Journal of Applied Data Sciences
Vol. 6, No. 4, December 2025, pp. 2907-2920

ISSN 2723-6471
2907

Classification of Batak Toba Ulos Motifs Based on Transfer Learning with MobileNetV2

Tonni Limbong[1,*], Gonti Simanullang[2], Parasian D.P. Silitonga[3], Donalson Silalahi[4]

[1]Department Information System, Faculty of Computer Sciences, Universitas Katolik Santo Thomas, Medan 20135, Indonesia

[2]Department Information technology, Faculty of Computer Sciences, Universitas Katolik Santo Thomas, Medan 20135, Indonesia

[3]Department philosophy, Faculty of philosophy, Universitas Katolik Santo Thomas, Medan 20135, Indonesia

[4]Department Master of Management, Faculty of Economics and Business, Universitas Katolik Santo Thomas, Medan 20135, Indonesia

(Received: March 20, 2025; Revised: June 15, 2025; Accepted: September 28, 2025; Available online: October 22, 2025)

Abstract

Indonesia possesses a rich cultural heritage, including the traditional Batak Toba Ulos textile, which is known for its diverse motifs and deep philosophical meanings. However, the preservation and visual recognition of Ulos remain challenging, particularly in terms of systematic documentation and automated classification. This study presents a visual recognition system for Batak Toba Ulos motifs using a transfer learning approach based on the MobileNetV2 architecture. The methodology involves the construction of a curated dataset of Ulos images, the application of data augmentation and preprocessing techniques, and model training utilizing ImageNet pre-trained weights. The system's performance was evaluated using accuracy, precision, recall, and F1-score metrics. Results show that the model is capable of accurately classifying all 12 Ulos classes, achieving F1-scores ranging from 0.93 to 0.97. These findings demonstrate that transfer learning is effective in overcoming the limitations of culturally specific, small-scale datasets. This research contributes to the development of artificial intelligence tools for cultural preservation and supports the digital documentation and promotion of Batak Toba Ulos to younger generations and broader audiences in an efficient and scalable manner.

Keywords: Batak Toba Ulos, Traditional Motifs, Image Classification, Transfer Learning, MobileNetV2, Cultural Preservation

1. Introduction

Indonesia is one of the most culturally diverse countries in the world, where traditional art plays an important role in representing local identity and values [1]. One example of this cultural richness is the traditional Batak Toba ulos cloth, which not only has aesthetic value but also carries deep philosophical, social, and spiritual meanings [2]. Historically, ulos has been integrated into various aspects of Batak Toba society, from traditional ceremonies and symbols of social status to expressions of Batak kinship values and cosmology. Each type of ulos has unique motifs, colors, and placements according to its meaning and function, ranging from Ragidup Ulos, Mangiring Ulos, to Sadum Ulos. However, amid the rapid pace of modernization and globalization, this traditional textile heritage faces the risk of losing its value and form due to the lack of digital documentation and the absence of an adaptive visual identification system [3]. The process of identifying and classifying ulos types manually still relies heavily on local knowledge and subjective experience, making it difficult to perform consistently, systematically, and sustainably. Therefore, a technology-based approach utilizing digital image processing and machine learning is needed to automatically, efficiently, and accurately recognize and classify the various types of ulos fabric.

Various studies have utilized image processing and machine learning for visual cultural preservation, but most still focus on scripts, paintings, or cultural monuments in general. [4] developed a Batak Toba script recognition system using Siamese Neural Network and one-shot learning to overcome data limitations, but it is still limited to letters and does not yet cover traditional textiles such as ulos. [5] proposed an ethnomathematics approach to generate Batak

*Corresponding author: Tonni Limbong (tonni.budidarma@gmail.com)

DOI: https://doi.org/10.47738/jads.v6i4.1036

ornamental patterns through symmetry-based geometric transformations, but this approach was not accompanied by machine-based classification processes. [6] introduced ResNet-NTS with triple attention to recognize painting styles with high accuracy, but its scope is still limited to Western and Central Asian artworks. [7] developed a mobile application for recognizing Chongqing's intangible cultural heritage using CNN and transfer learning, but it does not cover traditional textile objects from Indonesia. Meanwhile, [8] proposed the Vision Transformer (ViT-HVE) to evaluate cultural value based on deep learning in the Yellow River region, and [9] compared CNN and Transformer architectures for cultural objects, but the focus remains predominantly on monuments, not non-monumental textiles like Batak Toba ulos.

Previous studies have demonstrated the potential of machine learning technology in supporting the preservation of visual-based culture, but to date, no deep learning-based automatic classification system specifically designed for recognizing Batak ulos types from real digital images has been developed. Most approaches remain focused on script or painting, and have not yet addressed the richly meaningful local textile heritage such as ulos. Additionally, visual recognition of ulos faces technical challenges such as lighting variations, camera angles, pattern rotations, and complex fabric textures, which cannot be addressed using conventional approaches. Therefore, this study proposes the development of a visual recognition system for traditional Batak ulos fabric using a lightweight and efficient Convolutional Neural Network (CNN) architecture based on MobileNetV2. Using a transfer learning approach, the model can leverage visual representations from the pre-trained ImageNet model, then be fine-tuned to recognize various ulos images. The dataset used consists of ulos images from original visual documentation, processed through normalization, rotation- and lighting-based augmentation, and input dimension adjustment. It is hoped that this system can assist in the digital identification and preservation of ulos fabric, as well as open up opportunities for further application in cultural education platforms, digital museums, and mobile applications.

2. Literature Review

The application of artificial intelligence, particularly deep learning, has increasingly been adopted in the field of cultural heritage preservation, offering advanced capabilities for visual recognition, documentation, and classification. However, the majority of existing studies have focused on cultural artifacts such as monuments, paintings, and scripts, while relatively few have explored the automatic recognition of traditional textiles, especially those originating from Indonesia.

One study proposed the use of the ResNet-NTS network with triple attention for the classification of painting styles and achieved high accuracy in identifying Western and Central Asian artworks [7]. Another study developed a mobile application utilizing CNN and transfer learning to identify intangible cultural heritage elements in the Chongqing region, focusing on sculptures and decorative objects rather than textile patterns [8]. These studies demonstrate the potential of deep learning in cultural classification tasks but do not address the complex structures of woven fabric motifs.

In Indonesia, efforts related to Batak cultural heritage have mostly concentrated on character recognition. A study employed a Siamese Neural Network combined with one-shot learning to recognize Batak Toba characters from limited data [9]. Another approach used a geometric transformation method to generate traditional Batak ornamental motifs based on mathematical symmetry, although it lacked an automated classification component [10]. These works, while important for textual and ornamental preservation, have not extended their focus to fabric-based cultural artifacts such as Ulos.

Elsewhere, deep learning frameworks have been applied to monumental heritage. A Vision Transformer-based model was introduced to evaluate cultural values in the Yellow River region, although its application was limited to large-scale artifacts rather than everyday cultural textiles [11]. Comparative research has also been conducted on CNN and Transformer architectures for classifying cultural objects, showing that while deeper models may yield strong performance, they tend to overfit when trained on datasets with limited samples [12]. This characteristic is particularly relevant to traditional textile datasets, which often contain a small number of high-complexity images.

While these studies highlight important developments in cultural recognition, they generally rely on heavy-weight architectures or domain-specific adaptations that are not optimized for deployment on limited data or in low-resource

ISSN 2723-6471

environments. In contrast, MobileNetV2, a lightweight convolutional neural network model designed for mobile and embedded applications, uses inverted residual blocks and depthwise separable convolutions to achieve high accuracy with significantly reduced computational cost [13]. Its effectiveness has been demonstrated in diverse domains including medical imaging, agricultural diagnostics, and industrial object recognition [14], [15], particularly in cases involving small datasets and complex visual features.

Despite these advantages, there is a noticeable gap in research applying MobileNetV2 to the classification of traditional Indonesian textile motifs. To the best of current knowledge, no existing study has utilized this architecture for recognizing Batak Toba Ulos, which are characterized by intricate geometric patterns and rich cultural symbolism. The present research addresses this gap by introducing a MobileNetV2-based transfer learning model trained on a curated dataset of Batak Toba Ulos motifs. This approach not only demonstrates the architectural suitability of MobileNetV2 for cultural textile recognition but also contributes to the digital preservation of indigenous heritage through a replicable and efficient classification framework.

3. Methods and Materials

3.1. Research Approach

This study uses an experimental quantitative approach with the Research and Development (R&D) method [10] to develop a machine learning-based intelligent classification model for Batak Toba ulos. The model was developed using a Convolutional Neural Network (CNN) architecture [11], [12] based on MobileNetV2 [13], [15], [16] due to its computational efficiency and good visual generalization capabilities, with the research stages shown in figure 1.

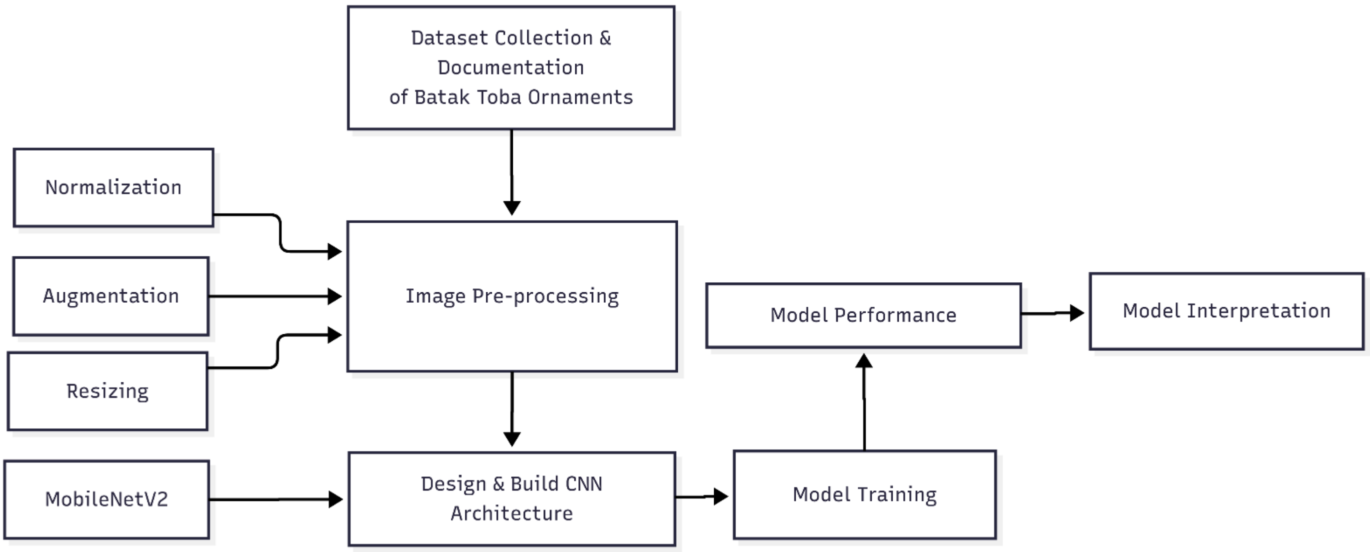

Figure 1. Research Procedure

The workflow begins with curated dataset collection and careful documentation of Batak Toba ornament images, ensuring each sample has clear class labels and provenance. A unified pre-processing pipeline then standardizes inputs: images are resized to the backbone's expected resolution, normalized to stabilize optimization, and augmented with label-preserving transforms to expand diversity and reduce overfitting. These steps produce clean, consistent tensors that flow into the chosen backbone—MobileNetV2—selected for its efficient inverted-residual architecture that delivers a strong accuracy–latency trade-off suitable for limited compute and edge deployment.

On top of the backbone, a lightweight classification head is designed and trained end-to-end using cross-entropy loss with standard practices such as learning-rate scheduling, early stopping, and checkpointing. After training, performance is quantified on a held-out split using accuracy, precision, recall, F1-score, and a confusion matrix to expose per-class behavior. Finally, interpretation methods (e.g., Grad-CAM and error analysis) are applied to visualize salient regions and diagnose failure modes, providing actionable insight to refine the dataset, tune augmentation, and iterate the model architecture.

3.2. Dataset

The dataset used in this study is a collection of Batak Ulos images that have been divided into three structured subsets, namely training, validation, and testing, with directory organization based on 12 different ulos classes as classification targets. The training subset consists of 2,221 images used to train the model, with the application of augmentation techniques to increase variation and generalization capabilities. The validation subset contains 399 images that have not undergone augmentation and are used to monitor model performance during the training process and detect possible overfitting. Meanwhile, the test subset includes 663 ulos images without augmentation, which are used for final evaluation and objective measurement of the model's overall performance. All images in the dataset have been resized to 224×224 pixels to meet the input requirements of the deep learning model used.

3.3. Model Architecture and Training

The model architecture used in this study adapts MobileNetV2 as the main model, with several adjustments for the task of classifying Batak Toba ornaments. Modifications were made to the output section by applying the softmax activation function according to the number of motif classes recognized. The model training process was carried out using a fine-tuning approach, which involves freezing approximately half of the initial layers of the default architecture and training the final half of the layers to adjust the model to the characteristics of the ornament data. For optimization, the Adam algorithm was used with an initial learning rate of 0.0001 to ensure stable convergence during training. The loss function applied is categorical cross-entropy, which is suitable for multi-class classification problems. Training was conducted for 20 epochs with a batch size of 32 to maintain a balance between efficiency and learning accuracy. During the training process, 20 percent of the dataset was used as a validation set to monitor model performance and avoid overfitting with the architecture in figure 2 below.

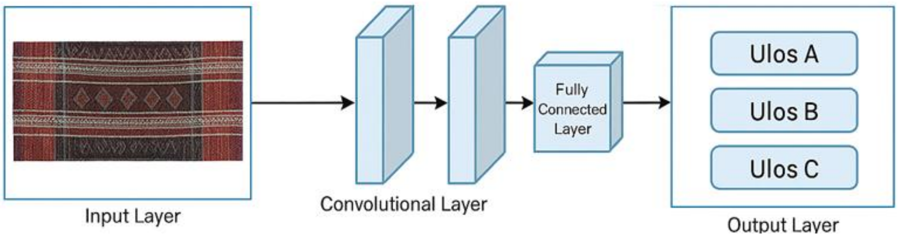

Figure 2. Architecture of the Training Model

3.4. Model Evaluation

Model evaluation is performed by measuring the percentage of correct predictions against the entire data set. Precision and recall are applied to evaluate the classification performance for each class specifically, while the F1-score is calculated as the harmonic mean of precision and recall to provide a balanced view of both [17], [18], [19]. A confusion matrix is used to analyze the distribution of classification results between classes [20].

3.5. Formulation of the Batak Toba Ulos Classification Model

The formulation of the Batak Toba Ulos classification model in this study was developed to utilize the visual characteristics of digital images in identifying distinctive patterns on Batak Toba Ulos fabrics. Image data representation is performed by converting digital images into numerical matrices of size H×W×C, where H is the image height, W is the image width, and C = 3 represents the three RGB color channels [21]. This structure becomes the main input in the feature extraction process by the machine learning model with the equation:

$$X \in \mathbb{R}^{(H\times W\times C)} \quad (1)$$

To ensure that the learning model formation process runs smoothly, pixel normalization is performed to adjust the intensity scale to the range [0,1], which is expressed as X_norm = X / 255, where X is the original image matrix and X_norm is the normalized image.

3.5.1. Augmentation

Augmentation is performed randomly to increase data diversity and model robustness against variations in image orientation and lighting [22]. The transformations applied include rotation, lighting, and horizontal flipping with the equation:

$$X' = \mathrm{rotate}(X_{norm}, \theta), \theta \in [-20^{o}, +20^{o}] \quad (2)$$

$$X'' = X' \cdot \alpha, \qquad \alpha \in [0.7, 1.3] \quad (3)$$

$$X''' = \mathrm{flip}(X'') \quad (4)$$

X_norm is the normalized image, θ is the random rotation angle, α is the lighting factor, and the flip operation is performed horizontally.

3.5.2. Convolution Layer

The convolution operation is a key component in the feature extraction process in convolutional neural networks. This process is performed by shifting the kernel on the image and calculating the corresponding element multiplication results, thereby producing a feature map that represents local patterns in the image [23]. The basic convolution operation is formulated in Formula (5)

$$S_{i,j}^{(k)} = (X * K^{(k)})_{i,j} = \sum_{m-0}^{M-1} \sum_{n-0}^{N-1} \sum_{c-1}^{C} X_{i+m,j+n,j} \cdot K_{m,n,c}^{(k)} \quad (5)$$

with X as the image input, K(k) as the k kernel, and S(k) as the convolution output at position (i,j). The values M and N indicate the kernel size, and C denotes the number of color channels.

3.5.3. Activation Function (ReLU)

The ReLU activation function is used to add non-linearity to the neural network. This function produces zero for negative inputs and retains the original value for positive inputs, formulated as $f(x) = \max(0, x)\mathrm{ReLU}$. ReLU is chosen because it is efficient and speeds up the model training process [24].

3.5.4. Depthwise Separable Convolution (MobileNetV2)

MobileNetV2 uses a depthwise separable convolution approach to improve feature extraction efficiency [25]. This process consists of two main stages: depthwise convolution and pointwise convolution. The first stage, depthwise convolution, applies filters separately to each input channel using formula (6)

$$Z_d^{(k)} = D^{(k)}(X) \quad (6)$$

where X is the image input or feature from the previous layer, D(k) is the convolution operation for the k filter on each channel independently. The result is $Z_d^{(k)}$. It is then further processed by pointwise convolution [26], [27], which is a 1×11 convolution to combine information between channels using formula (7).

$$Z_p^{(k)} = P^{(k)} Z_d^{(k)} \quad (7)$$

where P(k) is the kth convolution filter at the pointwise stage, and $Z_p^{(k)}$ is the result of feature combination. Next, the non-linear activation function f is applied to obtain the final output in formula (8).

$$Z^{(k)} = f Z_p^{(k)} \quad (8)$$

The output Z(k) is then passed on to the next layer in the network. This approach allows MobileNetV2 to generate feature representations efficiently, with much lower computational load than conventional convolutions, without sacrificing accuracy.

3.5.5. Feature Extraction to Loss Calculation

After the feature extraction process through the MobileNetV2 network, the classification stage begins with Global Average Pooling (GAP) [28]. The GAP function reduces the spatial dimensions of the extracted features to a single average value per feature channel [29] using equation (9).

$$z_k = \frac{1}{H \cdot W}\sum_{i=1}^{H}\sum_{j=1}^{W} S_{i,j}^{(k)} \tag{9}$$

H and W are the height and width of the feature, while $S_{i,j}^{(k)}$ is the activation value at position (i,j) for channel k. The result of GAP is then fed into the softmax activation function to generate the probability of each class:

$$\hat{y}_i = \frac{e^{zi}}{\sum_{j=1}^{C} e^{zi}} \quad \text{for } i = 1, \ldots, C \tag{10}$$

C is the number of classes, and $\hat{y}_i$ is the predicted probability for class i. To evaluate how well the predictions are, we use a categorical cross-entropy loss function as follows:

$$L = -\sum_{i=1}^{C} y_i \cdot \log(\hat{y}_i) \tag{11}$$

y_i is the actual label in one-hot encoding form, and $\hat{y}_i$ is the predicted probability. This process enables the model to learn the optimal representation for multi-class classification end-to-end.

3.5.6. Optimization with Adam Technique

Adam (Adaptive Moment Estimation) is an optimization algorithm widely used in deep learning because it combines the advantages of momentum and RMSProp [30]. In this process, the model weights are denoted as θ and updated using the loss gradient ∇L_t with a learning rate α. Adam calculates two types of exponential averages: m_t, which is the momentum of the gradient, and v_t, which is the momentum of the gradient squared. Both are controlled by decay coefficients β_1 and β_2, which are typically set to 0.9 and 0.999. To address initial bias, corrections are applied to produce $\hat{m}_t$ and $\hat{v}_t$.d using formula (12).

$$\theta_{t+1} = \theta_t - \alpha \cdot \frac{\hat{m}_t}{\sqrt{\hat{v}_t} + \epsilon} \tag{12}$$

where ε is a small constant (usually 1e-8) to prevent division by zero. Adam accelerates convergence and maintains model training stability, making it well-suited for lightweight architectures such as MobileNetV2.

4. Results and Discussion

4.1. Classification Model Architecture

In this study, the development of the Batak Toba Ulos motif recognition model was carried out using a Transfer Learning approach with a modified MobileNetV2 architecture. MobileNetV2 was chosen for its ability to balance model complexity and accuracy, as well as its efficiency in handling high-resolution image data that is limited in quantity, as is commonly found in the archiving of traditional cultural visuals. The model is designed to accept input in the form of 224 × 224-pixel images with 3 color channels (RGB), which are then processed through a series of lightweight convolution layers based on depthwise separable convolution. These layers are arranged in inverted residual blocks equipped with shortcut connections to maintain information continuity between layers. This structure enables the model to extract rich and contextual spatial features from the distinctive geometric motifs of Batak Toba Ulos, such as gorga, sihala, and simata ni ari. The features captured and processed by the convolutional layers are then summarized through a Global Average Pooling process, which aims to compress spatial information into a global feature representation without losing its essence. A Dropout layer is inserted afterward to reduce the risk of overfitting, given the limited amount of data. Finally, a Dense layer with a softmax activation function is used as the final classification to map the prediction results into the 12 predefined Batak Toba Ulos motif classes using the Batak Toba Ulos motif recognition structure in the following table 1.

Table 1. Deep Learning Model Layer Structure for Batak Toba Ulos Motif Classification

Layer	Output Shape	Parameters	Description
Input	(224, 224, 3)	0	RGB image input
Conv2D	(112, 112, 32)	864	Initial feature extraction
BatchNorm	(112, 112, 32)	128	Activation normalization
ReLU	(112, 112, 32)	0	Non-linear activation
DepthwiseConv2D	(112, 112, 32)	288	Convolution per channel
BatchNorm	(112, 112, 32)	128	Normalization
ReLU	(112, 112, 32)	0	Activation
Conv2D Project	(112, 112, 16)	512	Feature dimension reduction
BatchNorm	(112, 112, 16)	64	Normalization
Inverted Residual Block ×n	↓ spatial, ↑channel	1500000	Main MobileNetV2 convolution block
GlobalAvgPooling2D	(1280)	0	Spatial feature aggregation
Dropout	(1280)	0	Regularization
Dense (Softmax)	(12)	15372	class classification of Ulos motifs

4.2. Training Model

At this stage, the model is trained using the data prepared earlier. The goal is to optimize the weights in the network so that it can recognize patterns from the input data and produce accurate predictions. The training process lasts for 10 epochs, during which the training data and validation data are evaluated at each epoch to observe the progress of the model's performance.

Table 2. Model Performance per Epoch

Epoch	Accuracy (%)	Loss	Val Accuracy (%)	Val Loss	Time (s)
1	22.32	2.2907	26.57	2.1023	874
2	46.96	1.6533	48.12	1.6029	246
3	68.53	1.1803	57.89	1.2471	238
4	82.77	0.852	71.68	1.0093	252
5	88.16	0.6633	84.21	0.8304	246
6	92.0	0.5146	87.97	0.6952	252
7	95.11	0.4134	92.48	0.5845	238
8	96.36	0.3474	94.24	0.5085	240
9	96.6	0.2866	94.74	0.4563	254
10	97.91	0.2421	95.49	0.4061	239

Based on table 2, in the first epoch, the model showed low accuracy on the training data (22.32%) and validation data (26.57%), with high loss values (2.2907 for training and 2.1023 for validation), and a relatively long training time of 874 seconds. However, in the second epoch, there was a significant improvement, with training accuracy reaching 46.96% and validation accuracy 48.12%, accompanied by a consistent decrease in loss. The model's performance continued to improve in the third to fifth epochs, with validation accuracy increasing to 84.21% and validation loss decreasing to 0.8304, indicating that the model began to understand the patterns in the data more effectively. In the sixth to eighth epochs, training accuracy exceeded 90%, with validation accuracy reaching over 94%, and the continuously decreasing loss indicated a stable learning process without signs of overfitting. The training time per epoch was also relatively stable, ranging from 238 to 254 seconds. By the tenth epoch, the model achieved its highest accuracy of 97.91% on the training data and 95.49% on the validation data, with the lowest loss values (0.2421 and 0.4061). This indicates that the model successfully generalized well to new data. The accuracy per epoch graph is illustrated in figure 3 below:

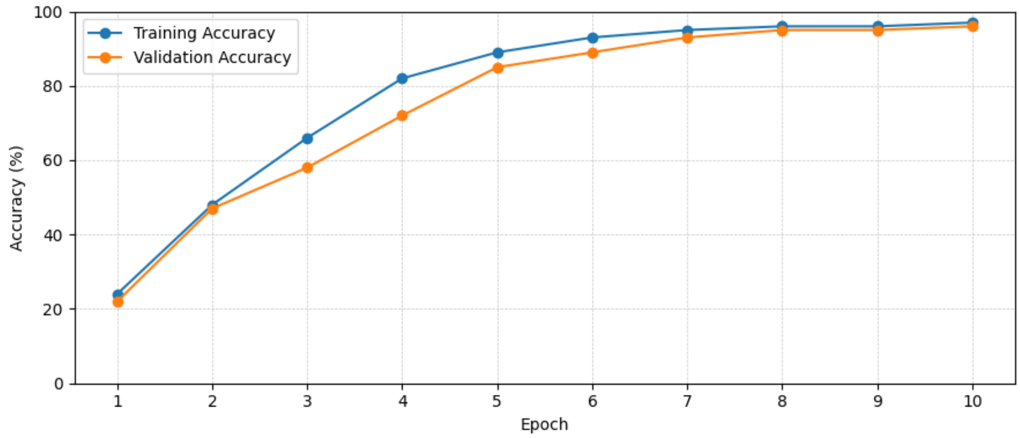

Figure 3. Accuracy Graph Vs Epoch

Based on figure 3, the model accuracy increased significantly during the 10 training epochs. At the beginning of training, the accuracy of the training and validation data was still low, but it increased sharply until the 5th epoch. Starting from the 6th epoch, the accuracy curve tended to stabilize, indicating that the model had reached convergence. At the end of training, the training accuracy reached 97.91% and the validation accuracy reached 95.49%, with a small difference, indicating that the model has good generalization and does not suffer from overfitting. The alignment of the two curves indicates an effective and stable training process, as well as the validity of the architecture and parameters used. The visualization of Loss Per Epoch is shown in figure 4.

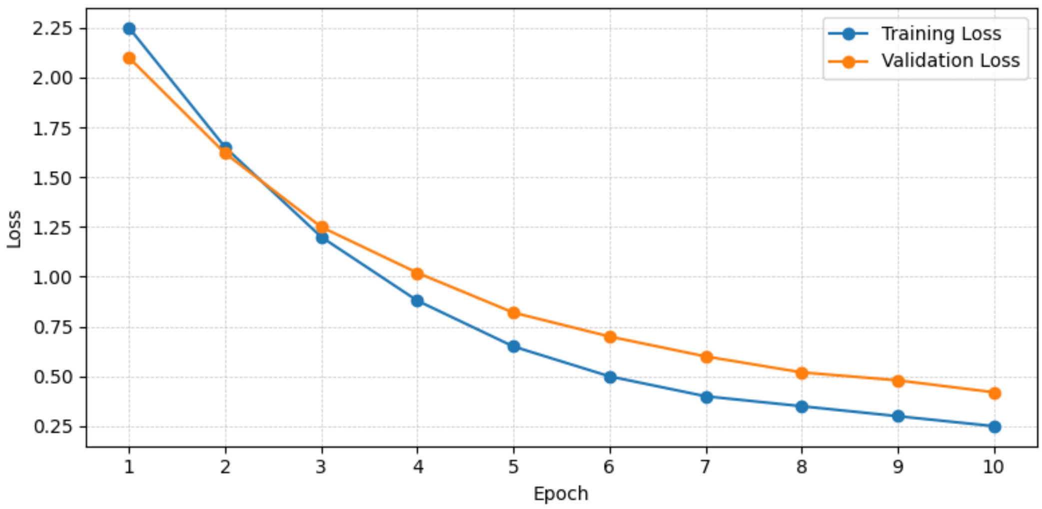

Figure 4. Graphic Loss Vs Epoch

Figure 4 shows a consistent decrease in loss in both training and validation data over 10 epochs. At the beginning of training, the loss value was still high (above 2.0), but it gradually decreased as the number of epochs increased. This decrease reflects the model's improved ability to minimize prediction errors. Starting from the 6th epoch, the rate of loss reduction began to slow down, indicating that the model was approaching convergence. At the end of training, the

ISSN 2723-6471

training loss reached 0.2421 and the validation loss reached 0.4061. The parallel trend of the graphs, which did not diverge from each other, indicates that the model did not experience overfitting and had stability in the learning process.

4.3. Model Evaluation Against the Dataset

The performance of the deep learning model was evaluated using several widely recognized classification metrics, namely precision, recall, and F1-score, for each individual class in the dataset. These metrics provide a comprehensive assessment of how accurately the model is able to identify and distinguish between the various types of Batak Toba Ulos motifs based on their manufacturing methods, which include both hand-weaving and machine-based production. The evaluation results clearly indicate that the model performs exceptionally well in learning and recognizing the intricate patterns associated with each class. This strong performance suggests that the model has successfully learned meaningful visual representations of the motifs during training, enabling it to generalize effectively to unseen validation data. To further understand the distribution of the model's predictions, a confusion matrix was generated and is presented in figure 5.

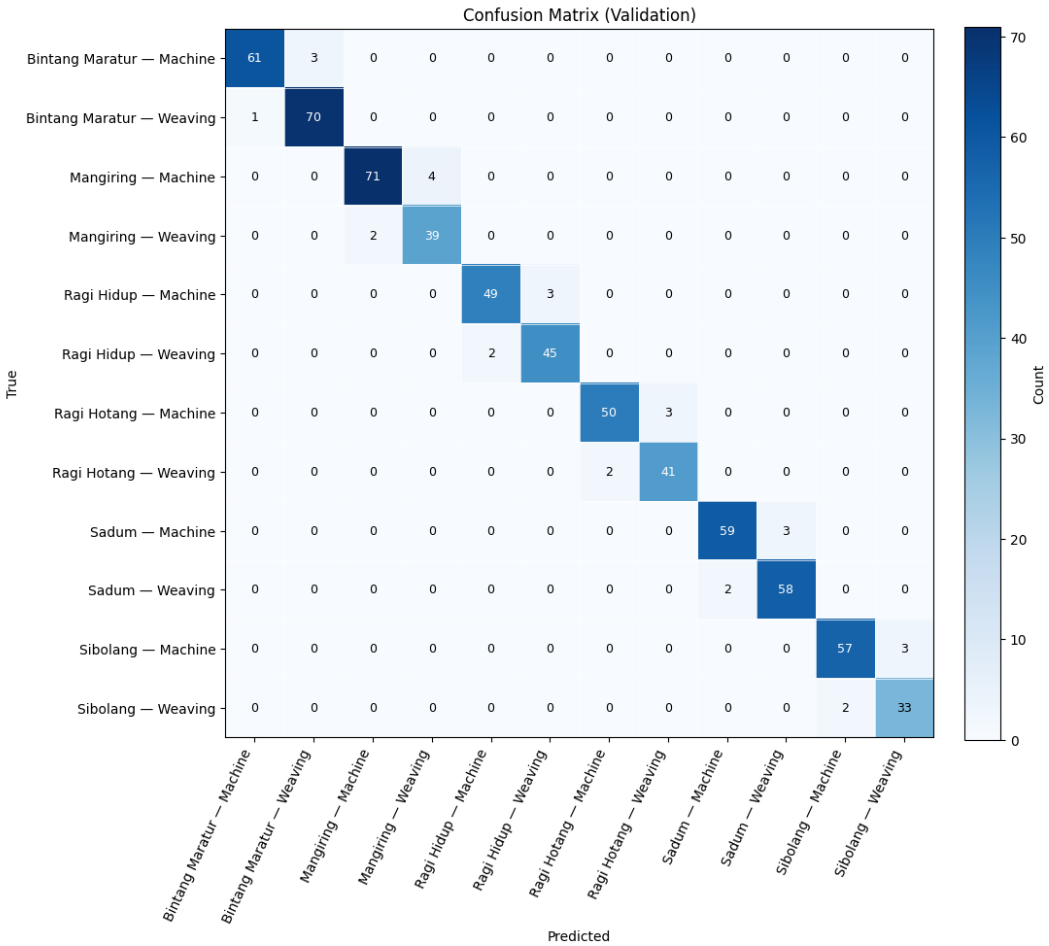

Figure 5. Confusion Matrix

Figure 5 illustrates the confusion matrix resulting from the model's classification outputs for twelve distinct classes. These classes are derived from combinations of six Ulos types, each represented in two different production methods: machine-made and hand-woven. In the matrix, the cells along the main diagonal represent the number of instances correctly predicted for each class. The number of correct predictions per class varies, with the highest reaching 71 and the lowest recorded at 33. The dominance of high values along the diagonal strongly indicates that the majority of predictions made by the model were accurate. This alignment of true labels with predicted labels serves as clear visual evidence of the model's high classification performance. Although the model achieved near-perfect accuracy in most classes, a few minor misclassifications were observed in several closely related categories.

In particular, the classes Mangiring – Machine and Mangiring – Weaving showed a small number of errors, where some samples were incorrectly predicted as their visually similar counterparts. Similar patterns of confusion were also present between the machine and weaving variants of Ragi Hidup and Ragi Hotang, though the misclassification rates remained relatively low. The classes Sadum and Sibolang, especially their machine and weaving variants, experienced two to three cases of incorrect predictions that reflected a modest decrease in precision and recall. Nevertheless, these misclassifications were limited in number and did not significantly impact the overall performance. The confusion observed in these cases can be attributed to the high degree of visual similarity shared between the motifs produced by different techniques within the same Ulos type. In some instances, even human evaluators may find it challenging to distinguish between these variations without close inspection.

To support the findings from the confusion matrix, table 3 presents the detailed classification metrics for each class, including precision, recall, and F1-score. These results further reinforce the model's strong performance. Most classes achieved F1-scores close to or above 0.95, reflecting an excellent balance between sensitivity and specificity in the model's predictions. Notably, classes such as Bintang Maratur – Machine and Bintang Maratur – Weaving reached F1-scores of 0.97, indicating highly reliable predictions. Other classes, including Ragi Hidup – Weaving and Sadum – Weaving, maintained similar performance levels. Although the classes Sibolang – Machine and Sibolang – Weaving demonstrated slightly lower values in comparison, their F1-scores remained high at 0.96 and 0.93, respectively. These results show that the model consistently delivers accurate predictions across all classes, even when subtle visual differences are present between them.

Table 3. Classification Evaluation

Class	Precision	Recall	F1-Score
Bintang Maratur – Machine	0.98	0.95	0.97
Bintang Maratur – Weaving	0.96	0.99	0.97
Mangiring – Machine	0.97	0.95	0.96
Mangiring – Weaving	0.91	0.95	0.93
Ragi Hidup – Machine	0.96	0.94	0.95
Ragi Hidup – Weaving	0.94	0.96	0.95
Ragi Hotang – Machine	0.96	0.94	0.95
Ragi Hotang – Weaving	0.93	0.95	0.94
Sadum – Machine	0.97	0.95	0.96
Sadum – Weaving	0.95	0.97	0.96
Sibolang – Machine	0.97	0.95	0.96
Sibolang – Weaving	0.92	0.94	0.93

The outstanding performance achieved in this study can largely be attributed to the modified MobileNetV2 architecture employed as the backbone of the classification system. This architecture integrates inverted residual blocks and depthwise separable convolutions, which together enhance the model's ability to extract spatial and geometric patterns while maintaining computational efficiency. Such capabilities are particularly valuable when working with traditional textiles like Ulos, which feature complex, repetitive patterns, gorga-like motifs, and symmetrical visual structures that are often difficult to detect using conventional convolutional techniques. Moreover, the model was trained using an effective combination of regularization techniques, such as dropout, and a customized data augmentation strategy that preserved the essential geometric properties of the textile patterns while enhancing generalization.

Throughout the training process, the model demonstrated stable learning behavior. The training accuracy improved steadily from 22.32 percent in the first epoch to 97.91 percent in the tenth epoch. At the same time, the validation accuracy reached 95.49 percent, which closely followed the training accuracy curve. The relatively small gap between the two metrics and the consistently decreasing trend in both training and validation loss values confirm that the model did not overfit the training data. This balance indicates that the applied regularization methods were effective in

ISSN 2723-6471

preventing the model from memorizing the training data and allowed it to generalize well to new, unseen examples. The learning curves also visually confirmed this finding, showing parallel progress between training and validation metrics that reflect a healthy and robust learning process.

Although none of the classes achieved perfect precision, recall, or F1-score, the model's consistently high values across all metrics demonstrate that it performs with a high degree of accuracy and reliability. The only modest performance drops were observed in the Sibolang classes, which likely result from their high intra-class similarity. These classes share nearly identical visual features across their machine and hand-woven variants, making them difficult to distinguish even under expert human observation. This finding suggests that additional strategies may be required to improve the model's sensitivity to fine-grained differences between such similar classes. Future work could explore advanced techniques such as fine-grained feature extraction or attention-based mechanisms to help the model focus more precisely on subtle but discriminative details within the motifs.

The combination of a modified MobileNetV2 architecture, pattern-aware data augmentation, and a carefully curated Batak Toba Ulos image dataset has resulted in a classification system with exceptional accuracy, stability, and generalization capability. This research highlights the potential of lightweight convolutional neural networks for applications beyond traditional image classification tasks, especially in domains where visual complexity is high but data availability is limited. Moreover, it contributes to the broader goal of digital preservation of cultural heritage by providing a reliable, automated approach to recognizing and categorizing traditional textile motifs. Compared to previous studies that relied on conventional CNNs or deeper models like ResNet50 and faced issues with overfitting and suboptimal accuracy, the results in this study show clear improvements in both performance and model efficiency. The model's ability to learn from relatively small but richly detailed datasets demonstrates the power of transfer learning when paired with architecture and training strategies tailored to the visual domain at hand. These findings support the hypothesis that MobileNetV2, when properly adapted, is well-suited for high-complexity, low-sample-size classification problems such as Ulos motif recognition. Most importantly, the model's success indicates that deep learning can be used not only for technological advancement but also for the meaningful task of preserving and celebrating cultural identity through visual heritage.

4.4. Discussion

The results of this study demonstrate that the use of a transfer learning approach with a modified MobileNetV2 architecture can effectively address the challenges in classifying Batak Toba Ulos motifs from image data. The model achieved consistently high classification metrics across all twelve target classes, with F1-scores ranging from 0.93 to 0.97. These results confirm that the combination of MobileNetV2's lightweight structure and depthwise separable convolutions is well-suited for tasks involving limited data availability but complex visual structures, such as traditional textile patterns.

One of the most notable findings in this study is the model's ability to generalize well despite the dataset's relatively small size and high visual complexity. The gradual and consistent increase in accuracy over the 10 training epochs, paired with steadily decreasing loss values, indicates a stable training process without signs of overfitting. This stability is attributed to several factors, including the use of data augmentation techniques that preserve geometric consistency, the application of dropout layers to prevent overfitting, and the decision to fine-tune only the later layers of the pre-trained MobileNetV2 network. These design choices reflect a careful balance between preserving the generalization capabilities of the base model and adapting it to the specific visual characteristics of Batak Toba Ulos.

Compared to previous studies in similar domains, this study marks a significant advancement. For instance, earlier research involving traditional cultural image recognition has primarily focused on character recognition, monument classification, or stylized paintings, with limited attention given to non-monumental, textile-based heritage. The use of conventional CNNs or deeper networks like ResNet50 in those contexts often resulted in overfitting or insufficient accuracy, particularly when datasets were small or lacked diversity. In contrast, the model developed in this study maintained validation accuracy of over 95 percent, indicating superior generalization even in the presence of subtle intra-class variations. This performance difference highlights the advantage of using MobileNetV2, not only due to its efficiency but also its capability to learn discriminative features in low-data environments.

The confusion matrix and class-level evaluation metrics revealed that the few misclassifications observed were primarily between visually similar pairs, such as the machine and weaving variants of Mangiring, Ragi Hidup, Ragi Hotang, Sadum, and Sibolang. This confusion is understandable given the high degree of similarity in texture, color, and motif layout between these variants. The close resemblance between these subclasses sometimes poses challenges even for human observers, which suggests that the model's near-perfect performance is already approaching the upper bound of what is practically achievable with current image-based recognition techniques. Nevertheless, the existence of these misclassifications points to an opportunity for future refinement. Incorporating fine-grained visual recognition techniques, such as attention mechanisms or localized feature learning modules, could further improve the model's sensitivity to subtle motif differences and reduce inter-class confusion.

The broader implications of this research are significant in both technological and cultural contexts. Technologically, the results affirm that lightweight convolutional networks like MobileNetV2 are not only viable but highly effective for domain-specific cultural applications. Culturally, the system provides a scalable and efficient framework for digitizing, preserving, and classifying indigenous textiles, which have historically been underrepresented in digital heritage research. The creation of a dedicated Batak Toba Ulos dataset further enhances the impact of this work by offering a reusable resource for future research and development in the field of cultural informatics.

However, the study is not without limitations. Although the model performs well on the current dataset, its performance on real-world images captured under diverse lighting, backgrounds, and device settings has yet to be evaluated. Future studies should consider expanding the dataset to include such variations, as well as exploring the integration of multi-modal data, such as historical context or regional usage information, to enrich the classification output. Furthermore, while this research has focused on classification, the potential extension of this system to object detection or segmentation tasks would be valuable for applications in Augmented Reality (AR) and virtual museum environments.

In summary, this study provides empirical evidence that transfer learning with MobileNetV2, when paired with pattern-aware augmentation and culturally grounded dataset construction, is highly effective for traditional textile classification. The system developed serves not only as a technological contribution to computer vision but also as a meaningful tool for preserving and promoting Indonesia's rich cultural heritage through digital innovation.

5. Conclusion

This study successfully developed a classification model for Batak Toba Ulos motif images using a transfer learning approach with a modified MobileNetV2 architecture. The model is designed to efficiently process high-resolution image data with limited sample sizes, which is a common characteristic of traditional cultural archives. Through the use of depthwise separable convolutions and inverted residual blocks, the model effectively captures spatial and contextual features found in the distinctive geometric patterns of Batak Toba Ulos, including motifs such as gorga, sihala, and simata ni ari.

The training process, conducted over 10 epochs, demonstrated consistent improvements in performance. Validation accuracy increased significantly from 26.57 percent in the first epoch to 95.49 percent in the tenth epoch, while validation loss decreased from 2.1023 to 0.4061. These results indicate a well-converged model with strong generalization capabilities. Evaluation using the confusion matrix and classification metrics such as precision, recall, and F1-score showed that all twelve classes were classified with high accuracy, achieving F1-scores between 0.93 and 0.97. Although minor misclassifications were observed, particularly among classes with high visual similarity between machine-made and handwoven variants, the model demonstrated robust and reliable performance across all categories.

Based on these findings, the model is well-suited to support automated visual recognition and digital documentation of Batak Toba Ulos motifs. Future work may focus on integrating the model into interactive platforms, such as virtual reality systems, to enhance user engagement and cultural education. Additional improvements could involve expanding the dataset, refining augmentation strategies, and exploring alternative model architectures that are more sensitive to fine-grained visual differences. Collaboration across disciplines, including contributions from cultural experts and traditional textile practitioners, will also be essential to ensure that the system remains technically accurate, culturally respectful, and aligned with the broader goals of heritage preservation through digital innovation.

6. Declarations

6.1. Author Contributions

Conceptualization: T.L., G.S., P.D.P.S.; Methodology: T.L.; Software: T.L. and D.S.; Validation: T.L., G.S., and P.D.P.S.; Formal Analysis: T.L., G.S., and P.D.P.S.; Investigation: T.L.; Resources: G.S.; Data Curation: T.L.; Writing Original Draft Preparation: T.L., G.S., and P.D.P.S.; Writing Review and Editing: T.L., G.S., and P.D.P.S.; Visualization: T.L. and D.S.; 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

The authors received no financial support for the research, authorship, and/or publication of this article.

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 or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] A. Marfu, “Harnessing emotional and cultural intelligence for corporate sustainability in Indonesia: Examining psychological contracts, task interdependence, and environmentally sustainable project performance,” *Sustain. Futur.*, vol. 10, no. 1, pp. 1-12, 2025.

[2] H. Hanan, D. Suwardhi, T. Nurhasanah, and E. S. Bukit, “Batak Toba Cultural Heritage and Close-range Photogrammetry,” *Procedia - Soc. Behav. Sci.*, vol. 184, no.1, pp. 187–195, 2015.

[3] J. Pourmahmoud, S. M. Hashemy Shahdany, and A. Roozbahani, “Practical drought risk assessment and management framework: A step toward sustainable modernization in agricultural water management,” *J. Hydrol.*, vol. 644, no. 1, pp. 1-21, 2024.

[4] Y. Pratama, S. T. N. Nainggolan, D. I. Nadya, and N. Y. Naipospos, “One-shot learning Batak Toba character recognition using siamese neural network,” *Telkomnika (Telecommunication Comput. Electron. Control*., vol. 21, no. 3, pp. 600–612, 2023, doi: 10.12928/TELKOMNIKA.v21i3.24927.

[5] M. Sinaga and D. Kartika, “The Application of Frieze Groups and Crystallographic Groups in Generating Batak Toba Ornament Motifs Using a Matlab Graphical User Interface,” *JTAM (Jurnal Teor. dan Apl. Mat.*, vol. 8, no. 1, pp. 72-86, 2024, doi: 10.31764/jtam.v8i1.17130.

[6] X. Zhang, “Oil painting image style recognition based on ResNet-NTS network,” *J. Radiat. Res. Appl. Sci*., vol. 17, no. 3, pp. 1-22, 2024, doi: 10.1016/j.jrras.2024.100992.

[7] Y. Liu, P. Cheng, and J. Li, “Application interface design of Chongqing intangible cultural heritage based on deep learning,” *Heliyon,* vol. 9, no. 11, pp. 1-12, 2023, doi: 10.1016/j.heliyon.2023.e22242.

[8] W. Xu, “Decoding Cultural Heritage Values: A Deep Learning Framework for Recognition and Weighted Evaluation,” *Research Square (Preprints),* vol. 2025, no. 1, pp. 0–23, 2025.

[9] T. Boyadzhiev, G. Lagani, L. Ciampi, G. Amato, and K. Ivanova, “Comparison of Different Deep Neural Network Models in the Cultural Heritage Domain,” *10th International Euro-Mediterranean Conference (EuroMed 2024),* vol. 10, no. 1, pp. 1–7, 2025.

[10] M. Hasso-Agopsowicz, “Identifying WHO global priority endemic pathogens for vaccine research and development (R&D) using multi-criteria decision analysis (MCDA): an objective of the Immunization Agenda 2030,” *eBioMedicine*, vol. 110, no. 1, pp. 1-24, 2024

[11] B. L. Nkuna, K. Abutaleb, J. G. Chirima, S. W. Newete, A. J. van der Walt, and A. Nyamugama, “Identification of maize leaf diseases using red, green, blue-based images with convolutional neural network (CNN) and residual network (ResNet50) models,” *Smart Agric. Technol*., vol. 12, no. 1, pp. 1-26, 2025.

[12] W. Gamaleldin, O. Attayyib, L. Mohaisen, N. Omer, and R. Ming, “Developing a hybrid model based on Convolutional Neural Network (CNN) and Linear Discriminant Analysis (LDA) for investigating anti-selection risk in insurance,” *J. Radiat. Res. Appl. Sci.*, vol. 18, no. 2, pp. 1-18, 2025.

[13] N. T J, “An Enhanced Deep Learning Framework for Prostate Cancer Detection Using Modified VGG16 and LeNet-MobileNetV2 Integration,” *Results Eng.,* vol. 2025, no. 1, pp. 1-18, 2025.

[14] L. Yin, N. Wang, and J. Li, “Electricity terminal multi-label recognition with a ‘one-versus-all’ rejection recognition algorithm based on adaptive distillation increment learning and attention MobileNetV2 network for non-invasive load monitoring,” *Appl. Energy*, vol. 382, no. 1, pp. 1-17, 2025.

[15] G. Zhou et al., “Optimizing MobileNetV2 for improved accuracy in early gastric cancer detection based on dynamic pelican optimizer,” Heliyon, vol. 10, no. 16, pp. 1-14, 2024.

[16] L. Peng, J. Zhang, Y. Li, and G. Du, “A novel percussion-based approach for pipeline leakage detection with improved MobileNetV2,” *Eng. Appl. Artif. Intell*., vol. 133, no. 1, pp. 1-17, 2024.

[17] F. Sarac and M. Yazici, “X-ray is not inferior to CT in terms of F1 score in the diagnosis of foreign body aspiration: a recall, precision and F1 score performance analysis based on bronchoscopically proven cases,” *J. Pediatr. (Rio. J)*., vol. 101, no. 5, pp. 1-20, 2025

[18] K. Krasnodębska, W. Goch, J. H. Uhl, J. A. Verstegen, and M. Pesaresi, “Advancing Precision, Recall, F-score, and Jaccard index: An approach for continuous, ratio-scale measurements,” *Environ. Model. Softw.*, vol. 193, no. 1, pp. 1-14, 2025.

[19] O. Peretz, M. Koren, and O. Koren, “Naive Bayes classifier – An ensemble procedure for recall and precision enrichment,” *Eng. Appl. Artif. Intell.,* vol. 136, no. 1, pp. 1-12, 2024.

[20] Y. Wang, Y. Jia, Y. Tian, and J. Xiao, “Deep reinforcement learning with the confusion-matrix-based dynamic reward function for customer credit scoring,” *Expert Syst. Appl*., vol. 200, no. 1, pp. 1-13, 2022.

[21] Y. Harnik, H. Shalit Peleg, A. H. Bermano, and A. Milo, “Data efficient molecular image representation learning using foundation models††Electronic supplementary information (ESI),” *Chem. Sci.*, vol. 16, no. 24, pp. 10833–10841, 2025.

[22] L. Isaza and K. Cepa, “Automation and augmentation: A process study of how robotization shapes tasks of operational employees,” *Eur. Manag. J.*, vol. 2024, no. 1, pp. 1-12, 2024

[23] P. Yuan, Z. Chen, Q. Jin, Y. Xu, and H. Xu, “Lightweight lotus phenotype recognition based on MobileNetV2-SE with reliable pseudo-labels,” *Comput. Electron. Agric.*, vol. 232, no.1, pp. 1-20, 2025.

[24] J. A. A. Opschoor and C. Schwab, “Deep ReLU networks and high-order finite element methods II: Chebyšev emulation,” *Comput. Math. with Appl.,* vol. 169, no. 1, pp. 142–162, 2024

[25] I. A. Chikwendu, “Attention-augmented and depthwise separable convolutional message passing for robust fraud detection in large-scale graphs,” *J. Adv. Res.*, vol. 2025, no. 1, pp. 1-12, 2025.

[26] K. Zhang, W. Wang, Z. Lv, J. Feng, H. Li, and C. Zhang, “LKDPNet: Large-Kernel Depthwise-Pointwise convolution neural network in estimating coal ash content via data augmentation,” *Appl. Soft Comput.,* vol. 144, no.1, pp. 1-21, 2023.

[27] Y. Zhang and J. De Smedt, “Index tracking using shapley additive explanations and one-dimensional pointwise convolutional autoencoders,” *Int. Rev. Financ. Anal.,* vol. 95, no. 1, pp. 1-17, 2024.

[28] S. Senthil Pandi, A. Senthilselvi, J. Gitanjali, K. ArivuSelvan, J. Gopal, and J. Vellingiri, “Rice plant disease classification using dilated convolutional neural network with global average pooling,” *Ecol. Modell.,* vol. 474, no. 1, pp. 1-16, 2022.

[29] R. Xu, Y. Zheng, X. Wang, and D. Li, “Person re-identification based on improved attention mechanism and global pooling method,” *J. Vis. Commun. Image Represent.,* vol. 94, no. 1, pp. 1-19, 2023.

[30] T. Le-Duc, H. Nguyen-Xuan, and J. Lee, “Sequential motion optimization with short-term adaptive moment estimation for deep learning problems,” *Eng. Appl. Artif. Intell.,* vol. 129, no. 1, pp. 1-13, 2024.