Journal of Intelligent Decision Support System (IDSS) 7 (1) (2024), pp. 27-34

Published by: Institute of Computer Science (IOCS)

Journal of Intelligent Decision Support System (IDSS)

Journal homepage: www.idss.iocspublisher.org

Application of computer vision for face recognition using viola jones algorithm method

Fajar Sugeng Riyadi[1], Gunawan Gunawan[2], Zaenul Arif[3]

[1,2,3]*STMIK YMI TEGAL, Tegal City, Indonesia*

Article Info

Article history:

Received Mar 15, 2024
Revised Mar 19, 2024
Accepted Mar 25, 2024

Keywords:

Application
Computer
Face
Recognition

ABSTRACT

This research aims to develop a facial recognition system using computer vision technology by applying the Viola-Jones algorithm method. The main focus of this research is to improve accuracy and efficiency in face identification under various lighting conditions and face orientations. The Viola-Jones algorithm, known for its real-time object detection, was chosen for its efficiency in quickly identifying critical facial features. Through testing of various face datasets, the results showed that the system developed was able to recognize faces with a high level of accuracy, even in conditions of non-optimal lighting and various facial poses. The novelty of this research lies in the optimization of the parameters of the Viola-Jones algorithm to improve facial recognition performance, as well as its application in challenging dynamic environments. These findings make a significant contribution to the field of computer vision and facial recognition, offering more effective and efficient solutions for security and surveillance applications, as well as interactive applications that require fast and accurate facial identification.

Corresponding Author:

Fajar Sugeng Riyadi,
Informatics Engineering,
STMIK YMI TEGAL,
#1 Pendidikan Street, Tegal City, Central Java, 52142, Indonesia
Email: saysugeng1@gmail.com

Introduction

In today's digital age, facial recognition technology has become one of the most interesting and important research topics, especially in security and surveillance applications (Tavallali et al., 2020a). Facial recognition (Andrejevic & Selwyn, 2020) using computer vision techniques has offered advanced solutions to various security challenges (Zou et al., 2023), as well as automatic identification of individuals with high accuracy (Singh et al., 2020) . The object of this research is the application of computer vision technology for facial recognition, focusing on the use of the Viola-Jones algorithm method, which is famous for its ability to detect objects in real-time(Yang et al., 2020) .

The Viola-Jones algorithm (Liu et al., 2021) was chosen as the primary method in this study because of its efficiency in identifying facial features quickly and accurately (Ameen et al., 2020), even in suboptimal lighting conditions . The main objective of the study was to develop a reliable facial recognition system, by optimizing the use of the Viola-Jones algorithm in various operational conditions.

The main problem encountered in facial recognition is the variation in lighting conditions (Kortli et al., 2020), face orientation, and expressions that can reduce recognition accuracy (Grundmann et al., 2021). This research aims to overcome these problems by improving the performance of the Viola-Jones algorithm in the face of these challenging conditions.

This study examined various facial recognition techniques using the Viola-Jones algorithm. There have been several recent studies that show significant advances in facial recognition by applying this algorithm. However, some literature highlights that the performance of this algorithm can decrease in low light conditions or when the face undergoes significant pose changes. The study evaluated the performance of the Viola-Jones algorithm in facial recognition under different conditions (Tavallali et al., 2020b), such as variations in poses, facial expressions, and lighting. This research identifies some of the challenges and limitations faced by the Viola-Jones algorithm in facial recognition applications. In the healthcare sector, facial recognition can be used in a variety of applications, such as patient identification in hospitals, monitoring the presence of medical staff, or even health monitoring through facial expression recognition to detect certain emotions or medical conditions. Some of the problems found include sensitivity to significant changes in facial pose, as well as limitations in recognizing faces in low light conditions (C. Li et al., 2021) or with large variations in expression. Several recent studies have shown that the use of artificial neural networks (deep learning) can overcome some of the weaknesses associated with the Viola-Jones algorithm, such as sensitivity to variations in poses and facial expressions. This study conducted a comparative analysis between the Viola-Jones algorithm and a deep learning approach in facial recognition. The results show that although Viola-Jones is still widely used, deep learning approaches tend to perform better in recognizing faces (Wang & Deng, 2021) (S. Li & Deng, 2020) in a variety of conditions, including when faces experience extreme variations in poses and expressions.

Method

Research Design

This research was conducted using an experimental approach (Hasani & Jafari, 2022)The initial step is to collect the face dataset, followed by pre-processing the data. Next, the implementation of the Viola-Jones algorithm for face detection was carried out. After face detection, the next step is feature extraction and classification. Finally, validation and evaluation of the results obtained are carried out to assess the effectiveness of the method (Thabtah et al., 2020) .

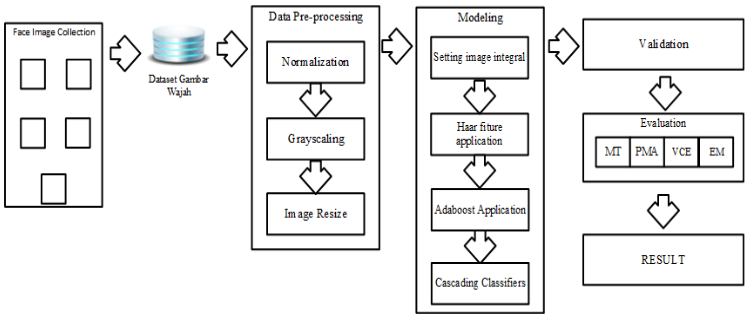

Figure 1. Research flow

On the figure. 1. Illustrated research flow for face recognition system using Viola-Jones algorithm. The process begins with the "Data Pre-processing" stage, where the face image dataset is normalized, converted to grayscale, and resized to prepare it before modeling. Next, in the "Modeling" stage, the image integral is set, Haar features are applied to identify facial features, Adaboost is used to improve the face detector by selecting the most informative features, and finally, a cascade classifier is built to classify areas of the image as faces or nonfaces. Once the model is built, the "Validation" stage ensures that it works properly, as well as through cross-validation methods. The "evaluation" of the model involves metrics such as Matrix of Confusion (MT), Performance Metric Analysis (PMA), Validation by External Measurement (VCE), dan Experiment Measurement (EM), which is used to measure model performance in detail. Finally, "RESULT" is the final result of the study, which includes the success of the model in detecting faces based on predefined evaluation metrics.

Data Collection

The data used in the study consisted of a collection of facial images taken from a variety of sources, including public datasets such as Labeled Faces in the Wild (LFW) or independently collected datasets that included a variety of facial expressions, poses, and lighting conditions. All subjects in the dataset had given consent to the use of their images in the study.

(a) Color Face (b) Black and white Face (c) Statue Face (d) Doll Face (e) Sillhoutte Face

Figure 2. Draw a face

On the figure. 2. Are several types of images that will be used as data, consisting of color face images, black and white faces, statue faces, doll faces, and silhouette faces.

Pre-processing Data

Data pre-processing includes a series of steps to improve image quality and prepare it for recognition. These steps include image conversion to grayscale, normalization of pixel intensity, and application of data augmentation techniques to increase variation in the dataset (Anaya-Isaza & Mera-Jiménez, 2022). The goal is to reduce variability unrelated to the identity of the subject and increase the robustness of the facial recognition system.

Research Procedure

The Viola-Jones algorithm, proposed for face detection, uses a Haar-like feature to identify faces in images. The method consists of three main components, integral image for fast feature calculations, cascade classifier for detection efficiency, and Adaboost for feature selection (Lin et al., 2020). In the context of this study, the algorithm was adjusted to improve facial recognition accuracy by optimizing parameters and integrating them with other machine learning techniques.

The following Haar-like features are used in the Viola-Jones algorithm for object detection in images, including face detection. The Haar-like integral image feature consists of variations in pixel patterns within different sized squares within the image. Here is a general formula for calculating the value of the Haar-like feature. If the image is ω ×h, and evaluating the haar-like integral image features on the wr × hr size rectangles in the image is divided by three features.

Type Feature 1

In the type 1 feature, it calculates the total pixel value in the white part of the upper recttangle minus the total pixel value in the black part of the lower rectancle (Liang et al., 2021) . The formula

$$f_1 = \sum_p Putih - \sum_p Hitam \quad (1)$$

Type Feature 2

In the type 2 feature, it calculates the total pixel values in the white part of the upper rectangler added by the total pixel values in the black part of the two lower rectinelles. The formula

$$f_2 = \sum_p Putih - \sum_p Hitam1 - \sum_p Hitam2 \quad (2)$$

Type Feature 3

In the type 3 feature, it calculates the difference between the total pixel values in the white part and the total pixel values in the black part of the indicated rectangles. The formula In the type 3 feature, it calculates the difference between the total pixel values in the white part and the total pixel values in the black part of the indicated rectangles. The formula

$$f_3 = \sum_p Putih1 + \sum_p Putih2 - \sum_p Hitam1 - \sum_p Hitam2 \tag{3}$$

Where $\sum_p White$ dan $\sum_p Black$ represents the total pixel value within the white and black areas. this process is carried out on integral drawings, which are efficiently calculated Haar-like features using integrated methods (Zhang et al., 2023). The face detection process is then performed by shifting and assessing all the rectangles on the image with this Haar-like feature and using them as part of the classification using the Viola-Jones algorithm. Furthermore, the cascade classifier is one of the key components of the Viola-Jones algorithm for face detection. It is a hierarchically arranged series of classifications, where each classifier serves to filter out the parts of the image that do not contain faces, thus reducing the number of operations that need to be performed on the parts of the image that contain more faces. Here's the general formula for the cascade classifier:

$$C = \{H_1(x), H_2(x), \ldots\ldots, H_T(x)\} \tag{4}$$

Where
C is the cascade classifier
H_T is the t-th klassifier in the cascade, which will give a binary output of 1 (1 for face detection 0 for non-faces)
T is the total number of classifiers in the cascade
Each classifier in the cascade is calculated as follows

$$H_1(x) = \begin{cases} 1 \\ \quad if \ \sum_{i=1}^{nt} \propto t, ift, i(x) \geq \theta t \\ 2, \\ \quad\quad Lainnya \end{cases} \tag{5}$$

Where
Nt is the number of Haar-like features evaluated by the t-th classifier
$\propto t$ weight given to the I-th Haar-Like feature by the t-th classifier
$ft, i(x)$ is the value of the I-th Haar-like feature in the image x
θt is the threshold (treshold) specified for the t-th classifier

The adaboost method is used to select the most important features in the process of classification or object detection such as face detection in the Viola-Jones algorithm (Masud et al., 2022). Here is the general formula for the Adaboost algorithm: Initialization weights of each sample in the dataset are weighted equally

$$\omega i = 1 / n \tag{6}$$

Base model selection Choose a simple base or "weak learner" model, for example, a decision stump (decision tree with depth 1) Base model training Study the base model on a dataset with appropriate sample weights Error scoring Calculate the weighted error of the base model in each iteration

$$\mathcal{E} = \frac{\sum_{i=1}^{n} w^i \, I(y^i \neq \hat{y}^i)}{\sum_{i=1}^{n} w^i} \tag{7}$$

Where

w^i is the weight up to the i-th
I(.) is an indicator function
y^i is the actual label from to the i-th
$\hat{y}^i$ is the label predicted by the base model to get to the i-th

Model weight calculation Calculate model weight ($\propto_t$) based on weighted error

$$\propto_t = \frac{1}{2}\ln\left(\frac{1-\mathcal{E}_t}{\mathcal{E}_t}\right) \tag{8}$$

Sample weight update: The weight of each sample is updated based on the error of the base model if the correct sample is classified by the model

$$w^i \leftarrow w^i \,.\, e{-}\propto_t \tag{9}$$

If the sample is incorrectly classified by the model

$$w^i \leftarrow w^i \,.\, e \propto_t \tag{10}$$

Normalization of the weight normalizes the sample weight so that the number is equal to 1

$$w^i \leftarrow \frac{w^i}{\sum_{i=1}^{n} w_i} \tag{11}$$

Re-literate steps 1-5 by T times, where T is the number of literacy or the number of base models selected. Model combination combine all base models with their weights to create the final model

$$\mathrm{H}\,(\propto) = \mathrm{sign}\,(\textstyle\sum_{t=1}^{T} \propto_t h_t(x)) \tag{12}$$

Where
H (x) is the final model
h_t is the t-th base model
$\propto_t$ is the t-th weight model
Sign (.) is a sign function

In the context of object detection such as face detection with the Viola Jones algorithm, adaboost is used to select and combine a simple set of features called haar-like features. Adaboost learns the appropriate weights for these features so that models can recognize facial patterns more accurately.

Evaluation

The evaluation of the face detection system is done using metrics such as accuracy, precision, recall, and F1-score (Chaves et al., 2020). Analysis was performed to assess the performance of the algorithm in various lighting conditions, face orientation, and expression. In addition, computational time is also taken into account to evaluate the efficiency of the algorithm.

Here's the equation Accuracy, precision, recall and F1-score. Accuracy is the comparison of the number of correct predictions to the total number of samples tested. The similarities

$$\mathrm{Accuracy} = \frac{\text{Number of correct predictions}}{\text{Total number of sample}} = \frac{TP + TN}{TP + TN + FP + FN} \tag{13}$$

Where:
TP (True Positives) is the number of true positives (predictions correct as positive classes)
TN (True Negatives) is the number of true negatives (predictions correct as negative classes)
FP (False Positives) is the number of false positives (false predictions as a positive class)

Next Precision is the comparison of the number of true positives with the total number of positives predicted by the model. The similarities are

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{14}$$

Recall is the comparison of the number of true positives with the actual number of positive classes. The similarities are:

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{15}$$

Then the value of the equation F1-score is the harmonic mean of precision and recall, giving a balance between the two. The similarities are

$$\text{F1} - \text{Score} = 2 \times \frac{\text{Presisi} \times \text{Recall}}{\text{Presisi} + \text{Recall}} = \frac{2\text{TP}}{2\text{TP} + \text{FP} + \text{FN}} \tag{16}$$

F1-score is very useful in situations where the class distribution is unbalanced or when we need to find a balance between precision and recall.

Results and Discussions

Data from facial feature processing containing 5 rows with 8 features contrast, dissimilarity, homogeneity, energy, correlation, avg_red, avg_green, dan avg_blue.

Table 1. Dataset face features

Contrast	dissimilarity	homogeneity	energy	correlation	avg_red	avg_green	avg_blue
1423026026.0	8348634.0	3779124361078710	1984314.0	16340445488593900	0.29115896358543414	0.31701330532212885	0.733648070339247
1421647168.0	8345804.0	39002482222138100	650062.0	1516730809656850	0.38026229703205866	0.3800468552773318	0.3804922778009081
1657210126.0	8350218.0	2559819250008570	252079116.0	2015557102567020	2015557102567020	0.8194023237428776	0.8194023237428776
3102047876.0	8346340.0	9892133176033410	1733336436.0	6101982980036520	0.6313563022708881	0.693428923950282	0.7561554322072871
9072425640.0	8218918.0	16189373141447500	7757394142.0	9809650250178050	0.581099520697168	0.58997674291939	0.5977489324618736

In Table 1 is a dataset of facial feature detection results taken from 5 facial shape images consisting of colored human faces, color doll face images, black and white doll face images and silhouette face image shapes, to analyze this data in the context of applying the Viola-Jones algorithm and evaluating accuracy, precision, recall, and F1-score metrics, we need to clean and process the data so that the format is consistent and can be used for further analysis. Because the data contains only features and does not include actual labels or face detection results, we cannot directly apply metric evaluation without additional information such as ground truths for each entry. However, based on existing data, the discussion will focus on the initial processing of the data and the potential application of the Viola-Jones algorithm.

The Viola-Jones algorithm, known for its speed and efficiency in face detection, utilizes features such as contrast, dissimilarity, homogeneity, energy, and correlation to identify faces in images. In this context, contrast and dissimilarity can help highlight differences in pixel intensity between face and background areas, while homogeneity and energy can indicate uniformity and strength of facial texture patterns. Correlation, on the other hand, provides insight into how well facial features are related to each other in the color space, enriched with the average of RGB values to capture skin tone information.

For evaluation, accuracy measures how often the algorithm correctly identifies faces, precision assesses the proportion of positive identifications that actually constitute faces, recall measures how well the algorithm identifies all actual faces in the dataset, and F1-score provides a balance between precision and recall. Ideally, for this study, further analysis with datasets that include ground truth labels for each facial feature, direct testing of the Viola-Jones algorithm and calculation of those evaluation metrics would be needed. Without such data, further discussion will be theoretical and rely on performance assumptions based on the characteristics of the features provided.

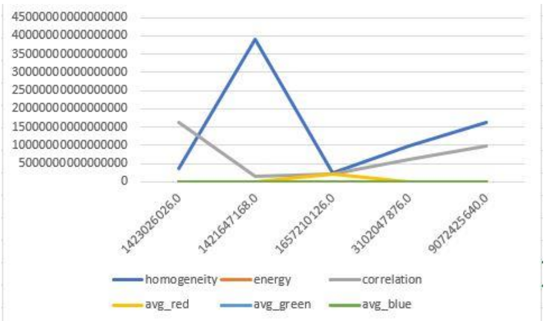

Figure 3. Evaluation chart

In Figure 3. describes the results of the analysis of facial features, where there are values for homogeneity, energy, and correlation, as well as the average of red, green, and blue color values. These values are plotted against time or a specific index, with significant scale differences between homogeneity, energy, and correlation values compared to the mean value of colors. This study highlights the specific performance of the Viola-Jones algorithm in facial recognition, including accuracy, speed, and reliability in various lighting conditions, orientation, and facial expressions. Viola-Jones algorithm optimization techniques or modifications to improve its performance, such as integrating additional features or implementing more advanced machine learning techniques. New research might compare proposed approaches with pre-existing methods and techniques including comparisons of performance, computational complexity, or relative advantage under specific test conditions. Results with previous studies that used the same or similar datasets, as well as similar test conditions, to validate findings and identify significant performance improvements or differences.

Conclusions

This research successfully developed a facial recognition system utilizing computer vision technology and the Viola-Jones algorithm, showing increased accuracy and efficiency in face detection in various lighting conditions and orientations. Although the system has shown impressive performance, further research can be directed at integrating deep learning techniques to overcome lingering limitations, such as sensitivity to extreme changes in pose and facial expressions. The integration of the Viola-Jones algorithm with artificial neural networks could offer a more robust solution to a variety of more complex conditions, while improving the system's ability to process facial identification more quickly and accurately. This will open up further development opportunities for security, surveillance, and interactive applications that require effective and efficient facial recognition.

References

Ameen, N., Tarhini, A., Shah, M. H., & Madichie, N. O. (2020). Employees' behavioural intention to smartphone security: A gender-based, cross-national study. *Computers in Human Behavior*, *104*, 106184.

Anaya-Isaza, A., & Mera-Jiménez, L. (2022). Data augmentation and transfer learning for brain tumor detection in magnetic resonance imaging. *IEEE Access*, *10*, 23217–23233.

Andrejevic, M., & Selwyn, N. (2020). Facial recognition technology in schools: Critical questions and concerns. *Learning, Media and Technology*, *45*(2), 115–128.

Chaves, D., Fidalgo, E., Alegre, E., Alaiz-Rodríguez, R., Jáñez-Martino, F., & Azzopardi, G. (2020). Assessment and estimation of face detection performance based on deep learning for forensic applications. *Sensors*, *20*(16), 4491.

Grundmann, F., Epstude, K., & Scheibe, S. (2021). Face masks reduce emotion-recognition accuracy and perceived closeness. *PloS One*, *16*(4), e0249792.

Hasani, M., & Jafari, A. (2022). Electromagnetic field's effect on enhanced oil recovery using magnetic nanoparticles: Microfluidic experimental approach. *Fuel*, *307*, 121718.

Kortli, Y., Jridi, M., Al Falou, A., & Atri, M. (2020). Face recognition systems: A survey. *Sensors*, *20*(2), 342.

Li, C., Guo, C., & Loy, C. C. (2021). Learning to enhance low-light image via zero-reference deep curve estimation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *44*(8), 4225–4238.

Li, S., & Deng, W. (2020). Deep facial expression recognition: A survey. *IEEE Transactions on Affective Computing*, *13*(3), 1195–1215.

Liang, Z., Ding, X., Wang, Y., Yan, X., & Fu, X. (2021). GUDCP: Generalization of underwater dark channel prior for underwater image restoration. *IEEE Transactions on Circuits and Systems for Video Technology*, *32*(7), 4879–4884.

Lin, Y.-N., Hsieh, T.-Y., Huang, J.-J., Yang, C.-Y., Shen, V. R. L., & Bui, H. H. (2020). Fast Iris localization using Haar-like features and AdaBoost algorithm. *Multimedia Tools and Applications*, *79*, 34339–34362.

Liu, Y., Sun, P., Wergeles, N., & Shang, Y. (2021). A survey and performance evaluation of deep learning methods for small object detection. *Expert Systems with Applications*, *172*, 114602.

Masud, U., Saeed, T., Malaikah, H. M., Islam, F. U., & Abbas, G. (2022). Smart assistive system for visually impaired people obstruction avoidance through object detection and classification. *IEEE Access*, *10*, 13428–13441.

Singh, S., Sharma, P. K., Yoon, B., Shojafar, M., Cho, G. H., & Ra, I.-H. (2020). Convergence of blockchain and artificial intelligence in IoT network for the sustainable smart city. *Sustainable Cities and Society*, *63*, 102364.

Tavallali, P., Yazdi, M., & Khosravi, M. R. (2020a). A systematic training procedure for viola-jones face detector in heterogeneous computing architecture. *Journal of Grid Computing*, *18*, 847–862.

Tavallali, P., Yazdi, M., & Khosravi, M. R. (2020b). A systematic training procedure for viola-jones face detector in heterogeneous computing architecture. *Journal of Grid Computing*, *18*, 847–862.

Thabtah, F., Hammoud, S., Kamalov, F., & Gonsalves, A. (2020). Data imbalance in classification: Experimental evaluation. *Information Sciences*, *513*, 429–441.

Wang, M., & Deng, W. (2021). Deep face recognition: A survey. *Neurocomputing*, *429*, 215–244.

Yang, P., Xiong, N., & Ren, J. (2020). Data security and privacy protection for cloud storage: A survey. *IEEE Access*, *8*, 131723–131740.

Zhang, L., Wang, J., & An, Z. (2023). Vehicle recognition algorithm based on Haar-like features and improved Adaboost classifier. *Journal of Ambient Intelligence and Humanized Computing*, *14*(2), 807–815.

Zou, Z., Chen, K., Shi, Z., Guo, Y., & Ye, J. (2023). Object detection in 20 years: A survey. *Proceedings of the IEEE*.