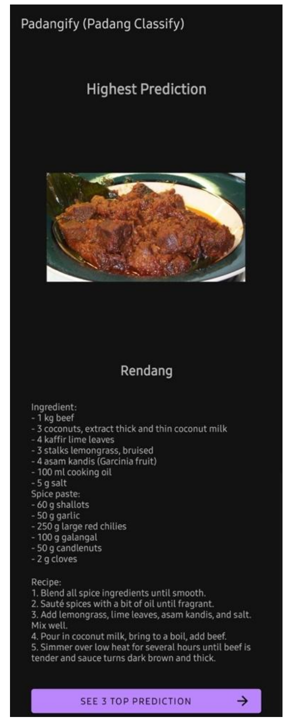
This study confirms the efficacy of transfer learning with deep CNN architectures for the fine-grained classification of Padang cuisine, a domain characterized by subtle inter-class visual differences that pose challenges even for advanced models.

While prior studies have predominantly addressed broad food categories or generalized cuisines [20], [21], [22], a notable gap remains in addressing region-specific dishes that exhibit high visual similarity and intra-class variability, such as those found in Padang cuisine. This niche, yet culturally significant, problem requires models capable of discerning subtle ingredients and presentation differences critical for accurate cultural documentation, culinary tourism, and AI-driven heritage preservation.

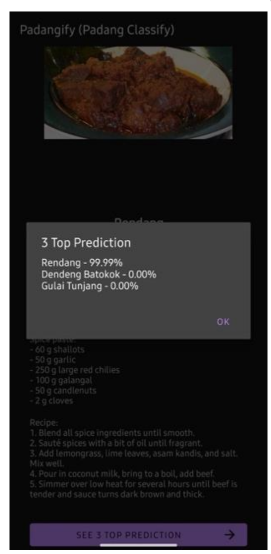
Experimental results demonstrated that when carefully tuned, deep CNNs can effectively capture these nuanced visual patterns. Among evaluated architectures, ResNet152V2 outperformed other models, achieving a test accuracy of 95.45% and improved generalization post-optimization with a test accuracy of 91.82% and validation loss of 0.4142.

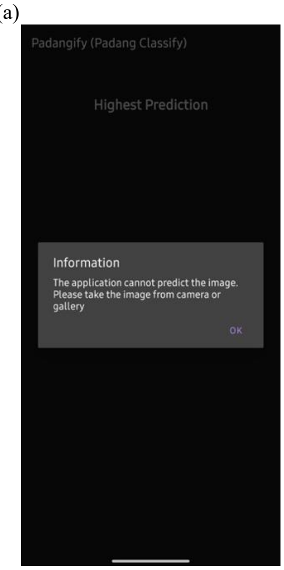
Misclassification predominantly occurred between visually similar dishes such as *gulai ikan* and *gulai tunjang*, reflecting intrinsic challenges in distinguishing closely related food items rather than model instability. In contrast, distinctly different dishes such as *ayam pop, dendeng batokok*, and *telur balado*, were classified with near-perfect accuracy, underscoring the model's strength in distinguishing more separable classes.

This study's methodological approach, which involves comparing multiple architectures, including VGG19 and MobileNetV2, highlights the necessity of a carefully selected and tuned model. For example, VGG19 achieved a low training loss but a high validation loss suggested overfitting, while the balanced performance of MobileNetV2 with smooth convergence and moderate generalization implied potential, but was outperformed by ResNet152V2.



(a)





(b) (c)

Fig. 12. Prediction result: (a) Prediction success, (b) top-3 prediction and (c) Prediction failure.

ResNet152V2, coupled with KerasTuner-based optimization, achieved greater stability and reduced validation loss, overcoming the limitations of generic training frameworks. common in many earlier studies [22], may be insufficient for fine-grained, real-world food classification.

The results demonstrated that deep learning models like ResNet152V2 were capable of capturing many of the subtle visual cues required to differentiate similar regional dishes, thereby providing practical proof of the model's suitability for culinary heritage preservation. However, some misclassifications among visually similar classes highlight inherent challenges and indicate that current models still face limitations in fully resolving all subtle intra-class differences. These findings suggest that while deep learning offers a promising foundation, future work should explore multi-modal data integration and advanced feature extraction techniques to further enhance discrimination in complex fine-grained culinary domains.

From a practical perspective, integrating the optimized ResNet152V2 into a TensorFlow Lite-based Android app highlights its real-world applicability. The model's low inference latency enables efficient on-device recognition suitable for field use, where internet access may be limited. This direct integration bridges academic research with applied technology, offering a scalable tool for promoting local cuisine awareness. The use of a mixed-source dataset by combining web-sourced images with original photos further enhances robustness, ensuring robustness across varied image qualities and contexts, directly addressing limitations in previous studies that relied heavily on studio-quality or limited datasets [20], [21].

These contributions extend beyond prior research by explicitly tackling dataset variability, fine-grained regional cuisine classification, and real-world mobile deployment, thereby addressing key gaps in the literature. This study can be considered as a substantive advancement in both the scope and practical applicability of automated regional food recognition.

G. Research Implication

This study advances both theoretical understanding and practical applications in fine-grained food recognition, with a specific focus on regional cuisine datasets that often favor of broader classification.

From a theoretical perspective, the benchmark evaluation establishes that for complex, localized visual tasks with subtle inter-class differences, deeper architectures such as ResNet152V2 can deliver superior generalization when paired with targeted hyperparameter tuning. This refines existing knowledge on model selection for cultural heritage datasets, demonstrating that depth and careful optimization are essential to prevent overfitting while preserving fine-grained visual cues. Furthermore, the use of mixed-source datasets strengthens theoretical models of robustness, showing that training with heterogeneous image sources improves performance in unconstrained, real-world settings.

From a practical standpoint, the development of the optimized ResNet152V2 in a TensorFlow Lite-based mobile application provides a working blueprint for translating research into an accessible, real-time recognition tool. This has direct implications for culinary tourism, cultural preservation, and digital heritage initiatives, as it enables on-device recognition without internet dependency. The methodological framework, spanning model selection, hyperparameter tuning, and mobile deployment, offers a transferable template for other fine-grained classification problems, ensuring that the approach can be adapted across diverse cultural and scientific domains. In doing so, this study bridges the gap between theoretical model development and socially meaningful, user-centric applications of artificial intelligence.

V. Conclusion

This study successfully automated the classification of traditional Indonesian Padang cuisine using a deep learning approach, thereby addressing a complex fine-grained image recognition challenge. Among the tested models, ResNet152V2 showed the best overall performance, demonstrating strong accuracy and balanced evaluation metrics across multiple classes. Class-wise analysis and the confusion matrix revealed that while some dishes were recognized with high accuracy, others remained challenging due to subtle visual similarities and the inherent limitations of the dataset's scope.

The optimized and trained model was successfully deployed as a mobile application using TensorFlow Lite, enabling real-time recognition with a latency of 1–5 seconds per image. To evaluate its usability, a post-deployment survey was conducted with 25 participants aged 18–30. The application received high average scores: 4.6 out of 5 for interface clarity, 4.5 for recognition accuracy, and 4.9 for processing speed. All respondents agreed that the application was effective in identifying Padang dishes. Suggestions included adding a usage tutorial, expanding the dish database, and providing short descriptions for each recognized dish before displaying ingredients and recipes.

Future work will address the observed limitations by integrating multi-modal data (such as image features and menu descriptions), employing attention-based mechanisms to enhance fine-grained feature extraction, applying advanced augmentation strategies, and expanding the dataset to encompass a broader range of Indonesian regional cuisines. Additionally, incorporating user feedback will improve the interface and overall usability, thereby enhancing real-world engagement.

In conclusion, this research lays the groundwork for intelligent culinary recognition systems that bridge artificial intelligence with local heritage. By delivering a scalable, real-time mobile application, this work advances technology in fine-grained visual recognition, demonstrating its applicability in culturally significant contexts. The contributions support cultural preservation, educational tools, and the promotion of Indonesian culinary tourism through accessible and innovative technology.

References

[1] UNESCO, "Basic texts of the 2003 Convention for the Safeguarding of the Intangible Cultural Heritage," Paris, 2018. Accessed: Mar. 10, 2025. [Online]. Available: https://digitallibrary.un.org/record/3841845?ln=en&v=pdf

[2] M. Monova-Zheleva, Y. Zhelev, and E. Nikolova, "Intangible cultural heritage presentation and preservation – Challenges and opportunities for museum specialists," *Digital Presentation and Preservation of Cultural and Scientific Heritage*, vol. 10, pp. 233–240, Sep. 2020, doi: 10.55630/dipp.2020.10.19.

[3] H. Yu, "Application model construction of intangible cultural heritage dissemination in film and television art creation in postepidemic era," *Wirel Commun Mob Comput*, vol. 2022, pp. 1–10, Sep. 2022, doi: 10.1155/2022/5448246.

[4] Q. Fan, C. Sun, and M. Zhang, "Research on the knowledge organization of intangible cultural heritage spatiotemporal data from a digital humanities perspective," *Knowledge Organization*, vol. 50, no. 8, pp. 526–541, 2023, doi: 10.5771/0943-7444-2023-8-526.

[5] T. Pistola *et al.*, "Creating immersive experiences based on intangible cultural heritage," in *2021 IEEE International Conference on Intelligent Reality (ICIR)*, IEEE, May 2021, pp. 17–24. doi: 10.1109/ICIR51845.2021.00012.

[6] D. Situmorang, F. Lubis, and F. Hakiem, "Strategi gastrodiplomasi indonesia melalui program indonesia spice up the world untuk meningkatkan nation branding indonesia di australia pada tahun 2022," *Jurnal Ilmiah Wahana Pendidikan*, vol. 10, no. 24.2, pp. 148–156, Dec. 2024, [Online]. Available: https://jurnal.peneliti.net/index.php/JIWP/article/view/11284

[7] I. Prastowo, A. A. Nurusman, H. K. E. P. Moro, Rizkianti, and C. Dewi, "Diversity of Indonesian offal-based dishes," *Journal of Ethnic Foods*, vol. 10, no. 1, Art. no. 15, Jun. 2023, doi: 10.1186/s42779-023-00181-8.

[8] P. Arsil, H. Le Dang, R. Wicaksono, and A. Hardanto, "Determinants of consumers' motivation towards ethnic food: evidence from Indonesia," *British Food Journal*, vol. 124, no. 10, pp. 3183–3200, Aug. 2022, doi: 10.1108/BFJ-05-2021-0605.

[9] I. Yuliana, W. I. F. Ningsih, and D. M. Sari, "Eksplorasi dan pengenalan makanan khas jawa dan sumatra pada generasi milenial melalui pembelajaran praktikum mata kuliah dasar kuliner," *Bubungan Tinggi: Jurnal Pengabdian Masyarakat*, vol. 4, no. 2, pp. 593–599, Jun. 2022, doi: 10.20527/btjpm.v4i2.5272.

[10] M. Nurmufida, G. H. Wangrimen, R. Reinalta, and K. Leonardi, "Rendang: The treasure of minangkabau," *Journal of Ethnic Foods*, vol. 4, no. 4, pp. 232–235, Dec. 2017, doi: 10.1016/j.jef.2017.10.005.

[11] T. Cheung, "Your pick: World's 50 best foods," CNN. Accessed: May 02, 2025. [Online]. Available: https://edition.cnn.com/travel/article/world-best-foods-readers-choice/index.html

[12] S. Waheed and S. Kumar, "Preserving traditional recipes and methods in the culinary world," in *Global Sustainable Practices in Gastronomic Tourism*, IGI Global, 2024, pp. 429–452. doi: 10.4018/979-8-3693-7096-4.ch025.

[13] H. M. Ozturk, "Approaches to conservation of gastronomic heritage in the process of industry 4.0 and climate change," in *Promoting Sustainable Gastronomy Tourism and Community Development*, 2024, pp. 95–120. doi: 10.4018/979-8-3693-1814-0.ch007.

[14] H. S. Nogay, N. H. Nogay, and H. Adeli, "Image-based food groups and portion prediction by using deep learning," *J Food Sci*, vol. 90, no. 3, Art. no. e70116, Mar. 2025, doi: 10.1111/1750-3841.70116.

[15] A. H. Rangkuti, J. M. Kerta, R. Y. Mogot, and V. H. Athala, "Identification of Indonesian traditional foods using machine learning and supported by segmentation methods," *JOIV: International Journal on Informatics Visualization*, vol. 8, no. 4, p. 2324, Dec. 2024, doi: 10.62527/joiv.8.4.2545.

[16] Q. Aini, F. H. Muhammad, E. Rustamaji, and Y. Thwe, "SISRES: Web-based culinary recommendation with collaborative filtering," *Applied Information System and Management (AISM)*, vol. 7, no. 1, pp. 1–8, Apr. 2024, doi: 10.15408/aism.v7i1.33798.

[17] F. Khairina Nst, I. Faisal, and K. Chiuloto, "Media pengenalan makanan khas daerah sumatera menggunakan teknologi augmented reality berbasis android," *ALGORITMA: Jurnal Ilmu Komputer dan Informatika*, vol. 06, no. 01, pp. 60–67, Apr. 2022.

[18] L. Alzubaidi *et al.*, "Review of deep learning: concepts, CNN architectures, challenges, applications, future directions," *J Big Data*, vol. 8, no. 1, p. 53, Mar. 2021, doi: 10.1186/s40537-021-00444-8.

[19] N. U. Gilal *et al.*, "Evaluating machine learning technologies for food computing from a data set perspective," *Multimed Tools Appl*, vol. 83, no. 11, pp. 32041–32068, Sep. 2023, doi: 10.1007/s11042-023-16513-4.

[20] A. Wibisono, H. A. Wisesa, Z. P. Rahmadhani, P. K. Fahira, P. Mursanto, and W. Jatmiko, "Traditional food knowledge of Indonesia: a new high-quality food dataset and automatic recognition system," *J Big Data*, vol. 7, no. 1, p. 69, Dec. 2020, doi: 10.1186/s40537-020-00342-5.

[21] M. A. Rasyidi, Y. S. Mardhiyyah, Z. Nasution, and C. H. Wijaya, "Performance comparison of state-of-the-art deep learning model architectures in Indonesian food image classification," *Bulletin of Electrical Engineering and Informatics*, vol. 13, no. 5, pp. 3355–3368, Oct. 2024, doi: 10.11591/eei.v13i5.7996.

[22] M. D. Darojat, Y. A. Sari, and R. C. Wihandika, "Convolutional neural network untuk klasifikasi citra makanan khas indonesia," *Jurnal Pengembangan Teknologi Informasi dan Ilmu Komputer*, vol. 5, no. 11, pp. 4764–4769, Oct. 2021, [Online]. Available: https://j-ptiik.ub.ac.id/index.php/j-ptiik/article/view/10096

[23] W. P. Kusumo and C. S. K. Aditya, "Klasifikasi citra makanan berdasarkan asal daerah menggunakan convolutional neural network," *Techno.Com*, vol. 23, no. 1, pp. 87–95, Feb. 2024, doi: 10.62411/tc.v23i1.9735.

[24] N. U. Gilal, K. Al-Thelaya, J. Schneider, J. She, and M. Agus, "SlowDeepFood: a food computing framework for regional gastronomy," in *Smart Tools and Apps for Graphics - Eurographics Italian Chapter Conference*, P. Frosini, D. Giorgi, S. Melzi, and E. Rodolà, Eds., The Eurographics Association, 2021. doi: 10.2312/stag.20211476.

[25] S. Priyadharshini and S. Shanthi, "A survey on detecting android malware using machine learning technique," in *2021 7th International Conference on Advanced Computing and Communication Systems (ICACCS)*, IEEE, Mar. 2021, pp. 1621–1627. doi: 10.1109/ICACCS51430.2021.9441712.

[26] K. S. Priya and R. D. Kumar, "Improving real time accuracy in image recognition for unmanned aerial vehicles (UAVS) with random forest algorithm comparing with support vector machine (SVM) algorithm," 2025, Art. no. 020142. doi: 10.1063/5.0270124.

[27] Moh. B. Tamam, H. Hozairi, M. Walid, and J. F. A. Bernardo, "Classification of sign language in real time using convolutional neural network," *Applied Information System and Management (AISM)*, vol. 6, no. 1, pp. 39–46, Apr. 2023, doi: 10.15408/aism.v6i1.29820.

[28] S. F. Mauludiah, Y. M. Arif, M. Faisal, and D. D. Putra, "Struggling models: An analysis of logistic regression and random forest in predicting repeat buyers with imbalanced performance metrics," *Applied Information System and Management (AISM)*, vol. 7, no. 2, pp. 31–38, Sep. 2024, doi: 10.15408/aism.v7i2.39326.

[29] F. F. Afrinanto, "Padang Cuisine (Indonesian Food Image Dataset)," 2022, *Kaggle*. doi: 10.34740/KAGGLE/DSV/4053613.

[30] I. Varlamis, "Messy data in education: enhancing data science literacy through real-world datasets in a master's program," *Educ Sci (Basel)*, vol. 15, no. 4, Art. no. 500, Apr. 2025, doi: 10.3390/educsci15040500.

[31] R. S. Mardiah and F. Fitrianingsih, "A comparative study of machine learning models for fashion product demand prediction: exploring algorithms, data splitting, and feature engineering," *Applied Information System and Management (AISM)*, vol. 8, no. 1, pp. 141–150, May 2025, doi: 10.15408/aism.v8i1.45600.

[32] J. P. S. Schuler, S. Romani, M. Abdel-Nasser, H. Rashwan, and D. Puig, "Grouped pointwise convolutions significantly reduces parameters in efficientnet," in *Frontiers in artificial intelligence and applications*, 2021, doi: 10.3233/faia210158.

[33] M. Di, Z. Peng, and S. Chen, "Performance comparison and application of deep-learning-based image recognition models," in *Second International Conference on Big Data, Computational Intelligence, and Applications (BDCIA 2024)*, Mar. 2025, doi: 10.1117/12.3059903.

[34] S. Cyriac, N. Raju, and S. Ramaswamy, "Comparison of full training and transfer learning in deep learning for image classification," in *Lecture notes in networks and systems*, 2021, pp. 58–67, doi: 10.1007/978-981-16-4486-3_6.

[35] R. Patel and A. Chaware, "Quantizing MobileNet Models for Classification Problem," in *2021 8th International Conference on Computing for Sustainable Global Development (INDIACom)*, 2021, pp. 348–351.

[36] N. A. A. Halim and A. W. Ismail, “Comparison on the metaverse space development in Spatial.io platform and Monaverse platform,” in *Lecture notes in networks and systems*, 2024, pp. 21–36, doi: 10.1007/978-981-97-2004-0_2.
[37] R. Ravindaran *et al.*, “Performance aalysis of a VGG based deep learning model for classification of weeds and crops,” in *2023 14th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, IEEE, Jul. 2023, pp. 1–5, doi: 10.1109/ICCCNT56998.2023.10307169.
[38] M. Shaha and M. Pawar, “Transfer learning for image classification,” in *2018 Second International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, IEEE, Mar. 2018, pp. 656–660, doi: 10.1109/ICECA.2018.8474802.
[39] M. Sangani, J. Gala, C. Patil, and S. Sankhe, “ResNet models in real world action: a comparative study on tomato disease recognition,” in *Technologies for Energy, Agriculture, and Healthcare*, London: CRC Press, 2024, pp. 536–545, doi: 10.1201/9781003596707-55.
[40] S. S. Mahamud, K. N. Ayve, A. H. Uddin, and A. S. M. Arif, “Tomato Leaf Disease Recognition with Deep Transfer Learning,” in *Lecture notes in networks and systems*, 2022, pp. 203–211, doi: 10.1007/978-981-19-4052-1_22.