Wiley
Journal of Electrical and Computer Engineering
Volume 2025, Article ID 3305430, 11 pages
https://doi.org/10.1155/jece/3305430

WILEY

Research Article

Hybrid Reinforcement Learning With Optimized SARSA for Improved Face Recognition Systems

Anil Kumar Yadav[1], **Purushottam Sharma**[2], **Xiaochun Cheng**[3] **and Nirmal Kumar Gupta**[4]

[1]*School of Computing & Artificial Intelligence, VIT Bhopal University, Sehore, Madhya Pradesh, India*
[2]*School of Computer Science & Engineering, Galgotias University, Greater Noida, India*
[3]*Computer Science Department, Swansea University, Bay Campus, Fabian Way, Swansea SA1 8EN, Wales, UK*
[4]*Department of Information Technology, Manipal University, Jaipur, Rajasthan, India*

Correspondence should be addressed to Purushottam Sharma; purushottam@galgotiasuniversity.edu.in and Xiaochun Cheng; xiaochun.cheng@swansea.ac.uk

Received 27 January 2025; Revised 28 July 2025; Accepted 19 August 2025

Guest Editor: Arpan Hazra

Face recognition is a key technique in modern image processing, yet it faces challenges such as achieving high accuracy, reducing computational time, and optimizing memory usage. This research proposes a hybrid model that integrates an enhanced State-Action-Reward-State-Action (SARSA) reinforcement learning (RL) framework to address these challenges in face recognition tasks. The model utilizes principal component analysis (PCA) for dimensionality reduction and initial feature extraction, followed by a SARSA-based online Q-learning algorithm to refine classification accuracy and resolve state overlap issues. During training, facial datasets are processed to extract critical features, and a state-action value table is constructed to guide decision-making during testing. This reinforcement-driven learning enables the system to dynamically update its policy based on the most rewarding actions, improving adaptability and performance. Experimental results demonstrate that the proposed approach enhanced traditional models in terms of recognition accuracy, classification efficiency, and training speed. Integrating optimized feature selection and policy learning mechanisms makes the model a promising solution for real-time and resource-efficient face recognition applications.

Keywords: classification; face recognition; feature extraction; learning agent; reward; SARSA

1. Introduction

The State-Action-Reward-State-Action (SARSA) algorithm is a pivotal reinforcement learning (RL) method designed to adjust action policies for specific environments. Extensive face datasets are essential to achieve a generalized SARSA model for facial recognition with improved accuracy and classification. SARSA is a decision-making system that stores state-action values in a lookup table. In the context of implementing SARSA for a face recognition task using an efficient hybrid model, the **states** can be defined as the extracted feature representations of input facial images, the **actions** correspond to selecting a potential identity or class label for the given face, and the **rewards** are determined based on the accuracy of the prediction—assigning a positive value for correct recognition and a negative or zero value for incorrect matches.

Face recognition is an inherent human skill crucial throughout a person's lifetime and plays a vital role in distinguishing between individuals. Modern computers leverage this ability across various fields for image recognition and identification. Several organizations are currently advancing Face Recognition Systems (FRSs). RL has been applied to recognize terrorist individuals, and smart city initiatives utilize facial recognition for identifying people and objects at traffic signals [1]. E-administration uses biometric facial detection for online attendance [2].

FRS are installed in diverse locations such as markets, railways, banks, airports, bus stops, and traffic areas for automatic facial analysis and identification. In healthcare, FRS helps detect individuals with depression and count people in crowds. Gradually, these systems have been used to gather information about the number of people entering offices. During pandemics, facial recognition has replaced fingerprint-based attendance systems for automatic attendance counting. FRSs are widely used for recognizing or verifying individuals for specific tasks [1–3]. When illegal activities occur, FRS acts as a safeguard to prevent fraudulent individuals from posing as legitimate users. They prevent unauthorized individuals from attempting to create problems in place of authentic users or objects.

Major challenges in FRS involve individual identification accuracy and reliability. In some instances, FRS effectively detects and recognizes individuals, minimizing the impact of false positives. If a face appears suspicious, the system may lock the device or system, requiring a PIN or password to prevent unauthorized access. Marko's decision process deals with a complex framework for making uncertain decisions with unbounded rewards, costs, and state-dependent discount factors [4]. Huge datasets can likely contain duplicate images and risk misuse [5]. Face detection and feature extraction measurement methods are used for accurate image analysis and deep learning for complex feature extraction [6]. Surveillance tactics are plans and techniques to monitor individuals, groups, or activities for security, intelligence, or law enforcement. The methods used vary according to goals, technology at hand, and applicable laws [7]. The unpredictability and diversity of real-world settings make face identification in unconstrained situations [8]. Using a nonconstant discount factor in discrete-time control introduces additional complexity and flexibility to traditional control frameworks [9]. To increase the effectiveness, generalization, or performance of machine learning models, pruning training sets for learning object categories involves reducing the size or improving the quality of the training dataset [10].

FRS is a manufactured technique for distinguishing or verifying a person from a digital image. The system contrasts the input image with stored images, displaying specific facial features. FRS are typically utilized in security frameworks and e-administration, incorporating biometrics like fingerprints and iris recognition. Early face recognition algorithms focused on extracting significant facial features by analyzing points across the face. The basic process involves breaking the input image into frames and converting them to grayscale. Efficient face recognition depends on training the system to recognize features such as the eyes, nose, and head [11–15].

Face recognition can be categorized into two main techniques. The first involves describing facial features by extracting them and defining a comprehensive view of the facial problem. The second method involves a complete view by removing the basic properties of the problem, along with training samples. Principal component analysis (PCA) is widely used for feature extraction by reducing the dimensionality of datasets. It transforms high-dimensional data into a smaller space using eigenvectors and covariance matrices. PCA simplifies the problem by grouping images and reducing dimensionality, though it struggles with low-intensity, noisy data during training.

FRS are crucial for security, providing layered protection for online and offline data [16–20]. Biometrics, including fingerprints and facial detection, determine whether individuals are authorized. Biometric frameworks integrate personal identifiers such as fingerprints, faces, irises, retinas, hand geometry, voices, and signatures. Consequently, face recognition has become increasingly integral to biometric mechanisms [21–24]. Early facial recognition systems were based on spot detection, while modern systems use zoom lenses for enhanced nose image recognition.

In 2001, the framework for FRS was established to facilitate various observations [25]. Initially, these systems were used in trials and yielded unsatisfactory results. However, extensive research eventually made face recognition more applicable in fields like navigation [26]. An example is automated systems that count individuals before they enter through smart gates.

Google has also explored integrating facial recognition for item searches and continues to develop this technology. In the contemporary context, face recognition plays a crucial role in data security, authentication, and verifying account holders in the banking sector, alongside traditional methods like PAN and other cards. Similarly, fingerprints are also extensively used for various types of verification [27].

Facial recognition systems are employed for authentication and electronic Know Your Customer (e-KYC) processes across various government schemes. For instance, in 2018, Patricia and Nancy Velasco utilized face recognition to count the number of people [28] on public transportation. Their system involved a single camera capturing input images, which were then processed using artificial neural networks. In 2021, Zhang introduced deep learning techniques for image classification, advancing the field further.

In 2020, Youhui Tian discussed the application of neural networks for image processing with a dual optimization model. This model aimed to integrate and optimize image processing tasks, convergence, and connection processes. In 2021, Yadav et al. described a learning approach where data was stored in a state-action Q-table, with positive or negative rew during each movement episode. This method evaluated discounting in specific episodes through trial-and-error interactions.

Further advancements include a 2022 study by Yadav et al., which introduced a hybrid machine learning (HML) model for face recognition using support vector machines (SVM) combined with PCA. This model, SVM-PCA, aims to improve face detection [29, 30]. Also, in 2022, Wang and Peng explored the application of RL for AI-based resource allocation in integrated sensing and communication systems.

1.1. The Proposed Plan. SARSA methods are applied in face data aggregation, training, testing, and facial prediction tasks. Experimental results indicate that when applied to the Yahoo the MegaFace (MF) dataset, the SARSA model effectively detects faces and determines whether they overlap collaboratively. SARSA excels in face prediction by improving accuracy, achieving higher classification rates, and reducing training time.

1.2. Issues With FRSs. While feature extraction and classification are crucial components of FRS, several factors can influence their performance. FRS must account for variations in illumination, poses, expressions, backgrounds, and occlusions, as well as rotation, scaling, and translation (RST). Table 1 outlines the factors affecting feature extraction in FRS projects. Systems commonly employed for feature extraction include mean, standard deviation, kurtosis, histograms, PCA, independent component analysis (ICA), and linear discriminant analysis (LDA).

2. Recognition Problems

Face identification and recognition are pivotal components of image processing techniques. Despite advances, computer vision and pattern recognition still struggle to achieve reliable face recognition. Key challenges include actual facial recognition, replacing fingerprint-based biometric attendance during pandemic situations, and achieving better accuracy, higher classification rates, and reduced training time.

The literature describes FRS focused on feature extraction and selecting overlapping subsets from each image. These challenges have prompted the current study, which introduces a detection system using two methodologies. First, a training model for feature extraction is developed using PCA to reduce image dimensions and extract features. Second, facial recognition utilizes the SARSA algorithm to minimize errors and address the overlapping problem.

2.1. Model Configuration. Facial recognition systems rely on image and video classifiers, where incoming samples and categories coincide within a high-dimensional region. The likelihood of an instance belonging to a particular class is maximized when the instance's similarity to the positive class is high while minimizing its similarity to overlapping classes. In this implementation, MATLAB serves as the programming language of choice. Both training and testing phases demand substantial image datasets and high-processing equipment, typically GPUs. Moreover, access to face databases is crucial for training and deploying classification and recognition models. Given the lack of comprehensive databases for face recognition, it becomes necessary to develop custom databases for training purposes. To address this need, the MF dataset, sourced from Yahoo's dataset, provides a substantial repository of facial images.

2.2. Proposed Model. The illustrated recognition framework, detailed in Figures 1 and 2, integrates a combined learning model for face detection utilizing the SARSA RL method. Utilizing two distinct databases is imperative to operating the facial recognition system effectively. One serves the purpose of training classification, while the other facilitates feature extraction for recognition. The input facial dataset is sourced from MF and obtained from Yahoo's dataset.

2.2.1. First Stage. In the initial stages, the focus lies on the segmentation and feature extraction of one or more faces. The objective is to determine whether a given feature overlaps within an image. This is achieved using a SARSA-based learning face recognition model, incorporating width, height, and location data. Additionally, PCA is employed for feature reduction when dealing with overlapping features within the facial dataset. This model is comprehensively delineated by the feature extraction and classification process, elucidated in Figure 1.

The hybrid face recognition approach illustrated in Figure 2 integrates two key algorithms to enhance overall performance. Specifically, it employs the SARSA model for classification and applies data preprocessing and feature engineering techniques to support face recognition. To reduce processing time, the input facial images are first resized. Feature extraction, as depicted in Figure 1, is then used to obtain key attributes from the images. The primary goal of this research is to boost the accuracy of facial recognition while keeping computational requirements low.

2.2.2. Second Stage. Moving on to the second stage, the operational procedure is depicted in Figure 2. During classifier training, the system can discern whether faces overlap or not. Similar to the previous stage, the facial dataset is obtained from MF and sourced from Yahoo's dataset. PCA is the training classifier for feature extraction and selecting overlapping subsets from each image. Initially, PCA is employed to reduce image dimensions and extract features. The well-defined steps of this process are illustrated in Figure 2. The fundamental premise of this learning technique lies in its ability to transform lower-level features (pixels) into higher-level features. PCA is utilized to reduce features during training for the first classifier. In contrast, the second classifier, SVM, employs the softmax function for feature classification, providing output alongside accuracy.

2.2.2.1. Methods. The proposed hybrid model operates through multiple stages of image noise reduction, beginning with resizing the images to half of their original dimensions. It then utilizes PCA-based feature extraction for classification, followed by SVM classification-based facial prediction. The overall workflow of the system is depicted in Figure 2.

2.3. Training Algorithm for Face Recognition. Moving forward to the training algorithm for face recognition, image preprocessing is conducted across various detection phases, as depicted in Figure 3. The recognition process is predicated

Table 1: Face identification and feature extraction face significant problems.

Factors	Clarification
Illumination	Different lighting conditions produce the illumination variety, which is expected to have a greater visual difference than the difference brought on by various personalities
Pose	Unique edges and places are in charge of posing variability during image collection. This variation modifies the spatial relationships between facial components and causes real distortion in face recognition algorithms like eigenfaces
Expression	Depending on their disposition and temper, people have different facial expressions. The range of expressions affects both the spatial relationship and the facial component
RST variation	The diversity in the picture-acquiring process also contributes to the RST (rotation, scaling, and translation) variety. It causes face recognition and discovery issues and may necessitate an extensive localization search across all conceivable RST factors
Cluttering	We must also consider how backgrounds and circumstances affect the people in the images. Face patches, which include this foundation, also constrain the execution of face recognition algorithms and impact face identification accuracy
Occlusion	Due to the variety, some facial features are kept a secret. Occultation is the term for it

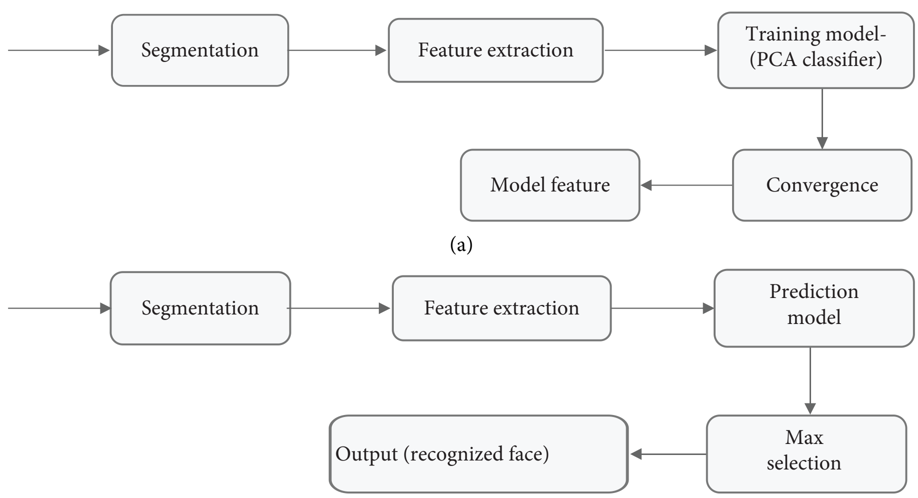

Figure 1: Structure of the hybrid feature extraction model. (a) Training. (b) Testing.

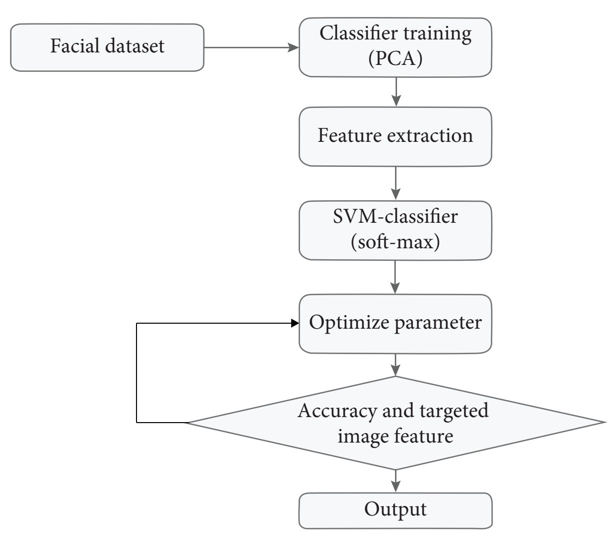

Figure 2: The hybrid PCA and SVM model-based face recognition framework.

on several steps, including image segmentation, feature extraction, and a SARSA-based predictive model for overlapping feature prediction. The methodology of face recognition is outlined through five distinct stages. Initially, detection is performed on the given data sample and converted into a grayscale image. Subsequently, preprocessing is executed to standardize the image. Feature extraction constitutes the third step, followed by the learning stage to train the framework. The final stage involves classification, wherein pertinent information is gathered.

In this approach, individuals are recognized through images, and following numerous episodes, our recognition model is trained. Therefore, the recognition model can be effectively trained if the dataset is available.

The database comprises a set of observations from MF (Yahoo's dataset) to detect whether faces overlap. The MF dataset includes over 1 million images from more than 690,000 persons and is primarily designed for large-scale face recognition challenges. As created from Flickr photos, the dataset contains wide pose, expression, lighting, and occlusion variations. Hence, thorough preprocessing is critical for reliable model performance. MF dataset requires several preprocessing steps before model training or evaluation to ensure consistent and accurate facial recognition results. These steps are face detection, face alignment, image resizing, normalization, duplicate and corrupted image removal, and filtering.

Various techniques are available to accomplish face recognition, including histogram techniques, the multiresolution approach, the information theory approach, and the eigenface approach. Figure 3 serves as a schematic representation of the SARSA model. The face recognition process integrates feature extraction, grayscale conversion into different segments, and facial recognition. Combined techniques prove to be more effective than individual methods.

Facial recognition techniques facilitate the recognition of images associated with real individuals. In feature extraction, faces are initially converted into grayscale contact blocks, considering width and height. The number of blocks retrieved from the observed sample and the overlapped feature is considered.

Algorithm 1 performs data preprocessing and feature engineering for face recognition, structured in alignment with the methodological framework discussed in RL with Optimized SARSA for face recognition.

2.4. Implementation Costs. The demand for high-performance equipment for deep learning and data science evaluation is considerable today. Several open-source options are accessible, including Google Colab, which Google provides. Our implementation utilizes MATLAB alongside various components, including:

- 8 GB RAM
- 40 GB disk space
- GPU for dataset processing
- 50 epochs
- 8 samples per face

Conventional computers are employed for MATLAB programming to facilitate face recognition and analyze parameters such as recognition rate, training time, and classification accuracy.

The SARSA algorithm, a model-free RL method, can be adapted to face recognition by defining appropriate states, actions, rewards, and transitions that guide the learning process toward accurate identification.

In face recognition, a *state* represents the extracted feature vector of a facial image. These features can be obtained through techniques like PCA, capturing essential facial characteristics. An *action* corresponds to the classification decision made by the agent; this means assigning the input face to one of the known identity labels in the database or classifying it as unknown.

The *reward* function is designed to reinforce correct decisions. A positive reward (e.g., +1) is given when the predicted label matches the identity. A negative reward (e.g., 1) is issued for incorrect predictions.

A *state transition* occurs when the system processes a new facial image or moves to the next step in feature refinement. The transition reflects how the agent updates its policy after receiving feedback (reward) based on the previous action.

SARSA updates its Q-values based on the tuple (s, a, r, s′, a′). Where s is the current state, a is the action, r is the received reward, s is the next state, and a′ is the next action.

3. Results and Analysis

The experiments outline the training and testing processes in detail in separate tables. The model is trained using a system of rewards and penalties: a penalty represents a false negative detected by the model, while a false positive indicates overlap in the given sample. To evaluate the facial data samples in the context of the problem, rewards are used in conjunction with PCA. The proposed model demonstrates improved accuracy in face detection, reduced testing time, and faster classification, as presented in Table 2.

3.1. Training With Face Classification. Face recognition is integral to our daily lives, with extensive research dedicated to enhancing image prediction to enable computers and machines to perform this task with precision and efficiency. Computerized face recognition has diverse applications, including biometric verification, surveillance, feature database indexing, and search functionalities. In biometric verification, various face recognition methods exist, such as the multiresolution approach, the information theory approach, the eigenface approach, and techniques primarily based on histograms. The process of face photograph enrolment, verification, and identification is illustrated in Figure 4. During the training phase, many standard datasets are used for self-training and execution. Figure 5 presents a graph showing the training over 60 epochs, achieving an accuracy of 97.67% for the entire training dataset.

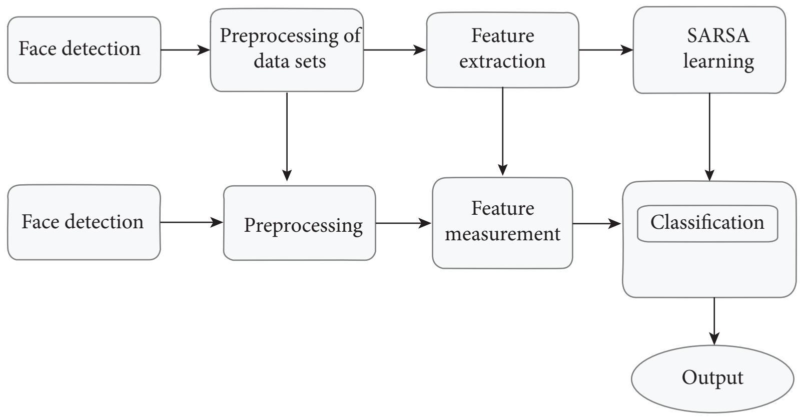

Figure 3: Training of a facial recognition model.

1. **Input**: facial data
2. **Output**: generate an accurate image
3. **Procedure**: Face recognition
   i. Extract feature
   ii. converted grey into different segments
   iii. Optimize with SoftMax () with a range between (0, 1)
4. **for** each segment
   i. face converted grey into contact blocks along with width and height
   ii. The number of blocks recovered from the observed sample and the overlapped feature.
5. Reshape image
6. **End Procedure**

Algorithm 1: Data preprocessing and feature engineering for face recognition.

Table 2: Training classifier parameters.

Parameter	Value
Learning rate	0.99
Epochs	50
Sample per face	8

The classification mechanism in the proposed approach is realized through Algorithm 2 integrated with the SARSA model, formulated on the principles described in hybrid RL with optimized SARSA for improved FRS.

3.1.1. Model Implementation. We extract facial features using a publicly accessible face matcher. We employ the scale hyperparameter and softmax loss to pre-train the user face recognition model on the server. The SARSA setup is used for training each user, utilizing the discrete cosine transform with a margin of 0.98 and minimal error loss. The model has trained over 50 episodes and 60 faces, with a discount rate 0.01. We use a low learning rate to ensure that training in a group context avoids overlapping paths in the given problem.

3.2. Testing With Facial Recognition. During the testing phase, 25% of the data are allocated for validation, while the remaining 75% is used to evaluate the overall recognition model. The graph in Figure 6 illustrates the testing results over nine epochs, achieving an accuracy of 97.88% for the entire training dataset.

To obtain training results, including accuracy, training time, testing time, detection, and classification rate, refer to Table 3. During the training sessions, facial data samples were varied: 10, 15, 20, 30, and 40 samples per session, with six examples per picture used in all experiments Table 3. Test 4 achieved an accuracy of 97.67%, surpassing all other analyses conducted on 180 images. It recorded a training time of 1.1212 s, notably longer than others. Testing time was as low as 0.0021 s in the performed tests. Notably, increasing the number of facial samples improved recognition rates, albeit at the cost of increased training time, while testing time decreased.

Various training methods, such as SARSA, SVM, and PCA, have been used for classification and recognition. Similarly, feature extraction is used to reduce overlapping features.

3.3. A Comparison of Facial Recognition Methods Using Various Episodes and Approaches. First, we presented the performance of different algorithms along with facial recognition toward various techniques under different episodes. The facial recognition rate evaluation is shown in Table 4 for HML [45], SARSA, PCA, and SVM algorithms. The formula used is a measure of recognition efficiency in

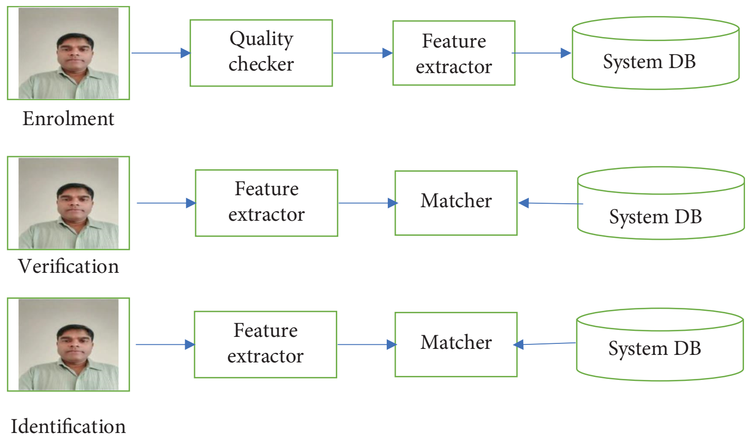

Figure 4: Classification model of face images.

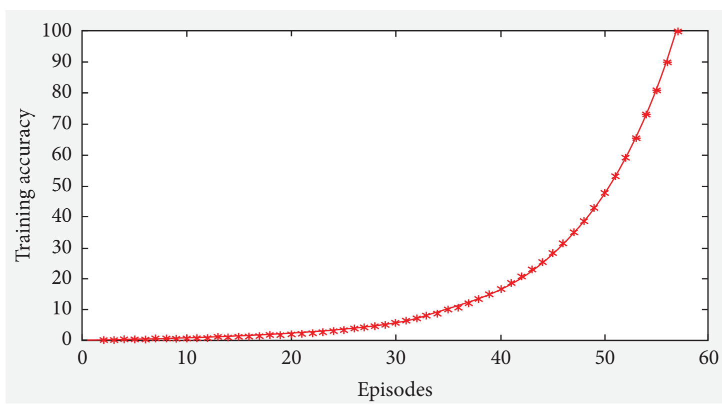

Figure 5: Training accuracy graph of the facial recognition model.

1. **Input:** Required to run Algorithm 1
2. **Output:** Identified an accurate image
3. **Procedure:** Generate the SARSA with the PCA model
   a. Initialize the input feature
   b. Quality checker
   c. Add matcher
   d. Reduce the overlapping
   e. Compile the model with SoftMax
4. Model 1 ← model compile
5. for < number of epochs > do
   The train and test data from the model are used to generate accuracy
6. end for
7. end procedure

Algorithm 2: SARSA MODEL for classification.

the given problem. Table Update with $R1 = R\gamma^{(x1-i)}$ where R1 is an immediate reward, R is the total reward, $x1$ is the current step, and I is the total step.

Figure 7 shows a comparison of total reward under the number of episodes for various algorithms, SARSA, HML [45], PCA, and SVM, with increased episodes. Figure 8 compares facial recognition rates for various algorithms under different numbers of episodes. We found that SARSA learning performed better on convergence than HML, PCA, and SVM algorithms. Therefore, the overall performance of SARSA learning is better than other algorithms.

3.4. Performance Comparison of Various Learning Techniques. Performance comparison of given facial data sample using PCA and SARSA and other machine learning techniques with SVM and DCT. We have made

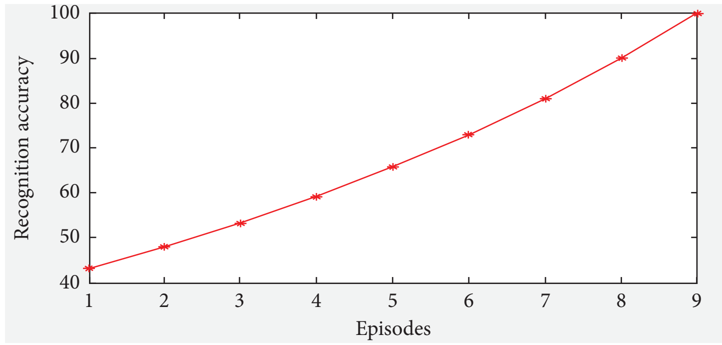

Figure 6: Recognition accuracy graph of the facial recognition model.

Table 3: Evaluation metrics for different possible solutions.

Features	Solution-1	Solution-2	Solution-3	Solution-4
Faces counts	14	25	34.5	45.55
Sample per face	8	8	8	8
Accuracy	97.27	90	95.67	97.67
Training time	2.35	1.202	2.01135	3.7025
Testing time	0.0456	0.1236	0.0021	0.0009
Detection rate	0.0255	0.10221	0.0246	0.0007
Classification rate	2.1325	1.0236	2.2323	3.445

Table 4: Comparison table between different recognition under total reward and episodes for SARSA and other algorithms.

Rate of recognition (R)	Episode 1000	2000	4000	6000	8000	10,000
R1	95.8	73.34	74.5	96.7	97.21	98.88
R2	71.59	72.43	64.62	77.7	82.43	84.93
R3	9.91	74.62	94.62	74.51	85.63	95.76
R4	5.62	82.36	96.32	822.36	86.82	83.93
R5	96.67	96.7	56.38	95.42	94.36	98.99
Average (%)	80.70	81.84	76.74	86.42	89.72	94.58

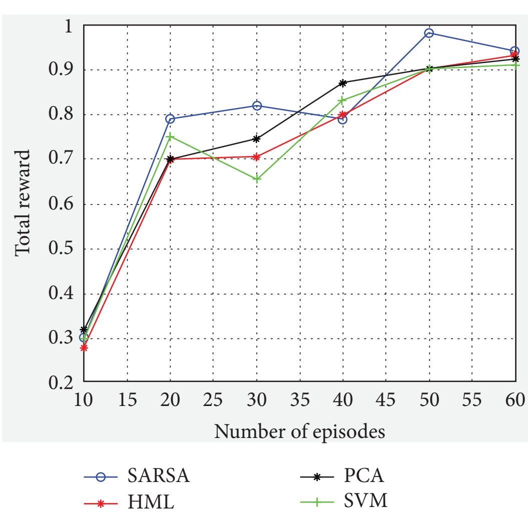

Figure 7: Comparison of the total reward of different methods.

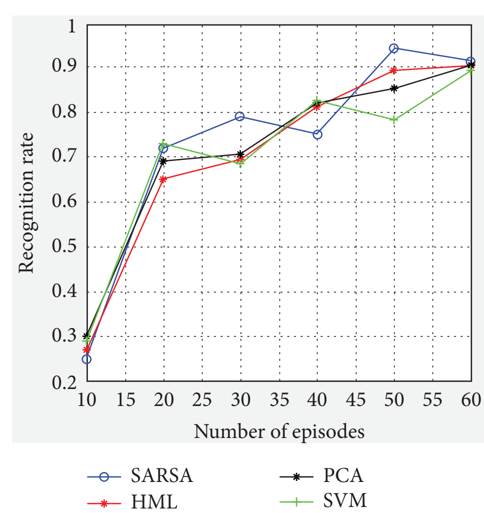

Figure 8: Comparison of the rate of recognition of different numbers of episodes.

four observations using the proposed model. Table 5 shows the experimental results. We found that the highest precision for the PCA-SARSA blend was 93.34% in experiment 3. The least training time was found in experiment No. 2. For the PCA with SARSA learning, the accuracy was 98.30% in experiment No. 4. The training time was 2.3121 in the second episode. Performance

Table 5: Accuracy comparison between various learning techniques.

Features	Solution-1		Solution-2		Solution-3		Solution-4	
Algorithms	**PCA-SARSA**	**PCA-SVM-**	**PCA SARSA**	**PCA-SVM**	**PCA SARSA**	**SVM-PCA**	**SARSA-PCA**	**PCA-SVM**
No. of faces	10	10	30	30	50	50	60	60
Sample per face	8	8	8	8	8	8	8	8
Accuracy	90.91	93.67	80.34	94.91	93.34	96.23	98.3	93.44
Training time	1.8324	2.5021	0.835	1.3101	2.3756	2.1421	2.744	3.8121
Testing time	0.812	0.074	0.3632	0.3006	0.823	0.0412	0.7862	0.0813
Detection rate	0.713	0.065	0.2513	0.30012	0.712	0.0311	0.6431	0.0741
Classification rate	1.2853	2.146	0.581	2.6512	2.372	2.802	2.231	2.855

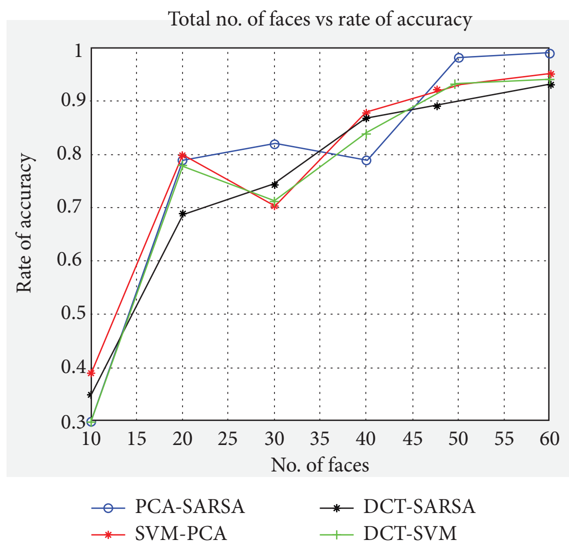

Figure 9: Total number of faces vs. rate of accuracy.

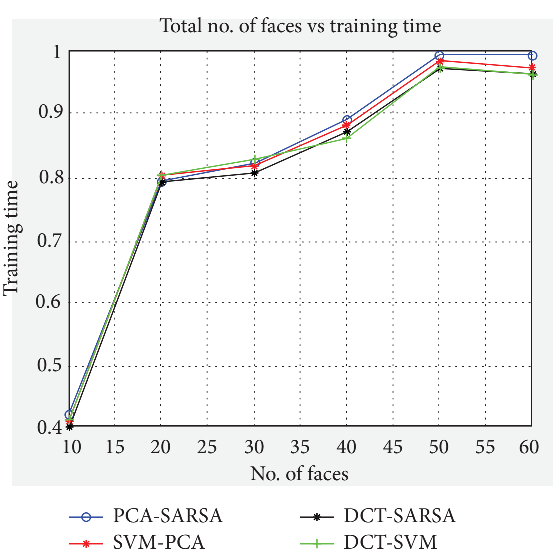

Figure 10: Total number of faces vs. training time.

comparison for better results of the proposed model is shown through Figures 9, 10, and 11 in the context of accuracy, training, and classification time.

Figure 9 shows a relative graph between the rate of accuracy and the number of faces, Figure 10 shows a relative graph between training time and the number of faces, and

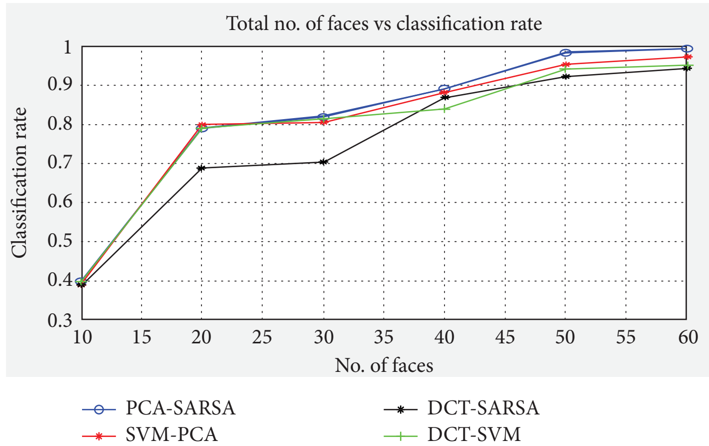

Figure 11: Total number of faces vs. classification rate.

Figure 11 shows a relative graph between classification rate and the number of faces. The accuracy of PCA with SARSA learning is relatively high at 98.83%.

Performance was evaluated using key metrics such as the rate of recognition, detection rate, and classification rate to assess the proposed approach's effectiveness comprehensively.

SARSA provides a theoretically grounded and practically effective solution for face recognition tasks where safety, uncertainty, real-time learning, and stability are critical. Its on-policy nature ensures alignment between education and decision-making, leading to better adaptation in real-world environments with high variability and limited labeled data. SARSA evaluates and improves the same policy used to make decisions, which makes it inherently safer and more stable during training. In face recognition, especially in real-time or sensitive applications such as surveillance or biometric authentication, making risky decisions (like misclassifying a person) can have serious consequences. SARSA's conservative update mechanism ensures that learning aligns with the current behavior of the agent, reducing the chances of overconfident misclassifications.

4. Conclusion

In this study, we tackle the challenge of developing a hybrid model for face recognition using SARSA. We introduce a SARSA learning framework to predict facial data in real-world scenarios, achieving an enhanced face recognition accuracy of up to 98.83%. Our approach significantly reduces both training time and classification duration. For applications in authentication and security, this model proves more effective than other biometric methods such as iris, fingerprint, and retina scans. Integrating collaborative learning techniques, SARSA, PCA, and DCT yields efficient outcomes, demonstrating superior accuracy, reduced training time, and enhanced classification rates.

Compared with DCT-SVM, the proposed combined learning model achieves lower time requirements while maintaining higher accuracy than earlier algorithms. DCT is a feature extractor and parameter optimizer within our integrated learning framework. Future research will explore enhancements to face recognition using ensemble learning, discrete wavelet transforms (DWT) paired with SVM, and PCA. Additionally, we plan to integrate intelligent algorithms like Q learning and ant colony optimization to improve feature selection from images and videos, particularly focusing on selfie images.

Data Availability Statement

The studies are conducted on already available data and materials for which consent is not required.

Ethics Statement

This is an observational study. This research includes no involvement of humans or animals, so no ethical approval is required.

Disclosure

All images within this manuscript are original works created by the authors unless otherwise stated. The authors retain all copyrights to these images.

Conflicts of Interest

The authors declare no conflicts of interest.

Funding

The authors was supported by the UKRI EPSRC Grant funded Doctoral Training Centre at Swansea University, through project RS718. The authors was also supported by the UKRI EPSRC Grant EP/W020408/1.

References

[1] S. Augenstein, H. B. McMahan, D. Ramage, et al., "Generative Models for Effective ML on Private, Decentralized Datasets" (2019).

[2] O. Vinyals, I. Babuschkin, W. M. Czarnecki, et al., "Grandmaster Level in Starcraft II Using Multi-Agent Reinforcement

Learning," *Nature* 575, no. 7782 (2019): 350–354, https://doi.org/10.1038/s41586-019-1724-z.

[3] D. Bertsekas, "Reinforcement Learning and Optimal Control," *Athena Scientific* 1 (2019): 1–315.

[4] Q. Wei and X. Guo, "Markov Decision Processes With State-Dependent Discount Factors and Unbounded Rewards/Costs," *Operations Research Letters* no. 5 (2011): 369–374, https://doi.org/10.1016/j.orl.2011.06.014.

[5] Y. Guo, L. Zhang, Y. Hu, X. He, and J. Gao, "Ms-celeb-1m: A Dataset and Benchmark for Large-Scale Face Recognition," in *Computer Vision–ECCV:14th European Conference* (Amsterdam, The Netherlands, 2016), 87–102.

[6] K. Bonsor and R. Johnson, "How Facial Recognition Systems Work, Howstuffworks. Com Np" (2001).

[7] M. J. Kumar, "Editorial Commentary: Face Recognition by Machines: Is It an Effective Surveillance Tactic?" *IETE Technical Review* 30, no. 2 (2013): 93–94, https://doi.org/10.4103/0256-4602.110546.

[8] G. B. Huang, M. Mattar, T. Berg, and E. Learned-Miller, "Labeled Faces in the Wild: a Database for Studying Face Recognition in Unconstrained Environments," in *Workshop on Faces in Real-Life Images: Detection, Alignment, and Recognition* (2008), 1–14.

[9] H. Jasso-Fuentes, J. L. Menaldi, and T. Prieto-Rumeau, "Discrete-Time Control with Non-constant Discount Factor," *Mathematical Methods of Operations Research* 92, no. 2 (2020): 377–399, https://doi.org/10.1007/s00186-020-00716-8.

[10] A. Angelova, Y. Abu-Mostafa, and P. Perona, "Pruning Training Sets for Learning of Object Categories," in *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, no. 1 (2005), 494–501, https://doi.org/10.1109/cvpr.2005.283.

[11] S. Pitis, "Rethinking the Discount Factor in Reinforcement Learning: A Decision Theoretic Approach," in *Proceedings of the AAAI Conference on Artificial Intelligence* (2019), 7949–7956, https://doi.org/10.1609/aaai.v33i01.33017949.

[12] Y. Lu, N. Zeng, Y. Liu, and N. Zhang, "A Hybrid Wavelet Neural Network and Switching Particle Swarm Optimization Algorithm for Face Direction Recognition," *Neurocomputing* 155 (2015): 219–224, https://doi.org/10.1016/j.neucom.2014.12.026.

[13] M. M. Kasar, D. Bhattacharyya, and T. H. Kim, "Face Recognition Using Neural Network: A Review," *International Journal of Security and Its Applications* 10, no. 3 (2016): 81–100, https://doi.org/10.14257/ijsia.2016.10.3.08.

[14] L. Shen, L. Bai, and Z. Ji, "A SVM Face Recognition Method Based on Optimized Gabor Features," in *Advances in Visual Information Systems: 9th International Conference* (2007), 165–174.

[15] A. Aydın and E. Sürer, "Using Generative Adversarial Nets on Atari Games for Feature Extraction in Deep Reinforcement Learning," in *28th Signal Processing and Communications Applications Conference* (2020), 1–4, https://doi.org/10.1109/siu49456.2020.9302454.

[16] B. F. Klare, B. Klein, and E. Taborsky, "Pushing the Frontiers of Unconstrained Face Detection and Recognition," in *IEEE Conference on Computer Vision and Pattern Recognition* (2015).

[17] Z. Liu, P. Luo, X. Wang, and X. Tang, "Deep Learning Face Attributes in the Wild," in *Proceedings of the IEEE International Conference on Computer Vision* (2015), 3730–3738, https://doi.org/10.1109/iccv.2015.425.

[18] M. Hasanbeig, A. Abate, and D. Kroening, "Cautious Reinforcement Learning With Logical Constraints" (2020).

[19] B. Maze, J. Adams, J. A. Duncan, N. Kalka, T. Miller, and C. Otto, "Iarpa Janus Benchmark-c: Face Dataset and Protocol," in *IEEE International Conference on Biometrics* (2018), 158–165.

[20] B. McMahan, E. Moore, D. Ramage, S. Hampson, and B. A. Arcas, "Communication-Efficient Learning of Deep Networks From Decentralized Data," in *Proceedings of the 20th International Conference on Artificial Intelligence and Statistics* (2017), 1273–1282.

[21] H. Wang, Y. Wang, and Z Zhou, "Cosface: Large Margin Cosine Loss for Deep Face Recognition," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2018), 5265–5274, https://doi.org/10.1109/cvpr.2018.00552.

[22] X. Zhang, Y. Guan, S. Wang, J. Liang, and T. Quan, "Face Recognition in Color Images Using Principal Component Analysis and Fuzzy Support Vector Machines," in *IEEE 1st International Symposium on Systems and Control in Aerospace and Astronautics* (2006), 1–4.

[23] D. Yi, Z. Lei, S. Liao, and S. Z. Li, "Learning Face Representation from Scratch" (2014).

[24] H. Wang, S. Yang, and W. Liao, "An Improved PCA Face Recognition Algorithm Based on the Discrete Wavelet Transform and the Support Vector Machines," in *International Conference on Computational Intelligence and Security Workshops* (2007), 308–311, https://doi.org/10.1109/cisw.2007.4425498.

[25] X. Li, J. Ma, and S. Li, "A Novel Face Recognition Method Based on Principal Component Analysis and Kernel Partial Least Squares," in *IEEE International Conference on Robotics and Biomimetics* (2007), 1773–1777.

[26] A. Aggarwal, M. Alshehri, M. Kumar, P. Sharma, O. Alfarraj, and V. Deep, "Principal Component Analysis, Hidden Markov Model, and Artificial Neural Network Inspired Techniques to Recognize Faces," *Concurrency and Computation: Practice and Experience* 33, no. 9 (2021): e6157, https://doi.org/10.1002/cpe.6157.

[27] R. Rioux and T. Simon, "Class Redundancy for Face Recognition by SVM," in *50th International Symposium*, no. 1 (2008), 117–120.

[28] P. Sharma, M. Alshehri, and R. Sharma, "Activities Tracking by Smartphone and Smartwatch Biometric Sensors Using Fuzzy Set Theory," *Multimedia Tools and Applications* 82, no. 2 (2023): 2277–2302, https://doi.org/10.1007/s11042-022-13290-4.

[29] A. K. Yadav, R. K. Pateriya, N. K. Gupta, P. Gupta, D. K. Saini, and M. Alahmadi, "Hybrid Machine Learning Model for Face Recognition Using SVM," *Computers, Materials & Continua* no. 72 (2022): 2698–2712.

[30] A. K. Yadav and S. Shrivastava, "Evaluation of Reinforcement Learning Techniques," in *Proceedings of the First International Conference on Intelligent Interactive Technologies and Multimedia* (2010), 88–92.