JURNAL MULTIDISIPLIN MADANI (MUDIMA)
Homepage: https://journal.formosapublisher.org/index.php/mudima
ISSN: 2808-5639 (Online)
Research Article

Volume 3, No 9, September (2023)	DOI: https://doi.org/10.55927/mudima.v3i9.5924	Page: 1855-1864

Real Time Face Recognition for Mobile Application Based on Mobilenetv2

I Made Subrata Sandhiyasa[1*], Devi Valentino Waas[2]
Institut Bisnis dan Teknologi Indonesia (INSTIKI)
Corresponding Author: I Made Subrata Sandhiyasa subrata.sandhiyasa@instiki.ac.id

ARTICLE INFO

Keywords: Face Recognition, Face Landmark, Deep Learning, MobileNetV2, Anti Spoofing

Received : *1 August*
Revised : *17 August*
Accepted : *18 September*

ABSTRACT

Real-time facial recognition is one of the technologies with significant applications in a variety of contexts, including supporting the process of employee attendance. Attendance is a crucial aspect of company administration that influences productivity and operational effectiveness. Traditional attendance mechanisms are susceptible to fraud and errors, so businesses must adopt digital attendance solutions. Face recognition with landmark-based anti-spoofing using MobileNetV2 on mobile devices is intended to be an innovative solution for attendance management. This system employs CNN with MobileNetV2 architecture to detect and identify employee faces in real time. MobileNetV2 is advantageous because it makes efficient use of mobile device resources without sacrificing precision. The research results demonstrate the extraction of eye and lip landmark points with the Blazeface model integrated in the Dart programming language using the Flutter framework. The implementation of the mobile system yields an application called FaceON that can aid in the prevention of potential fraud by employing anti-spoofing techniques. Before a visage can be verified, users must overcome obstacles such as winks and smiles. The contribution of this study is that this system is a dependable and innovative solution for employee attendance management in the digital age

Introduction

Real-time face recognition is one of the most important applications in the world of technology, with a wide range of potential applications, such as facilitating the employee attendance process, which is one of the most important aspects of company management affecting productivity and operational efficiency. Traditional attendance mechanisms are one of the vulnerabilities of companies attempting to implement the digital revolution; therefore, companies must implement modern and dependable attendance solutions that can meet the demands of the modern workplace. One intriguing solution is the use of real-time face recognition technology (Kocacinar et al., 2022; Ozdemir & Ugur, 2021; Ramos & Morales, 2020) to modify the face attendance business process via mobile applications (Kocacinar et al., 2022; Ozdemir & Ugur, 2021; Ramos & Morales, 2020).

Face attendance is an electronic attendance model that is widely used today(Prangchumpol, 2019; Surekha et al., 2017). The facial attendance currently being carried out focuses on analyzing the facial image that is captured by the attendance system camera when someone is going to take attendance(Putra et al., 2021). Several studies conducted have succeeded in analyzing a person's facial recognition using the Eigenface method [11], the Viola Jones method and the Local Binary Pattern Histogram (LBPH) algorithm (Buana, 2021). Existing research has limitations because it only recognizes a person's facial image without seeing whether it is really a human or just a person's facial image based on the interactions that occur when carrying out facial recognition. In the current implementation of attendance, recognition is only carried out from a person's face(Ramdhon & Febriya, 2021), where there is a problem if the face is a photo that has been printed.

Several related studies discussing facial recognition show that up to now, several methods have been used to explore Computer Vision, both in the form of face detection and even food recognition using the CNN method "Convolutional Neural Networks for Face Recognition" Real Time"(Zufar & Setiyono, 2016). In this research, it was proven that the use of the CNN method was able to detect faces in low light conditions. From this research, there is no doubt about the superiority of the CNN method.

Another study used CNN with the MobileNetV2 architecture for real-time face detection in industrial environments(Gururaj & Batra, 2021). In this study it was concluded that the use of MobileNetV2 is very accurate for detecting faces even when the face is wearing a mask. From the computing side, it is also stated that by using MobileNetV2 on devices that have limited computing, the architecture still has good speed and accuracy. Another study was conducted under the title Real-Time Mask Detection Based on SSD-MobileNetV2(Cheng, 2022), in this research it is explained that by using MobileNetV2 the resulting model is very light to run on mobile devices. Apart from that, the use of transfer learning in MobileNetV2 does not require much dataset as training data with high accuracy.

Based on previous research, to carry out facial recognition, researchers used Convolutional Neural Networks (CNN) with the MobileNetV2 architecture as the core of the facial recognition process. The use of this method is because, based on previous research, the CNN method with the MobileNetV2 architecture has been able to perform facial recognition with high accuracy, even in low lighting. This is certainly one of the advantages of the CNN method compared to other methods. Another thing that proves the superiority of the CNN method in image classification is the 2012 ImageNet Large Scale Visual Recognition Challenge won by Alex Krizhevsky with the application of CNN, this method has outperformed other Machine Learning methods such as SVM.

So to overcome this problem, before carrying out recognition, anti-spoofing validation can be carried out by detecting life on the face first. The scope of detection carried out is by checking whether the person making the attendance is a human or only a photo print. Aspects that can be validated are the presence of eye blinks and smiles within a certain period by utilizing facial contour landmarks. Where the value of the landmark will change according to a person's facial expression. To perform facial landmarks, in this study, the BlazeFace framework was used (C.O et al., 2022) in the process of recognizing interactions on the face of the absentee. BlazeFace is a face detection method with light processing and performs well on mobile GPUs.

The author's contribution to this research is the ability to detect facial vitality using blink and smile parameters based on facial landmarks. In addition, the visage is identified using MobileNetV2. This is done to prevent users from using photographs for facial recognition. Before the application verifies the face, the application will be able to detect winks and smiles as signs of vitality on the face.

METHODS

Convolutional Neural Network

Convolutional Neural Network (CNN) is a technique derived from the development of Multilayer Perceptron (MLP) for two-dimensional data processing (Wijaya et al., 2016). CNN is a member of the Deep Neural Network family due to the network's high depth and is preferable when applied to image data.

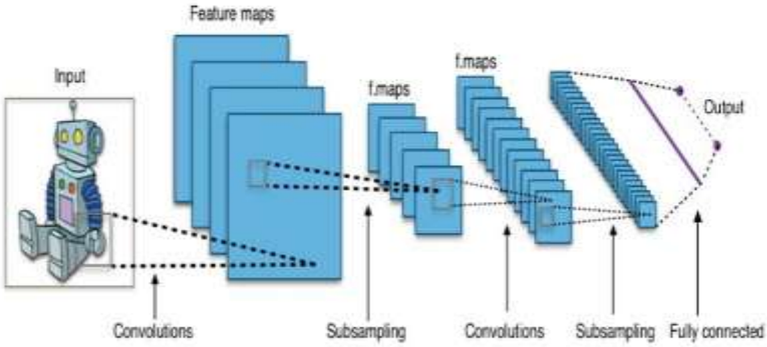

Figure 1. CNN Working Concept

CNN is a technique that was initially devised by a Japanese researcher named Kunihiko Fukushima under the name NeoCognitron. The concept created by Kunihiko Fukushima was then refined by an American researcher on behalf of LeChun. In research addressing the recognition of numbers and handwriting, LeChun created the initial CNN model, dubbed LeNet, with success. Alex Krizhevsky won the 2012 ImageNet Large Scale Visual Recognition Challenge by employing the CNN method. As a result, the implementation of the CNN method is becoming increasingly popular. After outperforming other Machine Learning methods such as SVM (Abroyan, 2017), this further establishes CNN as the superior image object classification method.

Mobilenetv2

MobileNet V2 is a convolutional neural network (CNN) architecture that can be used to circumvent the need for superfluous computing resources. MobileNet V2 represents an evolution of the MobileNet architecture. MobileNet architecture and CNN architecture in general utilize layers or convolution layers differently. MobileNetV2's convolution layer employs a filter with the same thickness as the input image. MobileNetV2 employs depthwise convolution, pointwise convolution, linear bottlenecks and shortcut connections between bottlenecks (Sandler et al., 2018).

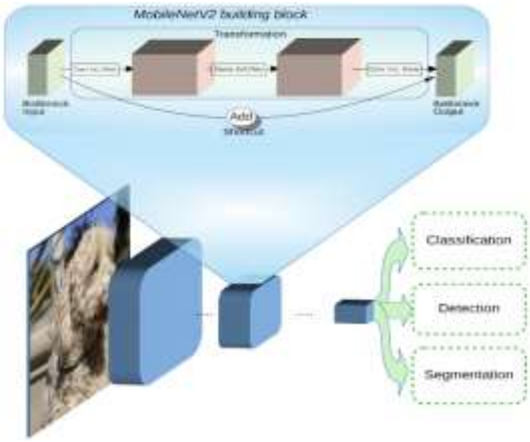

Figure 2. Mobilenet V2 Architecture

In the bottleneck there are inputs and outputs between models while the inner layer encapsulates the model's ability to change input from lower level concepts (pixels) to higher level descriptors. Ultimately, as with residual connections in traditional CNNs, shortcuts between bottlenecks allow for faster training and better accuracy. The model architecture in MobileNet V2 is that there are basic building blocks which are convolutions which can be separated in depth by residue. The MobileNetV2 architecture contains an initial full convolution layer with 32 filters, followed by 19 residual bottleneck layers. Researchers have adapted the architecture to different performance points, by using the input image resolution and width multiplier as tunable hyperparameters, which can be adjusted depending on the desired accuracy or performance trade-off. The main network (1.224 × 224 wide multiplier), has a computational cost of 300 million multiplication-additions and uses 3.4 million parameters. The computational cost of the network ranges from 7 additional multiplications to 585 million, while the model size varies between 1.7 million and 6.9 million parameters.

System Overview

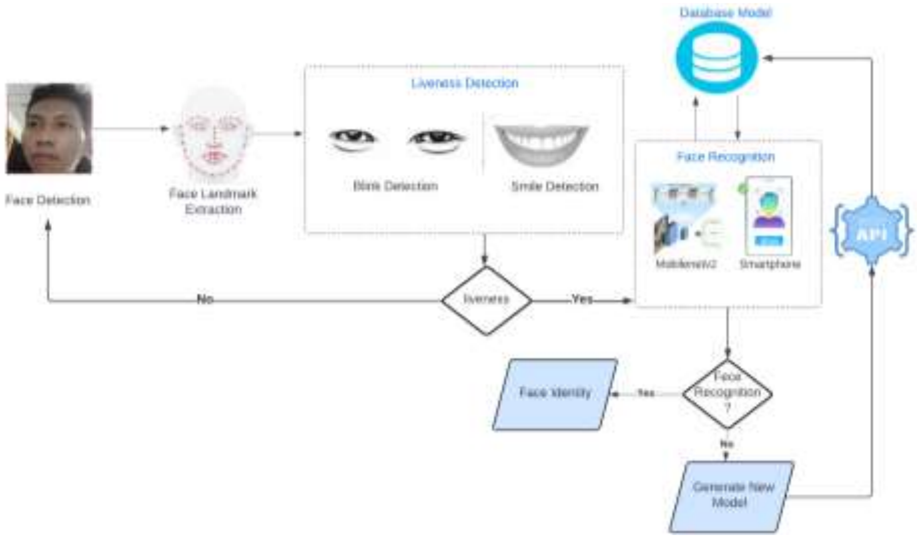

Figure 3. System Overview

As can be seen in Figure 3, the first stage carried out by the system is detecting the faces captured by the camera. If a face has been detected, the system will extract facial landmarks and separate parts of the landmarks. In this study, the landmarks used are the face and lips. If a face is detected with a wink and a smile, the next step is facial recognition using MobileNetV2. At this stage, the output generated by MobileNetV2 will be compared based on the closest distance generated, the face will be said to be the same if the proximity is below the threshold. If the faces are not the same, then the user can register a new face and enter the name of the face to be used as a model and stored in the face database

RESULTS AND DISCUSSION

System Implementation

Figure 4. FaceON Application

In Figure 4 it can be explained that there is an implementation of Mobilenetv2-based face detection which has been successfully implemented in a mobile application called FaceON. This application was built with the Dart programming language using the Flutter framework

Face Detection and Point Extraction of Eyes and Lips Landmarks

The point extraction of eye and lip landmarks in this study used the Blazeface model (Bazarevsky et al., 2019) where the model is integrated into the Dart programming language using the Flutter framework. facial ditector was tested with various backgrounds to test how accurate the facial ditector is in nature.

Figure 5. Face Detection Process

In extraction of eye landmarks, 32 landmark points were obtained consisting of 16 landmark points for the right eye and 16 landmark points for the left eye. Point landmarks of the eyes will be used to detect life on the face by detecting whether the eyes are blinking or not. Furthermore, extraction of lip landmark points was carried out, in that section a total of 34 landmark points were obtained consisting of 10 points on the upper lip landmark, 8 points on the upper lip landmark on the bottom, 8 points on the upper lower lip landmark and 8 points on the lower lip landmark. The lip landmarks obtained will be used to detect life on the face by detecting whether the lips change from normal to smiling or not.

(6a) Point Landmark Lips

(6b) Eye Landmark Point

(6c) Point Landmark Eyes & Lips

Figure 6. Extraction of Landmark Features of the Eyes and Lips

Facial Model Making

Making facial models is carried out using two collaborating library functions, namely Mlkit StandAlone to carry out face detection and preprocessing on images obtained from each frame from the streaming camera, and the MobileFaceNet model to process, classify and transform them into data structures that can be stored by a database in This is an array of numbers totaling 193 lines.

Figure 7. Face Diction

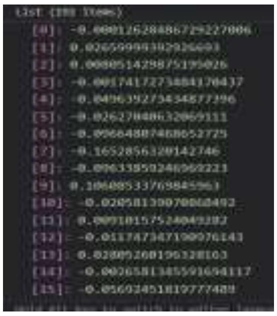

Figure 8. Result of Face Model Transformation

After the transformation is carried out, the transformation results are then stored in the smartphone's local database in a JSON file along with a label in the form of the name entered by the user when registering. At this stage the model is still stored in the smartphone's internal memory, then the face model will be stored on the cloud server using the REST API.

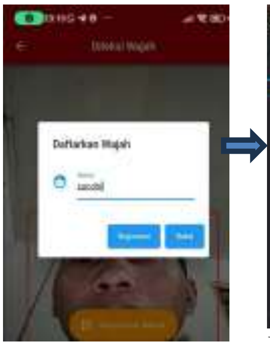

Figure 9. Face Registration Feature

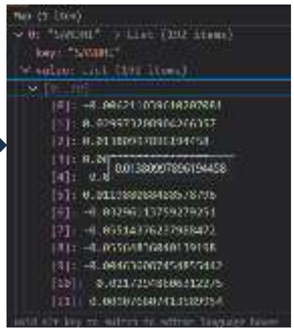
Figure 10. Model Results in JSON File

Life Detection on Face

Face life detection is carried out to verify the face captured by the camera. The scope of detection carried out is by checking whether the person making the attendance is a human or only a photo print. Aspects that can be validated are the presence of eye blinks and smiles within a certain period by utilizing facial contour landmarks. Where the value of the landmark will change according to a person's facial expression. This study uses 2 facial characteristics, namely blinking and smiling. After the eye and lip landmarks are obtained as previously explained, these landmarks will be calculated using the Eye Aspect Ratio (EAR) equation.(Souto Maior et al., 2018) as in Figure 11 below.

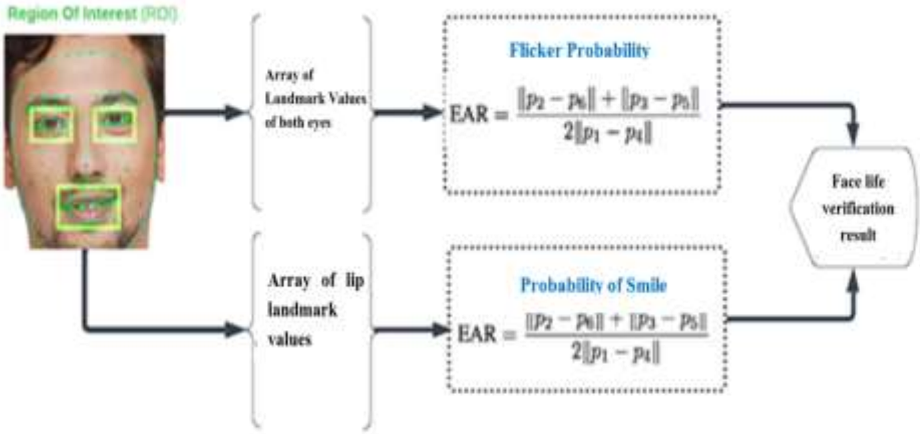

Figure 11. The EAR Equation

Based on Figure 11, the next result is that the application displays a notification for the user to blink at least 3 times and smile at least 3 times so that the face can be verified, which can be seen in Figure 12 below.

Figure 12. Life Detection Notification on Faces

Face Recognition

In the face recognition process, the distance between the transformation results of the face recorded by the camera and the values in the facial model database will be compared. If the value of the detected face exceeds the threshold set at 80%, then the face will refer to one of the stored models and the application will display the label of that model as in Figure 13 below.

(13a). Unrecognizable Face

(13b). Recognized Face

Figure 13. Face Recognition Process in Applications

In Figure 13 it can be explained that the facial recognition process by detecting winks and smiles on the face has been successfully carried out and displayed on the mobile application.

ACKNOWLEDGMENT

Thanks to the Ministry of Education, Culture, Research and Technology of the Republic of Indonesia for the Grant Funding provided through the Beginner Lecturer Research scheme (PDP) with the main contract number 184/E5/PG.02.00.PL/2023

CONCLUSION

The study found that the MobileNetV2-based face identification system in the " FaceON" mobile app supports digital employee attendance management. This application employs the Blazeface model for face detection, which is effectively integrated into Dart through the Flutter framework. The face identification capability is tested against various backgrounds to assure accuracy in natural circumstances. The software can also extract 32 eye landmark points and 34 lip landmark points to identify winks and grins. The face modeling approach uses Mlkit StandAlone for face detection and image preprocessing and MobileFaceNet for processing and classifying face data into a local database structure. The result and username label are stored in a JSON file after transformation. This app's face live detection ensures that the recognized face is a real person, not a photo. The program detects winks and smiles using facial contour markers. The software can recognize faces with high accuracy and notify users if they are recognized. The application compares identified face transformation results to the database model during face recognition. If the match rate reaches the threshold, the app recognizes the face and displays the label. This entire implementation shows how face recognition technology can facilitate mobile employee attendance.

REFERENCES

Abroyan, N. (2017). Convolutional and recurrent neural networks for real-time data classification. *2017 Seventh International Conference on Innovative Computing Technology (INTECH)*, *Intech*, 42–45. https://doi.org/10.1109/INTECH.2017.8102422

Bazarevsky, V., Kartynnik, Y., Vakunov, A., Raveendran, K., & Grundmann, M. (2019). *BlazeFace: Sub-millisecond Neural Face Detection on Mobile GPUs*. 3–6.

Buana, I. (2021). Penerapan Pengenalan Wajah Untuk Aplikasi Absensi dengan Metode Viola Jones dan Algoritam LBPH. *JURNAL MEDIA INFORMATIKA BUDIDARMA*, *5*, 1008. https://doi.org/10.30865/mib.v5i3.3008

C.O, A., A.O, A., O.O, O., & E.O, I. (2022). Fingerprint-Based Attendance Management System. *Journal of Computer Sciences and Applications*, *1*(5), 100–105.

Cheng, C. (2022). Real-Time Mask Detection Based on SSD-MobileNetV2. *2022 IEEE 5th International Conference on Automation, Electronics and Electrical Engineering, AUTEEE 2022*, 761–767. https://doi.org/10.1109/AUTEEE56487.2022.9994442

Gururaj, N., & Batra, K. (2021). InnovFaceNet: Deep Face Recognition for Industrial Environments. *Journal of Computational Vision and Imaging Systems*, *6*(1), 1–4. https://doi.org/10.15353/jcvis.v6i1.3553

Kocacinar, B., Tas, B., Akbulut, F. P., Catal, C., & Mishra, D. (2022). A real-time cnn-based lightweight mobile masked face recognition system. *Ieee Access*, *10*, 63496–63507.

Ozdemir, D., & Ugur, M. E. (2021). Model proposal on the determination of student attendance in distance education with face recognition technology. *Turkish Online Journal of Distance Education*, *22*(1), 19–32.

Prangchumpol, D. (2019). Face Recognition for Attendance Management System Using Multiple Sensors. *Journal of Physics: Conference Series*, *1335*(1), 12011. https://doi.org/10.1088/1742-6596/1335/1/012011

Putra, S. P., Fitri, I., & Ningsih, S. (2021). Absensi Pengenalan Wajah Menggunakan Menggunakan Algoritma Eigenface Berbasis Web. *Journal of Applied Informatics and Computing*, *5*(1 SE-Articles). https://doi.org/10.30871/jaic.v5i1.2711

Ramdhon, A. N., & Febriya, F. (2021). Penerapan Face Recognition Pada Sistem Presensi . *Journal of Applied Computer Science and Technology*, *2*(1 SE-Articles). https://doi.org/10.52158/jacost.v2i1.121

Ramos, L., & Morales, B. (2020). SwiftFace: real-time face detection. *ArXiv Preprint ArXiv:2009.13743*.

Sandler, M., Howard, A., Zhu, M., Zhmoginov, A., & Chen, L. C. (2018). MobileNetV2: Inverted Residuals and Linear Bottlenecks. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 4510–4520. https://doi.org/10.1109/CVPR.2018.00474

Souto Maior, C., Moura, M., Santana, J., Nascimento, L., Macedo, J., Lins, I., & Droguett, E. (2018, September). Real-time SVM Classification for Drowsiness Detection Using Eye Aspect Ratio. *Conference: Proceedings of Probabilistic Safety Assessment and Management PSAM 14At: Los Angeles*.

Surekha, B., Nazare, K. J., Viswanadha Raju, S., & Dey, N. (2017). Attendance recording system using partial face recognition algorithm. *Intelligent Techniques in Signal Processing for Multimedia Security*, 293–319.

Wijaya, A. Y., Soelaiman, R., Informatika, J. T., & Informasi, F. T. (2016). *Klasifikasi Citra Menggunakan Convolutional Neural Network (Cnn) Pada Caltech 101 Image Classfication Using Convolution Neural Network (Cnn) on Caltech 101*. *5*(1).

Zufar, M., & Setiyono, B. (2016). Convolutional Neural Networks untuk Pengenalan Wajah Secara Real - Time. *Jurnal Sains Dan Seni Its*, *5*(2), 72–77.