COMPARISON BETWEEN FACE AND GAIT HUMAN RECOGNITION USING ENHANCED CONVOLUTIONAL NEURAL NETWORK

Fatima Esmail Sadeq[*1], Ziyad Tariq Mustafa Al-Ta'i[2]
Department of Computer Science, College of Science, University of Diyala, Baquba, Iraq [12]
atomatotoatoma@gmail.com

ABSTRACT

Identifying people at distance is an important task in daily life Because of the increase in terrorism. Biometrics is a better solution to overcome personal identity problems, and this applies to soft biometrics also. Soft biometric are features that can be extracted remotely and do not require cooperation with people. This paper introduces a comparison between human face recognition and human gait recognition using soft biometric features. Nine face attributes and nine gait attributes are taken from a dataset built by researchers. The constructed dataset is composed from (66) videos for (33) persons. Features are extracted using Haar and MediaPipe methods. The extracted features are classified using enhanced convolutional neural network. This work achieves an accuracy of 95.832% in human face recognition and an accuracy of 89.583% in human gait recognition. From the above results it turns out that the proposed method achieved promising results with regard to Recognize people remotely.
Keywords: Soft biometrics, Face Recognition, Gait Recognition, MediaPipe, Convolutional Neural Network.

1. Introduction

Identification from a distance has become important due to the ever-increasing surveillance infrastructure that is being deployed in society (Reid *et al*., 2013). The use of biometrics in several surveillance applications such as human recognition or re-identification (Prakash *et al*., 2023). The difficulty of identifying people in varied positions is one of the main challenges modern person recognition techniques must overcome (Gharghabi *et al*., 2016). The development of a novel identification method known as soft biometrics is the result of substantial advancements made in the field of biometrics (Arigbabu *et al*., 2015). A need for biometric characteristics that enable the identification of people in difficult circumstances without demanding their cooperation has emerged (Reid & Nixon, 2013). Soft biometric features are defined as the characteristics that give distinctiveness and permanence to sufficiently differentiate any two individuals (Sikkandar & Thiyagarajan, 2020). Developing a standalone soft biometric system for recognition or retrieval is currently one of the biggest challenges due to certain critical factors directly affecting accuracy (Hassan *et al*., 2023).

Biometrics aims to recognize a person. Traditional biometrics has excellent accuracy and great versatility. But it is difficult to collect physical data from a distance, and also cooperation is often required like with lifting fingerprint (Kim *et al*., 2012). For this task, soft biometrics are valuable, since they provide useful information for identifying individuals (Romero *et al*., 2020). Soft biometrics is a new form of biometric identification which utilizes labeled physical or behavioral traits (Reid & Nixon, 2011). The term soft biometrics has been introduced in (2004) by Jain et. al. to describe a set of characteristics that provide some information about the individual (Jain *et al*., 2004). Soft biometric traits include characteristics such colors of eye, hair, beard and skin; shape and size of the head; general discriminators like height or weight and also description of indelible marks such as birth marks, scars or tattoos (Dantcheva *et al*., 2011). Anthropometric attributes are one type of soft biometrics that are refer to geometric and shape features of the face, body, and skeleton (Kang *et al*., 2018). The human face has various traits that have different levels of distinctiveness (Almudhahka *et al*., 2026). Usually, the overall accuracy of any soft biometrics system depends upon multiple factors, one of which is attribute correlation which is termed as supportive or non-supportive behavior of attributes towards each other during recognition process (Hassan & Izquierdo, 2022).

The link between soft biometrics and human description is one of the key benefits of using them; people intuitively utilize soft biometric qualities to identify and describe one another. Soft biometrics enable retrieval as well as identification. This is accomplished by bridging the semantic gap between human descriptions and biometric measures (Reid & Nixon, 2011). Systems that use soft biometrics have several advantages: they may be partially derived from the most widely used classical biometric identifiers; their acquisition is unobtrusive and requires no enrollment; training can be carried out beforehand on people who are not part of the target identification group (Dantcheva *et al.*, 2011). The evaluation of recognition performance in soft biometric recognition poses a challenge due to the absence of a standardized dataset, particularly when dealing with soft biometrics at varying distances. However, a notable advantage of soft biometrics is its flexibility regarding the resolution of captured images. Soft biometric labels can be collected even from a distance, eliminating the need for stringent image resolution requirements (Guo *et al.*, 2019). While face and gait are the primary biometrics applicable at a distance, they can encounter challenges such as low frame rate and/or resolution in surveillance situations. Nevertheless, it is noteworthy that even with low-resolution images, a comprehensive human description of the subjects can still be provided (Reid *et al.*, 2013). In contrast, soft biometrics present numerous advantages over alternative distance identification methods. They can be extracted from videos with low resolution and low frame rate, and possess robust attributes that remain unaffected by factors like camera viewpoint, sensor aging, and scene illumination (Tome *et al.*, 2014).

This paper uses soft biometric features an excellent accuracy was achieved in recognize person. Two different methods were used, the first method uses facial features. The face was detected using Haar classifiers. The trained Haar cascade analyses a picture and decides if the target item is there (Rudinskaya & Paringer, 2020). Then features were extracted using MediaPipe techniques. MediaPipe is an open-source framework with a hybrid platform that creates pipelines for processing perceptual data (Subramanian *et al.*, 2022). Nine soft biometric features for face (left eye distance, right eye distance, mouth distance, nose distance, area left, area right, area 2d mouth, area 2d region left, area 2d region right).

The second method used gait features extracted using MediaPipe techniques. Nine soft biometric features for gait (angle elbow, angle hip, length hip, angle mid hip, angle ankle, length ankle, distains ankle1, distains ankle2, distains mid pinky). Classification was done using convolutional neural network (CNN) for both methods. CNN is a supervised method, consists of an input and an output layer, as well as multiple hidden layers (Sakib *et al.*, 2019). Finally, the two methods were compared.

2. Literature Review

In the past years (2014- 2019), deep learning algorithms and specifically deep Convolutional Neural Networks (CNNs) have led to breakthroughs in many application domains including biometric recognition (Minaee *et al.*, 2023). Although deep learning research in biometrics has achieved good results, however, there is great room for improvement in different directions, such as using soft biometrics. Soft biometrics is extracted set of features from all human body (Garg *et al.*, 2018). Becerra-Riera *et al* (2019) presented an overview on face describable visual features and in particular of the so-called soft biometrics (e.g., facial marks, gender, age, skin color, and other physical characteristics). Terhörst *et al* (2020) proposed NFR (Negative Face Recognition), a new face recognition approach that enhances the soft-biometric privacy on the template-level by representing face templates in a complementary (negative) domain. Negative templates describe facial properties that does not exist for this person, while ordinary templates characterize facial properties of person. Sun *et al* (2023) they used convolutional neural networks (CNNs) for facial point detection. Global high-level features are extracted over the whole face region at the initialization stage, which help to locate high accuracy keypoints. Fard *et al* (2021) proposed a ASMNet, a lightweight Convolutional Neural Network (CNN) structure with multi-tasking learning to detect facial landmark points and to estimate head position. Vukadinovic & Pantic (2005) they proposed a robust, highly accurate method for detecting 20 facial points using GentleBoost classifier learned on features extracted with Gabor filters. Arca *et al* (2006) they introduced a completely automatic face recognition system. This

method works on color images they extracted (24) facial feature points, then extracted the feature of the points by Gabor. The recognition of the faces by compare the similarity of points' feature. Saleem *et al* (2023) they proposed a system to identify a person from a video clip using facen*et al*gorithm. This system analyzed the facial features of a person and accordingly classified the images.

In addition to facial features, there are soft biometric features that are extracted from human gait. Nithyakani *et al* (2019) proposed a scheme for gait recognition. They used deep convolutional neural network to extract the gait features of a person by training the neural network architecture with Gait Energy Image (GEI). Sharif *et al* (2020) they presented method for human gait recognition based on extracted shape features using histograms of oriented gradients (HOGs), geometric features, and texture features with local binary patterns (LBPs). Principal component analysis (PCA) was applied on the features for feature reduction. A support vector machine (SVM) was employed to perform the classification. Wang & Yan (2020) they introduced a novel gait classifier which takes full use of deep learning (DL) technology and proposed a novel gait-recognition method using a convolutional LSTM approach named Conv-LSTM. Elharrouss *et al* (2021) they presented a method to gait recognition for person re-identification. They used multitask CNN and extracted gait energy images (GEIs). Based on the GEI, features were extracted that were later classified using CNN. Chao *et al* (2019) they gave a new network named GaitSet to learn identity information from a set of silhouettes. This method exploited deep learning model to compute the cross-view gait representation using a set of independent frames. Sahak *et al* (2017) they presented method of human recognition based on oblique and frontal gait using features extracted from Kinect, orthogonal least square (OLS) is used for feature selection and multi-layer perceptron for classification. Then the optimized Multi-Layer Perceptron (MLP) with two feature sets is used for the recognition of gait and estimated its effectiveness by using neural network classifier that provides better classification results. Zhang *et al* (2019) presented an autoencoder-based method termed GaitNet that can disentangle appearance and gait feature representation from raw RGB frames, and utilize a multi-layer LSTM structure to further explore temporal information to generate a gait representation for each video sequence.

3. Dataset

The authors are locally constructed the dataset for this work. This dataset contains a video of various people that was recorded in real time as they passed in front of and next to the camera. The people photographed in this data set are university students, their ages range between (18-22) and they number 33 people. A camera was used for filming and it was its type Nikon d7200. The camera is positioned "130 cm" above the ground and (450 cm) away from the subject. Two videos are taken for each person, therefore sixty-six videos are used in this research. There are two types of data set (Front and Side dataset), table (1) shows information of the dataset.

Table 1 - Information of Dataset

No.	Dataset Details	Number
1	Number of sample (person)	33
2	The number of videos per sample	2
3	Frame Rate	1frame per sec.
4	Camera type	Nikon d7200
5	File Format	MP4

4. Research Methods
4.1 Face Recognition

Human face recognition includes four main steps: face detection, features extraction, pre-processing features, and face recognition. Face detection aims to locate the face regions in the images. Haar Cascade classification method is used to detect the face of person in videos. The Haar Cascade method is implemented in python code using Open CV's Cascade Classifier function.

In features extraction step, nine features are extracted by detecting region of interest (ROI) using Media Pipe Model by area of triangle face. For ROI determination, the Media Pipe gets the

bounding box and the main points, then draws a line between each two points to get the area of the triangle face. The output of the media pipe model is the major nine-coordinate for every face which are presented in figure (1). Four features are extracted by computing the distance of the four 2D landmark faces. These landmark faces are shown in figure (1), which are (left eye distance, right eye distance, mouth distance, and nose distance). The distance is calculated using Euclidian distance equation.

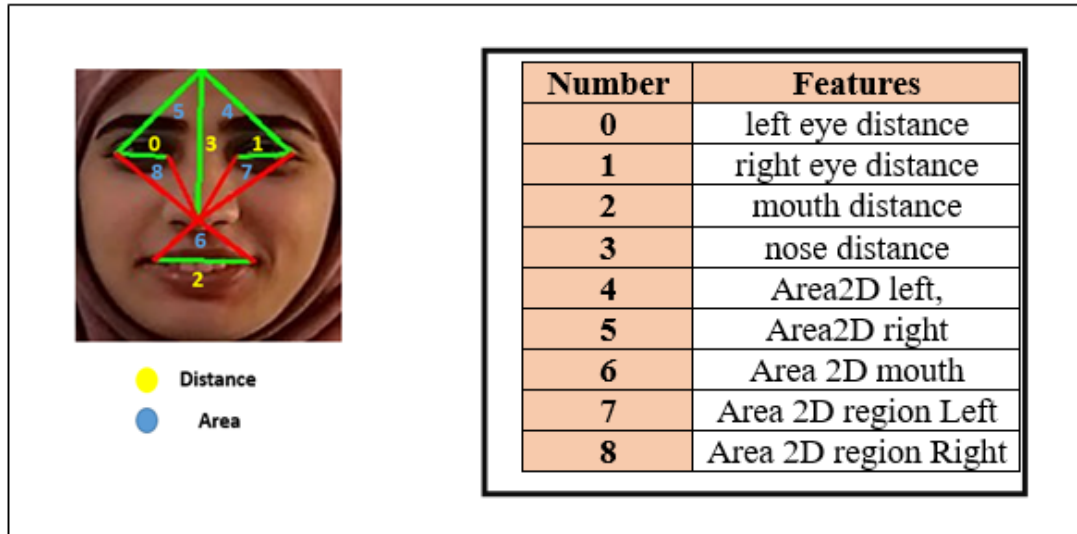| Number	Features
0	left eye distance
1	right eye distance
2	mouth distance
3	nose distance
4	Area2D left,
5	Area2D right
6	Area 2D mouth
7	Area 2D region Left
8	Area 2D region Right

Fig. 1. Face Features using Media Pip Model.

The area of polygon is determined by the area of the region it occupies. Figure (2) shows an example of the area of a polygon on a two-dimensional plane.

A regular polygon is a polygon that has equal sides and equal angles, the proposed model used regular polygon. Thus, the technique to calculate the value of the area of regular polygons is based on the formulas associated with each polygon.
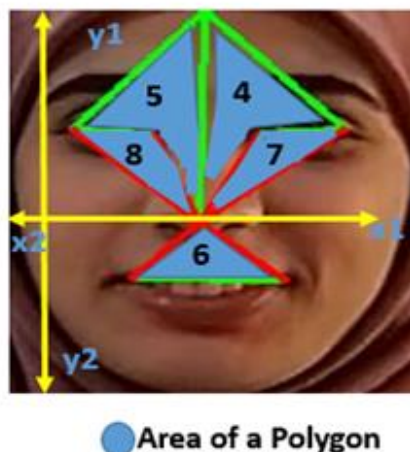

Fig. 2. Area of Polygon in a Two-Dimensional Plane

The pre-processing step consists of two sub-steps which are cleaning data using Exploratory Data Analysis (EDA) and balance data using Synthetic Minority Oversampling (SMOTE) technique. Cleaning Data involves analyzing and removing the outliers from the dataset. Exploratory data analysis is a method of sifting through large amounts of information to find outliers. The EDA handles outliers in the data using algorithm (1).

Algorithms (1) Removing Outliers

Input: Data Set
Output: Cleaned Dataset

Begin
Step 1: For each column in a dataset:
a. Retrieve the lower range and upper range values from the column.
b. Sort the items of the column in ascending order.
c. Calculate the quantiles for the column:
- Q1 = column.quantile(0.25)
- Q3 = column.quantile(0.75)
d. Determine the lower range and upper range values based on the interquartile range (IQR):
- IQR = Q3 - Q1
- lower range = Q1 - (1.5 * IQR)
- upper range = Q3 + (1.5 * IQR)
e. Update the data in the column with the new range values.
Step 2: For each value in the column:
a. If the value is less than the lower range:
- Set the new value as the lower range value.
b. Update the value in the column.
c. If the value is greater than the upper range:
- Set the new value as the upper range value.
d. Update the value in the column.
End

In this paper, SMOTE over-sampling technique is used on the dataset. The SMOTE technique is clarified in algorithm (2).

Algorithm (2) SMOTE Algorithm

Input: dataset**,** Smote Percentage call (N)**,** Number of nearest neighbors (k)
Output: Balanced training dataset

Begin
Step1: for each minorities data in the dataset, do
{ get data and name class }
Step2: for each row [X] in the data minorities Class, do
        { **a.** Find the k nearest neighbors of the row
         b. per= N/100
         c. while Per $\neq$ 0 do
             { Select one of k nearest neighbors, call Y
               Select a random number R [0,1]
                Y=X+R(Y-X)
                 append Y to New Sample
                 Per= Per -1
        } }
Step 3: add a new sample to data minorities Class call Synthetic data
Step 4: return Synthetic data
End

Face recognition is the last step in human face recognition methodology. This step depends on Enhanced Convolutional Neural network (ECCN). The enhancement performance of the CNN is done through the use of the CNN one dimension four times and the global average pooling one dimension as shown in figure (3).
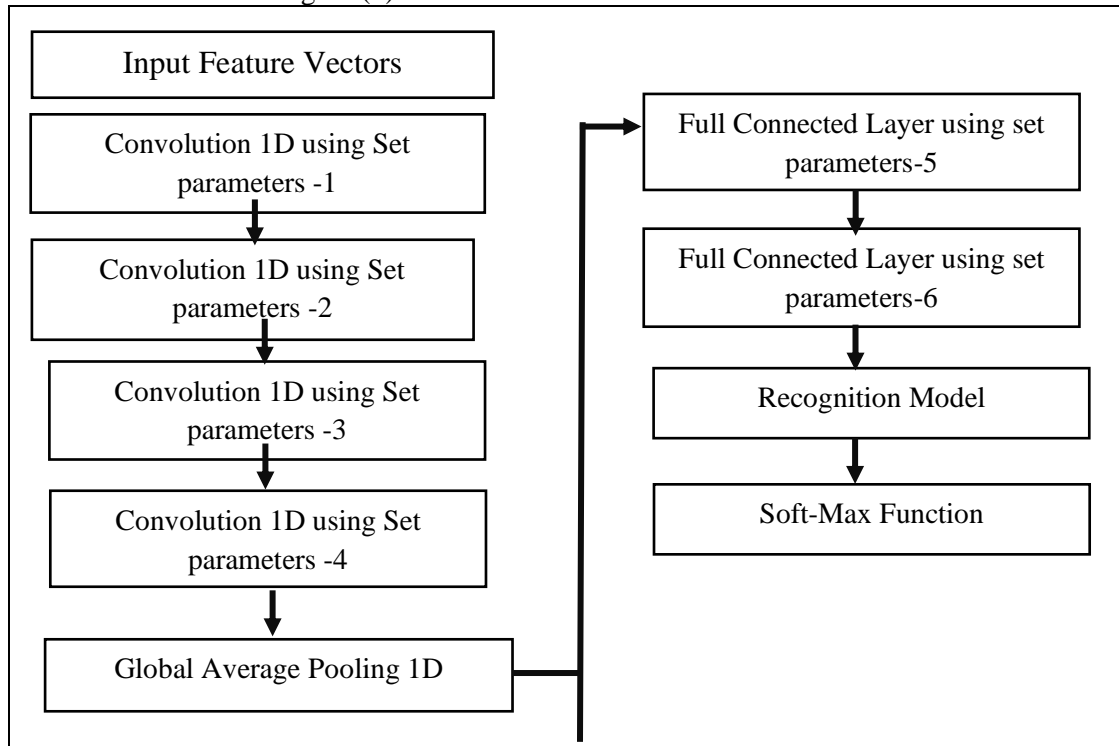

Fig. 3. The Block Diagram of the Proposed ECNN

The CNN algorithm implements a highly effective approach known as global average pooling, which replaces the traditional Flatten layers commonly used in conventional CNN architectures. In the last Convolution layer of the CNN, it generates a feature map specifically for each class involved in the classification task. Instead of employing fully connected layers on top of these feature maps, the CNN takes the average of each individual feature map. This technique aids in addressing the issue of overfitting. To further enhance the model's performance, a Dropout technique with a dropout rate of 0.5% is applied after the global average pooling layer. Subsequently, two dense layers are incorporated into the CNN, utilizing the 'ReLU' and 'SoftMax' activation functions respectively, to facilitate the identification of the target variable. For a comprehensive overview of the CNN model's architecture, please refer to Figure (4).
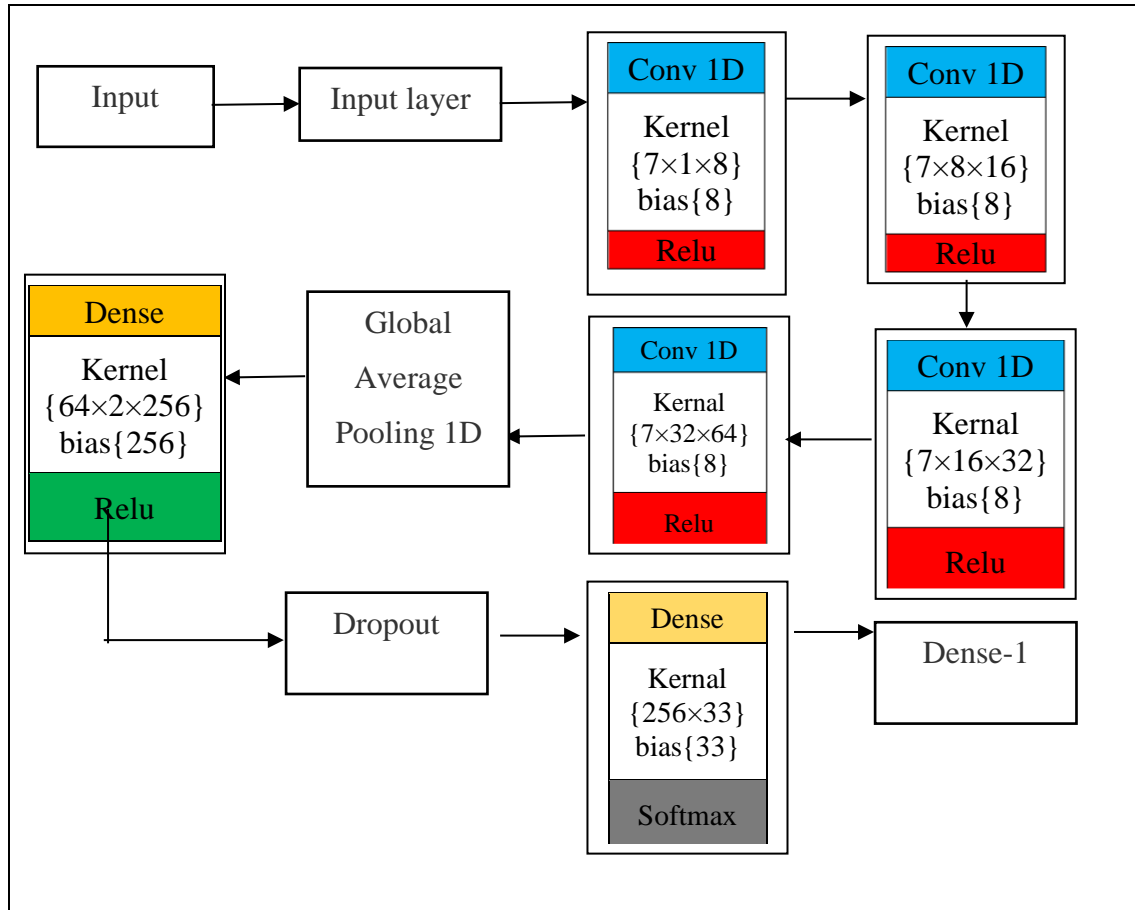
Fig 4. The Detailed Proposed ECNN

4.2. Gait Recognition

Human gait recognition includes three main steps: features extraction, pre-processing features, and gait recognition. In feature extraction step, the Media Pipe model is used to find the 3D key joint positions for each frame of the video. Dynamic gait feature, a method of feature extraction, is used in this paper. Figure (5) illustrates the crucial joint positions that are necessary for this study.

Number	Names of key joint positions
0	angle elbow
1	angle hip
2	length hip
3	angle mid hip
4	angle ankle
5	length ankle
6	distains ankle1
7	distains ankle2
8	distains mid pinky

Fig. 5. Names of the Key Joint Position

The extraction of the left and right leg angles is explained in algorithm (3).

Algorithm (3) Calculation of the"Knee "and "Elbow Angles"
Inputs:video
Outputs: Features of angle

Begin
Step 1: Provide a video as input.
Step 2: For each frame in the video:
a. Retrieve the width and height of the frame.
b. Calculate the angle of the knee.
c. Utilize a pose landmark model to obtain the left-side landmarks [lip, knee, ankle].
d. Utilize a pose landmark model to obtain the right-side landmarks [hip, knee, ankle].
e. Calculate the feature angle of the left knee by determining the angle between the left hip, knee, and ankle.
f. Calculate the feature angle of the right knee by determining the angle between the right hip, knee, and ankle.
Step 3: For each frame:
a. Calculate the angle of the elbow.
b. Employ a pose landmark model to extract the left-side landmarks [shoulder, elbow, wrist].
c. Employ a pose landmark model to extract the right-side landmarks [shoulder, elbow, wrist].
d. Calculate the feature angle of the left elbow by determining the angle between the left shoulder, elbow, and wrist.
e. Calculate the feature angle of the right elbow by determining the angle between the right shoulder, elbow, and wrist.
Step 4: Save the features including the left angle of the knee, the right angle of the knee, the left angle of the elbow, and the right angle of the elbow.
End.

Preprocessing step of gait features includes cleaning data as done in face features preprocessing, and balancing data using SMOTE algorithm as in algorithm (2).

Human gait recognition is based on the pattern of the gait obtained in the preceding steps, such that features which are utilized as input to the proposed enhancement CNN (ECCN) algorithm as shown in figure (4)

5. Results and Discussions
Results of the face and gait recognition are given separately.
5.1. Results of Face Recognition
Figure (6) and (7) clarifies the accuracy and loss for both training and testing for the face front features.
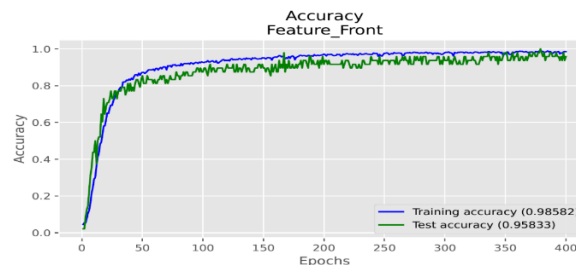

Fig. 6. Accuracy in Training and Testing of the Proposed ECNN with Front Features
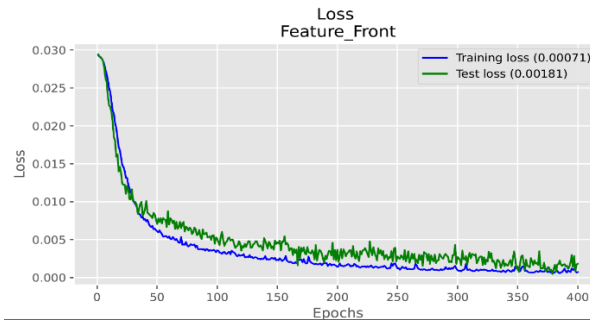
Fig. 7. Loss in Training and Testing of the Proposed ECNN with Front Features

Figure (6) shows the training accuracy about (0.98582), while the testing accuracy is around (0.95833) with 400 epochs. However, figure (7) shows the training loss value reaches approximately (0.0071), and the testing loss, which is around (0.00181) with (400) epochs.

5.2 Results of Gait Recognition

Figures (8 and 9) describes the accuracy and loss for both training and testing of the gait side features.
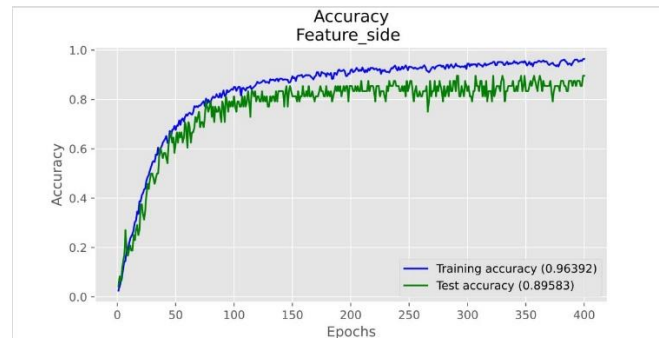

Fig. 8. Accuracy of the Proposed ECNN Algorithm with Gait Side Features

Fig. 9. Loss of the Proposed ECNN Algorithm with Gait Side Features

Figure (8) shows the training accuracy value is about (0.96392) and the testing accuracy is about (0.89583) with (400) Epoch. Figure (9) shows the training loss value is about (0.00179) and the Testing loss is about (0.00519) with 400 Epoch.

Figure (10) describes values of accuracy for the proposed ECNN algorithm based on face front features and gait side features.
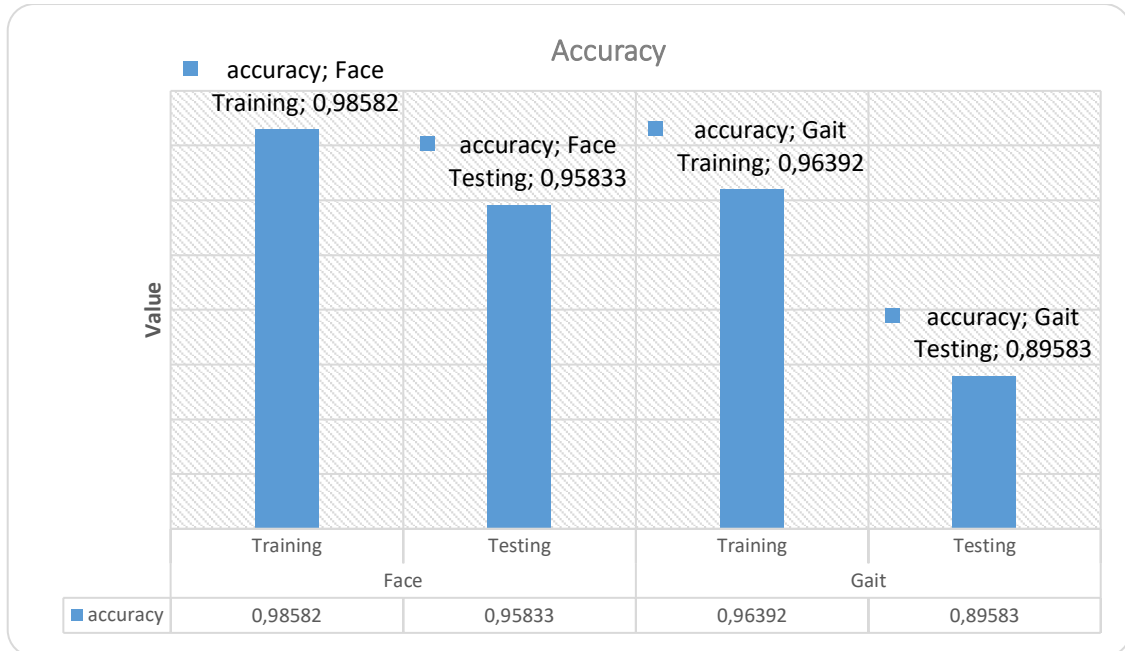
Fig. 10. Values of Accuracy of the Proposed ECNN Algorithm based on Face Front Features and Gait Side Features

6. Comparison with Previous Works

Several studies have been concerned with identification human based on soft biometric features using different methods and techniques adopted in previous years.

For performance evaluation, the proposed system achieved an accuracy higher than (Terhörst *et al*., 2020; Fard *et al*., 2021; Vukadinovic & Pantic, 2005) for facail features. For gait, the accuracy of previous works was higher than our accuracy, but our method is distinguished by using different features than previous methods.

This section shows a comparison between proposed classification system of identification humane and related methods, as illustrated in Table (2).

Table 2 - Comparison between the Proposed System and Related Methods

Ref.	Classification Approach	Classification algorithm	Dataset Name or type of biometric	Accuracy value
Terhörst *et al* (2020)	Machine learning	CSN-Transformation	ColorFeret dataset	90%
Sun *et al* (2013)	Deep learning	three-level convolutional networks	BioID and LFPW datasets	50%
Fard *et al* (2021)	Deep learning	Convolutional Neural Network (CNN)	WFLW dataset	50%
Vukadinovic & Pantic (2005)	Ensemble learning	GentleBoost feature templates	Cohn-Kanade database	93%
Arca *et al* (2006)	Machine learning	Support Vector Machine (SVM)	1180 image of the XM2VTS dataset	97.7%
			400 picture of the UniMiDb database	98.8%
Saleem *et al* (2023)	Machine learning	Support Vector Machine (SVM)	96 images	%82.432
Nithyakani *et al* (2019)	Deep learning	Deep Convolutional Neural Network (DCNN)	TUM- GAID dataset	94.7%
Sharif *et al* (2020)	Machine learning	Multi-class support vector machine (MSVM)	1.CASIA-A dataset	98.6%
			2.CASIA-B dataset	93.5%
			3.CASIA-C dataset	97.3%
Wang & Yan (2020)	Deep learning	convolutional Long Short-Term Memory (Conv-LSTM)	OU-ISIR dataset	99%
	Deep learning		CASIA-(B) dataset	99%

Elharrouss *et al* (2021)		Convolutional Neural Network (CNN)	OU-ISIR and OU-MVLP dataset	98%
Chao *et al* (2019)	Deep learning	proposed a GaitS*et algorithm*	CASIA-B gait dataset	95.0%
			OU-MVLP gait dataset	87.1%
Sahak *et al* (2017)	Deep learning	Multi-Layer perceptron (MLP)	New dataset	90.6%
Parashar *et al* (2022)	Machine learning	Convolutional Neural Network (CNN)	New dataset	95.23%
Zhang *et al* (2019)	Deep learning	Multi-layer LSTM structure	New dataset	81.8%
Our proposed system	Deep learning	ECNN	New Front dataset	0.95%
			New Side dataset	0.89%

Although our results are valid, we cannot conclude much about Soft Biometrics. Further evaluation on a much larger and most-varied database is still needed. In previous research, face templates comparison was used they used the entire face. In this research, features extracted from the face were used. Therefore, time is less and storage is less. In gait recognition, geometric features were used that were not used by previous researchers. This research also proves that recognition of a person through his face is more reliable, while recognition through gait is less reliable.

7. Conclusion

Recognition of people from a distance is one of the most requested topics these days due to the complexity of life and the evolution of nature and global crime, in addition to the emergence of diseases such as the Corona pandemic. This research proves that the human face and the gait can be used to identify people from a distance, but in different accuracies. This research shows that recognition of a person through his face is more reliable, while recognition through gait is less reliable. The research also opens the way to develop a better method in recognition the human gait.

References

Almudhahka, N., Nixon, M., & Hare, J. (2016). Human face identification via comparative soft biometrics. In *2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)* (pp. 1-6). IEEE. https://doi.org/10.1109/ISBA.2016.7477246.

Arca, S., Campadelli, P., & Lanzarotti, R. (2006). A face recognition system based on automatically determined facial fiducial points. *Pattern recognition*, *39*(3), 432-443. https://doi.org/10.1016/j.patcog.2005.06.015 .

Arigbabu, O. A., Ahmad, S. M. S., Adnan, W. A. W., & Yussof, S. (2015). Integration of multiple soft biometrics for human identification. *Pattern Recognition Letters*, *68*, 278-287. https://doi.org/10.1016/j.patrec.2015.07.014.

Becerra-Riera, F., Morales-González, A., & Méndez-Vázquez, H. (2019). A survey on facial soft biometrics for video surveillance and forensic applications. *Artificial Intelligence Review*, *52*(2), 1155-1187. https://doi.org/10.1007/s10462-019-09689-5 .

Chao, H., He, Y., Zhang, J., & Feng, J. (2019, July). Gaitset: Regarding gait as a set for cross-view gait recognition. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 33, No. 01, pp. 8126-8133). https://doi.org/10.1609/aaai.v33i01.33018126.

Dantcheva, A., Velardo, C., D'angelo, A., & Dugelay, J. L. (2011). Bag of soft biometrics for person identification: New trends and challenges. *Multimedia Tools and Applications*, *51*, 739-777. https://doi.org/10.1007/s11042-010-0635-7.

Dantcheva, A., Velardo, C., D'angelo, A., & Dugelay, J. L. (2011). Bag of soft biometrics for person identification: New trends and challenges. *Multimedia Tools and Applications*, *51*, 739-777. https://doi.org/10.1007/s11042-010-0635-7 .

Elharrouss, O., Almaadeed, N., Al-Maadeed, S., & Bouridane, A. (2021). Gait recognition for person re-identification. *The Journal of Supercomputing*, *77*, 3653-3672. https://doi.org/10.1007/s11227-020-03409-5 .

Fard, A. P., Abdollahi, H., & Mahoor, M. (2021). ASMNet: A lightweight deep neural network for face alignment and pose estimation. In *Proceedings of the IEEE/CVF Conference on computer vision and pattern recognition* (pp. 1521-1530). https://doi.org/10.1109/CVPRW53098.2021.00168 .

Garg, R., Arora, A., Singh, S., & Saraswat, S. (2018). Biometric authentication using soft biometric traits. In *2018 Fifth International Conference on Parallel, Distributed and Grid Computing (PDGC)* (pp. 259-264). IEEE. https://doi.org/10.1109/PDGC.2018.8745766 .

Gharghabi, S., Azari, B., Shamshirdar, F., & Safabakhsh, R. (2016). Improving person recognition by weight adaptation of soft biometrics. In *2016 6th International Conference on Computer and Knowledge Engineering (ICCKE)* (pp. 36-40). IEEE. https://doi.org/10.1109/ICCKE.2016.7802112.

Guo, B. H., Nixon, M. S., & Carter, J. N. (2019). Soft biometric fusion for subject recognition at a distance. *IEEE Transactions on Biometrics, Behavior, and Identity Science*, *1*(4), 292-301. https://doi.org/10.1109/TBIOM.2019.2943934 .

Hassan, B., & Izquierdo, E. (2022, October). Rsfs: A soft biometrics-based relative support features set for person verification. In *Fourteenth International Conference on Digital Image Processing (ICDIP 2022)* (Vol. 12342, p. 1234202). SPIE. https://doi.org/10.1117/12.2644457.

Hassan, B., Sherazi, H. H. R., Ali, M., & Bashir, A. K. (2023). A multi-channel soft biometrics framework for seamless border crossings. *EURASIP Journal on Advances in Signal Processing*, *2023*(1), 65. https://doi.org/10.1186/s13634-023-01026-x

Jain, A. K., Dass, S. C., & Nandakumar, K. (2004). Can soft biometric traits assist user recognition?. In *Biometric technology for human identification* (Vol. 5404, pp. 561-572). SPIE. https://doi.org/10.1117/12.542890.

Kang, W., Lu, Y., Li, D., & Jia, W. (2018). From noise to feature: Exploiting intensity distribution as a novel soft biometric trait for finger vein recognition. *IEEE transactions on information forensics and security*, *14*(4), 858-869. https://doi.org/10.1109/TIFS.2018.2866330.

Kim, M. G., Moon, H. M., Chung, Y., & Pan, S. B. (2012). A survey and proposed framework on the soft biometrics technique for human identification in intelligent video surveillance system. *Journal of Biomedicine and Biotechnology*, *2012*. https://doi.org/10.1155/2012/614146.

Minaee, S., Abdolrashidi, A., Su, H., Bennamoun, M., & Zhang, D. (2023). Biometrics recognition using deep learning: A survey. *Artificial Intelligence Review*, 1-49. https://doi.org/10.1007/s10462-022-10237-x .

Nithyakani, P., Shanthini, A., & Ponsam, G. (2019). Human gait recognition using deep convolutional neural network. In *2019 3rd international conference on computing and communications technologies (ICCCT)* (pp. 208-211). IEEE. https://doi.org/10.1109/ICCCT2.2019.8824836 .

Parashar, A., Parashar, A., & Shekhawat, R. S. (2022). A robust covariate-invariant gait recognition based on pose features. *IET Biometrics*, *11*(6), 601-613. https://doi.org/10.1049/bme2.12103.

Prakash, A., Nambiar, A., & Bernardino, A. (2023). Multimodal Adaptive Fusion of Face and Gait Features using Keyless attention based Deep Neural Networks for Human Identification. *arXiv preprint arXiv:2303.13814*.

Reid, D. A., & Nixon, M. S. (2011). Using comparative human descriptions for soft biometrics. In *2011 International Joint Conference on Biometrics (IJCB)* (pp. 1-6). IEEE. https://doi.org/10.1109/IJCB.2011.6117513

Reid, D. A., & Nixon, M. S. (2011, October). Using comparative human descriptions for soft biometrics. In *2011 International Joint Conference on Biometrics (IJCB)* (pp. 1-6). IEEE. https://doi.org/10.1109/IJCB.2011.6117513 .

Reid, D. A., & Nixon, M. S. (2013). Human identification using facial comparative descriptions. In *2013 International Conference on Biometrics (ICB)* (pp. 1-7). IEEE. https://doi.org/10.1109/ICB.2013.6612962.

Reid, D. A., Nixon, M. S., & Stevenage, S. V. (2013). Soft biometrics; human identification using comparative descriptions. *IEEE Transactions on pattern analysis and machine intelligence*, *36*(6), 1216-1228. https://doi.org/10.1109/TPAMI.2013.219 .

Reid, D. A., Samangooei, S., Chen, C., Nixon, M. S., & Ross, A. (2013). Soft biometrics for surveillance: an overview. *Handbook of statistics*, *31*, 327-352. https://doi.org/10.1016/B978-0-444-53859-8.00013-8.

Romero, M., Gutoski, M., Hattori, L. T., Ribeiro, M., & Lopes, H. S. (2020). Soft biometrics classification in videos using transfer learning and bidirectional long short-term memory networks. *Learning and Nonlinear Models*, *18*(1), 47-59. https://doi.org/10.21528/lnlm-vol18-no1-art4.

Rudinskaya, E., & Paringer, R. (2020). Face detection accuracy study based on race and gender factor using haar cascades. In *CEUR Workshop Proceedings* (Vol. 2667, pp. 238-242).

Sahak, R., Tahir, N. M., Yassin, A. I., & Kamaruzaman, F. (2017). Human gait recognition based on frontal view using kinect features and orthogonal least square selection. In *2017 IEEE International Symposium on Signal Processing and Information Technology (ISSPIT)* (pp. 388-393). IEEE. https://doi.org/10.1109/ISSPIT.2017.8388674 .

Sakib, S., Ahmed, N., Kabir, A. J., & Ahmed, H. (2019). An overview of convolutional neural network: Its architecture and applications. *Preprints*, 2018110546. https://www.preprints.org/manuscript/201811.0546/v4 .

Saleem, S., Shiney, J., Priestly Shan, B., & Kumar Mishra, V. (2023). Face recognition using facial features. *Materials Today: Proceedings*, *80*, 3857–3862. https://doi.org/10.1016/j.matpr.2021.07.402 .

Sharif, M., Attique, M., Tahir, M. Z., Yasmim, M., Saba, T., & Tanik, U. J. (2020). A machine learning method with threshold based parallel feature fusion and feature selection for automated gait recognition. *Journal of Organizational and End User Computing (JOEUC)*, *32*(2), 67-92. https://doi.org/10.4018/JOEUC.2020040104 .

Sikkandar, H., & Thiyagarajan, R. (2020). Soft biometrics-based face image retrieval using improved grey wolf optimisation. *IET Image Processing*, *14*(3), 451-461.

Subramanian, B., Olimov, B., Naik, S. M., Kim, S., Park, K. H., & Kim, J. (2022). An integrated mediapipe-optimized GRU model for Indian sign language recognition. *Scientific Reports*, *12*(1), 11964. https://doi.org/10.1038/s41598-022-15998-7 .

Sun, Y., Wang, X., & Tang, X. (2013). Deep convolutional network cascade for facial point detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 3476-3483). https://doi.org/10.1109/CVPR.2013.446 .

Terhörst, P., Huber, M., Damer, N., Kirchbuchner, F., & Kuijper, A. (2020). Unsupervised enhancement of soft-biometric privacy with negative face recognition. *arXiv preprint arXiv:2002.09181*.

Tome, P., Fierrez, J., Vera-Rodriguez, R., & Nixon, M. S. (2014). Soft biometrics and their application in person recognition at a distance. *IEEE Transactions on information forensics and security*, *9*(3), 464-475. https://doi.org/10.1109/TIFS.2014.2299975 .

Vukadinovic, D., & Pantic, M. (2005). Fully automatic facial feature point detection using gabor feature based boosted classifiers. In *2005 IEEE International Conference on Systems, Man and Cybernetics* (Vol. 2, pp. 1692-1698). IEEE. https://doi.org/10.1109/icsmc.2005.1571392 .

Wang, X., & Yan, W. Q. (2020). Human gait recognition based on frame-by-frame gait energy images and convolutional long short-term memory. *International journal of neural systems*, *30*(01), 1950027. https://doi.org/10.1142/S0129065719500278 .

Zhang, Z., Tran, L., Yin, X., Atoum, Y., Liu, X., Wan, J., & Wang, N. (2019). Gait recognition via disentangled representation learning. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 4710-4719). https://doi.org/10.1109/CVPR.2019.00484 .