www.ijecs.in
International Journal of Engineering and Computer Science
Volume 14 Issue 04 April 2025, Page No. 27096- 270105
ISSN: 2319-7242 DOI: 10.18535/ijecs/v14i04.5082

Gender Recognition from Standard Facial Images

Esraa Farouk Abou-Elmagd
Faculty of Commerce South Valley University Qena, Egypt

Abstract
Human gender plays an imperative role as a social construct and an essential form of an individual's personality. This is highly reflected in social communication, forensic science, surveillance, and targeted marketing. The task has always been practiced in conjunction with face recognition, as it yields high performance compared to other features, e.g., gait and sound. This is attributed to its richness with diverse useful clues for gender inspection. However, facial features were always used separately without any integration between them, and a definitive solution has not yet been found yet. We propose a robust feature fusion method that integrates various facial features, in addition to a set of new features. Furthermore, Haar-like features are used to extract the essential features, while the new features used are, i.e., beard, moustache, cheeks, forehead, and face. The scientific novelty of the research is the robust fusion of features, to achieve high recognition accuracy. The proposed fusion method is a priority-based that arranges classifiers' responses from the strongest to weakest according to their importance to yield a final decision. When the proposed technique compares the recent and standard benchmarks, it achieves high recognition accuracy. The experiments on various standard and challenging datasets widely adopted in the scientific community, namely LFW, Data Hub, NIST, and Caltech Web Faces demonstrated a robust performance of 99.1% accuracy.

Keywords: Gender Recognition, Standard Images, Facial Features, Face Detection.

1. Introduction

Humans use various static and dynamic characteristics to successfully identify and interact with other people. Gender is one of these features and is defined as a personal conception of oneself as male or female. However, if the task of gender classification can be fully automated, it will be very helpful in many applications [1], [2]. For example, improving the intelligence of the surveillance system to find criminal suspects and serving customers based on their gender. Looking at a person's face, no longer solely determine who he is, but also different essential information, for example, human gender [3], [1]. Recognizing gender based on facial features is something that most people do automatically and unconsciously [4], [2]. However, this task has remained a challenging problem in the field of computer vision, and many years of research is just starting to yield beneficial technological solutions. Automated gender recognition from facial features has various ideal properties driven into sensible techniques. To the best of our knowledge, this challenge has been based entirely on internal face features, e.g., hairstyle, chin, and jawline [4].

The gender classification was first introduced as a psychophysical issue that focused on the efforts of understanding human visual processing and identifying key discriminative features. Psychophysical researchers have concluded that the gap between facial masculinity and femininity can also be used to improve the efficiency of automated facial-based gender recognition systems. Gender recognition is strongly dependent on face recognition. A typical system for recognizing gender that starts with the face detection phase followed by feature extraction, as depicted in Figure 1. Thus, the main task is to use Computer Vision (CV) techniques that extract a robust facial feature to identify gender from facial images.

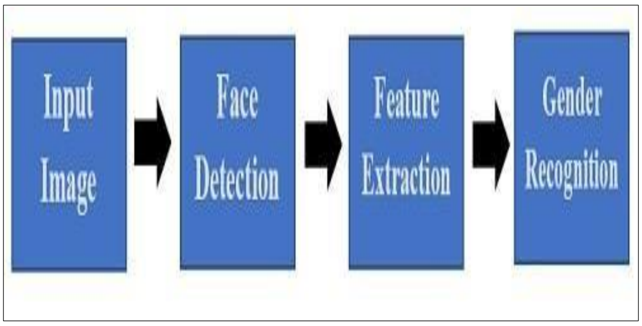

Figure 1. Flowchart of the Proposed Gender Recognition Framework

Generally, it is difficult to find a classifier that works best with the chosen feature extractor to achieve the optimal classification performance. Also, merging the features is a core process that determines how the features will work seamlessly together. Finally, the domain affects the recognition performance as well, where any change requires a complete re-designing of the system. This paper contributes and solves many problems by introducing a robust priority-based fusion method that merges classifiers' responses according to their weight to formulate a final classification response. Furthermore, a group of external features are used to provide more discriminative power, i.e., beard, moustache, cheeks, forehead, and face.

The remainder of this paper is structured as follows: Section 2 presents the related literature work. Section 3 introduces the proposed method, while Section 4 presents the experiments and analysis of the proposed method. Finally, Section 5 concludes the paper.

2. Literature Review

Gender is the most widely studied facial demographic attributes in the CV field [6], [7]. Existing methods generally address the problem in two phases: (1) face detection and (2) feature extraction.

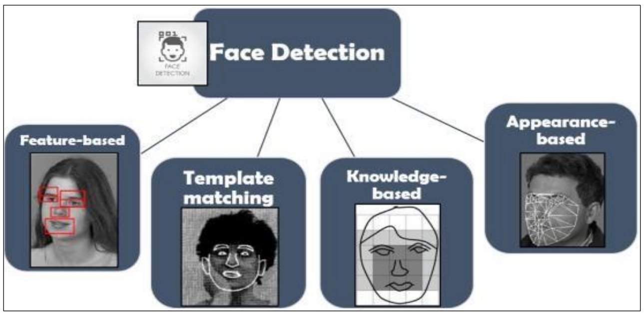

Figure 2. Classification of face detection techniques. Inner images credits are for [16,13]

Regarding the face detection phase, throughout the literature, numerous hand-crafted features were proposed and fused in different ways. Generally, those algorithms could be classified into four categories, as depicted in Figure 2, and detailed as follows:

- Knowledge-based. Encode what comprises a normal face, e.g., the link between facial features [10].
- Feature invariant based. Aim to find features of a face's structure that exists even when the position, the viewpoint, or the illumination change [6].
- Template matching. Several standard patterns are saved to characterize the face as a whole or individual facial aspect [11].
- Appearance-based. The models are trained using a collection of training photos that reflect the variety of faces [7], [12].

Throughout the literature, regarding the feature extraction and selection stage, a wide range of features were evaluated. These include values of intensities from grayscale images [13], LBP [14], facial-strips [15], Haar-like features [16], HOG [17], and SIFT features [18]. These features/descriptors could be extracted at a global level from the entire face image or on a local level from a defined sub-region inside the face image. The internal features are composed of the eyes, the nose, and the mouth and the external features are located in the head, ears, and chin. Some researchers are conducted in the face and neck regions, hair, and upper body clothing and Head shoulder-based gender recognition. Later, a series of psychological experiments showed that subface individual features, e.g., eyes, nose, mouth, and chin, carry much gender information seen in isolation [19].

The Haar-like features [7] were among the most famous features that were widely used in recognizing the gender because they capture important characteristics of the facial area, especially the contrast features. Haar-like features were used in conjunction with the Adaboost classifier in [20] and were used with a probabilistic boosting tree.

Neural Networks (NN) were employed as well in various research attempts to recognize the gender, based on mostly frontal face images. As an example, SEXNET, a fully connected two- layer neural network, was trained to identify gender from 30×30 face images. Another NN configuration is presented in [21] that employs a serious of Gabor filters to extract the required responses and train the proposed NN. It is worth mentioning that the majority of the previous research, standard images were mostly used for the gender recognition problem.

Conclusively, some serious problems appeared in the previous related literature. Some studies only depended on one feature, which could result in low recognition accuracy, and others used misleading classification features [22]. Also, the order of features affects the recognition accuracy [22]. These problems will be mostly tackled in the proposed approach, as will be detailed in the next section.

3. Proposed Method

The structure of the proposed technique is depicted as shown in Figure 3. The technique is mainly composed of three phases, in addition to the initial face detection step. The first phase detects Haar-like features from an input facial image. The second phase is where the new features are extracted. Finally, a set of Ada boost classifiers are applied individually to each extracted feature, while their final output is fused using the priority-based approach. Each component of the framework is explained shortly, in subsections 3.1, 3.2 and 3.3 respectively.

Any gender-based recognition system's performance is mainly determined by the selected features set. Thus, an extra features set is proposed and used, i.e., moustache, beard, forehead, cheek, and face shape, to increase the system performance. The presence strength of each feature is then evaluated using a set of Ada boost classifiers. Finally, the proposed priority-based method applied to classifier output to recognize the gender of the input image gender.

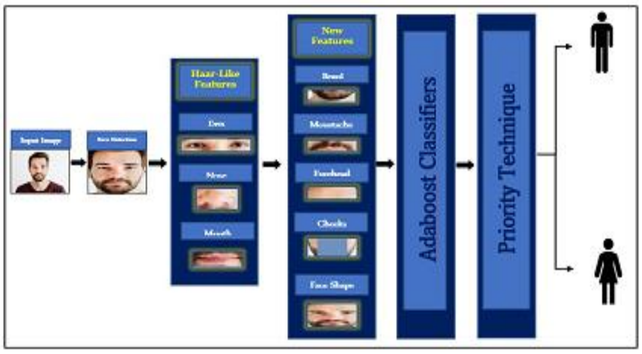

Figure 3. Illustration of the Proposed Technique.

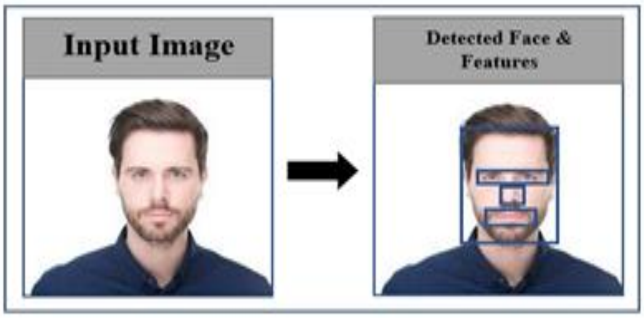

Figure 4. Face detection and Initial Feature Extraction Phases. The depicted image is a real example from the LFW dataset [14]

A. Face Detection and Feature Extraction (Phase 1)

The initial face detection step is performed using the Ada boost face detection algorithm, then respective main facial components (face, eyes, nose, and mouth) are located using the Active Shape Model (ASM) [23] and cropped to fixed window sizes for later processing. The used ASM is considered a template-based method, as it depends on a prior model of expected features' locations to find the best match position between the model and the input image. This is a critical step as it affects the efficiency of the second feature extraction phase. Figure 4 shows an illustrative example of the features that will be generated by the proposed method.

Practically, the Viola Jones algorithm is used to detect the face along with its internal features. The algorithm mainly uses five Haar-like features, that is, a scalar product between the image and some templates (exactly five), as depicted in Figure 5. More precisely, let I and P denote an input image and a pattern, both same size N×N. The feature associated with pattern P of image I is defined by the following equation 1:

$$\sum_{i=1}^{N}\sum_{j=1}^{N} I(i,j)1_{P(i,j) is white} - \sum_{i=1}^{N}\sum_{j=1}^{N} I(i,j)1_{P(i,j) is black} \quad (1)$$

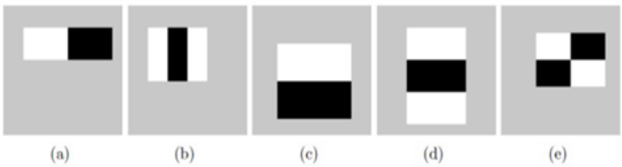

Figure 5. The Five Haar-like patterns

B. Feature Extraction (Phase 2)

The next step is to extract the new proposed features, i.e., beard, moustache, forehead, cheek, and face shape. Each of these features is critical in gender identification [25]. Gender recognition relies heavily on the beard area because anyone with a beard is highly likely to be a male [1]. Another discriminative male feature is the moustache area [26]. Furthermore, the forehead area is also an important feature in recognizing the gender because anyone with hair over the forehead is mostly a female.

The steps for extracting the previous three features are the same as illustrated in algorithm 1 (Figure 6). These features were extracted as follows:

1. The grayscale or binary picture is subjected to morphological bottom-hat filtering. Bottom- hat filtering calculates the picture's morphological closure and then subtracts the original image from the result.
2. A dilation process is then used to gradually increase the borders of foreground pixel areas, which are generally white pixels. As a result, foreground pixel areas expand in size, while gaps inside those regions shrink. Finally, the black area in the input image is computed, which indicates if the feature is found or not.

```
Algorithm 1: Facial Feature extraction algorithm.
Input: Detected Feature as image (B);
Output: Feature Found or Not;
Filter the image: B1←imbothat (B);
Dilate the image:B2←imdilate (B1);
Compute the black area in the image: Black-Area ←(B2 == 1);
Compute: Output←Sum (Black-Area (:));
Compute the size of image B1:M←size (B1, 1) × size (B1, 2);
if output > M / 2 then
    result ← Feature Found;
    else
        result ← Feature Not Found;
    end
end
return result;
```

Figure 6. Facial feature extraction algorithm

For the cheek area, it is examined as well to detect any hair presence on either the left or right cheek. If both cheeks have hair, the input image is almost certainly for males. On the other hand, if only one cheek has hair, the image seems more for a female. However, if there is no hair at all over the two cheeks, this means that the result will be inconclusive, i.e., the image may be a male or a female. The equation for this procedure is as follows:

$$V = F(R) + F(L),\ V \in \{Male, Female\} \quad (2)$$

where R and L are right and left cheeks, respectively, and F(x) is a binary function illustrated in Algorithm 2 (Figure 7) and defined as followed in Equation 3:

$$F(x) = \begin{cases} 1 & \text{Hair exisits on cheek.} \\ 0 & \text{Hair does not exisit on cheek.} \end{cases} \quad (3)$$

```
Algorithm 2: Cheek Extraction Algorithm
Input: Right_Cheek, Left_Cheek;
Output: Type of gender;
if (Right_Cheek == True) & (Left_Cheek == True) then
    Gender ← Male;
end
if (Right_Cheek == True) & (Left_Cheek == False) then
    Gender ← Female;
end
if (Right_Cheek == False) & (Left_Cheek == True) then
    Gender ← Female;
end
if (Right_Cheek == False) & (Left_Cheek == False) then
    Gender ← Male or Female;
end
return Gender;
```

Figure 7. Cheek extraction algorithm

Figure 8 illustrates the three different scenarios of hairs on the cheeks and how they affect the process of gender recognition.

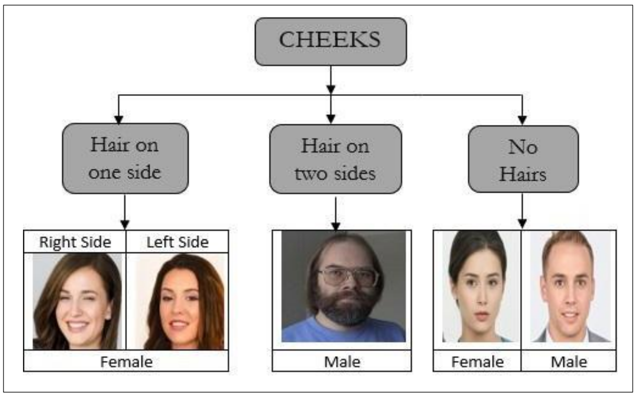

Figure 8. Illustration of the three different scenarios of hairs on the cheeks and how they affect the process of gender recognition. The depicted images are real examples from the LFW dataset

Finally, regarding the face shape, a rectangular face shape is highly likely to yield a male, while an oval shape is more to yield a female [22].

C. Gender Recognition

The final step in the proposed gender recognition is the classification stage. This is done using an Adaboost algorithm. The Adaboost algorithm is a collection of strong/weak classifiers, with a different classifier for each feature. At the end classifiers output are fused and combined to recognize the gender of the input image. Generally, there are two types of classifiers based on the features used in the classifier: strong and weak classifiers. An extracted feature yields a strong classifier if its performance accuracy is high. On the other hand, a weak classifier results from a low performance feature. The types of classifiers related to the aforementioned features are listed below:

1. Mustache is a strong classifier as it results in male.
2. Beard is a strong classifier as it results in male.
3. Hair on the forehead is a strong classifier as it results in female.
4. Hair on the cheek is a weak classifier as it results in female or male.
5. Face shape is a weak classifier as it results in female or male.

Mustache and beard are the most important facial features. If the input image has any of these features, the gender must be reported as male. Therefore, moustache and beard features correspond to strong classifiers, and there is no need to further check other features to verify the result. The second step is to check the third facial feature, i.e., the forehead, which corresponds to a strong female classifier. Finally, the two other features (cheeks and face shape) correspond to weak classifiers as they depict three possible classification states, i.e., male, or female or not.

To determine the best feature order, 120 experiments were run across two benchmark datasets. The analysis begins with a single feature and subsequently expands. Increasing the number of characteristics and varying their permutations. The accuracy metric was used at this investigation stage to report on the overall performance. (The datasets and metrics are introduced in the next section). Furthermore, the reference benchmark in this stage was the normal method that classifies the features without applying any priority for individual features. Figures 9, Figures 10, Figure 11, and Figure 12 depict the sequence of experiments. The depicted charts conclude that five features with the order (M, B, C, F and FS) achieve the highest performance, as finally depicted in Figure 13.

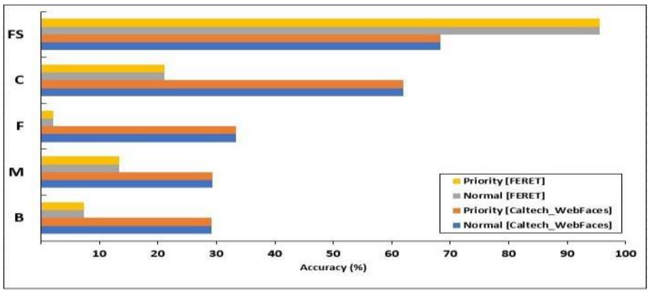

Figure 9. Performance comparison between the normal method versusthe priority-based method using one feature over FERET and Caltech Web Faces datasets

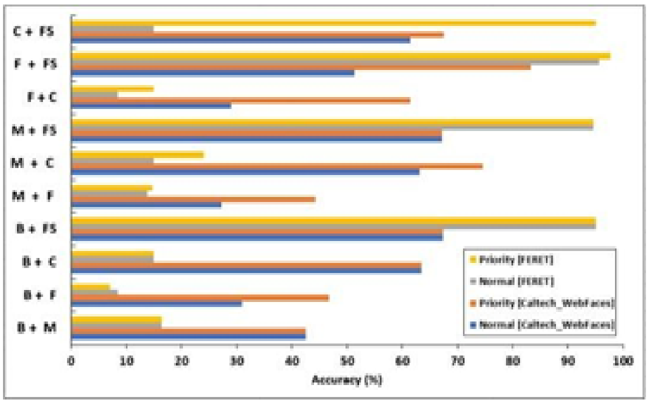

Figure 10. Performance comparison between the normal method versus priority-based method using two features different combinations over FERET and Caltech Web Faces datasets

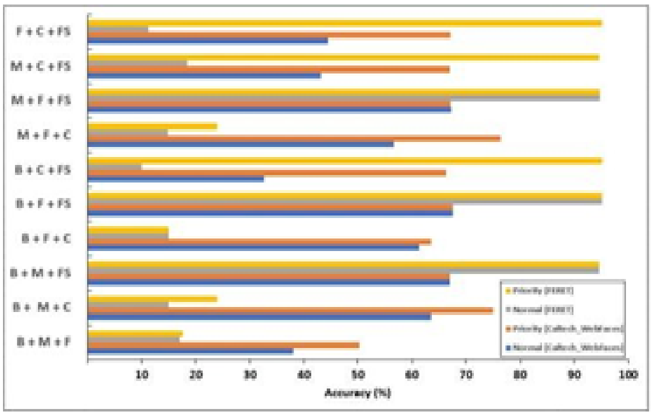

Figure 11. Performance comparison between the normal method versus the priority-based method using three features different combinations over FERET and Caltech Web Faces datasets

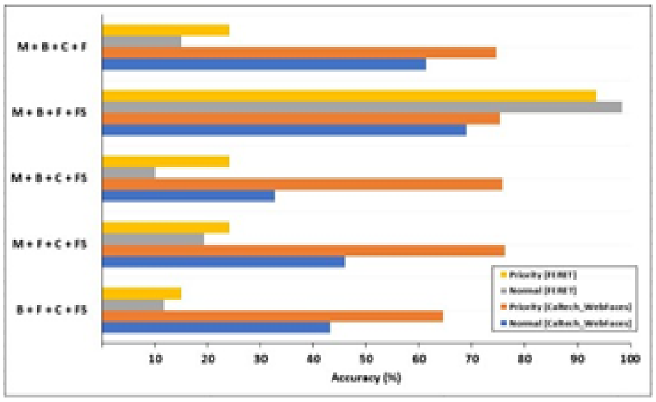

Figure 12. Performance comparison between the normal method versus the priority-based method using four features different combinations over FERET and Caltech Web Faces datasets

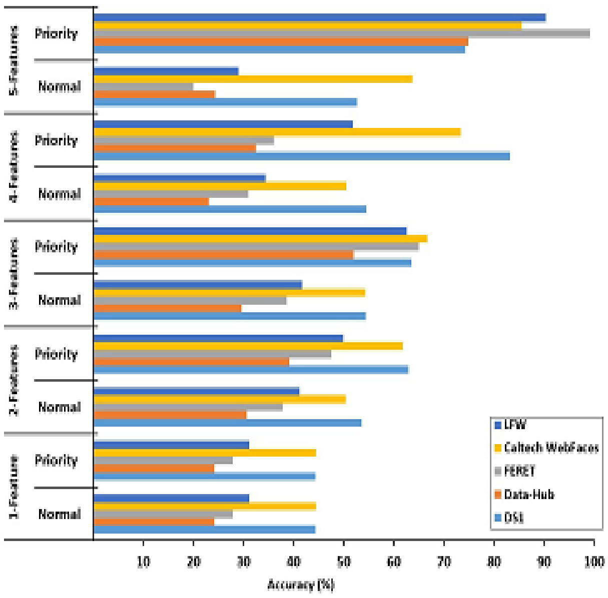

Figure 13. Performance comparison between the normal method versus the priority-based method using the five features (M+B+C+F+FS) over the FERET and Caltech Web Faces datasets respectively

4. Experiments and Results

This section introduces and discusses the proposed gender recognition system's performance. The datasets used are presented in section 4.1, while the results are introduced and discussed in section 4.2.

A. Datasets

A collection of challenging and various benchmark datasets was used in this paper. The first is the LFW [5] dataset, a public face analysis benchmark that combines 13233 images for 5749 humans. The second dataset is the FERET dataset [27], which is a standard facial recognition dataset with 14126 images, 1199 individuals, and 365 duplicate image sets. The third is the Caltech Web Faces [28] that has a total of 10524 faces in 7092 images gathered from the internet. The fourth is the Data-hub dataset [29], which is composed of 214 images and a common facial analysis dataset. The fifth and final dataset is a selective group of 500 images, that were chosen at random from the internet, called DS1. All datasets are very challenging and depict various resolutions and in different settings, e.g., portrait images. In Table 1, the datasets are summarized and presented.

Table 1. Summary of the test datasets

Name	Size	Males	Females	Type
DS1	500	237	263	Unconstrained
Data-hub	214	184	30	Unconstrained
FERET	1216	1123	93	Constrained
Caltech Web Faces	10524	6288	4236	Constrained
LFW	13233	10183	3050	Unconstrained

5. Results and Discussion

According to the standard experimental design, all datasets were randomly split by assigning 70% of the images to training and 30% to validation [1]. Furthermore, an accuracy measure is used, as it reviews the percentage of successfully recognized images with respect to the total dataset. The following equation depicts the accuracy metric's formula

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}}$$

where TP is the number of recovered relevant matches, FP is the number of retrieved irrelevant matches, FN is the number of missing relevant matches, and TN is the number of not retrieved relevant matches. Experimentally, to validate the priority-based method performance, it is being compared against the normal method (as a first step). The normal method for recognizing gender depends on counting the classifiers' output male/female and voting for the majority gender. However, if the classifiers vote equally for male and female features, the result is inconclusive. The experiment was repeated five times using: one feature, two features, three features, four features, and five features. This is to emphasis the significance and contribution of the various features to overall recognition accuracy. The results in Figure 14 indicate that the proposed priority-based approach is effective with 46±25.1% average higher accuracy using the five features over the five datasets.

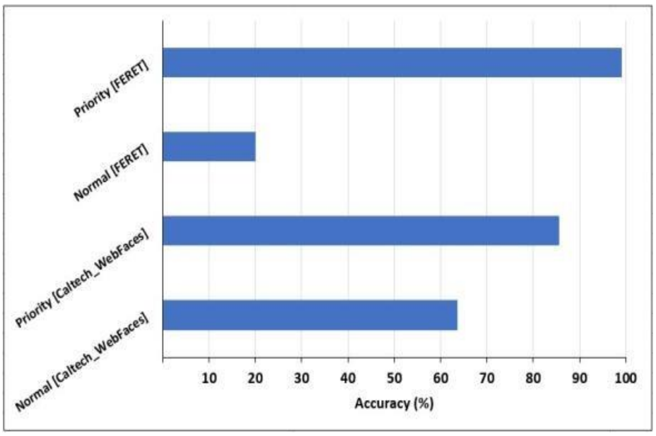

Figure 14. Performance comparison between the normal method versus the priority-based method using a different number and combination of features over five datasets. The proposed priority-based method was more effective

Furthermore, the performance of the suggested priority-based approach is compared with two baselines to ensure its robustness. The selected baselines represent the current research themes in gender recognition deep learning and hand facial features based. The baselines are Shunting Inhibitory CNN [30] and Pattern-based texture [31]. Finally, regarding the qualitative performance of the proposed gender recognition model, Figure 15 shows a collection of challenging images from the LFW, Data-Hub, Caltech WebFaces, and FERET datasets. Its results demonstrate the priority-based technique's classification efficiency. Table 2 summarizes and presents the results.

Table 2. Performance comparison of the proposed priority-based gender recognition technique versus benchmark baselines

Baseline	DS1	LFW	Data-Hub	FERET	Caltech WebFaces
Proposed Technique	74.2%	90.4%	74.8%	99.1%	85.6%
Shunting Inhibitory CNN	-	85.7%	-	-	-
Pattern-based Texture	-	90%	-	-	-
PCA	-	-	-	85%	-
LDA	-	-	-	90%	-
A Convolutional Neural Network	-	98.73%	-	-	
DSIFT	-	-	-	-	85%

Figure 15. Sample images from the test datasets that were classified using the proposed priority-based technique

6. Conclusion

A robust strategy for gender recognition was introduced in this study, which is based on a priority order fusion for the output of a group of weak/strong classifiers. The technique uses a group of facial features, i.e., beard, moustache, forehead, cheeks, and face shape. This study concludes the importance of facial features in achieving better accurate results as it depends on strong features that classify an image as female or male [1]. The proposed technique has achieved high accuracy over several challenging datasets, i.e., LFW, data-Hub, Caltech WebFaces, and FERET datasets. Furthermore, the suggested technique's generality has been proven by running an experiment on an Internet-harvested dataset from Google, which yielded state-of-the-art results. Some studies work in gender recognition based on facial features on the same datasets. However, these studies achieved lower accuracy than the proposed work. Regarding future work, more features from the upper/lower body will be used in the future to improve accuracy. The work will also be extended to better handle unconstrained images

References

1. Bekhet, S., Alghamdi, A. M., & Taj-Eddin, I. (2022). Gender recognition from unconstrained selfie images: a convolutional neural network approach. International Journal of Electrical and Computer Engineering (IJECE), 12(2), 2066-2078.
2. Bekhet, S., & Alahmer, H. (2021). A robust deep learning approach for glasses detection in non‐standard facial images. IET Biometrics, 10(1), 74-86.
3. Choudhury, S., Chattopadhyay, S. P., & Hazra, T. K. (2017, August). Vehicle detection and counting using haar feature-based classifier. In 2017 8th annual industrial automation and electromechanical engineering conference (IEMECON) (pp. 106-109). IEEE.
4. Egorov, A. D. (2018). Algorithm for optimization of Viola–Jones object detection framework parameters. In Journal of Physics: Conference Series (Vol. 945, No. 1, p. 012032). IOP Publishing.
5. Bhardwaj, A., Karger, D., Subramanyam, H., Deshpande, A., Madden, S., Wu, E., ... & Zhang, R. (2015, August). Collaborative data analytics with DataHub. In Proceedings of the VLDB Endowment International Conference on Very Large Data Bases (Vol. 8, No. 12, p. 1916).
6. Upchurch, P., Gardner, J., Pleiss, G., Pless, R., Snavely, N., Bala, K., & Weinberger, K. (2017). Deep feature interpolation for image content changes. In Proceedings of the IEEE conference on computer vision and pattern recognition (pp. 7064-7073).
7. Hermosilla, G., Ruiz-del-Solar, J., & Verschae, R. (2017). An enhanced representation of thermal faces for improving local appearance-based face recognition. Intelligent Automation & Soft Computing, 23(1), 1-12.

8. Dou, J., Qin, Q., & Tu, Z. (2018). Robust image matching based on the information of SIFT. Optik, 171, 850-861.
9. Sharma, D., Zafar, S., Babaie, M., & Tizhoosh, H. R. (2018, November). Facial recognition with encoded local projections. In 2018 IEEE Symposium Series on Computational Intelligence (SSCI) (pp. 1261-1268). IEEE.
10. Calder, A. J., & Young, A. W. (2016). Understanding the recognition of facial identity and facial expression. Facial Expression Recognition, 41-64.
11. Sensors, J. O. (2023). Retracted: Face Detection and Recognition Algorithm in Digital Image Based on Computer Vision Sensor.
12. Nguyen, D. T., & Park, K. R. (2016). Enhanced gender recognition system using an improved histogram of oriented gradient (HOG) feature from quality assessment of visible light and thermal images of the human body. Sensors, 16(7), 1134.
13. Bruno, D. R., Berri, R. A., Barbosa, F. M., & Osório, F. S. (2023). CARINA Project: Visual perception systems applied for autonomous vehicles and advanced driver assistance systems (ADAS). IEEE Access, 11, 69720-69749.
14. Hossam, M., Afify, A. A., Rady, M., Nabil, M., Moussa, K., Yousri, R., & Darweesh, M. S. (2021, July). A comparative study of different face shape classification techniques. In 2021 International Conference on Electronic Engineering (ICEEM) (pp. 1-6). IEEE.
15. Dadi, H. S., & Pillutla, G. M. (2016). Improved face recognition rate using HOG features and SVM classifier. IOSR Journal of Electronics and Communication Engineering, 11(04), 34-44.
16. Richmond, S., Howe, L. J., Lewis, S., Stergiakouli, E., & Zhurov, A. (2018). Facial genetics: a brief overview. Frontiers in genetics, 9, 462.
17. Khan, K., Attique, M., Syed, I., & Gul, A. (2019). Automatic gender classification through face segmentation. Symmetry, 11(6), 770.
18. Fekri-Ershad, S. (2020). Developing a gender classification approach in human face images using modified local binary patterns and tani-moto based nearest neighbor algorithm. arXiv preprint arXiv:2001.10966.
19. Anwarul, S., & Dahiya, S. (2019). A comprehensive review on face recognition methods and factors affecting facial recognition accuracy. Proceedings of ICRIC 2019: Recent Innovations in Computing, 495-514.
20. Azzopardi, G., Greco, A., & Vento, M. (2016, August). Gender recognition from face images with trainable COSFIRE filters. In 2016 13th IEEE international conference on advanced video and signal based surveillance (AVSS) (pp. 235-241). IEEE.
21. Althnian, A., Aloboud, N., Alkharashi, N., Alduwaish, F., Alrshoud, M., & Kurdi, H. (2020). Face gender recognition in the wild: an extensive performance comparison of deep-learned, hand-crafted, and fused features with deep and traditional models. Applied Sciences, 11(1), 89.
22. Ng, C. B., Tay, Y. H., & Goi, B. M. (2015). A review of facial gender recognition. Pattern Analysis and Applications, 18, 739-755.
23. Li, Z., Jiao, Z., & He, A. (2020). Knowledge‐based artificial neural network for power transformer protection. IET Generation, Transmission & Distribution, 14(24), 5782-5791.
24. Raji, I. D., & Fried, G. (2021). About face: A survey of facial recognition evaluation. arXiv preprint arXiv:2102.00813.
25. Tolba, M., & Moustafa, M. (2016). GAdaBoost: accelerating adaboost feature

selection with genetic algorithms. arXiv preprint arXiv:1609.06260.

26. Kaddar, B., Fizazi, H., & Boudraa, A. O. (2018). Texture features based on an efficient local binary pattern descriptor. Computers & Electrical Engineering, 70, 496-508.
27. Akhmet, M., Tleubergenova, M., & Zhamanshin, A. (2022). Dynamics of shunting inhibitory cellular neural networks with variable two-component passive decay rates and Poisson stable inputs. Symmetry, 14(6), 1162.
28. Adnan, S., Ali, F., & Abdulmunem, A. A. (2020, November). Facial feature extraction for face recognition. In Journal of Physics: Conference Series (Vol. 1664, No. 1, p. 012050). IOP Publishing.
29. Rustam, Z., & Ruvita, A. A. (2018, November). Application support vector machine on face recognition for gender classification. In Journal of Physics: Conference Series (Vol. 1108, No. 1, p. 012067). IOP Publishing.
30. Kumar, S., Singh, S., & Kumar, J. (2017, May). A study on face recognition techniques with age and gender classification. In 2017 International Conference on Computing, Communication and Automation (ICCCA) (pp. 1001-1006). IEEE.

Author Profile

Esraa Farouk received the M.Sc. degree and PhD in Computer Science from Assiut University, Assiut, Egypt. She currently works as a professor at Faculty of Commerce, South Valley University, Qena, Egypt.

Email: esraa.mohamed11@com.svu.edu.eg