# IMPLEMENTATION OF DEEP LEARNING TECHNOLOGY IN THE ATTENDANCE SYSTEM AT PT HARPA OCEAN BERSAMA

**Arya Rizky Tri Putra [1*], Muhammad Firza Pahlevi[2,] Moh. Wahyu Abrory[3]**
[1,2,3]*Universitas Pembangunan Nasional "Veteran" Jawa Timur (Indonesia)*
**) email: 22082010067@student.upnjatim.ac.id*

**Abstract**

Employee attendance integrity constitutes a fundamental pillar of organizational management. Nevertheless, many companies encounter inefficiencies and vulnerabilities associated with manual or contact-based recording systems. PT Harpa Ocean Bersama is currently grappling with analogous challenges, as its reliance on conventional methods has resulted in administrative inefficiencies and a vulnerability to fraudulent practices, such as "buddy punching." The objective of this study is to develop a contactless Face Recognition Attendance System to automate and secure the workforce verification process. The system is constructed using a Convolutional Neural Network (CNN) framework, specifically leveraging the MobileNetV2 architecture combined with Transfer Learning. This approach was selected to optimize detection accuracy while minimizing computational costs, making it suitable for real-time application. The development methodology encompasses a rigorous pipeline, which includes the following steps: data acquisition, preprocessing using Haar Cascades for precise face isolation, data augmentation, and model fine-tuning using the TensorFlow library. The resulting system boasts robust employee identification capabilities, achieving a testing accuracy of 99.3% and a precision of 100%. This solution has been demonstrated to enhance data reliability and operational speed in comparison with conventional manual methods. It is anticipated that the implementation of this system will modernize human resource operations at PT Harpa Ocean Bersama, thereby ensuring a transparent, efficient, and fraud-resistant attendance environment.

**Keywords:** Face Recognition, MobileNetV2, Transfer Learning, Attendance System, Convolutional Neural Network.

## 1. INTRODUCTION

Employee attendance management constitutes a foundational element of organizational governance, exerting a direct influence on operational efficiency and human resource accountability. In the contemporary industrial epoch, precise attendance data is paramount for calculating payroll, evaluating performance, and maintaining workplace discipline (Putri et al., 2025). Conventional methods, such as manual logbooks or paper-based signatures, are increasingly regarded as outdated due to their vulnerability to human error, data manipulation, and the laborious nature of manual replication (Sharma & Pal, 2022). Moreover, while hardware-based solutions such as fingerprint scanners have gained widespread adoption, they have also given rise to concerns regarding hygiene and maintenance, particularly in environments with high foot traffic. Consequently, there is an increasing demand for contactless, automated biometric systems that ensure both security and fluidity in workforce management.

PT Harpa Ocean Bersama, a dynamic entity in the maritime industry, requires a robust system to manage its workforce effectively. Presently, the attendance process is confronted with challenges pertaining to efficiency and data integrity. Conventional recording methods and generic biometric tools frequently result in administrative inefficiencies and the potential for inaccuracies in the tracking of employee presence. These limitations impede the capacity of the HR department to generate real-time reports and create vulnerabilities for fraudulent activities, such as "buddy punching" (attendance fraud). In order to maintain competitive operational standards, a transition toward an intelligent, automated identification system is necessary.

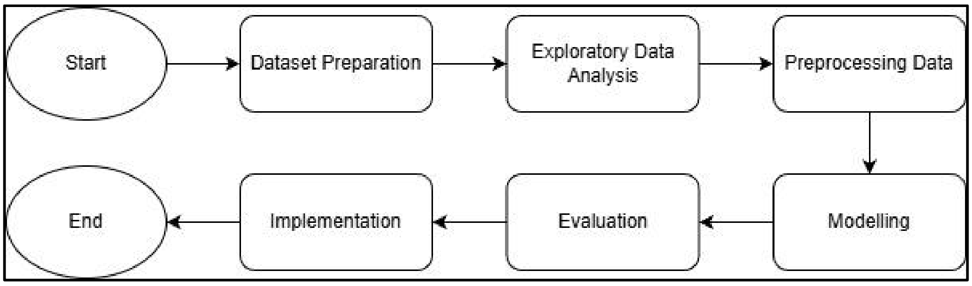
In the domain of Artificial Intelligence, Convolutional Neural Networks (CNN) have emerged as the state-of-the-art technique for computer vision tasks, including facial recognition (Setiawan & Lukman, 2023). However, training deep CNN architectures from scratch necessitates substantial datasets and considerable computational resources, which is often impractical for specific corporate applications with limited data. Transfer Learning provides a strategic solution to this challenge by leveraging pre-trained models on extensive datasets (e.g., ImageNet) and fine-tuning them for specific tasks. Among the various architectures, MobileNetV2 is distinguished by its lightweight structure and efficiency, utilizing depthwise separable convolutions to reduce computational cost without compromising significant accuracy (Hu & Ge, 2020).

In light of these considerations, this study proposes the development of a Face Recognition Attendance System for PT Harpa Ocean Bersama. The system has been designed to implement the MobileNetV2 architecture through a transfer learning approach, with the objective of accurately identifying employees from the hob_dataset (training, validation, and testing sets) and processing real-time predictions. The objective of this research is to enhance the efficacy of manual systems by providing a contactless, efficient, and precise attendance solution. The successful implementation of this system is expected not only to streamline administrative workflows but also to serve as a model for the adoption of AI-driven efficiency tools within the company.

## 2. METHODOLOGY

This study applies a structured research methodology consisting of several sequential stages. The process begins with dataset preparation, followed by exploratory data analysis to examine the characteristics of the collected facial image data. The next stage involves data preprocessing to prepare facial images for model training. Subsequently, a Convolutional Neural Network (CNN) model is developed and trained during the modelling stage (Dini Nurul Azizah et al., 2024). The trained model is then evaluated to assess its performance before being implemented in the attendance system to recognize employee identities based on facial images (Amri, 2025). The overall research methodology applied in this study is illustrated in Figure 1.

**Figure 1. Research Methodology**




## A. Dataset Preparation

The dataset used in this study consists of facial images (selfies) of employees at PT Harpa Ocean Bersama. The dataset includes 18 employees, where each employee is represented as one class in the face recognition process. Facial images were captured directly with a frontal face orientation to ensure that facial features were clearly recorded. All images were stored in JPG format and organized in a structured manner to support subsequent data processing stages.

To enhance dataset diversity and reduce the risk of overfitting, data augmentation was applied to each original facial image. The augmentation techniques employed include brightness adjustment, horizontal flipping, image rotation, and conversion to grayscale. These augmentation methods were intended to simulate variations in lighting conditions and facial orientations that commonly occur during real-world attendance recording.

The prepared dataset was then divided into three subsets: training data, validation data, and testing data. This data partitioning was conducted to facilitate effective model training, performance monitoring during the training process, and objective evaluation of the final Convolutional Neural Network (CNN) model.

## B. Exploratory Data Analysis

Exploratory Data Analysis (EDA) is conducted as an initial step to understand the structure and characteristics of the facial image dataset prior to preprocessing and model development (Akal et al., 2020). EDA is performed separately on the training, validation, and testing datasets to ensure that each subset exhibits consistent data distribution and characteristics. This stage aims to identify potential data-related issues that may affect the performance of the Convolutional Neural Network (CNN).

During the EDA stage, the analysis includes examining the number of labels and the distribution of images across each class within the dataset directories. In addition, the characteristics of facial images are inspected to ensure adequate visibility of facial regions and to identify variations in lighting conditions and facial orientations that need to be considered in subsequent processing stages.

EDA also involves analyzing image size and resolution, pixel value distribution, and checking for duplicate images across the training, validation, and testing datasets. These analyses are intended to determine the necessity of image resizing and pixel normalization, as well as to ensure that no data duplication or data leakage occurs between dataset subsets before the CNN training process is performed.

## C. Preprocessing

The pre-processing stage is conducted to prepare the facial image dataset so that it conforms to the input requirements of the Convolutional Neural Network (CNN). All pre-processing steps are applied consistently to the training, validation, and testing datasets to ensure uniformity and reliability during the model development process.

The pre-processing pipeline begins with face detection and face cropping using the Haar Cascade Classifier to extract the relevant facial region and remove unnecessary background information. If multiple faces are detected within a single image, the face with the largest bounding area is selected as the primary face. A margin is added around the detected facial region during cropping to preserve important facial features.

Following face cropping, image size and resolution normalization is performed. Each facial image is resized to 224 × 224 pixels using proportional resizing with padding to maintain the original aspect ratio and prevent facial distortion. This step ensures that all images have a consistent input dimension suitable for the CNN model.

Subsequently, data augmentation is applied to increase dataset diversity and reduce the risk of overfitting (Syahrul Gunawan Ramdhani & Enny Itje Sela, 2023). The augmentation techniques include horizontal flipping, image rotation, brightness adjustment, and grayscale conversion. To maintain input consistency, grayscale images are stored using three color channels. Augmentation is performed in a controlled manner to prevent repeated augmentation of previously augmented images.

## D. Modelling

The modelling stage is conducted to design and train a Convolutional Neural Network (CNN) model for facial recognition within the attendance system. Pre-processed facial image data are prepared using a data generator mechanism to load images in batches from the dataset directories. Real-time data augmentation is applied exclusively to the training dataset to enhance data variability, while the validation and testing datasets undergo preprocessing without augmentation to maintain objective performance monitoring.

The CNN model is developed using a transfer learning approach by adopting MobileNetV2 as the base architecture (Sandler et al., 2019). The MobileNetV2 model pre-trained on the ImageNet dataset is utilized as a feature extractor, with all base model parameters frozen during the initial training stage. Additional classification layers are appended to the base model, consisting of a Global Average Pooling layer, a fully connected layer with ReLU activation, and a dropout layer to regulate model complexity. The output layer employs a softmax activation function to support multi-class classification according to the number of employee identities. The model is compiled using the Adam optimizer and the categorical cross-entropy loss function.

The training process is performed using the training dataset, while the validation dataset is used to monitor the learning process across epochs. To control the training process, several callback mechanisms are implemented, including early stopping, adaptive learning rate adjustment, and model checkpointing. These mechanisms are intended to regulate training duration, manage learning dynamics, and retain the model state with the best validation performance prior to further evaluation and system implementation.

**E. Evaluation**

The evaluation stage is conducted to assess the performance and learning behavior of the Convolutional Neural Network (CNN) model during the training process. Model evaluation focuses on monitoring the model's ability to learn from the training data and to generalize to unseen data using validation data. This approach ensures that the model performance can be analyzed objectively throughout the training phase.

At this stage, model performance is evaluated by measuring accuracy and loss metrics on both the training and validation datasets across multiple epochs (Brockman et al., 2016). These metrics are used to observe the convergence behavior of the model, identify potential overfitting or underfitting, and assess the stability of the learning process. The evaluation results are visualized in the form of training and validation accuracy and loss curves to provide a clear representation of model performance over time.

The outcomes of this evaluation stage serve as the basis for determining whether the trained CNN model has achieved sufficient performance and stability before being deployed in the implementation stage of the attendance system.

**F. Implementation**

The implementation stage describes the deployment of the trained and evaluated Convolutional Neural Network (CNN) model into an employee attendance system at PT Harpa Ocean Bersama. The system is implemented in the form of a Progressive Web Application (PWA), which enables flexible access across devices while supporting real-time facial recognition during the attendance process.

At this stage, new facial images that are not part of the training, validation, or testing datasets are processed by the system. Each input image undergoes the same pre-processing pipeline applied during model training, including face detection and cropping, image resizing, and pixel normalization, to ensure consistency between training and operational conditions. The pre-processed images are then forwarded to the trained CNN model to perform facial recognition based on probability scores generated by the softmax activation function.

The output of the recognition process is subsequently used as the basis for attendance recording within the system. This implementation stage focuses on integrating the CNN model with the attendance mechanism and preparing the system for validation and performance evaluation under real-world operational conditions..

## 3. FINDINGS AND DISCUSSION

This section presents the experimental results obtained from the implementation of the Convolutional Neural Network (CNN)–based face recognition attendance system and discusses its performance, reliability, and practical implications within the operational environment of PT Harpa Ocean Bersama.

### A. Dataset and Preprocessing Results

The facial image dataset used in this study consists of 20 employee classes, with each class representing an individual employee. After the data acquisition process, preprocessing was performed to ensure uniformity and relevance of the input images. Haar Cascade Classifier successfully detected and isolated facial regions from the original images, effectively eliminating background noise and irrelevant visual information.

**Figure 2. Example of pre-processing dataset result.**

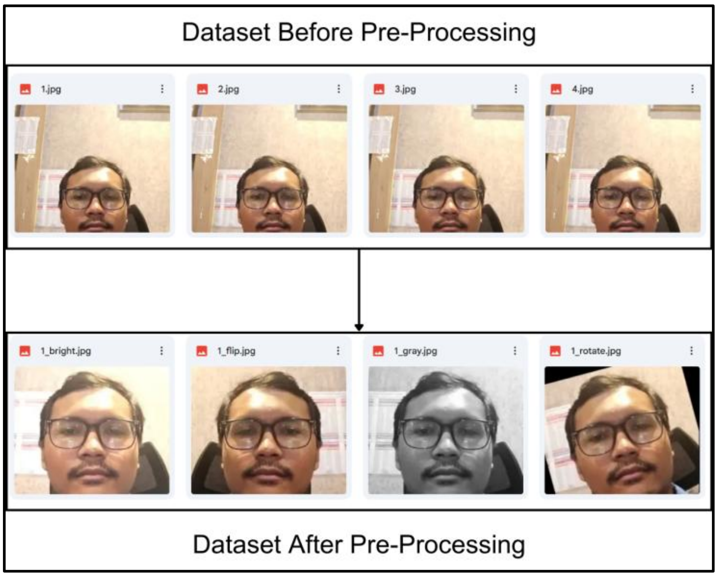
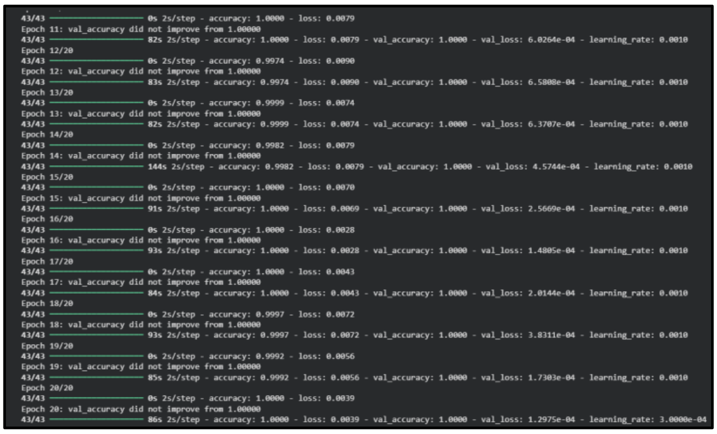



Image resizing to 224 × 224 pixels ensured compatibility with the MobileNetV2 input requirements, while maintaining the original aspect ratio to prevent facial distortion. Data augmentation techniques, including horizontal flipping, rotation, brightness adjustment, and grayscale conversion, increased dataset diversity and improved the robustness of the model against variations in lighting conditions and face orientation.

## B. Model Training Performance

The performance of the proposed CNN model during the training phase demonstrates a stable and well-converged learning process. As shown by the training logs, the model achieved high accuracy values from the early epochs and consistently maintained performance close to 100% accuracy throughout the training process. The validation accuracy reached 1.0000 and did not experience further improvement in subsequent epochs, indicating that the model had already learned optimal feature representations for the given dataset. The decreasing loss values across epochs further confirm that the optimization process successfully minimized classification error without introducing instability.

**Figure 3. Training and validation accuracy and loss during the model training process.**

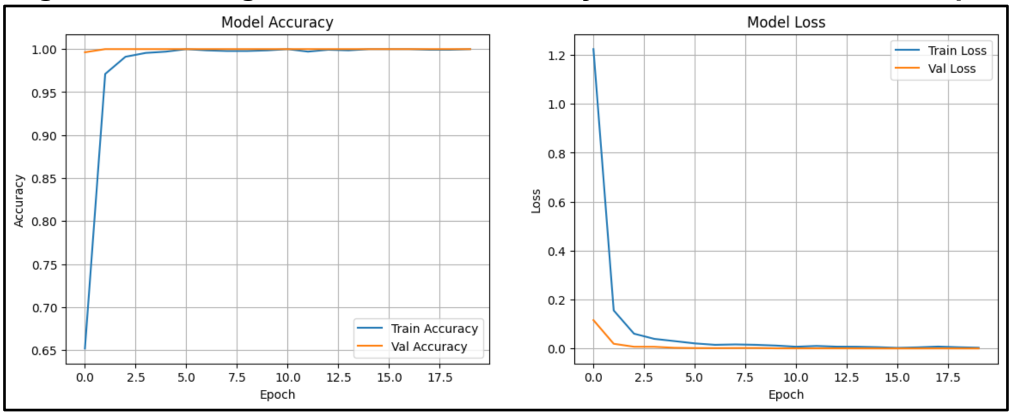

```
43/43 ━━━━━━━━━━━━━━━━━━━━ 0s 2s/step - accuracy: 1.0000 - loss: 0.0079
Epoch 11: val_accuracy did not improve from 1.00000
43/43 ━━━━━━━━━━━━━━━━━━━━ 82s 2s/step - accuracy: 1.0000 - loss: 0.0079 - val_accuracy: 1.0000 - val_loss: 6.0264e-04 - learning_rate: 0.0010
Epoch 12/20
43/43 ━━━━━━━━━━━━━━━━━━━━ 0s 2s/step - accuracy: 0.9974 - loss: 0.0090
Epoch 12: val_accuracy did not improve from 1.00000
43/43 ━━━━━━━━━━━━━━━━━━━━ 83s 2s/step - accuracy: 0.9974 - loss: 0.0090 - val_accuracy: 1.0000 - val_loss: 6.5808e-04 - learning_rate: 0.0010
Epoch 13/20
43/43 ━━━━━━━━━━━━━━━━━━━━ 0s 2s/step - accuracy: 0.9999 - loss: 0.0074
Epoch 13: val_accuracy did not improve from 1.00000
43/43 ━━━━━━━━━━━━━━━━━━━━ 82s 2s/step - accuracy: 0.9999 - loss: 0.0074 - val_accuracy: 1.0000 - val_loss: 6.3707e-04 - learning_rate: 0.0010
Epoch 14/20
43/43 ━━━━━━━━━━━━━━━━━━━━ 0s 2s/step - accuracy: 0.9982 - loss: 0.0079
Epoch 14: val_accuracy did not improve from 1.00000
43/43 ━━━━━━━━━━━━━━━━━━━━ 144s 2s/step - accuracy: 0.9982 - loss: 0.0079 - val_accuracy: 1.0000 - val_loss: 4.5744e-04 - learning_rate: 0.0010
Epoch 15/20
43/43 ━━━━━━━━━━━━━━━━━━━━ 0s 2s/step - accuracy: 1.0000 - loss: 0.0070
Epoch 15: val_accuracy did not improve from 1.00000
43/43 ━━━━━━━━━━━━━━━━━━━━ 91s 2s/step - accuracy: 1.0000 - loss: 0.0069 - val_accuracy: 1.0000 - val_loss: 2.5669e-04 - learning_rate: 0.0010
Epoch 16/20
43/43 ━━━━━━━━━━━━━━━━━━━━ 0s 2s/step - accuracy: 1.0000 - loss: 0.0028
Epoch 16: val_accuracy did not improve from 1.00000
43/43 ━━━━━━━━━━━━━━━━━━━━ 93s 2s/step - accuracy: 1.0000 - loss: 0.0028 - val_accuracy: 1.0000 - val_loss: 1.4805e-04 - learning_rate: 0.0010
Epoch 17/20
43/43 ━━━━━━━━━━━━━━━━━━━━ 0s 2s/step - accuracy: 1.0000 - loss: 0.0043
Epoch 17: val_accuracy did not improve from 1.00000
43/43 ━━━━━━━━━━━━━━━━━━━━ 84s 2s/step - accuracy: 1.0000 - loss: 0.0043 - val_accuracy: 1.0000 - val_loss: 2.0144e-04 - learning_rate: 0.0010
Epoch 18/20
43/43 ━━━━━━━━━━━━━━━━━━━━ 0s 2s/step - accuracy: 0.9997 - loss: 0.0072
Epoch 18: val_accuracy did not improve from 1.00000
43/43 ━━━━━━━━━━━━━━━━━━━━ 93s 2s/step - accuracy: 0.9997 - loss: 0.0072 - val_accuracy: 1.0000 - val_loss: 3.8311e-04 - learning_rate: 0.0010
Epoch 19/20
43/43 ━━━━━━━━━━━━━━━━━━━━ 0s 2s/step - accuracy: 0.9992 - loss: 0.0056
Epoch 19: val_accuracy did not improve from 1.00000
43/43 ━━━━━━━━━━━━━━━━━━━━ 85s 2s/step - accuracy: 0.9992 - loss: 0.0056 - val_accuracy: 1.0000 - val_loss: 1.7303e-04 - learning_rate: 0.0010
Epoch 20/20
43/43 ━━━━━━━━━━━━━━━━━━━━ 0s 2s/step - accuracy: 1.0000 - loss: 0.0039
Epoch 20: val_accuracy did not improve from 1.00000
43/43 ━━━━━━━━━━━━━━━━━━━━ 86s 2s/step - accuracy: 1.0000 - loss: 0.0039 - val_accuracy: 1.0000 - val_loss: 1.2975e-04 - learning_rate: 3.0000e-04
```

The training process was regulated using an early stopping mechanism, as indicated by repeated notifications that the validation accuracy did not improve beyond its optimal value. This mechanism effectively prevented overfitting by halting unnecessary weight updates once the model performance stabilized. Additionally, the adaptive learning rate strategy reduced the learning rate toward the final epochs, enabling finer weight adjustments and contributing to the very low validation loss values observed. Overall, these results indicate that the MobileNetV2-based transfer learning approach provided efficient convergence, strong generalization capability, and suitability for real-time face recognition implementation.

## C. Model Evaluation Results

The performance of the proposed classification model was evaluated using accuracy and loss metrics on both training and validation datasets. This evaluation aims to analyze the learning behavior of the model during the training process and to assess its ability to generalize to unseen data.

**Figure 4. Training and validation accuracy and loss curves across epochs.**

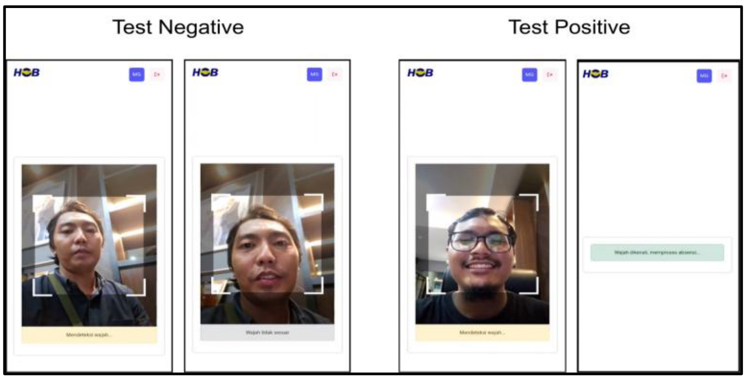



As shown in Figure 4, the model demonstrates a rapid improvement in accuracy during the early epochs, reaching nearly 100% accuracy for both training and validation datasets. The accuracy curves remain stable and closely aligned throughout the training process, indicating that the model does not suffer from overfitting. In addition, the loss values for both training and validation decrease sharply and converge toward zero, which suggests that the model learns effectively and achieves a well-optimized and stable performance.

### D. Implementation at PT Harpa Ocean Bersama

The implementation of the proposed face recognition attendance system was conducted at PT Harpa Ocean Bersama using a Progressive Web Application (PWA) deployed at app.hobbunker.com. The trained CNN model was integrated into the system to perform real-time facial recognition during the attendance process. System validation was carried out through two testing scenarios, namely positive and negative testing. In the positive test scenario, the registered user Mirza GA was successfully recognized by the system, and the attendance record was correctly generated, demonstrating accurate identification of authorized employees. Conversely, in the negative test scenario, individuals whose facial data were not registered in the system were not recognized, and no attendance data were recorded, indicating the system's ability to prevent unauthorized or fraudulent attendance. These results confirm that the implemented PWA-based system operates effectively and reliably under real-world conditions at PT Harpa Ocean Bersama.

**Figure 5. Implementation model in the employee attendance application.**



Furthermore, the deployment of this system supports the digital transformation of human resource management at PT Harpa Ocean Bersama by introducing an automated and contactless attendance mechanism. The integration of facial recognition technology within a PWA platform ensures operational flexibility and ease of access across multiple devices, while maintaining high accuracy and reliability. This approach not only strengthens attendance transparency and accountability but also demonstrates the practical applicability of convolutional neural network–based solutions as a sustainable and scalable framework for modern industrial workforce management.

## 4. CONCLUSION

This research resulted in the development of a Face Recognition Attendance System designed using the MobileNetV2 architecture and Transfer Learning to support workforce management at PT Harpa Ocean Bersama. This system integrates automated facial detection, preprocessing, and real-time classification, thereby improving the efficiency, accuracy, and integrity of attendance data recording. Through the implementation of contactless verification, the system is expected to eliminate fraudulent practices such as "buddy punching," reduce administrative workload, and support the company's transition toward digital operational excellence.

## REFERENCES

Akal, I., Kusuki, Y., Takayanagi, T., & Wei, Z. (2020). Codimension two holography for wedges. *Physical Review D*, *102*(12), 126007. https://doi.org/10.1103/PhysRevD.102.126007

Amri, H. (2025). Real-Time Face Attendance System Using CNN Mobilenet and MTCNN. *JURNAL TEKNOLOGI DAN OPEN SOURCE*, *8*(1), 164–179. https://doi.org/10.36378/jtos.v8i1.4403

Brockman, G., Cheung, V., Pettersson, L., Schneider, J., Schulman, J., Tang, J., & Zaremba, W. (2016). *OpenAI Gym* (arXiv:1606.01540). arXiv. https://doi.org/10.48550/arXiv.1606.01540

Dini Nurul Azizah, Raisa Mutia Thahir, Luthfi Dika Chandra, Naufal Ardhani, Endang Purnama Giri, & Gema Parasti Mindara. (2024). Implementation of Automatic Attendance System Based on Face Recognition Using CNN Method in IPB University Vocational School Environment. *International Journal of Multilingual Education and Applied Linguistics*, *1*(4), 01–10. https://doi.org/10.61132/ijmeal.v1i4.94

Hartanto, B., Yudanto, B. W., Nugroho, D., & Tomo, S. (2024). IMPELEMENTASI CONVOLUTIONAL NEURAL NETWORK MENGGUNAKAN MODEL MOBILENET DALAM APLIKASI PRESENSI BERBASIS PENGENALAN

WAJAH. *Biner : Jurnal Ilmiah Informatika dan Komputer*, *3*(1), 22–26. https://doi.org/10.32699/biner.v3i1.6607

Hu, L., & Ge, Q. (2020). Automatic facial expression recognition based on MobileNetV2 in Real-time. *Journal of Physics: Conference Series*, *1549*(2), 022136. https://doi.org/10.1088/1742-6596/1549/2/022136

Putri, I. M., Ramadhani, D. P., Indriyani, P., Aidah, E. N., & Cahyani, A. P. (2025). Predictive Analytics in Attendance Systems for Employee Productivity and Accountability. *International Transactions on Education Technology (ITEE)*, *3*(2), 104–113. https://doi.org/10.33050/itee.v3i2.718

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L.-C. (2019). *MobileNetV2: Inverted Residuals and Linear Bottlenecks* (arXiv:1801.04381). arXiv. https://doi.org/10.48550/arXiv.1801.04381

Setiawan, R. B., & Lukman, N. (2023). Attendance System Face Recognition Using Convolutional Neural Network (CNN). *CoreID Journal*, *1*(3), 116–121. https://doi.org/10.60005/coreid.v1i3.16

Sharma, R., & Pal, R. K. (2022). A Comparative Analysis of Biometric and Manual Attendance Measures using Queueing Theory. *International Journal of Advanced Research in Science, Communication and Technology*, *2*(8), 886–892.

Syahrul Gunawan Ramdhani & Enny Itje Sela. (2023). Implementasi Face Recognition Untuk Sistem Presensi Universitas Menggunakan Convolutional Neural Network. *Indonesian Journal of Computer Science*, *12*(6). https://doi.org/10.33022/ijcs.v12i6.3498