IEEE Access
Multidisciplinary ⋮ Rapid Review ⋮ Open Access Journal

Received 20 February 2024, accepted 10 March 2024, date of publication 18 March 2024, date of current version 22 March 2024.
Digital Object Identifier 10.1109/ACCESS.2024.3377564

RESEARCH ARTICLE

Automatic Player Face Detection and Recognition for Players in Cricket Games

MAHMOOD UL HAQ [1], MUHAMMAD ATHAR JAVED SETHI [1],
SADIQUE AHMAD [2,3], (Member, IEEE), MOHAMMED A. ELAFFENDI [2],
AND MUHAMMAD ASIM [2,4]
[1]Department of Computer System Engineering, UET Peshawar, Peshawar 25000, Pakistan
[2]EIAS: Data Science and Blockchain Laboratory, College of Computer and Information Sciences, Prince Sultan University, Riyadh 11586, Saudi Arabia [3]Department of Computer Sciences, Bahria University, Karachi Campus, Karachi 75260, Pakistan
[4]School of Computer Science and Technology, Guangdong University of Technology, Guangzhou 510006, China Corresponding author: Sadique Ahmad (ahmad01.shah@ieee.org)

The authors would like to acknowledge the support of Prince Sultan University for the Article Processing Charges (APC) of this publication.

ABSTRACT In this paper, we have developed an augmented reality cricket broadcasting application that uses player face recognition during play to display player personal data. The system utilizes the AdaBoost algorithm for player and player face detection, employing a PAL based face recognition model to recognize the faces of players on the field. The system is trained on a large dataset of cricket game footage and achieves high accuracy in detecting and recognizing players' faces even with several conditions such as occlusion, non-uniformillumination,expressionandposevariation.Thesystemhasthepotentialtoenhancetheviewing experience of cricket games by providing real-time player identification and statistics. The system can also be used in other sports to provide similar benefits. The paper discusses the system's methodology, results, and implications for the future of sports broadcasting. Overall, the system provides a promising solution for automatic player face detection and recognition in sports broadcasting.

INDEX TERMS Face recognition, player detection, cricket.

I. INTRODUCTION

Face recognition technology is a subset of computer vision that focuses on identifying and verifying the identity of an individual based on their facial features [1]. Although the technology has been there since the 1960s, major advancements in the sector were made in the 1990s. Face recognition technology has become an increasingly important tool in the entertainment industry in recent years [2]. From movie theaters to sports stadiums, it is being used to enhance the experience of fans and viewers alike. In the context of sports, automatic player face recognition technology has the potential to revolutionize the way that fans interact with the game, providing real-time information on players and their performances [3].

Early face recognition systems were based on simple geometric features, such as the distance between the eyes and the width of the nose [4]. However, these systems were prone to errors, particularly when faced with variations in pose, expression, and lighting conditions. In the 2000s, deep learning techniques emerged as a powerful tool for face recognition, leading to significant improvements in accuracy and efficiency [5]. Deep learning algorithms, such as convolutional neural networks (CNNs) [6], can learn complex representations of facial features from large data sets, enabling them to recognize faces under a wide range of conditions [7].

Today, face recognition technology is used in a wide range of applications, from security and surveillance [8] to entertainment [9] and social media. In the entertainment industry, face recognition is being used to enhance the fan experience by providing real-time information on celebrities, athletes, and performers. In the context of sports, face statistics on a player's performance in a particular game or

The associate editor coordinating the review of this manuscript and approving it for publication was Joewono Widjaja.

2024 The Authors. This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 License.
For more information, see https://creativecommons.org/licenses/by-nc-nd/4.0/

recognition technology can be used to provide fans with information on players and their performances. This information can include tournament, as well as biographical information, such as their age, height, and weight [10], [11].

Cricket is a sport that is particularly well-suited to the application of face recognition technology [12]. With so many players to keep track of, it can be difficult for fans to know who is who on the field. Automatic player face recognition technology can help to solve this problem by providing fans with real-time information on the players they are watching [13].

One of the key advantages of our automatic player face recognition system is that it can be used in real-time during live cricket matches, providing fans with up-to-date information on the players they are watching. This can enhance the overall fan experience, making it easier for fans to follow the action on the field and providing them with insights into the performances of their favorite players. However, there are also limitations to our approach that must be considered. For example, the system may struggle to recognize players who are wearing helmets or who have their faces partially obscured by other players. In addition, the accuracy of the system may be impacted by variations in camera angle and distance [14]. To address these limitations, future research could focus on developing more advanced deep learning-based approaches that can handle a wider range of occlusions and lighting conditions [15]. Additionally, integrating multiple camera angles and using 3D face models could further improve the accuracy of the system. In conclusion, automatic player face recognition technology has the potential to enhance the fan experience in cricket and other sports. Our research has demonstrated the feasibility of using machine learning-based approaches to develop accurate and efficient face recognition systems for cricket. With further research and development, this technology could revolutionize the way fans engage with sports, providing them with new insights and enhancing their overall enjoyment of the game. The following are the research's contributions in this paper:

- In this work, we created an AR based sports broadcasting application that display real time player personal data in cricket games images.
- A variety of player images, including face expressions, lighting, occlusion, and game scenarios, were gathered from the internet to test the robustness of the proposed algorithm.
- Proposed algorithm does not impose any restriction on the input images and can handle the following aspects
  a) **Image Resolution:**The proposed algorithm efficiently recognizes player face images of 100 × 140, 50 × 70, and 20 × 30 pixels.
  b) **Posevariation:**Several pose changes are successfully recognized by the suggested facial recognition system.
  c) **Non-uniform illuminations:** In various illumination circumstances, the accuracy of the proposed face recognition system is significantly greater.
  d) **Occlusion:**The proposed algorithm gives better resultinrecognizingoccludedplayerfaceimages.
- The developed technique is free from issues such as player motion and texture degradation. The only thing viewers need to do is take a picture with their camera or smartphone; players will be automatically detected/recognized and personal information of the player will then be displayed automatically.
- The uniqueness of our model lies in its novelty, as there is presently no existing application especially for this particular task.

Overall, our research demonstrates the potential of automatic player face recognition technology to enhance the experience of fans and viewers in the context of cricket. By providing real-time information on players and their performances, this technology can help fans to better understand and appreciate the game. We hope that our work will inspire further research in this area, and ultimately contribute to the continued growth and development of the sport of cricket.

II. LITERATURE REVIEW

The issue of detecting and identifying faces in broadcast videos is a well-researched subject, and a comprehensive review of the extensive literature on face detection and recognition has been presented in [16] and [17]. Many face recognition techniques are tested on controlled environments and for a limited number of faces and poses, such as in serials or movies. However, these techniques may not be suitable for recognizing faces in sports videos, including soccer videos, due to the unpredictable variations in pose, illumination, settings, scale, and occlusion that may occur in uncontrolledenvironments.Recently,studieshaveshownthat SIFT local features, typically used for object recognition and classification tasks [18], can be utilized to detect and group faces of the same person in video shots [19], [20]. In previous work [21], textual cues from superimposed captions and jersey numbers for player recognition, and SIFT descriptors focused on the eyes for face matching were utilized. In [22] and [23], the authors investigated the automatic labeling of characters' appearances in TV videos and demonstrated that high accuracy can be achieved by combining visual and textual information. These studies focused on detecting faces in TV or film video sequences and finding instances of the same person among all detected faces. Similarity of faces was measured by computing the distances between local descriptors based on statistical parameters, such as the $\chi 2$ statistic.

To date, research on the advancement of Augmented Reality (AR) applications remains limited. While it has been utilized in some entertainment applications utilizing vision-based tracking techniques [24], there have been shortcomings in other attempts to develop AR technology. For example, one technique presented in [25] aimed to create a visual enhancement for TV-broadcasted court net sports, but it generated redundant virtual scenes that negatively affected system performance. Similarly, in [26], virtual scenes were generated using multiple synchronous video

accuracy, even in the presence of occlusions and changes in pose and lighting.

However, there are still challenges to be addressed, such as the need for large amounts of labeled training data and the potential for biases in the data. Future research can focus on improving the accuracy and efficiency of face recognition in sports videos, as well as exploring the integration of face recognition with other technologies, such as action recognition, to provide more accurate player identification. Overall, automatic player face recognition for sports holds great promise for enhancing the sports viewing experience and providing valuable insights for sports analysis.

sequences of a given sports game, but this technique faced synchronization problems in real-time transmitted videos and images. Another proposed technique, a graphics-based animation AR technique [27], produced unsatisfactory results due to the complete loss of texture and motion of the players. In [28], authors attempted to insert and manipulate 3D virtual content into broadcasted tennis videos, but this technique failed in certain situations, such as when one net was invisible. In [29], a developed system performed a camera calibration algorithm to establish a mapping between the soccer court field in the image and that of the virtual scene, but it had the major drawback of player pose restriction, as the user could only choose from three defined poses (walk, stop, and run). This confinement to only three poses limited the system's performance improvement. Despite these previous efforts, an acceptable level of reality in AR applications has yet to be achieved.

In summary, Published work has use a variety of techniques and methodologies to achieve high accuracy in face recognition in sports videos. Advantages include real-time performance, handling faces from different angles [30] and lighting conditions [31], high accuracy, and robustness to changes in facial expression [32]. Limitations include the need for large amounts of training data [33] computational intensity, and limitations on recognition from specific angles or lighting conditions.

In conclusion, automatic player face recognition for sports has the potential to enhance the viewing experience and improve the analysis of sports events. The literature review highlights that traditional methods of face recognition in sports videos, such as template matching and feature-based methods, have been surpassed by modern deep learningbased approaches. These approaches utilize convolutional neural networks (CNNs) and recurrent neural networks (RNNs) to extract features from frames in a video sequence and classify them based on the identity of the player. Table 2 presents the review of several face recognition algorithms. The reviewed papers show that machine learning-based approaches for player detection and recognition in videos can achieve high

III. METHODOLOGY

This section details the proposed methodology's framework, which can be seen in Figure 1. After conducting a comparative analysis of various player detection and face detection techniques, we chose to utilize AdaBoost with Haar-Like features due to its real-time performance and high accuracy. For face recognition, we employed PAL based face recognition model, as it is adept at capturing hierarchical relationships between features, which is especially beneficial in tasks such as object recognition.

A. IMAGE ACQUISITION

Images of several cricket players were gathered in this step. For this kind of search, Google is utilized. Various factors, including poses, facial expressions, illumination, and occlusion, were taken into consideration when choosing the photos for the suggested collection. These images were collected from different cricket teams having different game format such as test, odi and t20. We have a wide range of face images of cricket players in our collection, representing various teams, playing styles, and backgrounds. It includes images from practices, games, and team portraits that provide a variety of facial expressions and facial recognition challenges. A panel of three specialists in image processing was recommended for this task. From the gathered images, these specialists have chosen the players who can be seen to be included in the dataset.

These images were then resized to 100 × 140, 50 × 70 and 20 × 30 pixels. Figure 2 presents the images of three player in different situations while table 2 provides the specification along with the introduced challenging conditions added in these images.

B. PLAYER DETECTION AND FACE DETECTION

For player detection, four baseline human detection algorithms such as AdaBoost [51], HOG [52], faster R-CNN [53] and CNN [54] has been tested based on accuracy, occlusion, computational complexity and image resolution to select the best algorithm. These algorithms were tested on the above collected human (Player) and non-human images taken from different cricket ground as shown in figure 3. To test the robustness of these algorithms, the original images were resized to 100 × 140, 50 × 70, and 20 × 30 pixels. Table 3 and Table 4 presents the player detection and face detection accuracy of these algorithms. From these tables, it can be concluded that AdaBoost, or Adaptive Boosting, is a machine learning-based algorithm that can achieve moderate accuracy in detecting faces. It is known for its fast training and

TABLE 1. Comparative analysis of several published FR algorithms.

inference time. So AdaBoost based player detection model and face detection model has been selected for our proposed model.

The AdaBoost algorithm enhances accuracy by utilizing a sequence of weak classifiers at different stages to create a strong classifier, which is a linear combination of the weak ones. In our framework for player and face detection, we adopt the widely used Viola and Jones [55] method. By applying supervised AdaBoost learning to a sample set

<table>
<tr><th>Ref</th><th>Year</th><th>Technique</th><th>Dataset</th><th>Accuracy</th><th>Advantages</th><th>Limitations</th></tr>
<tr><td>[34]</td><td>2017</td><td>SphereFace</td><td>LFW, YTF, MegaFace</td><td>99.42%</td><td>1. High accuracy.
2. Good performance in large-scale face recognition tasks.</td><td>1. Requires a large amount of training data.
2. Computationally expensive.</td></tr>
<tr><td>[35]</td><td>2015</td><td>VGG-Face</td><td>LFW, YouTube Faces</td><td>98.95%</td><td>1. High accuracy.
2. Robustness against variations in facial expressions and lighting conditions.</td><td>1. High computational cost.
2. Requires a large amount of data for training.</td></tr>
<tr><td>[36]</td><td>2016</td><td>3D Face Morphable</td><td>BU-3DFE, Bosphorus, 3DFAW</td><td>99.05%</td><td>1. High accuracy.
2. Can handle pose and expression variations.</td><td>1. Requires 3D models of the face.
2. Computationally expensive.</td></tr>
<tr><td>[37]</td><td>2014</td><td>DeepFace</td><td>LFW, YouTube Faces</td><td>97.35%</td><td>1. High accuracy.
2. Can handle variations in facial expressions and lighting conditions.</td><td>1. Requires a large amount of data for training.
2. Computationally expensive.</td></tr>
<tr><td>[38]</td><td>2015</td><td>Fast R-CNN</td><td>PASCAL VOC, COCO, WIDERFACE</td><td>94.90%</td><td>1. Good accuracy.
2. Can handle real-time face detection tasks.</td><td>1. Limited to face detection tasks.
2. Requires a large amount of data for training.</td></tr>
<tr><td>[39]</td><td>2015</td><td>FaceNet</td><td>LFW, MegaFace, Sports videos</td><td>99.63%</td><td>1. High accuracy.
2. Can handle large-scale datasets.</td><td>1. High computational cost.
2. Requires a large amount of data for training.</td></tr>
<tr><td>[40]</td><td>1991</td><td>Haar Cascade Classifier, and Max-Margin object detection</td><td>Sports videos</td><td>85%</td><td>1. Low computational cost.
2. Can handle small datasets.</td><td>-
2. Cannot handle variations in pose and illumination well.</td></tr>
<tr><td>[41]</td><td>2018</td><td>Convolutional Neural Networks</td><td>Self-collected dataset of surveillance videos</td><td>96.70%</td><td>1. High accuracy with deep learning methods</td><td>1. Limited dataset size</td></tr>
<tr><td>[42]</td><td>2015</td><td>Viola Jones with AdaBoost LDA</td><td>Self-collected dataset of player images</td><td><table><tr><th>No of faces/ image</th><th>Accuracy (%)</th></tr><tr><td>1-3</td><td>100</td></tr><tr><td>4</td><td>88.88</td></tr><tr><td>5</td><td>83.33</td></tr><tr><td>6</td><td>-</td></tr></table></td><td>1. High accuracy
2. Low computational complexity</td><td>1. Unable to recognize more than five faces per image</td></tr>
<tr><td>[43]</td><td>2019</td><td>PAL</td><td>LFW, CMU Multi-PIE</td><td><table><tr><th>Dataset</th><th>Accuracy (%)</th></tr><tr><td>LFW</td><td>94</td></tr></table></td><td>1. High accuracy.
2. Only use one training image for each class</td><td>1. High computational time</td></tr>
</table>

TABLE 1. *(Continued.)* Comparative analysis of several published FR algorithms.

				CMU Multi-PIE: 96	3. Suitable for low resolution. 4. Gives promising results even in occlusion	
[44]	2019	ArcFace	LFW, AgeDB-30, MegaFace	99.83% (LFW)	1. High Accuracy	1. Requires a large amount of training data
[45]	2019	Capsule Network	LFW dataset.	93.7%	-	Computationally expensive
[46]	2023	QMagFace	AgeDB, XQLFQ, LFW and CFP-FP.	Dataset / Accuracy (%): XQLFQ: 83.95 LFW: 99.83 AgeDB: 98.50 CFP-FP: 98.74	1. High Accuracy	1. Occlusion and illumination are not discussed
[47]	2022	Texture Feature + Viola jones + SVM	Own collected records	96.8 %	1. Can recognize multiple faces in a single frame.	1. Occlusion and illumination are not discussed
[48]		CNN with MeDiConv layer	LFW, CP-LFW, CA-LFW, and YTF	Data set / Accuracy (%): LFW: 72.65 YTF: 76.54 Note: Gaussian_blur+MG +Scaling+Compression were added in images of these datasets	1. More robust to gray scale variations, rotation changes, and noise	-
[49]	2023	CNN with preprocessing	VGGFACE2.	94.1%	-	-
[50]	2022	FaceMask Net-21	Own collected dataset, RMRD.	Dataset / Accuracy (%): Own dataset: 88.92 RMRD: 82.2	1. High accuracy 2. Less execution time	3. Fail to recognize multiple faces in a single frame

of training data {xi, yi}, we derive a robust player classifier. a collection of Haar-like rectangle features and merges them The algorithm selects a group of weak classifiers (hj(x)) from into a strong classifier. The resulting strong classifier g(x) is

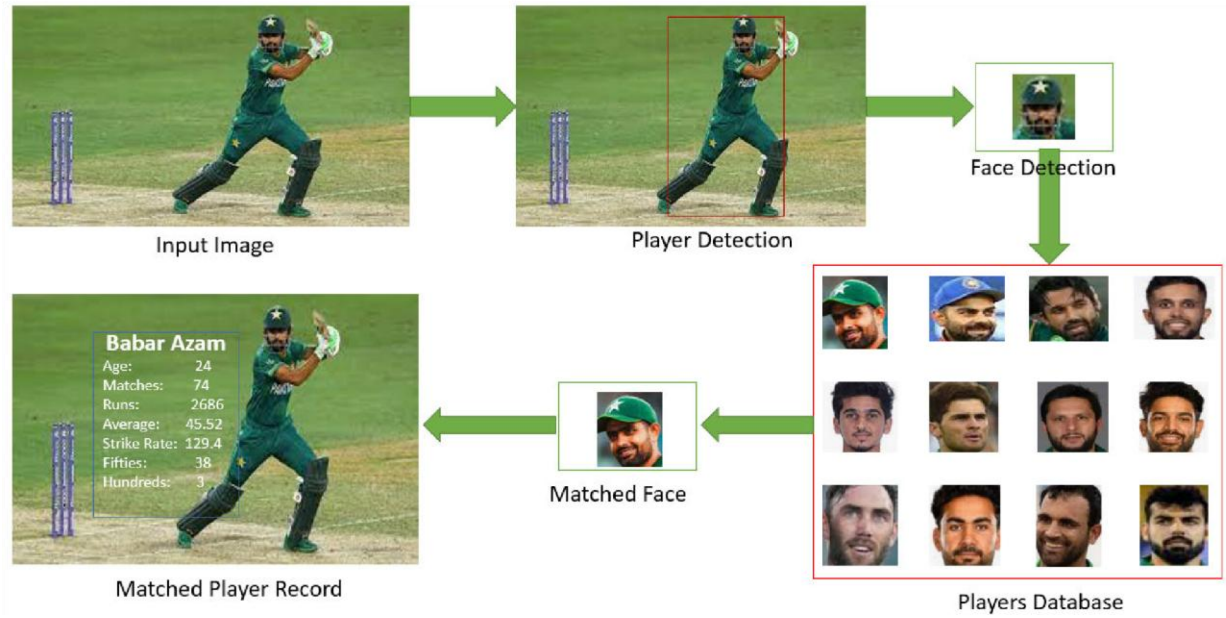


FIGURE 1. Framework of our proposed system.

TABLE 2. An explanation of images collected under difficult circumstances.

No of images	No of subjects	No of teams	Images per subject/ team	Conditions					
				Face Pose	Occlusion	Illuminatio n	Resolution	Facial Expression	Tilted Face
850	50	7	Varies	Yes	hat, helmet, glasses, sun block	Day, night, sunny, cloudy	100 × 140, 50 × 70 and 20 × 30 pixels	Smile, angry, sad	Yes

TABLE 3. Comparison of detection rate of player detection algorithms.

No of players per image	Image Resolution (Pixel)	Algorithms			
		AdaBoost	HOG	Faster R-CNN	CNN
1	100 × 140	98	84	98	98
	50 × 70	98	80	98	98
	20 × 30	98	78	94	90
2 or more	100 × 140	98	88	98	98
	50 × 70	98	76	92	92
	20 × 30	96	66	90	90

The final strong classifier, represented by g(x), is generated using an input image x, h weak classifiers, and a decision threshold of h. Algorithm 1 presents the Pseudo code of AdaBoost

defined as follows:

$$g(x) = \begin{cases} 1, & \sum_{i} \alpha_i h_i(x) \geq 0 \\ 0, & Otherwise \end{cases} \quad (1)$$

C. FACE ALIGNMENT

Facial alignment typically refers to the process of adjusting the position and orientation of facial features, such as the eyes, nose, and mouth, so that they are centered and

TABLE 4. Comparison of detection rate of face detection algorithms.

No of faces per image	Image Resolution (Pixel)	Algorithms			
		AdaBoost	HOG	Faster R-CNN	CNN
1	100 × 140	100	88	100	100
	50 × 70	100	88	100	100
	20 × 30	100	72	100	100
2 or more	100 × 140	100	90	100	100
	50 × 70	92	80	100	96
	20 × 30	92	70	96	90

FIGURE 2. Images of three players in different situations.

FIGURE 3. Example images for our proposed model.

symmetrical. This can be done manually by a human, or with the help of computer software and algorithms.

Facial alignment is an important step in many computer vision and machine learning tasks that involve analyzing or recognizing faces, such as facial recognition systems or emotion detection. It can also be useful in cosmetic surgery, where precise adjustments to facial features can improve the appearance and symmetry of a person's face.

The paper in question involves identifying variations in pose in facial images. To achieve this, we use the FLE method (facial landmark estimation method) [56] to locate 68 specific points on the face. This method is known for its robustness in handling variations in pose in real-time situations, making it a useful tool for analyzing the shape and position of facial features and detecting changes in expression and other factors [57].

Algorithm 1 Pseudo Code Of Adaboost Algorithm

- Given example images $(x_1,y_1),\ldots\ldots,(x_n,y_n)$ where $y_i = 0,1$ for negative and positive examples respectively.
- Initialize weights $w_{1,i} = \frac{1}{2m}, \frac{1}{2l}$ for $y_i = 0,1$ respectively, where m and l are the number of negatives and positives respectively.
- For $t = 1,\ldots\ldots T$:

1. Normalize the weights,

$$W_{t,i} \leftarrow \frac{W_{t,i}}{\sum_{j=1} W_{t,j}}$$

So that w_t is a probability distribution.

2. For each feature, j, train a classifier h_j which is restricted to using a single feature. The error is evaluated with respect to w_t, is $\epsilon_j = \sum_i w_i \,|\, h_j(x_i) - y_i \,|$
3. Choose the classifier, h_t, with the lowest error ϵ_t.
4. Update the weights: $W_{t+1,i} = w_{t,i}\beta_t^{1-e_i}$ where $e_i = 0$ if example x_i is classified correctly, $e_i = 1$ otherwise, and

$$\beta_t = \frac{\epsilon_t}{1 - \epsilon_t}$$

5. The final strong classifier is:

$$g(x) = \begin{cases} 1 & \sum_{t=1}^{T} \alpha_t h_t(x) \geq \frac{1}{2} \sum_{t=1}^{T} \alpha_t \\ 0 & \text{otherwise} \end{cases}$$

Where $\alpha_t = \log 1/\beta_t$

D. FACE RECOGNITION

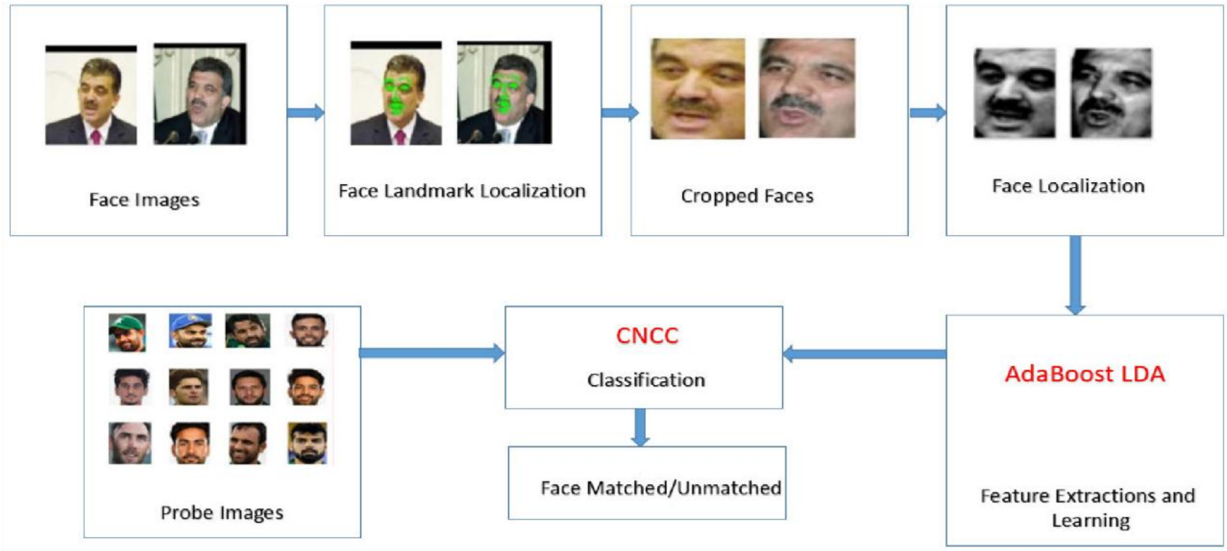


FIGURE 4. PAL based face recognition model.

TABLE 5. Accuracy of the proposed model.

Image Resolution (Pixels)	No of persons/ image	Player Detection Accuracy (%)	Face Detection Accuracy (%)	Face Recognition Accuracy (%)	System Accuracy (%)
100 × 140	1	98	100	100	98
	2	100	100	100	100
	3	98	100	96	94
	4	100	100	93	93
	5	92	100	72	64
	Above than 5	Not Processed			
50 × 70	1	98	100	100	98
	2	100	100	100	100
	3	98	100	96	94
	4	100	90	93	93
	5	94	89	83	66
	Above than 5	Not Processed			
20 × 30	1	98	100	100	98
	2	100	100	100	100
	3	98	100	87	85
	4	96	93	88	77
	5	88	88	73	49
	Above than 5	Not Processed			

For face recognition, PAL algorithm has been used because of its robustness to pose variation, occlusion, illumination variation and low resolution.

Initially the PAL face recognition algorithm involves computing the mean and standard deviation of the facial images for each subject, followed by an adjustment using a formula specified in Equation (2). This technique aims to mitigate errors arising from differences in lighting conditions.

$$I_{new} = \frac{(I - \bar{i}) \times \sigma_d}{\sigma_i} + i_d \qquad (2)$$

In this approach, $\bar{i}$ and σ_i represents the mean and standard deviation of each facial image, while i_d and σ_d represent predefined mean and standard deviation values, respectively. In the FLE method, a single training image (mean image) is utilized for each subject/class. The mean image is obtained by calculating the mean value of multiple facial images of each subject, and detailed under the second bullet of Table 4. This step is partially inspired by applying PCA to generate the mean image ($\bar{I}$), which is computed using Equation (3).

$$\bar{I}^{s} = \frac{\sum_{j}^{J} I_{jn}^{s}}{J} \quad (3)$$

In the given equation, I_{jn}^{s} represents the jth normalized training image of subject s, where J denotes the total number of training images for subject s. The FLE approach is employed to first estimate and normalize both probe and gallery images using Equation (1). These normalized images are then input into the Adaptive Boosting (AdaBoost) algorithm, which is combined with LDA for recognition. To facilitate the learning process, the proposed PAL face recognition algorithm utilizes AdaBoost with LDA. The training set (Z) comprises multiple classes (C), with each class (Zi) containing samples (zij) and their corresponding labels (yij). The sample space (Z) is defined as zijϵ Z, and the label set (Y) is represented by yij = iϵY. Given an unseen face sample (z), the aim of the learning process is to estimate a function or classifier h(z): Z→Y. AdaBoost applies a given weak learner (ht) repeatedly to the training set with weighted version in several rounds (T), and then combines all weak classifiers (ht) constructed in each round into a strong classifier $hf(z)$. The final strong PAL classifier can be expressed as follows:

$$h_f(z) = \sum_{y \in Y} \log \frac{1}{\beta_t} h_t(z,y) \quad (4)$$

Finally the Classic Nearest Centre Classifier (CNCC) is used for final face classification. The CNCC algorithm utilizes a normalized Euclidean distance for its calculations. By using the nearest center rule, it determines the class label y(z) for an input face (z) based on equation 5.

$$y(z) = argmax_i d(z,i,l) \quad (5)$$

where the classification score d(z, I, L) falls within the range of 0 to 1. The NCC produces two outputs, namely the classification score d(z, I, L) and the class label y(z). For the sake of differentiation, h(z) is defined as y(z) and h(z, I) is defined as d(z,I,L).Thepseudocodeandmodelof PALfacerecognition algorithm are presented in figure 4 and algorithm 2.

E. PLAYER STATISTICS

The ultimate objective of this project is to present significant statistics related to a player or players once an image is captured or acquired from a video. The accuracy of the detection and recognition results plays a crucial role in achieving this goal. The personal details of each player stored in our database are obtained by utilizing the face matching outcome. This application has the capability to exhibit various information to the viewers about the players, such as their name, and sports record. Although this system was designed for a cricket game, it can be adapted for use in any sport where the audience can capture live videos or images using their smartphones.

IV. EXPERIMENTAL RESULTS AND DISCUSSION:

The system proposed in this research is comprised of three modules: player detection, face detection, and face recognition. To fully evaluate the performance of the system, it is essential to analyze the results of each module separately, in addition to considering the system's overall performance as a 'black box'.

The research utilizes a dataset of 850 diverse images that ~~contain varying numbers of players and are captured under~~

Algorithm 2 Pseudo Code of PAL Based Face Recognition ~~Model~~

1. **Input**: A set of input images A = $a^{J}_{i=1}$ $^{I}_{i=1}$ with I = {1,2,……,I} classes and J images of each class.
2. **for** i = 1,…,I
   Convert RGB images to grey.
   Estimate and crop face ($I_{cropped}$.
   Update mean and standard deviation of each image,
   $I_n = \frac{I_{cropped} - \bar{X}}{\sigma_i} \times \sigma_{def} + \bar{X}_{def}$
   Calculate mean image of each class, $tr_i = \frac{\sum_{j=1}^{J} a_{ij}}{J}$.
   Final Training images of each class, T_r = {tr_1,……,tr_1}.
   end for
3. for J = 1. window size (5): ▶ for our simulations we set 5 = 10 to get the recognition result.
   Initialize mislabelled distribution over $m, D_l(i) = \frac{1}{m} = \frac{1}{N(C-1)}$
4. **Do for** t = 1,…,T: ▶ for our simulations we set T = 30 to get the recognition result.
   If t = 1, choose i samples per class for learner.
   Train the LDA feature extractor (L). Build a PAL Classifier h_t. ▶ h_t is a strong and an ensemble of 1800 weak classifiers. Calculate pseudo loss, e_t
   Calculate $\beta_t = e_t/(1 - e_t)$
   If β_t = 0, abort the loop and Update the distribution
   end for
5. Final PAL Classifier of training image, $hf(\mathbf{z}) = argmax \sum \log \frac{1}{\beta_t} h_t(\mathbf{z},y)$.

6. Compare output of PAL classifier with testing image and generate a matching score.
7. Check the maximum matching score for each window:

   $(M_{score}\ (J)), I_{recog} = argmax\ (M_{score}\ )$, ▶ maximum matching score= test of images.
8. Find labels
9. **Output**: Use CNNC for each window to recognize face image.

---

different lighting conditions. The dataset also includes small pose variations of the players' faces from the frontal view. A total of 50 different subjects (players) are included in the dataset,asalreadyshowninFigure3.Table5providesaquantitative summary of the performance of each module used in the research. The number of players per image is an important factor as it reflects the size of the detected player. Images with fewer players tend to be captured from a shorter distance and have higher resolution, which can improve the subsequent steps' performance. In the following sections, each module developed in the research is discussed and analyzed in detail.

A. PLAYER DETECTION

The images in Figure 5 illustrate examples of players that have been detected using the methodology employed by the research. These examples demonstrate the feasibility of the proposed approach for real-world scenarios. Additionally, the research performs face detection and recognition on images containing multiple players, even when there is variation in lighting conditions and deviations from the frontal pose.

B. FACE DETECTION

The player face detection module receives detected players as its input and is responsible for detecting the faces of the players. Figure 6 provides some examples of the results obtained from the player face detection module. It can be observed that the module performs well in most cases, except when the detected players are very small in size. Face detection module is unable to process extremely blurred and invisible player face images, which is typically the case when even a human eye is unable to locate such faces.

FIGURE 5. Player detection.

C. FACE RECOGNITION

The database was randomly divided into two subsets: a training set (Ttrain) and a test set (Ttest). The Face Recognition system was first trained using the Ttrain training images, and the resulting face recognizer was then applied to the Ttest test images to determine the classification error rate (CER). Figure 7 presents few examples of output of module. When the input face is matched to a face in the database, personal information of the matched player, such as age, score, and nationality, is retrieved from the database of players' statistics. The developed application demonstrates the feasibility of computer vision-based approaches for enhancing sports broadcasting.

The proposed module was implemented in low-level language. Implementing the entire system in a low-level language would make it usable in real-time applications, such as processing images from a video. Figure 8shows the execution time of each module. A Dell Precision Tower 7810 with a 16 GB RAM is used for the experiments. With the exception of the 100 × 140 pixel images having five players per image, all situations have execution times that are almost exactly real time. Nonetheless, there are a number of variables that can improve the suggested model's accuracy and execution time.

- One way to decrease the execution time is to use a workstation that is more powerful and has a larger RAM and GPU.

FIGURE 6. Player face detection.

FIGURE 7. Output of proposed module.

- The utilization of CNN for player and face detection could potentially enhance the precision of the suggested model. But this tactic will lengthen the training period and raise the parameters during execution.

In certain instances, face recognition failed and only player and facial detection worked. In other instances, both the subsequent face detection and recognition attempts failed, with the exception of player detection. Table 6 displays a few instances of these outcomes.

The face recognition system for players works flawlessly even with a small sample size of 20 × 30 and just one training/testing sample available. Our experiments included player face databases of various sizes, ranging from 100 × 140 to 20 × 30 image size, and the system proved to be robust and accurate in all scenarios. This implies that the proposed player face recognition system we used is highly suitable for real-world applications, such as identifying players in live sports games where the image captured by a smartphone is no smaller than 20 × 30.

The accuracy ofthe face recognition module isdisplayed in figure 9 when the quantity of training images is changed. The accuracy of the proposed model could be affected by altering the training images. It is also clear that accuracyrises and may even reach 100% as the number of training images increases. The key points of figure 9 are

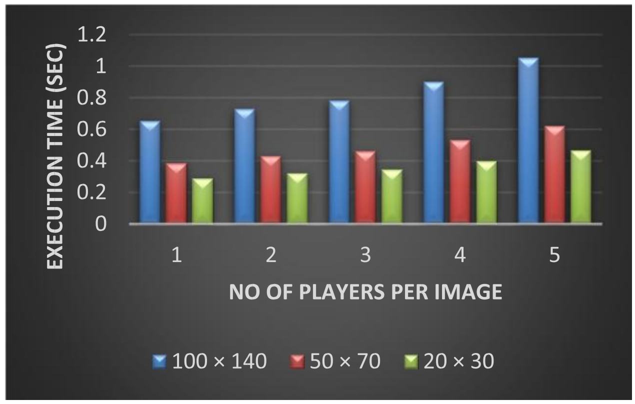


FIGURE 8. Execution time of proposed model.

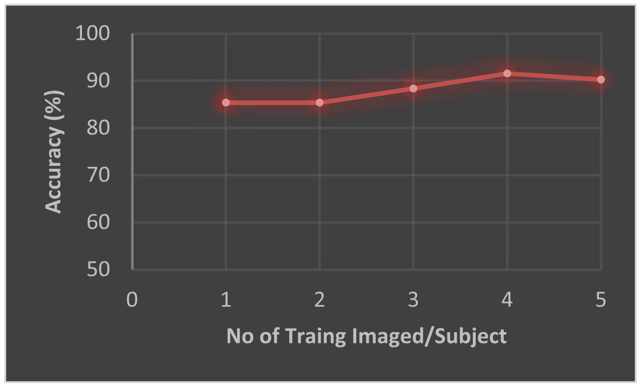


FIGURE 9. Accuracy in terms of each subject's usage of training images.

- Themoretrainingimagesamodelhas,themoreaccurate it becomes.
- A training sample that has similarities with several participantsmayresultinamodel'saccuracybeingreduced.
- A model may occasionally get overfit with too many training examples.

Proposed model are designed to operate in real-time, with minimum latency and high accuracy, making them suitable for various sports applications especially cricket. These systems typically use advanced algorithms to analyze and match features in real-time, providing instantaneous and reliable results. Table 6 presents the results when player are not recognized under several conditions.

D. ROBUSTNESS OF THE MODEL UNDER EXTREME CONDITIONS

The suggested model was evaluated on images with each of these conditions separately in order to assess the robustness of the system under extreme conditions, such as changing lighting, occlusions, and position fluctuations. In order to achieve this, two images (100 × 140 pixel) of each person in each of the three conditions (occlusion, stance variation, and light variation) were obtained. These images were used to evaluate the suggested model. Samples of photos acquired for this purpose are shown in Figure 10. Table 7 presents the accuracy of the proposed model for each of the three mentioned conditions. The key points of table 7 are

- The proposed model effectively identified images with light variation having accuracy of 97 %.
- The proposed model also promises to recognize images with pose variation
- Images with occlusion have the highest identification error rates.

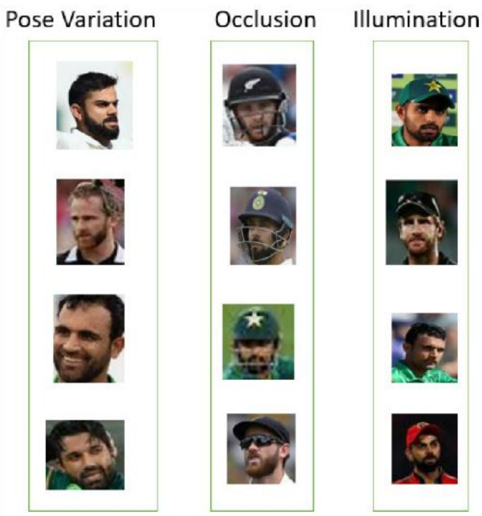


FIGURE 10. Sample of images under several conditions.

V. COMPARISON WITH CUTTING-EDGE ALGORITHMS

An experiment was evaluated in which the collected player images were recognized and tested on state-of-the-art face recognition algorithms, including PCA [58], AdaBoostLDA [43], CNN [59], and Capsule Network used for character recognition [60], in order to assess the robustness of the face recognition algorithm used in this model.

The state of certain algorithms caused images of faces to be transformed to 128 × 128 pixels. As was previously said, there are multiple facial expressions, lighting, poses, and occlusion in these pictures. The results of these algorithmsare displayed in Figure 11.

TABLE 6. Error Performance with their reasons of error.

Image	Player Detection	Face Detection	Face Recognition	Reason for Error Output
	4/5 players detected	4/4 faces detected	4/4 faces recognized	High occlusion
	Player detected	Face detected	Face not recognized	High occlusion and pose variation
	4/4 players detected	2/4 faces detected	2/2 faces recognized	High pose variation
	Player detected	Face detected	Face not recognized	High similarity with another subject
	5/6 players detected	4/5 faces detected	2/4 faces recognized	High occlusion and pose variation
	6/9 players detected	5/6 faces detected	3/5 faces recognized	High occlusion and pose variation
	5/5 players detected	3/5 faces detected	2/3 faces recognized	High occlusion

TABLE 7. Accuracy of model under extreme conditions.

Conditions	Accuracy (%)
Pose Variation	93
Occlusion	67
Illumination	97

The main explanations from Figure 11 are as follows.

- When it comes to face images, the CNN algorithm has the highest recognition rate (95.6%), followed by the

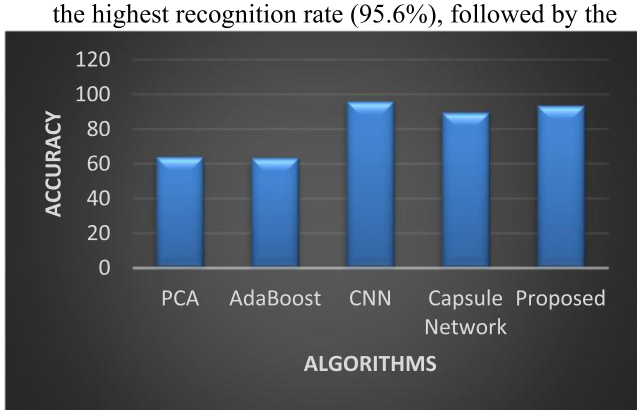


FIGURE 11. Comparison of recognition accuracy.

suggested network (93.33%) and the capsule network (87.6%).

- CNN, however, is not appropriate for this kind of application because to its lengthy training period and sporadic addition of fresh subjects and photos.
- After adding fresh images/subjects, the suggested model can be trained with ease.
- The concept that has been proposed is now in its early stages, and it is intended to be expanded in the future to include all cricket teams and players.

VI. IMPLICATION OF PROPOSED MODEL IN CRICKET INDUSTRY

In the cricket industry, face recognition technology can be a useful tool for a variety of purposes, such as player identification, security, fan engagement, and more. To ensure a face recognition model's efficacy and moral application, implementation in this situation calls for careful thought and simplification. Consider the following implications:

A. PRIVACY ISSUES

The collection and storage of player face data gives rise to privacy issues. Players can feel uneasy about their biometric information being saved and even shared without their knowledge or agreement.

B. ETHICAL GUIDELINES

When using player facial recognition, it is important to follow ethical guidelines such accountability, openness, and permission. It's crucial to prevent the misuse or exploitation of player data.

C. PLAYER BUY-IN

Players' concerns and opinions should be taken into account after they have been informed about the usage of face recognition technology. For deployment to be effective, it is crucial to guarantee that players feel at ease using the technology.

D. TECHNICAL DIFFICULTIES

A robust face recognition system demands technical know-how and money to develop and maintain. To maintain the system's accuracy and security, regular updates and maintenance are required.

Using a facial recognition model to identify players in the cricket sector can have advantages like easily storing and accessing player record, effectiveness, and personalization. It also presents privacy, security, ethical, and technical challenges, all of which demand careful consideration. Responsible use of such technology in the cricket industry requires observing legal and ethical guidelines, respecting players' privacy, and obtaining their consent.

VII. CONCLUSION AND FUTURE WORK

In this paper we proposed a computer vision-based system that can automatically identify and label players in real-time, providing viewers with relevant information such as player statistics, past performances, and current form. This work demonstrates the viability of creating a real-time augmented reality application that solely uses image processing and computer vision techniques to improve users' experiences. One straightforward use case for this kind of device is when spectators are seated in a stadium or other public area without access to a TV that provides crucial player and game status information. Viewers can begin shooting videos with their smartphones or cameras, boosting the quality of the video or image by automatically detecting and recognizing players and displaying statistics with each frame in real time.

Future research can focus on improving the accuracy and efficiency of the proposed system and exploring its potential integration with other technologies such as action recognition, player tracking, and analytics. Overall, the proposed system has the potential to revolutionize the way cricket matches are watched and analyzed, and can pave the way for similar systems in other sports and entertainment domains.

ACKNOWLEDGMENT

The authors would like to thank Prince Sultan University for their valuable support.

REFERENCES

[1] H. Ullah, M. U. Haq, S. Khattak, G. Z. Khan, and Z. Mahmood, "A robust face recognition method for occluded and low-resolution images," in *Proc. Int. Conf. Appl. Eng. Math. (ICAEM)*, Taxila, Pakistan, Aug. 2019, pp. 86–91, doi: 10.1109/ICAEM.2019.8853753.

[2] F. Munawar, U. Khan, A. Shahzad, M. U. Haq, Z. Mahmood, S. Khattak, and G. Z. Khan, "An empirical study of image resolution and pose on automatic face recognition," in *Proc. 16th Int. Bhurban Conf. Appl. Sci. Technol. (IBCAST)*, Islamabad, Pakistan, Jan. 2019, pp. 558–563, doi: 10.1109/IBCAST.2019.8667233.

[3] G. Rajeshkumar, M. Braveen, R. Venkatesh, P. J. Shermila, B. G. Prabu, B. Veerasamy, B. Bharathi, and A. Jeyam, "Smart office automation via faster R-CNN based face recognition and Internet of Things," *Measurement: Sensors*, vol. 27, Jun. 2023, Art. no. 100719, doi: 10.1016/j.measen.2023.100719.

[4] A. Rehman, T. Saba, M. Mujahid, F. S. Alamri, and N. ElHakim, "Parkinson's disease detection using hybrid LSTM-GRU deep learning model," *Electronics*, vol. 12, no. 13, p. 2856, Jun. 2023, doi: 10.3390/electronics12132856.

[5] M. U. Haq, M. A. J. Sethi, R. Ullah, A. Shazhad, L. Hasan, and G. M. Karami, "COMSATS face: A dataset of face images with pose variations, its design, and aspects," *Math. Problems Eng.*, vol. 2022, pp. 1–11, May 2022, doi: 10.1155/2022/4589057.

[6] M. A. Hamza, S. Ben Haj Hassine, S. Larabi-Marie-Sainte, M. K. Nour, F. N. Al-Wesabi, A. Motwakel, A. Mustafa Hilal, and M. Al Duhayyim, "Optimal bidirectional LSTM for modulation signal classification in communication systems," *Comput., Mater. Continua*, vol. 72, no. 2, pp. 3055–3071, Jan. 2022, doi: 10.32604/cmc.2022.024490.

[7] M. U. Haq, M. A. J. Sethi, and A. U. Rehman, "Capsule network with its limitation, modification, and applications—A survey," *Mach. Learn. Knowl. Extraction*, vol. 5, no. 3, pp. 891–921, Aug. 2023, doi: 10.3390/make5030047.

[8] J. Wu, C. Peng, W. Tan, and Z. Wu, "FaceEncAuth: Face recognition privacy security scheme based on FaceNet and SM algorithms," *Comput. Eng. Appl.*, vol. 58, no. 11, pp. 93–99, 2022.

[9] R. Pena, F. Ferreira, F. Caroli, L. Schirmer, and H. Lopes, "Globo face stream: A system for video meta-data generation in an entertainment industry setting," in *Proc. 22nd Int. Conf. Enterprise Inf. Syst.*, 2020, pp. 1–9, doi: 10.5220/0009380203500358.

[10] H. Li, A. Manickam, and R. D. J. Samuel, "RETRACTED ARTICLE: Automatic detection technology for sports players based on image recognition technology: The significance of big data technology in China's sports field," *Ann. Oper. Res.*, vol. 326, no. S1, p. 97, Jan. 2022, doi: 10.1007/s10479-021-04409-1.

[11] H. Qiu, D. Gong, Z. Li, W. Liu, and D. Tao, "End2End occluded face recognition by masking corrupted features," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 10, pp. 6939–6952, Oct. 2022, doi: 10.1109/TPAMI.2021.3098962.

[12] V. Ellappan and R. Rajkumar, "Classification of cricket videos using finite state machines," *Int. J. Inf. Technol. Manage.*, vol. 20, nos. 1–2, p. 83, Jan. 2021, doi: 10.1504/ijitm.2021.114156.

[13] X. Li, Y. Xiang, and S. Li, "Combining convolutional and vision transformer structures for sheep face recognition," *Comput. Electron. Agricult.*, vol. 205, Feb. 2023, Art. no. 107651, doi: 10.1016/j.compag.2023.107651.

[14] E. Bermejo, E. Fernandez-Blanco, A. Valsecchi, P. Mesejo, O. Ibáñez, and K. Imaizumi, "FacialSCDnet: A deep learning approach for the estimation of subject-to-camera distance in facial photographs," *Expert Syst. Appl.*, vol. 210, Dec. 2022, Art. no. 118457, doi: 10.1016/j.eswa.2022.118457.

[15] S. Arya, N. Pratap, and K. Bhatia, "Future of face recognition: A review," *Proc. Comput. Sci.*, vol. 58, pp. 578–585, Jan. 2015, doi: 10.1016/j.procs.2015.08.076.

[16] D. Mamieva, A. B. Abdusalomov, M. Mukhiddinov, and T. K. Whangbo, "Improved face detection method via learning small faces on hard images based on a deep learning approach," *Sensors*, vol. 23, no. 1, p. 502, Jan. 2023, doi: 10.3390/s23010502.

[17] S. Hangaragi, T. Singh, and N. N, "Face detection and recognition using face mesh and deep neural network," *Proc. Comput. Sci.*, vol. 218, pp. 741–749, Jan. 2023, doi: 10.1016/j.procs.2023.01.054.

[18] D. Vaithiyanathan and M. Manigandan, "Real-time-based object recognition using SIFT algorithm," in *Proc. 2nd Int. Conf. Electr., Electron., Inf. Commun. Technol. (ICEEICT)*, Trichirappalli, India, Apr. 2023, pp. 1–5, doi: 10.1109/ICEEICT56924.2023.10157675.

[19] O. Arandjelovic and A. Zisserman, "Automatic face recognition for film character retrieval in feature-length films," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, San Diego, CA, USA, Jun. 2005, pp. 860–867, doi: 10.1109/CVPR.2005.81.

[20] J. Sivic, M. Everingham, and A. Zisserman, "Person spotting: Video shot retrieval for face sets," in *Proc. Int. Conf. Image Video Retr.*, Singapore. Berlin, Germany: Springer, Jul. 2005, pp. 226–236.

[21] M. Bertini, A. Del Bimbo, and W. Nunziati, "Automatic detection of player's identity in soccer videos using faces and text cues," in *Proc. 14th ACM Int. Conf. Multimedia*, Oct. 2006, pp. 663–666.

[22] D. Ozkan and P. Duygulu, "A graph based approach for naming faces in news photos," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2006, pp. 1477–1482.

[23] M. Everingham, J. Sivic, and A. Zisserman, "'Hello! My name is...buffy'—Automatic naming of characters in TV video," in *Proc. BMVC*, vol. 2, no. 4, p. 6, Sep. 2006.

[24] J. Looser, R. Grasset, and M. Billinghurst, "A 3D flexible and tangible magic lens in augmented reality," in *Proc. 6th IEEE ACM Int. Symp. Mixed Augmented Reality*, Nara, Japan, Nov. 2007, pp. 51–54, doi: 10.1109/ISMAR.2007.4538825.

[25] J. Han, D. Farin, and P. H. N. de With, "A real-time augmented-reality system for sports broadcast video enhancement," in *Proc. 15th ACM Int. Conf. Multimedia*, Sep. 2007, pp. 337–340.

[26] N. Inamoto and H. Saito, "Free viewpoint video synthesis and presentation of sporting events for mixed reality entertainment," in *Proc. ACM SIGCHI Int. Conf. Adv. Comput. Entertainment Technol.*, Sep. 2004, pp. 42–50.

[27] D. Liang, Y. Liu, Q. Huang, G. Zhu, S. Jiang, Z. Zhang, and W. Gao, "Video2Cartoon: Generating 3D cartoon from broadcast soccer video," in *Proc. 13th Annu. ACM Int. Conf. Multimedia*, Nov. 2005, pp. 217–218.

[28] X. Yu, X. Yan, T. T. Phuong Chi, and L. F. Cheong, "Inserting 3D projected virtual content into broadcast tennis video," in *Proc. 14th ACM Int. Conf. Multimedia*, Oct. 2006, pp. 619–622.

[29] K. Matsui, M. Iwase, M. Agata, T. T. Tanaka, and N. Ohnishi, "Soccer image sequence computed by a virtual camera," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit.*, Jun. 1998, pp. 860–865.

[30] Y. Tao, W. Zheng, W. Yang, G. Wang, and Q. Liao, "Frontal-centers guided face: Boosting face recognition by learning pose-invariant features," *IEEE Trans. Inf. Forensics Security*, vol. 17, pp. 2272–2283, Jan. 2022, doi: 10.1109/TIFS.2022.3183410.

[31] P. Li, L. Prieto, D. Mery, and P. J. Flynn, "On low-resolution face recognition in the wild: Comparisons and new techniques," *IEEE Trans. Inf. Forensics Security*, vol. 14, no. 8, pp. 2000–2012, Aug. 2019, doi: 10.1109/TIFS.2018.2890812.

[32] R. K. Tripathi and A. S. Jalal, "Novel local feature extraction for age invariant face recognition," *Expert Syst. Appl.*, vol. 175, Aug. 2021, Art. no. 114786, doi: 10.1016/j.eswa.2021.114786.

[33] I. Adjabi, A. Ouahabi, A. Benzaoui, and A. Taleb-Ahmed, "Past, present, and future of face recognition: A review," *Electronics*, vol. 9, no. 8, p. 1188, Jul. 2020, doi: 10.3390/electronics9081188.

[34] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj, and L. Song, "SphereFace: Deep hypersphere embedding for face recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jul. 2017, pp. 6738–6746.

[35] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *Proc. Brit. Mach. Vis. Assoc.*, 2015, pp. 1–12, doi: 10.5244/c.29.41.

[36] X. Zhu, Z. Lei, X. Liu, H. Shi, and S. Z. Li, "Face alignment across large poses: A 3D solution," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 146–155.

[37] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "DeepFace: Closing the gap to human-level performance in face verification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Columbus, OH, USA, Jun. 2014, pp. 1701–1708, doi: 10.1109/CVPR.2014.220.

[38] R. Girshick, "Fast R-CNN," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Santiago, Chile, Dec. 2015, pp. 1440–1448, doi: 10.1109/ICCV.2015.169.

[39] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.(CVPR)*, Boston, MA,USA, Jun. 2015, pp. 815–823, doi: 10.1109/CVPR.2015.7298682.

[40] A. Y. Oguzhan and M. Mihai, "Face detection and recognition in organic video: A comparative study for sport celebrities database," *IS&T Int. Symp. Electron. Imag. Sci. Technol.*, vol. 34, no. 10, pp. 355–356, Jan. 2022, doi: 10.2352/ei.2022.34.10.ipas-355.

[41] X. Zhang, X. Zhou, M. Lin, and J. Sun, "ShuffleNet: An extremely efficient convolutional neural network for mobile devices," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit.*, Salt Lake City, UT, USA, Jun. 2018, pp. 6848–6856, doi: 10.1109/CVPR.2018.00716.

[42] Z. Mahmood, T. Ali, S. Khattak, L. Hasan, and S. U. Khan, "Automatic player detection and identification for sports entertainment applications," *Pattern Anal. Appl.*, vol. 18, no. 4, pp. 971–982, Sep. 2014, doi: 10.1007/s10044-014-0416-4.

[43] M. U. Haq, A. Shahzad, Z. Mahmood, A. A. Shah, N. Muhammad, and T. Akram, "Boosting the face recognition performance of ensemble based LDA for pose, non-uniform illuminations, and low-resolution images," *Ksii Trans. Internet Inf. Syst.*, vol. 13, no. 6, Jun. 2019, doi: 10.3837/tiis.2019.06.021.

[44] J. Deng, J. Guo, J. Yang, N. Xue, I. Kotsia, and S. Zafeiriou, "ArcFace: Additive angular margin loss for deep face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 10, pp. 5962–5979, Oct. 2022, doi: 10.1109/TPAMI.2021.3087709.

[45] A. Chui, A. Patnaik, K. Ramesh, and L. Wang. (2019). *Capsule Networks and Face Recognition*. [Online]. Available: Lindawangg.github.io

[46] P. Terhörst, M. Ihlefeld, M. Huber, N. Damer, F. Kirchbuchner, K. Raja, and A. Kuijper, "QMagFace: Simple and accurate qualityaware face recognition," in *Proc. IEEE/CVF Winter Conf. Appl. Comput. Vis.*

(WACV), Waikoloa, HI, USA, Jan. 2023, pp. 3473–3483, doi: 10.1109/WACV56688.2023.00348.

[47] U. Gawande, P. Joshi, S. Ghatwai, S. Nemade, S. Balkothe, and N. Shrikhande, "Efficient attendance management system based on face recognition," in *ICT Systems and Sustainability*, vol. 321. Singapore: Springer, 2022, pp. 113–121, doi: 10.1007/978-981-16-5987-4_12.

[48] J. Zhang, Z. Su, and L. Liu, "Median pixel difference convolutional network for robust face recognition," 2022, *arXiv:2205.15867*.

[49] T.-H. Tsai, J.-X. Lu, X.-Y. Chou, and C.-Y. Wang, "Joint masked face recognition and temperature measurement system using convolutional neural networks," *Sensors*, vol. 23, no. 6, p. 2901, Mar. 2023, doi: 10.3390/s23062901.

[50] R. Golwalkar and N. Mehendale, "Masked-face recognition using deep metric learning and FaceMaskNet-21," *Int. J. Speech Technol.*, vol. 52, no. 11, pp. 13268–13279, Feb. 2022, doi: 10.1007/s10489-021-03150-3.

[51] K. H. Walse, R. V. Dharaskar, and V. M. Thakare, "A study of human activity recognition using AdaBoost classifiers on WISDM dataset," *Inst. Integrative Omics Appl. Biotechnol. J.*, vol. 7, no. 2, pp. 68–76, 2016.

[52] Y. Pang, Y. Yuan, X. Li, and J. Pan, "Efficient HOG human detection," *Signal Process.*, vol. 91, no. 4, pp. 773–781, Apr. 2011, doi: 10.1016/j.sigpro.2010.08.010.

[53] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137–1149, Jun. 2017, doi: 10.1109/TPAMI.2016.2577031.

[54] H. Qin, J. Yan, X. Li, and X. Hu, "Joint training of cascaded CNN for face detection," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, Jun. 2016, pp. 3456–3465, doi: 10.1109/CVPR.2016.376.

[55] W.-Y. Lu and M. Yang, "Face detection based on viola-jones algorithm applying composite features," in *Proc. Int. Conf. Robots Intell. Syst. (ICRIS)*, Haikou, China, Jun. 2019, pp. 82–85, doi: 10.1109/icris.2019.00029.

[56] H. Ouanan, M. Ouanan, and B. Aksasse, "Facial landmark localization: Past, present and future," in *Proc. 4th IEEE Int. Colloq. Inf. Sci. Technol. (CiSt)*, Tangier, Morocco, Oct. 2016, pp. 487–493, doi: 10.1109/CIST.2016.7805097.

[57] A. S. Almasoud, T. A. E. Eisa, F. N. Al-Wesabi, A. Elsafi, M. Al Duhayyim,I.Yaseen,M.A. Hamza,andA.Motwakel,"Parkinson's detection using RNN-Graph-LSTM with optimization based on speech signals," *Comput., Mater. Continua*, vol. 72, no. 1, pp. 871–886, Jan. 2022, doi: 10.32604/cmc.2022.024596.

[58] K. Almazrouei and T. Rabie, "Survey of face recognition using eigenfaces," in *Proc. 5th Int. Conf. Robot. Comput. Vis. (ICRCV)*, Nanjing, China, Sep. 2023, pp. 169–174, doi: 10.1109/icrcv59470.2023.10329085.

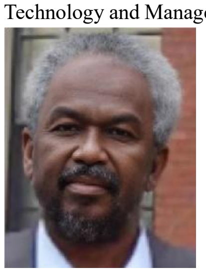
[59] *Face Recognition Using CNN*. Accessed: Oct. 27, 2023. [Online]. Available: https://www.kaggle.com/code/sifboudjellal/face-recognition-usingcnn

[60] K. Sreekala, C. P. D. Cyril, S. Neelakandan, S. Chandrasekaran, R. Walia, and E. O. Martinson, "Capsule network-based deep transfer learning model for face recognition," *Wireless Commun. Mobile Comput.*, vol. 2022, pp. 1–12, Jul. 2022, doi: 10.1155/2022/2086613.

MAHMOOD UL HAQ received the B.S. degree in electrical and electronics engineering and the M.S. degree in electrical engineering from COMSATS University Islamabad (CUI), Abbottabad Campus, Pakistan, in 2016 and 2018, respectively. He is currently pursuing the Ph.D. degree in computer system engineering with the University of Engineering and Technology Peshawar, Pakistan. He has published numerous manuscripts in reputable journals and conferences. He was a recipient of the Prime Minister Fee Refunding Scheme and the Government of Pakistan Scholarship Award for M.S. studies. His research expertise encompasses topics, such as object recognition, face recognition, image segmentation, and natural language processing.

MUHAMMAD ATHAR JAVED SETHI received the Bachelor of Science degree (Hons.) in computer information systems engineering and the Master of Science degree in computer systems engineering from the University of Engineering and Technology (UET) Peshawar, Pakistan, in 2004 and 2008, respectively, and the Ph.D. degree from the Department of Electrical and Electronic Engineering, Universiti Teknologi PETRONAS (UTP), Malaysia, in 2016. He is an Assistant Professor with the Department of Computer Systems Engineering, UET Peshawar. He has published numerous manuscripts in reputable journals, conferences, and books. He also wrote a book *Bio-Inspired Fault-Tolerant Algorithms for Network-on-Chip*. He is actively involved in technical program committees of various international conferences. He is serving as an Associate Editor for *EAI Endorsed Transactions on Context-Aware Systems and Applications* and *EAI Endorsed Transactions on Bioengineering and Bioinformatics*.

SADIQUE AHMAD (Member, IEEE) received the master's degree from the Department of Computer Sciences, IMSciences University, Peshawar, Pakistan, in 2015, and the Ph.D. degree from the Department of Computer Sciences and Technology, Beijing Institute of Technology, China, in 2019. He is working as a Researcher at Prince Sultan University, Riyadh. Previously, he worked on Cognitive Computing, Deep Cognitive Modeling for Students' Performance Prediction, and Cognitive Modeling in object detection using remote sensing images. Currently, he is focusing on Deep Cognitive Modeling for Trust Management in social cybersecurity, IoT, and blockchain technologies. He has achieved above 58 research articles in peer-review journals and conferences including top journals such as Information Sciences, Science China Information Sciences, Computational Intelligence and Neuroscience, Physica-A, and IEEE ACCESS. He has reviewed over 180 scientific research articles for various well-known journals, including Information Sciences, IEEE ACCESS JOURNAL, Knowledge-Based Systems, Education and Information Technologies, Information Technology and Management, ICEEST Conference, and ICONIP Conference.



MOHAMMED A. ELAFFENDI is currently a Professor of computer science with the Department of Computer Science, Prince Sultan University; the former Dean of CCIS, AIDE; the Rector, the Founder, and the Director of the Data Science Laboratory (EIAS); and the Founder and the Director of the Center of Excellence in Cybersecurity. His current research interests include data science, intelligent and cognitive systems, machine learning, and natural language processing.

MUHAMMAD ASIM received the M.S. degree in mathematics from the University of Peshawar, Peshawar, Pakistan, in 2013, the M.Phil. degree in mathematics from Kohat University of Science and Technology, Kohat, Pakistan, in 2016, and the Ph.D. degree in computer science and technology from Central South University, Changsha, China, in 2022. Currently, he is conducting a Postdoctoral Researcher with the EIAS Data Science Laboratory, College of Computer and Information Sciences, Prince Sultan University, Riyadh, Saudi Arabia. He has been awarded as an Outstanding International Graduate of Central South University, in 2022. His current research interests include artificial intelligence, computational intelligence techniques, cloud computing, edge computing, 5G/6G communication systems, and autonomous vehicles.