Journal of Theoretical and Applied Information Technology

15th September 2023. Vol.101. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

INTELLIGENT SECURITY SYSTEM BASED ON BIOMETRIC FACE RECOGNITION

IDA MULYADI[1], MUHAMMAD FAISAL[2*], KOMANG ARYASA[3], SAMSURIAH[4], MUSDALIFA THAMRIN[5], DARNIATI[6], MUHAMMAD ADNAN[7]

[1,2,7]Academian, STMIK Profesional, Department of Computer Science ,Makassar, Indonesia

[3]Academian, Department of Informatics Engineering, Universitas DIPA, Makassar, Indonesia

[4]Academian, Department of Management Informatics, STMIK Profesional, Makassar, Indonesia

[5,6]Academian, Department of Information System, STMIK Profesional, Makassar, Indonesia

E-mail: [1]idamulyadi@stmikprofesional.ac.id, [2]muhfaisal@stmikprofesional.ac.id, [3]komang.aryasa@undipa.ac.id, [4]samsuriahagus@stmikprofesional.ac.id, [5]nonongthamrin@gmail.com, [6]darniati@stmikprofesional.ac.id, [7]adnanahsann1212@gmail.com

ABSTRACT

Biometric face recognition is used in various applications, including security systems, access control systems, time and attendance systems, and public safety applications. The application system also requires a non-physical system that can be applied to certain parts to support creating a complete security system. As an essential layer of the security system in the mixing process, authentication techniques for mixing participants may be required during mixing. Currently, the application of facial recognition is still being developed through the development of methods to increase the recognition rate of facial recognition based on facial position and expression. Based on the comparison of the algorithms used, it is known that the CAMSHIFT algorithm has the best accuracy value, with an accuracy of 99.51%. Based on the experimental results, information was obtained that the F-measure value for FASTER R-CNN and CAMSHIFT was 0.96. Therefore it can be concluded that the Faster R-CNN method is a method that can be used to detect large numbers of objects. At the same time, the CAMSHIFT algorithm is ideal for use to support the authentication process on the face. In the future, a biometric-based security system can be implemented using the extended method to get better accuracy, resulting in speed and accuracy of detection and use in more dynamic conditions.

Keywords: *Faster-RCNN, CAMSHIFT, SURF, HAAR, Face Recognition, Confusion Matrix*

1. INTRODUCTION

Facial biometric recognition technology is one of the most convenient and practical because it allows authentication remotely without requiring manual authentication operations [1]. Biometric face recognition is used in various applications, including security systems, access control systems, time and attendance systems, and public safety applications. However, there are concerns about privacy and the potential for misuse of this technology, so it is essential to use it ethically and with appropriate safeguards in place. The importance of a comprehensive security system for the judicial court trial has been recognized, with the construction of bulletproof buildings receiving widespread support. However, it is crucial to understand that physical measures alone are insufficient to ensure adequate protection. To conduct the judicial trial solemnly and securely, all suitable elements, both physical and non-physical, should be integrated into the security system. In addition to physical security measures, non-physical systems tailored for specific areas are necessary to develop a complete and robust security infrastructure.

Biometric recognition technology has significantly progressed over the past decade and is now used in several services and applications [2], one application is an intelligent security system based on biometric facial recognition in the courtroom is a technology designed to improve security in the court environment using biometric-based facial recognition technology that aims to identify and authenticate individuals by comparing their unique facial characteristics with biometric data that has been stored in a database. Face recognition has grown in popularity as an authentication method due to its advantages over other biometric features

15th September 2023. Vol.101. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

used for identity matching [3], which promises an authentication mechanism already widely available in the era of mobile computing [4].
Face detection is a fundamental process that determines the presence and position of a face in images and videos, referred to as a bounding box. Despite the development of new techniques and methods, the system must be enhanced to distinguish the actual object from object spoofing with greater accuracy [5]. Therefore, in this research, a hybrid Faster R-CNN and CAMSHIFT were conducted to authenticate the participants of the judicial trial in real time so that the monitoring process through the computer system continues to be carried out when a trial is being carried out.

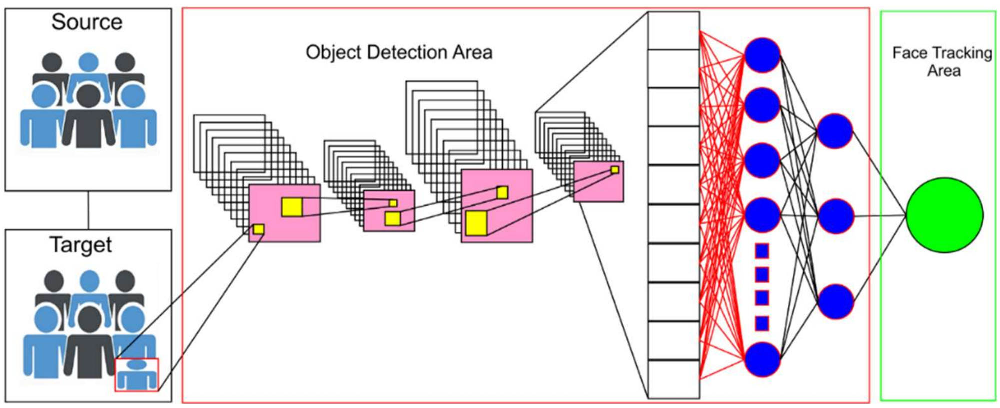

Figure 1. The research design for face recognition

Figure 1 shows an illustrative example illustrating the recognition process. The face authentication process is carried out in real-time in stages designed using a combination of Faster R-CNN and CAMSHIFT methods to obtain object detection results. In this case, the identification process uses the deep learning concept that aims to get the accuracy of face identification results, where personal profile data is stored in the database while saving face images to a file to be used as comparison material in the face identification process through the hidden layer.

2. LITERATURE REVIEW

The authors summarize several studies on face recognition using different methods in Table 1.

Table 1. Literature Review

Publications	Contribution	Comments
Elmahmud, Deep face recognition using imperfect facial data (2019)	founding that the parts of the face, including the eyes, nose, and cheeks, are the parts that have low recognition but are still able to produce a faster recognition rate when each part of the face is combined [6].	It is interesting to apply that before face authentication is performed, the image of the object on the face will be labeled first.
Kortli, Face recognition systems: A survey (2020)	From several experiences by researchers, it was found that the database can be used to support the face recognition process [7].	Interesting to apply the technique of using databases to support the face recognition process.
Adjabi, Taleb-Ahmed, Past, present, and future of face recognition: A review (2020)	The deep learning approach provides a reality for facial recognition research as a reference point for topics worth considering [8].	Lighting conditions and facial expressions are said to have an effect on face detection quality performance. As a result, in this study, lighting parameters and facial expressions will be used to evaluate the quality of face recognition.
Xiaochun, Faster R-CNN with Classifier Fusion for Small Fruit Detection (2018).	The Faster R-CNN Algorithm has the role of Five convolution screens that can be implemented to detect almonds [9].	This study will investigate the effectiveness of Faster R-CNN in improving object classification as a source of information.

15th September 2023. Vol.101. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Kollapudi, A Novel Faster RCNN with ODN-Based Rain Removal Technique (2022).	The foundation of the Densely Connected Networks Optimal Rain removal technique is the application of FRCNN-ODN is Faster R-CNN [10].	To produce high-quality regional proposals and assist the human object detection process, the Region Proposal Network (RPN) and Fast R-CNN model are elements that are also applied to FRCNN.

Based on the information collected from previous research in Table 1. Deep learning models for object detection can be implemented through the Faster R-CNN method or hybrid techniques with different methods to increase the accuracy value.

3. METHODOLOGY

3.1 Convolutional Neural Network (CNN)

A convolutional Neural Network is a development based on a human artificial neural network commonly used to detect and recognize objects in images. It consists of neurons with weights, biases, and activation functions. The illustration process is shown in Figure 2. Modularity in deep learning frameworks used in CNN generally allows flexibility for adaptation to diverse architectures.

The R-CNN and Fast R-CNN methodologies have flaws, one of which is the complexity of calculating the proposal region of the RPN, which cannot match the computational speed of CNN [12]. The Regional Proposal Network(RPN) is a deep convolutional neural network that builds the area to be detected in Faster R-CNN. In the object detection process using Faster R-CNN, an RPN module is responsible for utilizing the proposed area [13].

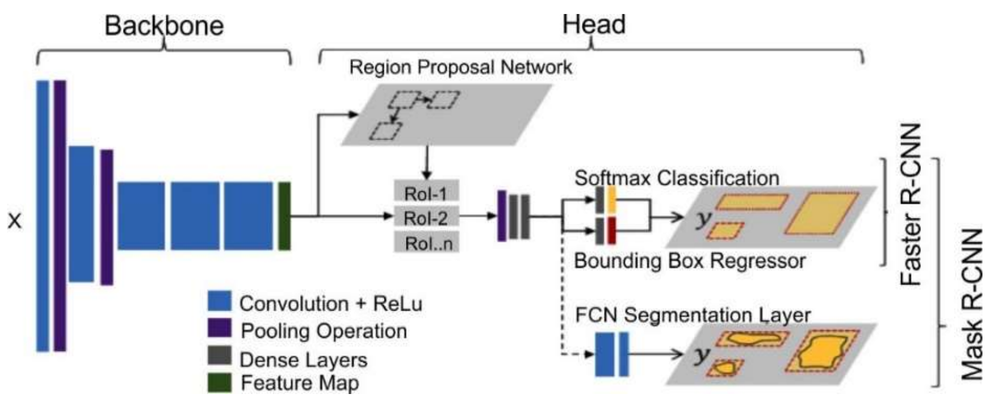

Figure 2. Faster-R-CNN and Mask-RCNN, respectively [13]

The primary layer in Faster R-CNN is the convolutional layer component, which performs convolution operations on the kernel matrix or filter matrix to extract the element from an image. Edge detection and corner detection are two examples of image feature values. The specific flows are shown in Figure 3.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

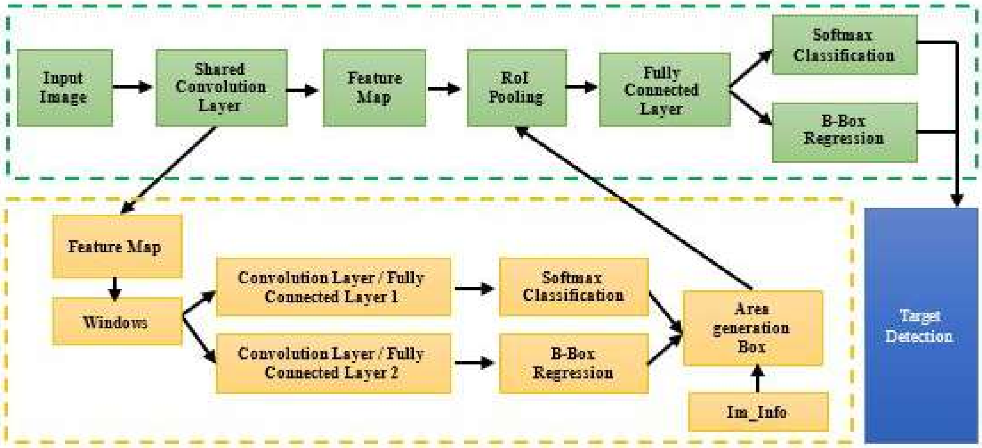

Figure 3. The Faster R-CNN algorithm framework [16]

The number and size of filters can be changed, and the value of the filter matrix can be randomized to produce different convolution results. Faster RCNN combines the offered RPN and Fast R-CNN networks into a sole network with shared convolution features [14] to improve the object tracking process's performance [15]. In general, the training process of Faster R-CNN employs random weight initialization by sampling from the Gaussian distribution and all layers with an initial learning level of 0.0001, which is accomplished through four stages. In the first step, an initialized network trains the RPN to generate a proposal boundary box as a candidate mass. The second step is to separate the classifier network training process using the RPN bounding box. The third step initializes the RPN, but the shared convolution layer parameter is frozen with the learning speed set to 0, and the RPN has trained again by updating the RPN unique layer. The fourth step involves performing a joint layer refinement to train the classifier network using the bounding proposal box obtained by updating the unique layer contained in the Faster R-CNN [17].

3.2 Haar Cascade

The Haar Cascade is based on a convolutional neural network, so it can detect an object quickly and in real-time [18]. Each stage classifier in the haar cascade detects whether the image sub-window contains an object of interest [18]. The decision rules in the Haar Cascade filter sub-images from the main image for faster detection using the pixel value formula to detect and identify objects based on image features:

$$\text{Pixel} = (\text{SDP} / \text{NDPs}) - (\text{SLPs} / \text{NLPs}) \quad (1)$$

Description :

PV = Pixel.
SDPs = Sum Dark Pixels
NDPs = Accumulating The Dark Pixels
SLPs = Sum Light Pixels
NLPs = Accumulating The Light Pixels

The process of calculating each pixel in the image is carried out through an algorithmic algorithm that involves each pixel as a variable p and its eight neighbours compared to the pixel variable p, where each neighbour will be assigned a value of 1 if it contains a variable x greater than or equal to the variable p [20], as written in the formula:

$$\text{LBP}(XC, YC) = \sum_{p-0}^{p-1} 2ps(IP - IC) \quad (2)$$

XC and YC variable is the centre pixel, the IC variable is the brightness, and the IP variable is the brightness of the adjacent. The function defined is written: $s(x)=1$ if $x \geq 0$ and $s(x) = 0$ otherwise.
In this study, the face authentication process is based on spoofing, where the case of facial spoofing is significant in preventing security breaches in the face recognition system [21]. The researcher stated that facial spoofing data could be presented naturally in a video stream format that uses temporal consistency to consolidate pseudo-label reliability for specific images [22]. Based on previous research, this study will perform facial authentication experiments on selected objects with the object selection process using Faster R-CNN. In contrast, the haar cascade is used in the facial spoofing process on selected objects.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

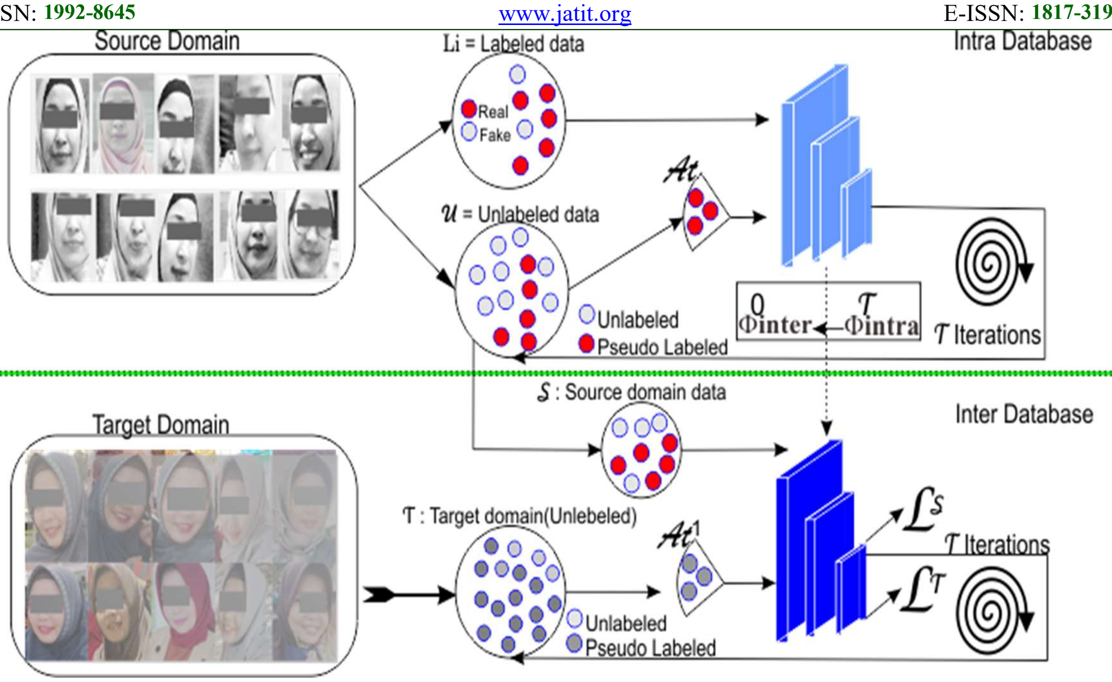

Figure 4. Progressive Transfer Learning for Face Anti-Spoofing

Figure 4 depicts an application of anti-spoofing technology.

3.3 CAMSHIFT(Continuously Adaptive Mean Shift)

CamShift stands for Continuously Adaptive Mean Shift, a facial recognition algorithm that adapts or adjusts to the colour probability distribution that always changes every frame change of a video sequence. The user's head may be outside the image display, the user's facial colour may change under different lighting conditions, and the user may differ in facial features such as moustaches and glasses [22], where The CamShift algorithm can be used for nose template matching for front facial posture recognition [23].

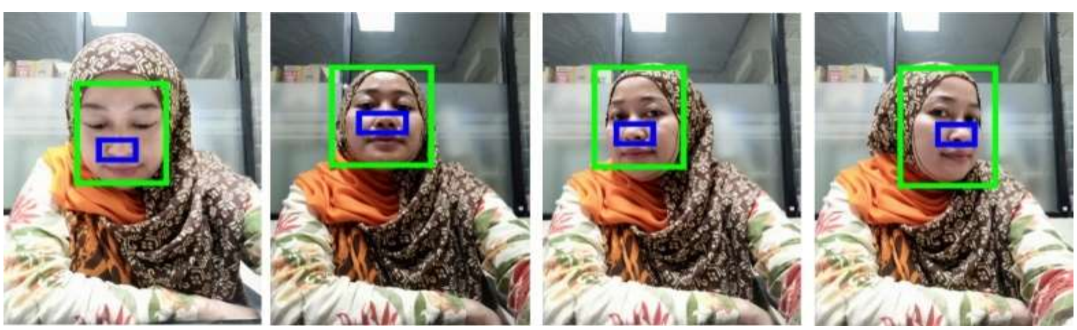

Figure 5. Nose Template Matching Method

In the CamShift algorithm, after the process of determining the search for a face frame has been successfully carried out, the next step is to perform image processing of the image colour probability distribution, where the mean area of the framed object can be calculated using the formula :

Zeroth Moment :

$$\text{Moo} = \sum_x \sum_z I(x,y) \tag{3}$$

First Moment x,y

$$M_{1o} = \sum_x \sum_z I(x,y) \ \ and \ \ M_{01} = \sum_x \sum_z I(x,y) \tag{4}$$

Mean Location in Centroid

$$X_c = \frac{M_{10}}{M_{00}} \ Y_c = \frac{M_{10}}{M_{00}} \tag{5}$$

Second Moment :

$$M_{2o} = \sum_x \sum_z I^2(x,y)$$

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

$$M_{02} = \sum_x \sum_z y^2(x,y) \quad (6)$$

Object Orientation :

$$\theta = \frac{arctan\left(\frac{2\left(\frac{M_{11}}{M_{00}}-x_c y_c\right)}{\left(\frac{M_{20}}{M_{00}}-x_c^2\right)-\left(\frac{M_{02}}{M_{00}}-y_c^2\right)}\right)}{2} \quad (7)$$

$$L = \frac{\sqrt{(a+c)+\sqrt{b^2+(a-c)^2}}}{2} \quad (8)$$

$$W = \frac{\sqrt{(a+c)-\sqrt{b^2+(a-c)^2}}}{2} \quad (9)$$

$$a = \left(\frac{M_{20}}{M_{00}} - x_c^2\right) \; b = \left(\frac{M_{11}}{M_{00}} - x_c y_c\right) \; c = \left(\frac{M_{02}}{M_{00}} - y_c^2\right) \quad (10)$$

All stages of the CAMSHIFT algorithm produce values x, y, object rotation, length, and width of the z region, which are the distance between the face and the camera position. The CAMSHIFT algorithm on face tracking uses nose matching, used in Figure 5.

3.4 SURF (Speeded Up Robust Feature)

The SURF algorithm is a development of the SIFT algorithm where SURF utilizes the computational speed of square filters by using an integral image, a matrix image in which the value of each pixel is the accumulation of the top pixel values and left. Besides having data robustness, the SURF algorithm also introduces data aggregation and filter boxes in calculations, which increases the registration time[24]. The SURF Algorithm formula, namely :
Citra Integral :

$$I(x,y) = \sum_{x=0}^{x} \sum_{y=0}^{y} N(x',y') \quad (11)$$

Matrix Hessian :

D = (A+B+C+D)-(A+B)-(A+C)+A (12)

Determinant Hessian:

$$det\left(H_{approx} = D_{(xx)}D_{(yy)} - (0{,}9D_{xy})^2\right) \quad (13)$$

Scale-space :

$$D(x) = D + \frac{\partial D^T}{\partial x} x + \frac{1}{2} x^T \frac{\partial^2 D}{dx^2} \quad (14)$$

Extreme Space :

$$\hat{X} = \frac{\partial^2 D^{-1}}{\partial x^2} x \frac{\partial D}{\partial x} \quad (15)$$

In the SURF Algorithm, the integral is the cumulative sum of the pixel values , and the zero padding adds zeros to the rows and columns of the image. The Hessian matrix consists of second-order Gaussian partial derivatives calculated from various filter box sizes using the octave scale as the image input. The SURF algorithm implemented employs the local maxima technique to detect interest points based on image pixel positions obtained by computing the determinant of the Hessian matrix. The next stage is carried out after the exciting point is detected, and a comparison is made between the threshold and the exciting point if there is a difference between the octave scales.

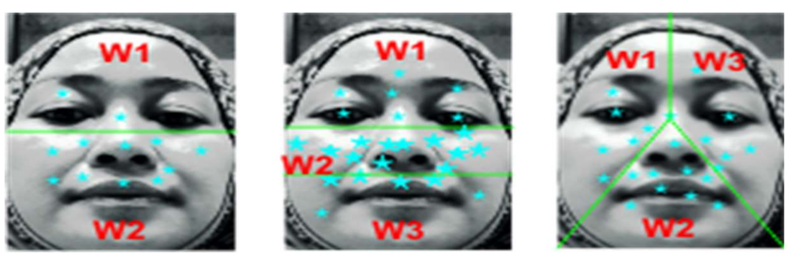

Figure 6. Face images of different clustering sub-regions in the SURF Algorithm

In Figure 6. The left image illustrates that the horizontal line divides the feature points. The middle image The feature points are cluster-tired into three equal parts. The right image The feature points are clustered and centred on the left eye, right eye, and the centre of the mouth.

3.5 Confusion Matrix

The confusion matrix is a machine-learning idea containing information about the actual class and the predicted classification[26]. The confusion matrix generates information about classification predictions and actuals using correct answer data. The results of testing the object detection and face recognition process using the Confusion Matrix calculation to obtain the values of rate accuracy, rate loss, precision, recall, and F1-Score were reported in this study. The explanation of the confusion matrix [11] is shown in Table 2.

15th September 2023. Vol.101. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 2. Confusion Matrix

Predicted	**Actual**	
	Negative (null detection)	**Positive (alternate detection)**
Negative (null detection)	TN	FN
Positive (alternate detection)	FP	TP

TP = Number of human objects detected.
TN = Number of other objects detected.
FP = Number of other objects detected as human objects.
FN = Number of human objects detected as other objects.

4. DISCUSSION AND RESULTS

The first step in this study is to collect judicial participant data into a database. The database contains two types of data: biodata and biometrics of judicial participants. Then, high-resolution cameras are used to produce high-quality images. When the trial is carried out, the camera will send video files in real-time, allowing the person's object to be identified during the authentication process. The selected face will then be compared to the facial image in the database. Several comparisons with commonly used facial recognition methods demonstrate the proposed approach's effectiveness and superiority [26]. The authentication process is repeated with different facial expressions to obtain the correct image-matching results.

4.1 Detection Quality

Object detection is about detecting the bounding box with the highest detection score for a given input image [27]. This study used The Faster R-CNN technique to implement object detection with video data converted into images. A text file containing information about the figure name, the bounding box size, and the class is provided before beginning the training process. Nine anchors are used based on the Faster R-CNN default anchor value. In this study, the video file becomes the input as an initial stage, where the video will be processed by convolution, and a pre-trained image will be made. A feature map is also created during this process to collect all information related to the vector representation of the input data. After gathering all information on the feature map, the Proposed Region Network will process the data to predict the image areas considered person objects and bounding boxes in the image area considered person objects. After that, the initial feature map information and the feature map information uploaded to the RPN will go to RoI Pooling. This study uses the procedure for collecting RoI Pooling, where a new data set from real-time video footage is used for the analysis phase. The data set consisted of a video sequence shot during the trial. Participants moved around to create various patterns. Video is shot at 50 frames per second in 1920 x 1080 resolution. To address the computational complication problem in video handling, video sequences captured for investigation are 20 to 50 seconds long and contain vital factors that may be present in longer video sequences.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Figure 7. Results of the Face Identification

According to Figure 7, using training results in an accuracy rate of 95.3%, indicating that the Faster RCNN is suitable for object detection processing. The test results on 10 sample videos are shown in Table 3. Then Table 4 shows the results of object participant detection: TP is worth 248, TN is worth 0, FP is 0, FN is worth 12, and then the value of the confusion matrix calculation is obtained.

Table 3. Result of Confusion Matrix

Accurate	95.3
Precicion	1
Recall	0.95
F1-Score	0.97
Rate Accurate	95.3
Rate Loss	4.70

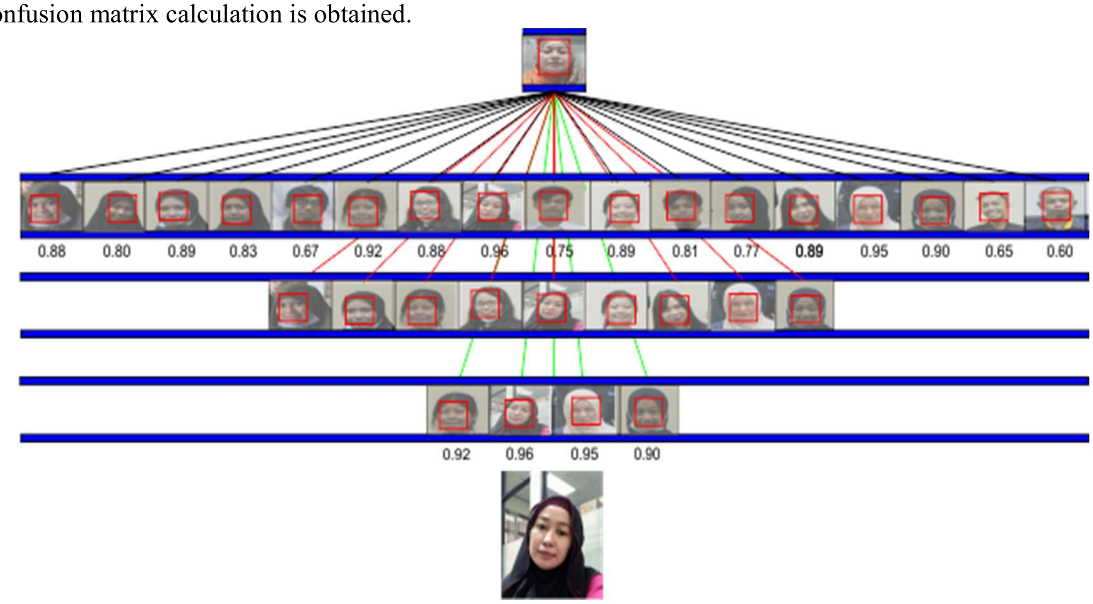

Figure 8. Result tracking of face recognize

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3. shows the overall average accuracy of all image types tested was 95.3%. Due to lighting factors, the image quality of a participant in the back position may result in less accuracy.

4.2 Face Tracking Quality

Facial biometric data has been used as a method of human identification and authentication with a high level of security in various applications, where developments continue to be made through advances in computer vision technology and pattern recognition technology [28]. As in this study, facial recognition is used to aid in identifying trial participants, with the results shown in Figure 8.
In Figure 8. The facial data tracking process is shown based on the selected input image, where each dataset will be compared with the input until the greatest similarity value is obtained.
In this study, a comparison was made using two methods, namely Faster R-CNN + CAMSHIFT, Faster R-CNN + SURF and Faster R-CNN + HAAR-LBP, to detect and face recognize in an inclined position in the image. The test results in Figure 7 show that both methods can identify the target face when compared to the image at the source. Furthermore, in Figure 8, the target face in an oblique position was also successfully recognized even though it was given a blur effect. However, it required a relatively longer time in the recognition process. Tests are carried out on each algorithm that is determined by using time and accuracy as variables, where the results are shown in Table 4.

Table 4. Result of Comparison Algorithm

OBJECT	TIME (sec/frame)	Faster RCNN+ CAMSHIFT	Faster RCNN+ SURF	Faster RCNN+ HAAR-LBP
			Accurate (%)	
1	0.35084	99.53	89.70	94.58
2	0.35282	99.42	90.50	96.35
3	0.47714	99.61	89.87	94.18
4	0.43022	99.05	90.88	96.02
5	0.47453	99.98	89.82	93.28
6	0.36288	99.57	92.89	95.95
7	0.35387	99.86	89.99	93.07
8	0.47680	99.48	91.25	96.74
9	0.44890	99.53	92.27	94.10
10	0.35156	99.10	89.60	93.47
Average	0.407956	99.51	90.68	94.77

Based on a comparison of the algorithm results, it is known that the CAMSHIFT algorithm has the best accuracy value is 99.51%. The novelty of this research is combining the FASTER-RCNN algorithm with the CAMSHIFT algorithm for facial recognition processes. This approach offers a unique and effective solution to improve accuracy and efficiency in identifying and tracking objects under various conditions. By harnessing the power of both algorithms, the researchers aim to achieve better results in detection and recognition tasks, making the fusion of these algorithms a promising and innovative contribution to computer vision and security systems.

4.3 Experimental Results

In this study, testing was carried out through an iteration process to determine the greatest accuracy value and the ideal time to use. The resulting learning rate value between 0-1 can be used as a measure of the speed of the ongoing training process. Another factor to consider is that if the learning rate is too large, the training process may exceed the optimal state when it reaches the minimum error value. In this study, testing was carried out to see the results obtained based on 150 with a learning rate of 0.0001. The test results can be seen in Figures 9 and 10.

15th September 2023. Vol.101. No 17

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

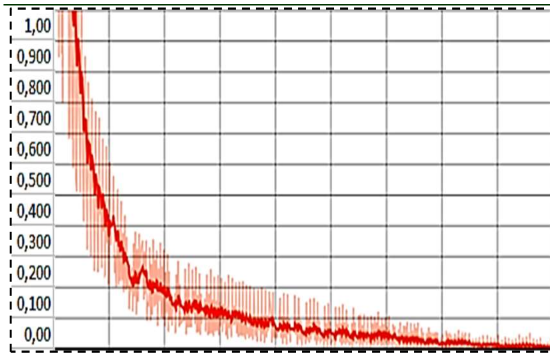

Figure 9. Training Loss 150 iterations

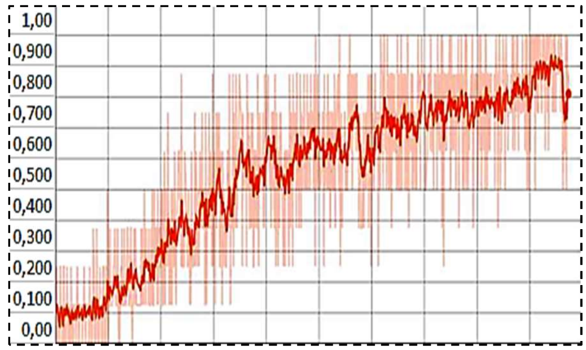

Figure 10. Training Accuracy 150 iterations

After obtaining the object detection process results, the next step is calculating the value of the selected object to be used as the comparison accuracy value between the participant data in the database and image extraction.

We considered many of the adapted datasets in Table 5 when conducting face identification and labelling trials. The main data set involves all photos taken at the Makassar District Court Office. The second data set consists of images with various facial orientations consisting of several photos containing images last data set comprises photos of some population data in the downtown area of Makassar City.

Table 5. The data set for detection and tagging

Data Set Name	Total Images
Participants in the District Court office	800
Custom Different Oriented Facess	1019
Randomly Selected Makassar People	1500

To investigate the efficacy of the planned method, we compute precision, recall, and F-measure values[29], which are defined as:
Precision (P) is the section of rescued documents that are relevant.

$$\text{Precision(P)} = \frac{\text{TruePositives}}{\text{TruePositives} + \text{FalsePositives}} \quad (16)$$

Recall (R) is the section of pertinent documents that are retrieved.

$$\text{Recall(R)} = \frac{\text{TruePositives}}{\text{TruePositives} + \text{FalseNegatives}} \quad (17)$$

The F-measure (F) is clear as a harmonical mean of Precision (P) and Recall (R).

$$\text{F-measure(F)} = \frac{2PR}{\text{Precision} + \text{Recall}} \quad (18)$$

The notions can be made clear by examining Table 6.

Table 6. Results for face tagging

Data Set Name	TP	TN	FP	FN	P	R	F
Participants in the District Court office	700	50	30	20	0.9589041	0.9722222	0.9655172
Custom Different-Oriented Faces	796	80	73	70	0.9159954	0.9191685	0.9175792
Randomly Selected People	1176	150	101	73	0.9209083	0.9415532	0.9311163

Based on the data in Table 6, information is obtained that the F-measure value of the results of the FASTER R-CNN and CAMSHIFT experiments is 0.96%. Therefore it was stated that the system could be used to support a security software-based security system at the district court office of Makassar City.

Based on Previous research, as shown in Table 1, has introduced several effective methods, algorithms, and databases for seamless facial recognition. These approaches have shown impressive performance in controlled environments with consistent lighting and viewing angles. However, their effectiveness may decrease significantly under specific conditions, mainly when dealing with variations in lighting and

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

facial angles. In this study, we aim to address this limitation by employing the Faster RCNN + CAMSHIFT method as an alternative solution to enhance facial recognition performance in challenging scenarios. Nonetheless, one of the limitations of this study is the real-time facial identification process, as rapid object movement can lead to slow-motion effects, causing difficulties in recognizing and tracking objects accurately. Contributions to this study are shown in Table 7.

Table 7. Contributions of the research

Feature	Contribution
Security system for the district court office of Makassar City	Security system berbasis bimetrik wajah at the district court office of Makassar City
Tracking Technique	Real-Time
Methodology	Faster RCNN + CAMSHIFT
Object	Participants at the district court office of Makassar City
Dataset	Load from database

5 CONCLUSION

This study concludes that the security system layer at the court office can use facial recognition technology in real-time. The test results show that Faster R-CNN + CAMSHIFT is an ideal method for face recognition on large datasets with an accuracy of 99.51%. Experimental results obtained the value for Participants in the District Court office of F-Measures 0.96. The limitation of this study is the difficulty of the real-time facial recognition process when objects move quickly. Future work of this research is building face recognition techniques using hybrid algorithms and applied to IoT Cloud Computing technology.

REFERENCES:

[1] H. Imaoka *et al.*, "The future of biometrics technology: from face recognition to related applications,"*APSIPA Trans. Signal Inf. Process.*, vol. 10, no. 1, 2021, doi: 10.1017/ATSIP.2021.8.

[2] B. Meden *et al.*, "Privacy–Enhancing Face Biometrics: A Comprehensive Survey," *IEEE Trans. Inf. Forensics Secur.*, vol. 16, pp. 4147–4183, 2021, doi: 10.1109/TIFS.2021.3096024.

[3] H. Pranoto and O. Kusumawardani, "Real-time Triplet Loss Embedding Face Recognition for Authentication Student Attendance Records System Framework," *JOIV Int. J. Informatics Vis.*, vol. 5, no. 2, pp. 150–155, May 2021, doi: 10.30630/joiv.5.2.480.

[4] B. Rexha, G. Shala, and V. Xhafa, "Increasing Trustworthiness of Face Authentication in Mobile Devices by Modeling Gesture Behavior and Location Using Neural Networks," *Futur. Internet*, vol. 10, no. 2, p. 17, Feb. 2018, doi: 10.3390/fi10020017.

[5] K. Mohan, P. Chandrasekhar, and K. V. Ramanaiah, "Object-specific face authentication system for liveness detection using combined feature descriptors with fuzzy-based SVM classifier," *Int. J. Comput. Aided Eng. Technol.*, vol. 12, no. 3, p. 287, 2020, doi: 10.1504/IJCAET.2020.106213.

[6] A. Elmahmudi and H. Ugail, "Deep face recognition using imperfect facial data," *Futur. Gener. Comput. Syst.*, vol. 99, pp. 213–225, 2019, doi: 10.1016/j.future.2019.04.025.

[7] Y. Kortli, M. Jridi, A. Al Falou, and M. Atri, "Face recognition systems: A survey," *Sensors (Switzerland)*, vol. 20, no. 2, 2020, doi: 10.3390/s20020342.

[8] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, present, and future of face recognition: A review," *Electron.*, vol. 9, no. 8, pp. 1–53, 2020, doi: 10.3390/electronics9081188.

[9] X. Mai, H. Zhang, and M. Q. H. Meng, "Faster R-CNN with classifier fusion for small fruit detection," *Proc. - IEEE Int. Conf. Robot. Autom.*, pp. 7166–7172, 2018, doi: 10.1109/ICRA.2018.8461130.

[10] P. Kollapudi *et al.*, "A Novel Faster RCNN with ODN-Based Rain Removal Technique," *Math. Probl. Eng.*, vol. 2022, pp. 1–11, May 2022, doi: 10.1155/2022/4546135.

[11] M. Faisal *et al.*, "Faster R-CNN Algorithm for Detection of Plastic Garbage in the Ocean: A Case for Turtle Preservation," *Math. Probl. Eng.*, vol. 2022, pp. 1–11, 2022, doi: 10.1155/2022/3639222.

[12] A. Z. Syaharuddin, Z. Zainuddin, and Andani, "Multi-Pole Road Sign Detection

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Based on Faster Region-based Convolutional Neural Network (Faster R-CNN)," *AIMS 2021 - Int. Conf. Artif. Intell. Mechatronics Syst.*, 2021, doi: 10.1109/AIMS52415.2021.9466014.

[13] T. Kattenborn, J. Leitloff, F. Schiefer, and S. Hinz, "Review on Convolutional Neural Networks (CNN) in vegetation remote sensing," *ISPRS J. Photogramm. Remote Sens.*, vol. 173, pp. 24–49, Mar. 2021, doi: 10.1016/j.isprsjprs.2020.12.010.

[14] L. Qi *et al.*, "Ship target detection algorithm based on improved faster R-CNN," *Electron.*, vol. 8, no. 9, 2019, doi: 10.3390/electronics8090959.

[15] Y. Ren, C. Zhu, and S. Xiao, "Object Detection Based on Fast/Faster RCNN Employing Fully Convolutional Architectures," *Math. Probl. Eng.*, vol. 2018, 2018, doi: 10.1155/2018/3598316.

[16] M. Fan, Y. Li, S. Zheng, W. Peng, W. Tang, and L. Li, "Computer-aided detection of mass in digital breast tomosynthesis using a faster region-based convolutional neural network," *Methods*, vol. 166, pp. 103–111, 2019, doi: 10.1016/j.ymeth.2019.02.010.

[17] A. Rastogi and B. S. Ryuh, "Teat detection algorithm: YOLO vs. Haar-cascade," *J. Mech. Sci. Technol.*, vol. 33, no. 4, pp. 1869–1874, Apr. 2019, doi: 10.1007/s12206-019-0339-5.

[18] L. T. H. Phuc, H. Jeon, N. T. N. Truong, and J. J. Hak, "Applying the Haar-cascade Algorithm for Detecting Safety Equipment in Safety Management Systems for Multiple Working Environments," *Electronics*, vol. 8, no. 10, p. 1079, Sep. 2019, doi: 10.3390/electronics8101079.

[19] A. B. Shetty, Bhoomika, Deeksha, J. Rebeiro, and Ramyashree, "Facial recognition using Haar cascade and LBP classifiers," *Glob. Transitions Proc.*, vol. 2, no. 2, pp. 330–335, 2021, doi: 10.1016/j.gltp.2021.08.044.

[20] A. Liu *et al.*, "Cross-ethnicity face anti-spoofing recognition challenge: A review," *IET Biometrics*, vol. 10, no. 1, pp. 24–43, Jan. 2021, doi: 10.1049/bme2.12002.

[21] R. Quan, Y. Wu, X. Yu, and Y. Yang, "Progressive transfer learning for face anti-spoofing," *IEEE Trans. Image Process.*, vol. 30, pp. 3946–3955, 2021, doi: 10.1109/TIP.2021.3066912.

[22] R. Bankar and S. Salankar, "Face Tracking Performance in Head Gesture Recognition System," *Int. J. Eng. Adv. Technol.*, vol. 9, no. 5, pp. 1096–1099, Jun. 2020, doi: 10.35940/ijeat.E1043.069520.

[23] P. Jia, H. H. Hu, T. Lu, and K. Yuan, "Head gesture recognition for hands-free control of an intelligent wheelchair," *Ind. Robot An Int. J.*, vol. 34, no. 1, pp. 60–68, Jan. 2007, doi: 10.1108/01439910710718469.

[24] Y. Chen, M. Zhang, B. Hyng, and R. Wang, "SURF Algorithm-Based Data Aggregation Method and Digital Sharing Economy," *Mob. Inf. Syst.*, vol. 2022, 2022, doi: 10.1155/2022/1513129.

[25] X. Deng, Q. Liu, Y. Deng, and S. Mahadevan, "An improved method to construct basic probability assignment based on the confusion matrix for classification problem," *Inf. Sci. (Ny).*, vol. 340–341, pp. 250–261, May 2016, doi: 10.1016/j.ins.2016.01.033.

[26] P. Lu, B. Song, and L. Xu, "Human face recognition based on convolutional neural network and augmented dataset," *Syst. Sci. Control Eng.*, vol. 9, no. S2, pp. 29–37, 2021, doi: 10.1080/21642583.2020.1836526.

[27] P. Mohandas, J. S. Anni, R. Thanasekaran, K. Hasikin, and M. M. Azizan, "Object Detection and Movement Tracking Using Tubelets and Faster RCNN Algorithm with Anchor Generation," *Wirel. Commun. Mob. Comput.*, vol. 2021, 2021, doi: 10.1155/2021/8665891.

[28] P. Hu, H. Ning, T. Qiu, Y. Xu, X. Luo, and A. K. Sangaiah, "A unified face identification and resolution scheme using cloud computing in Internet of Things," *Futur. Gener. Comput. Syst.*, vol. 81, pp. 582–592, 2018, doi: 10.1016/j.future.2017.03.030.

[29] C. Manning, P. Raghavan, and H. Schütze, "edition (c) 2009 Cambridge UP An Introduction to Information Retrieval," no. c, pp. 1–581, 2009, [Online]. Available: http://www-nlp.stanford.edu/IR-book/.