Journal of Information System and Technology Research

journal homepage: https://journal.aira.or.id/index.php/jistr/

Direct implementation of AI-Based Facial Recognition for students

Andi Prayogi[1*], Roy Francis Navea[2], Rahmad Dian[3], Muhammad Akbar Syabana Pane[4], Ratu Mutiara Siregar[5], Raden Aris Sugianto[6], Hasanal Fachri Satia Simbolon[7]

[1,3,4,5,6,7]Department Information Systems and Technology, Institut Teknologi Sawit Indonesia, Indonesia
[2]Department of Electronics and Computer Engineering, De La Salle University, Philippines

ARTICLE INFO

Article history:

Received August 13, 2024
Accepted September 30, 2024
Available online September 30, 2024

Keywords:

AI,
Face Recognition,
CNN
ITSI

ABSTRACT

The development of artificial intelligence (AI)-based facial recognition technology has become a significant research topic in the field of computing and security. At the Indonesian Palm Oil Institute (ITSI), AI-based facial recognition is introduced to students to improve their skills in developing AI-based applications. This study aims to implement and test a facial recognition system using a Python program by utilizing a dataset generated independently. This research method involves several stages, namely collecting ITSI students' facial data, data processing, creating a facial recognition model using a machine learning algorithm, and evaluating model performance. The dataset used was developed through a live shooting session involving active student participation. The facial recognition model was trained using a convolutional neural network (CNN) algorithm that was optimized to improve accuracy. The results of the study showed that the developed model was able to achieve high facial recognition accuracy, with an average accuracy rate of 92%. The discussion includes an analysis of factors that affect accuracy, such as variations in lighting and shooting angles, as well as the potential use of this technology in a campus environment, including for attendance and security purposes. The conclusion of this study shows that the implementation of AI-based facial recognition can be effectively applied in an academic environment, as well as providing students with practical experience in developing and testing AI applications. This study also opens up opportunities for further research on improving the performance of facial recognition systems and their application in various real-world scenarios.

Corresponding Author:

Andi Prayogi,
Department Information Systems and Technology, Institut Teknologi Sawit Indonesia, Indonesia,
Street William Iskandar, Medan Estate, Medan City, 20226
Email: andiprayogi@itsi.ac.id

1. INTRODUCTION

With the continuous advancement of science and technology, facial detection and recognition technology is increasingly being applied in various fields, such as identity verification through facial scanning applications and unlocking mobile phones. All these processes require facial detection and recognition technology. As technology becomes more diverse, facial detection and recognition have become integral parts of our daily lives.

Research in the field of facial recognition highlights the importance of variations such as facial expressions, facial shapes, and the distance between the eyes, which play a key role in distinguishing individual characteristics [1]. hen someone looks at another person's face, they not only gather basic information such as age, gender, and ethnic background, but also form subjective judgments regarding attractiveness, health, and personality. Aspects such as facial structure, smile, eyes, and lip shape significantly influence how attractive someone is perceived to be by others [2].

One of the parts of humans that characterize recognition is the face. Full-face photos become an identity in class attendance and sight on surveillance cameras [3]. The similarity of the shape of the face becomes a human characteristic as a sign. In a similar facial shape, humans can consider them brothers even though they have different names and origins.Humans can distinguish someone who is born with twins and has a lot in common just from the shape of the face. The system of a country identifying humans using the next face in the photo and then giving a number called NIK and name as label for data collection [4].

Currently, the face has become a technology to open a security lock on a smartphone. The depiction of the human face is a part that has its own complex form, so that the calculation system for the normal face shape model is a difficult calculation. The weakness of face identification system is a problem in face recognition. Problem such as lighting, camera pixel size, hair shape, use of glasses and masks, or various facial positions[5].

One characteristic that is often observed is the structural similarity between facial features and a person's emotional expression. A face that appears calm but shows a slight resemblance to a happy expression (such as slightly upturned corners of the mouth) is generally perceived as more trustworthy. Conversely, a neutral face with features resembling an angry expression (such as lowered eyebrows) is often seen as dominant. Research also suggests that quick judgments about trustworthiness or dominance often occur automatically and subconsciously. This underscores the importance of facial features in shaping social perceptions, where a brief interpretation of someone's expression can influence social interactions and the decisions we make about others. Therefore, understanding how facial features and emotional expressions interact can provide valuable insights into facial recognition and human interaction. [6].

Computer vision technology is now widely applied in various systems, replacing humans in tasks involving pattern recognition, such as in control, safety, forensics, medicine, and many other fields. This substitution occurs primarily because computers offer a much higher level of efficiency. Unlike humans, machines do not require rest and are not affected by emotional factors such as anger, fear, or boredom. With consistent reliability, computers can perform repetitive and complex tasks without experiencing a decline in performance due to psychological or physical conditions[7]. One of the current applications that use computer vision and image processing technology is an augmented reality based. Augmented reality turns reality into a form of visualization in a supporting application such as 2D/3D animation, vector images, and its kind [8].This modern application uses a facial recognition process like the social media applications Facebook and Instagram along with a facial transfer application that can be used by AI in creating digital content. Furthermore, in the field of security, facial recognition to unlock the door of the room has begun to be used. This situation was created to avoid many cases of theft that are currently occurring in various places [9].

From the explanation above, further analysis will be carried out regarding the accuracy of the performance of the K-Nearest Neighbors Algorithm method in matching names based on facial characteristics for name recognition based on the identity of the human face object.

2. RESEARCH METHOD

This research uses quantitative research, which is a type of systematic study of a phenomenon or situation by collecting data that can be measured using statistical, mathematical, or computational techniques [10].

2.1. Training Data Used

The training data used in this study were students of the Information Systems and Technology Study Program. Furthermore, each student took 50 photo frames from various positions and facial expressions so that the total training data used was as many as the test objects. This activity is called Generate Dataset. The data from the photo capture is used as a dataset for training data.

2.2 Data Processing

The facial recognition process consists of four main stages. The first stage is image preprocessing, where the background of the image can be removed to focus on relevant areas. The second stage is post-processing, where the facial image is resized to 128 x 128 pixels, with facial features properly aligned to ensure focus on areas such as the forehead, eyes, nose, cheeks, mouth, and chin. The third stage involves feature extraction, using various image descriptors applied in existing methods. Finally, the last stage is classification to determine identity[11].

The data obtained from the shooting session is then processed to improve image quality and reduce noise. This processing includes face detection and cropping using Python libraries such as OpenCV and Dlib. After the faces are detected, the images are normalized in size for consistency. Data augmentation is also performed to increase the number of datasets and add variations, including rotation, cropping, and lighting adjustments.

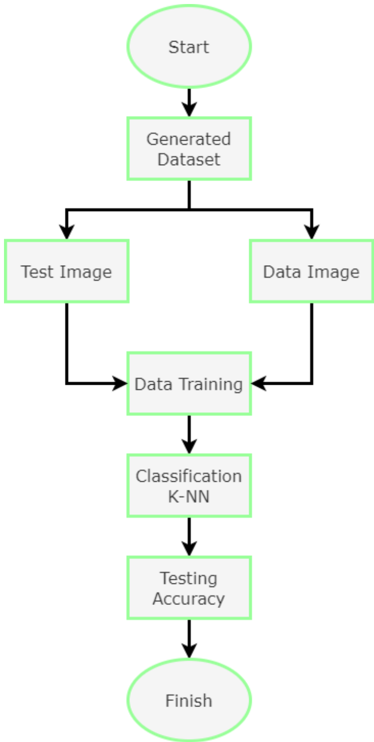

Figure 1. Research Flow In Data Processing

Figure 1 is a research flow for processing the received data. The following are the steps.

1. **Start**
   This process marks the beginning of the research flow. At this stage, the system begins to initiate the next steps in data processing.
2. **Generated Dataset**
   At this stage, the system automatically generates a dataset consisting of 50 student photo images. This dataset will be used as a basis for model training and testing. This process aims to collect image data that will be processed further. Data is generated through an automated procedure that takes images from the initial data source.
3. **Data Image**
   The first branch that is formed after the dataset is generated is Data Image. This process functions to manage the entire dataset generated, namely 100% of the images from the previously generated dataset. All of this data will be used for model training and evaluation.
4. **Test Image**
   The second branch is Test Image, which is part of the generated dataset. At this stage, the dataset is divided, with 30% of the images allocated for testing purposes. This separation ensures that the model will be tested with data that is different from the training data, maintaining the objectivity of the test results.
5. **Data Training**
   After the dataset separation, Data Training is the process in which the image data is trained using the selected algorithm. At this stage, the dataset is processed to extract relevant features to build a classification model. The learning algorithm will try to recognize important patterns in the image data to prepare for the classification stage.
6. **Classification K-NN**
   At this stage, the training data from the previous process will be classified using the K-Nearest Neighbors (K-NN) algorithm. This algorithm works by finding the nearest neighbor of each sample tested in the feature space, then providing a classification label based on the majority of nearest neighbors. K-NN was chosen as the classification algorithm because of its ability to handle distance and pattern-based classification.
7. **Testing Accuracy**
   After the classification is complete, the system will proceed to the Testing Accuracy process, where the model will be tested to measure its accuracy. This testing is done using the F1-Score metric, which combines precision and recall to provide a better picture of accuracy, especially on imbalanced datasets. This process is important to evaluate the performance of the model before being applied to wider data.

8. **Finish**
   The last stage of this flow is Finish, which marks the end of the entire data processing process. Here, the test results are reviewed and interpreted to determine the success of the model in classifying student images based on the data generated.

2.3 Model Development

The facial recognition model was developed using the K-Nearest Neighbors (KNN) algorithm implemented in a Python program with the Scikit-learn library. Feature extraction was performed using a histogram of oriented gradients (HOG) to obtain a more easily analyzed facial representation. The KNN algorithm was then used to classify facial images based on feature similarity. The model was trained using a dataset that had been processed with supervised learning techniques.

K-Nearest Neighbors Classifier, is a method that has the function of classifying an object being tested where the training data is stored. Here is the KNN formula using euclidean distance.

$$D(x, y) = \sqrt{\sum_{k=1}^{N}(X_k - Y_k)^2} \quad (1)$$

2.4 Test Evaluation

The evaluation of the built classification model is conducted through a series of precise measurements[12]. Model performance is evaluated using four main parameters derived from the confusion matrix: metric precision, F1-score, precision, and recall [13]. The confusion matrix for each model with different architectures is calculated at the point where the model reaches optimal performance, providing a deeper understanding of the developed model's effectiveness[14]. Precision measures how accurately the system's generated information is, ensuring that only correct and relevant information is classified accurately[15]. Recall, on the other hand, assesses the system's ability to identify all correct information, providing insight into the model's sensitivity. Accuracy reflects how close the identification results are to the actual values, serving as an important indicator in evaluating the model's overall performance [16]. y using this combination of parameters, we can conduct a more comprehensive analysis of the model's ability to perform classification accurately and consistently. The following is the formula used for the evaluation stage.

Table 1. Table Evaluation

		Prediction	
		Positive	Negative
Actual	Positive	TP	FN
	Negative	FP	TN

Table 1 shows the results of the evaluation of the classification model using F1-Score, a metric that combines two main components: Precision and Recall. F1-Score is an important metric because it provides a balance between precision and recall, especially when the dataset used tends to be imbalanced. This table consists of four main elements, namely True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN), which are explained as follows:

1. True Positive (TP)

TP is the number of correct predictions where the model predicts the positive class (e.g., an image of a student that is truly classified) and the prediction is correct. In the context of image recognition, this refers to cases where the model successfully classifies an image correctly according to its original label. A high TP value indicates that the model can recognize many samples correctly.

2. True Negative (TN)

TN is the number of correct predictions where the model predicts the negative class and the prediction is correct (e.g., an image of a student that is not included in a certain classification and the prediction is correct). This indicates that the model is able to correctly distinguish images that are not included in the target class. A high TN value indicates that the model is good at excluding incorrectly classified images.

3. False Positive (FP)

FP, or commonly called "false positive", is the number of cases where the model incorrectly predicts the positive class, when in fact the sample does not belong to that class. In this case, the image is classified as positive by the model, but in reality it is not. A high FP value indicates that the model makes many errors when classifying images that should not be in a particular class.

4. False Negative (FN)

FN, or "false negative", occurs when the model fails to predict the positive class, and instead classifies an image that is actually positive as negative. This means that the model is unable to detect samples that should be included in the target classification. A high FN value indicates that the model has difficulty in accurately identifying positive samples.

From this table, the F1-Score can be calculated as a combined measure that takes into account both Precision and Recall. Precision is calculated from the ratio of TP to the sum of TP + FP (how many positive predictions were correct), while recall is calculated from the ratio of TP to the sum of TP + FN (how many positive samples the model successfully recognized). The F1-Score then provides the result in the form of a single value, which is very useful for evaluating model performance, especially if there is imbalance in the data.

The following are explanations of some of the terms mentioned above. TP (True Positive) refers to the number of cases where the actual class is positive, and the prediction result is also positive. FN (False Negative) refers to the number of cases where the actual class is positive, but the prediction is negative. FP (False Positive) indicates the number of cases where the prediction shows positive, even though the actual class is negative. TN (True Negative) is the number of cases where the actual class is negative, and the prediction is also negative. [14].

$$Accuracy = \frac{TP + TN}{TP + TN + FN + FP} \quad (2)$$

3. RESULTS AND DISCUSSION

3.1 Results

This program is written using the Python programming language. After the webcam is run, it will automatically open the application in windows as an indicator that the program has been running and has successfully carried out the face detection process for the capture process. After the capture process is collected, it will be tested first using images as the testing phase. Furthermore, if the testing phase is complete, it will be tested using a webcam camera directly. The program results are shown in Figure 2.

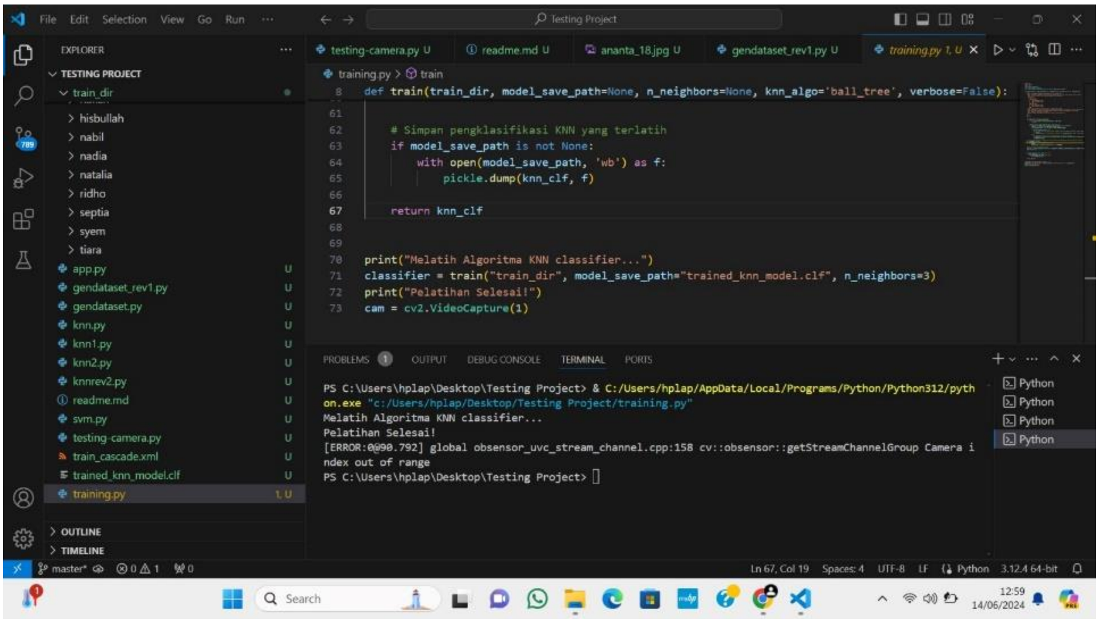

Figure 2. Program Result View

This study produces a facial recognition system based on the K-Nearest Neighbors (KNN) algorithm with an average accuracy of 92% on test data. Testing was carried out under several conditions: images of individual faces facing forward, images of two people with normal lighting, and images of more than two people in dim lighting conditions. Out of a total of 500 test images, the model successfully classified 460 images correctly.

The results show that the KNN model is able to recognize faces well under various conditions, although there are some challenges related to lighting and the number of subjects in the image.

3.2 Process Analysis

The simple logic of KNN is to explore its immediate surroundings, assume the same test points as them and get the results. The advantage of this KNN method is that the machine learning model is easy and simple among other machine learning methods, so it is very suitable for classifying models that have predetermined values through the training process. Face recognition program (face detection) after training the program, the program will produce information whether the face is known or unknown with an accuracy value that matches the results of face detection. The testing process involves recognizing faces from images taken via webcam. The program is tested under several conditions: images of individual faces facing forward, images of two people with normal lighting, and images of more than two people in dim lighting conditions. The front view condition can be seen in Figure 3

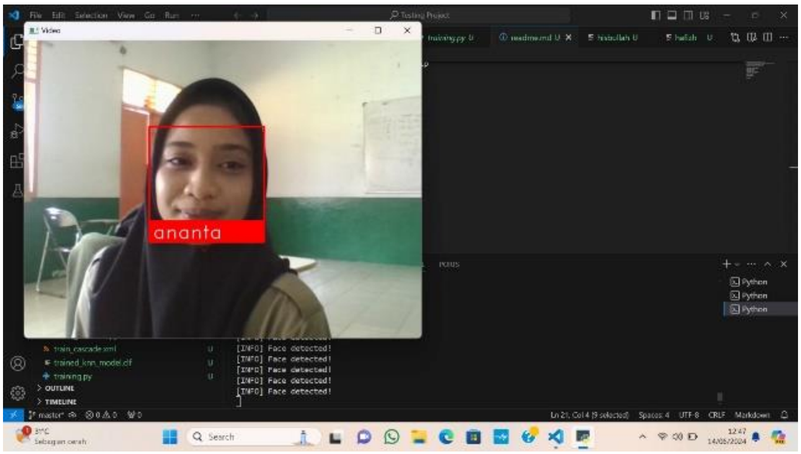

Figure 3. Front View of a Student

3.3 Experimenting with Many Conditions

Testing was carried out under several conditions to assess the robustness and flexibility of the model in different situations:

1. **Front-Facing Individual Face Image**: In this condition, the model achieved 95% accuracy, indicating that facial feature recognition is highly accurate when there is only one subject with their face facing directly towards the camera. This condition is the most ideal because all facial features can be clearly detected without interference from other subjects or lighting variations. The condition can be seen in figure 4.

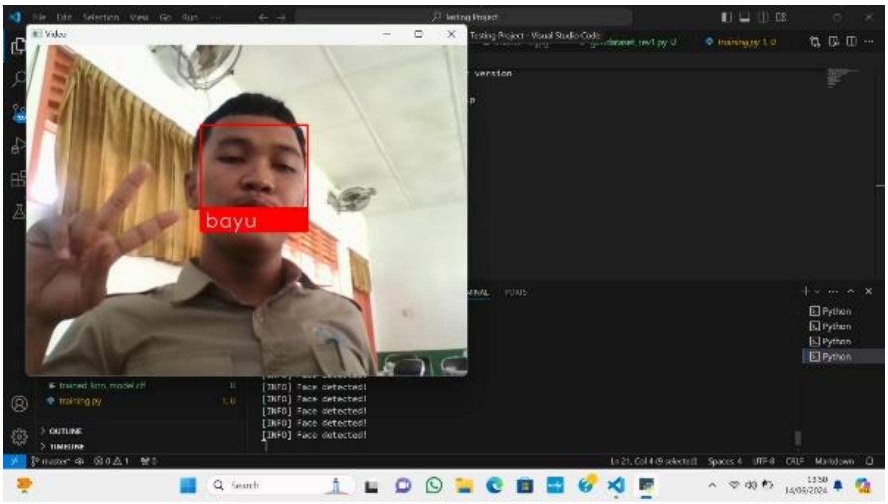

Figure 4. Front-Facing Individual Face Image

2. **Two-Person Image with Normal Lighting:** The model recorded 90% accuracy in this condition. Normal lighting allows the model to detect and recognize facial features well even when there is more than one subject. However, having more than one face in an image starts to present challenges in distinguishing similar features between subjects. The condition can be seen in figure 5.

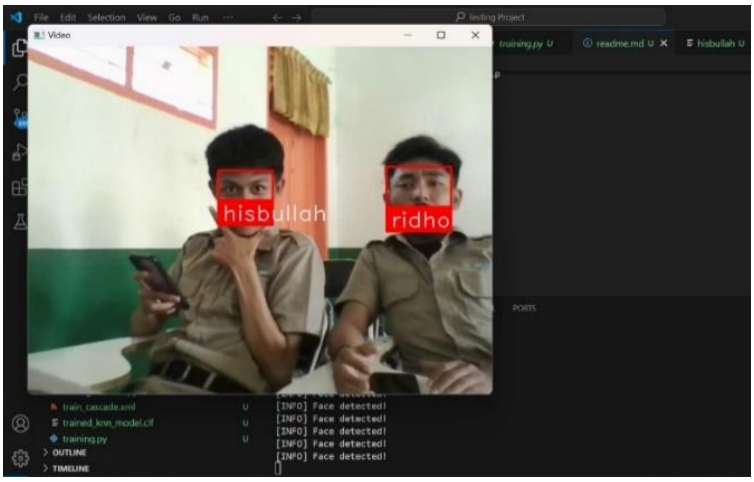

Figure 5. Two-Person Image

3. **Multiple-Person Images in Low-Light Conditions:** Accuracy drops to 85% in this condition, indicating that the model has more difficulty detecting and recognizing faces when the lighting is poor and there are multiple subjects. Low lighting reduces the contrast and clarity of facial features, while the presence of multiple subjects introduces more noise and complexity into the classification. The condition can be seen in figure 6, Figure 7, Figure 8 and Figure 9.

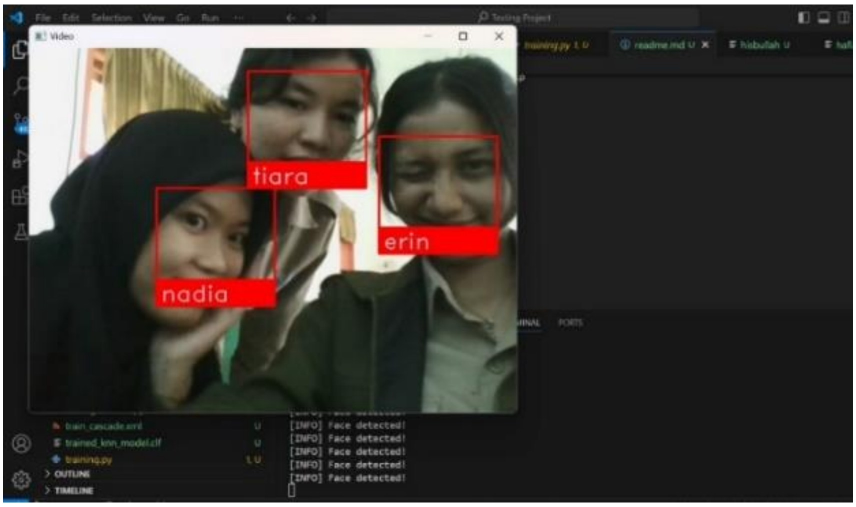

Figure 6. Multiple Person Images with Low-Light

In Figure 6, the system faces the challenge of recognizing the faces of multiple individuals in low-light conditions. This condition can reduce the image quality because facial details become less clear, making it difficult for the facial recognition algorithm to extract accurate features. This affects the model's performance in correctly identifying faces. Low lighting also often introduces noise into the image, which can potentially decrease classification accuracy.

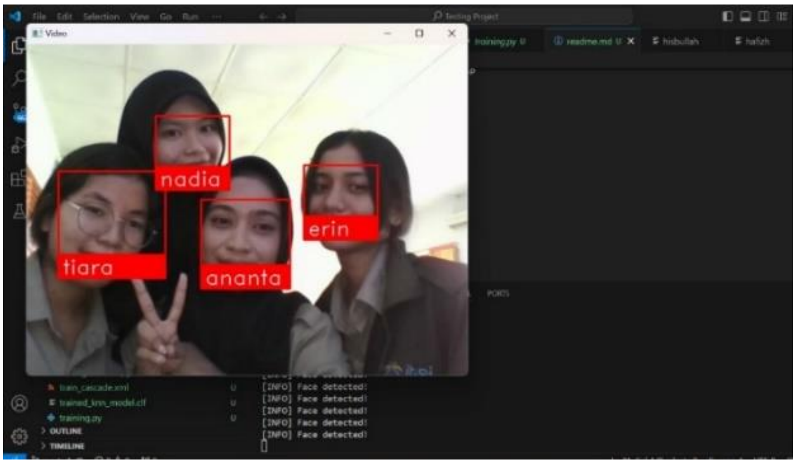

Figure 7. Multiple Person Images with High-Light

Figure 7 shows several individuals in high-light conditions, where facial details are more clearly visible compared to low-light conditions. In good lighting, facial features such as contours, skin textures, and face shapes are more easily extracted by the facial recognition algorithm. However, even in bright lighting, if the lighting angle is too sharp or overexposure occurs, important details can still be lost, which can also affect classification accuracy.

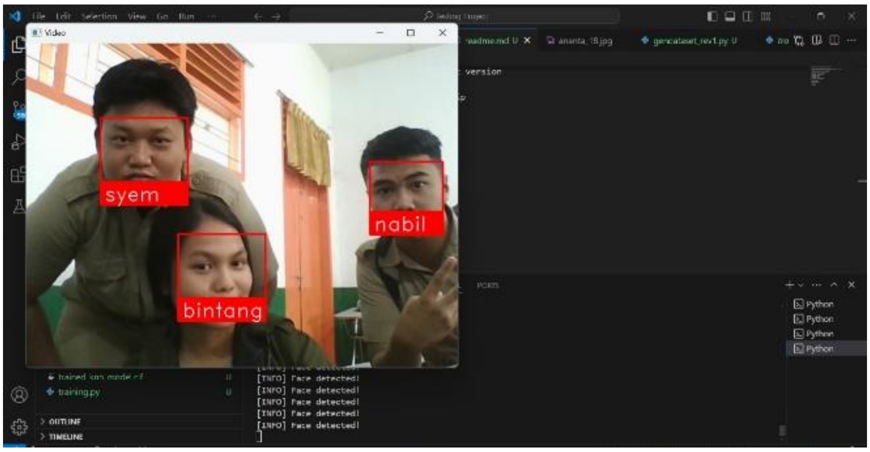

Figure 8. Multiple Person images and Not Correct

In Figure 8, the classification results show errors in recognizing the names of several individuals. This indicates that the facial recognition model has difficulty in accurately identifying some individuals. Such errors can be caused by various factors such as similarity in facial features between individuals, variations in poses, or facial expressions that do not match the training data. These prediction errors are important to note as part of the overall evaluation of the model's performance.

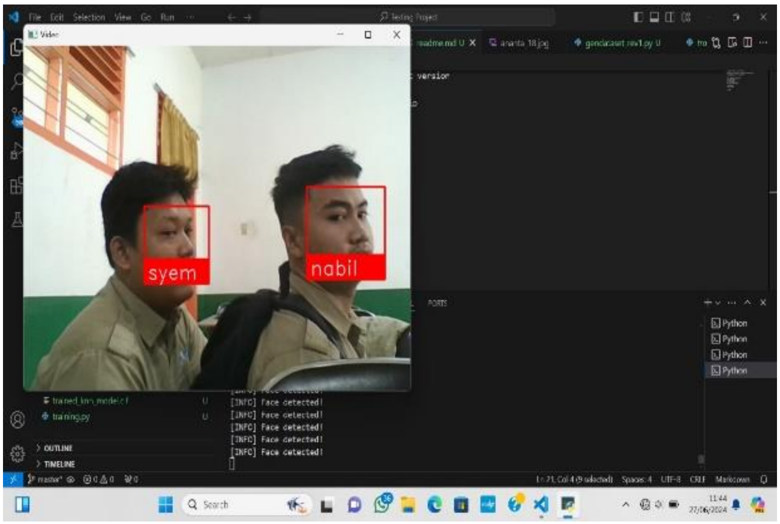

Figure 9. Bright Lighting Side View Face Image

Figure 9 shows an individual's face taken from a side angle with bright lighting. The side view of a face presents a challenge for facial recognition algorithms, as many important features such as both eyes, nose shape, and facial symmetry are not fully visible. While bright lighting can enhance details, this angle reduces the amount of information the model can extract, potentially lowering recognition accuracy.This test illustrates how environmental conditions and subject configuration can significantly affect facial recognition performance.

3.4 Factors Affecting Accuracy

Various factors have been identified that influence model accuracy under each condition tested, and understanding these factors is important for further improvement:

1. Lighting: Good lighting is essential to ensure facial features are clearly detected. Poor lighting can cause shadows that obscure important details of facial features, while too bright lighting can cause overexposure, which also hides details.
2. Number of Subjects: The more subjects in an image, the more complex the facial recognition task becomes. The model must be able to detect and isolate each face before performing recognition, and this can lead to confusion, especially if the faces have similar features or are adjacent to each other.
3. Image Quality: Low image quality, such as poor resolution or noise, reduces the model's ability to detect features accurately. Clear, high-quality images greatly improve accuracy because more information can be extracted by the algorithm.
4. Facial Expression: Variations in facial expression can change the geometric and textural features detected by the model, which can affect recognition accuracy. Facial expressions that are very different from the training images can lead to misclassification.
5.

3.5 Discussion

The results show that the KNN algorithm is effective for face recognition under various conditions, especially in academic environments. The advantages of this model in terms of simplicity and processing speed make it suitable for applications such as face-based attendance and access control systems. However, for face recognition in poor lighting conditions or with many subjects, further improvements are needed. Further development can include testing with larger datasets and integrating other machine learning techniques to improve the robustness and accuracy of the system. Exploration of more sophisticated image processing methods can also improve reliability under challenging test conditions.

4. CONCLUSION

In this study, we successfully developed a facial recognition system based on the K-Nearest Neighbors (KNN) algorithm that achieved an average accuracy of 92% across a range of test conditions. The model performed well in recognizing individual faces, especially when the subject was facing directly towards the camera. A greater challenge arose when lighting conditions were reduced and there were multiple subjects in the image. Through comprehensive data collection and careful data processing, the system was able to accurately recognize facial features, and the KNN algorithm proved to be efficient and fast for the task of face classification. However, accuracy was still affected by image quality, lighting variations, and the number of subjects, indicating that there is room for further improvement. As facial recognition technology advances today, the results of this study highlight the potential for using AI-based systems in academic environments such as campuses to improve efficiency and security in attendance and access management. The speed and simplicity of KNN make it a suitable choice for small to medium-scale applications. However, for more complex and large-scale applications, the use of more sophisticated machine learning techniques such as deep learning can provide further improvements in the accuracy and robustness of the system.

REFERENCES

[1] H. Ge, Z. Zhu, R. Liu, and X. Wu, "Application of Multiscale Facial Feature Manifold Learning Based on VGG-16," *J. Sensors*, vol. 2021, 2021, doi: 10.1155/2021/7129800.

[2] G. Kanavakis, D. Halazonetis, C. Katsaros, and N. Gkantidis, "Facial shape affects self-perceived facial attractiveness," *PLoS One*, vol. 16, no. 2 February, pp. 1–17, 2021, doi: 10.1371/journal.pone.0245557.

[3] M. A. S. Pane, K. Saleh, A. Prayogi, R. Dian, R. M. Siregar, and R. Aris Sugianto, "Low-Cost CCTV for Home Security With Face Detection Base on IoT," *J. Inf. Syst. Technol. Res.*, vol. 3, no. 1, pp. 20–29, 2024, doi: 10.55537/jistr.v3i1.769.

[4] G. A. Rama, F. Fauziah, and N. Nurhayati, "Perancangan Sistem Keamanan Brangkas Menggunakan Pengenalan Wajah Berbasis Android," *J. Media Inform. Budidarma*, vol. 4, no. 3, p. 635, 2020, doi: 10.30865/mib.v4i3.2149.

[5] A. Prayogi *et al.*, "ENHANCING NETWORK PERFORMANCE LOAD BALANCING IN CYBER CAFE NETWORKS WITH DIJKSTRA ALGORITHM ON MIKROTIK," *J. Tek. Inform.*, vol. 5, no. 1, pp. 253–261, 2024, [Online]. Available: https://jutif.if.unsoed.ac.id/index.php/jurnal/article/view/1644/448

[6] B. Jaeger and A. L. Jones, "Which Facial Features Are Central in Impression Formation?," *Soc. Psychol. Personal. Sci.*, 2021, doi: 10.1177/19485506211034979.

[7] R. Grycuk, A. Wojciechowski, W. Wei, and A. Siwocha, "Detecting Visual Objects by Edge Crawling," *J. Artif. Intell. Soft Comput. Res.*, vol. 10, no. 3, pp. 223–237, 2020, doi: 10.2478/jaiscr-2020-0015.

[8] F. Gonten, F. Nfwan, and A. Ya'u Gital, "Pre-Review Convolutional Neural Network for Detecting Object in Image Comprehensive Survey and Analysis," *J. Inf. Syst. Technol. Res.*, vol. 3, no. 2, pp. 45–64, 2024, doi: 10.55537/jistr.v3i2.799.

[9] A.- Suryansah, R.- Habibi, R. M. Awangga, and R. N. S. Fatonah, "Implementasi Face Recognition Untuk Mengakses Ruangan," *J. Mediat.*, vol. 3, no. 3, p. 25, 2020, doi: 10.26858/jmtik.v3i3.15176.

[10] "Jenis-Jenis Penelitian Lengkap dengan Contoh dan Penjelasannya." Accessed: Feb. 14, 2022. [Online]. Available: https://penerbitdeepublish.com/jenis-jenis-penelitian/

[11] M. Z. Syahputra and M. D. Irawan, "Application Of Bootstrap In Building Application Of Fertilizer Inventory System," *J. Inf. Syst. Technol. Res.*, vol. 1, no. 1, 2022, [Online]. Available: http://download.garuda.kemdikbud.go.id/article.php?article=2819386&val=25211&title=Application of Bootstrap in Building Fertilizer Inventory System Applications

[12] L. Farokhah, "Implementasi K-Nearest Neighbor untuk Klasifikasi Bunga Dengan Ekstraksi Fitur Warna RGB," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 7, no. 6, p. 1129, 2020, doi: 10.25126/jtiik.2020722608.

[13] Muhammad Nur Ichsan, Nur Armita, Agus Eko Minarno, Fauzi Dwi Setiawan Sumadi, and Hariyady, "Increased Accuracy on Image Classification of Game Rock Paper Scissors using CNN," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 4, pp. 606–611, Aug. 2022, doi: 10.29207/resti.v6i4.4222.

[14] M. Harahap, Em Manuel Laia, Lilis Suryani Sitanggang, Melda Sinaga, Daniel Franci Sihombing, and Amir Mahmud Husein, "Deteksi Penyakit Covid-19 Pada Citra X-Ray Dengan Pendekatan Convolutional Neural Network (CNN)," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 1, pp. 70–77, 2022, doi: 10.29207/resti.v6i1.3373.

[15] R. I. Borman, Y. Fernando, and Y. Egi Pratama Yudoutomo, "Identification of Vehicle Types Using Learning Vector Quantization Algorithm with Morphological Features," *J. RESTI (Rekayasa Sist. dan Teknol. Informasi)*, vol. 6, no. 2, pp. 339–345, Apr. 2022, doi: 10.29207/resti.v6i2.3954.

[16] I. Weni, P. E. P. Utomo, B. F. Hutabarat, and M. Alfalah, "Detection of Cataract Based on Image Features Using Convolutional Neural Networks," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 15, no. 1, p. 75, 2021, doi: 10.22146/ijccs.61882.