STM JOURNALS

Trends in
Machine Design

ISSN: 2455-3352
Volume 11, Issue 2, 2024
May–August
DOI (Journal): 10.37591/TMD

http://engineeringjournals.stmjournals.in/index.php/TMD/index

Review TMD

AI-Powered Face Detection and Recognition Using Machine Learning

Saurabh Shandilya[1,*], Shubham Prajapati[2], Vishal Agrawal[3], Rahul Arora[4], Sohaib Nasir[5]

Abstract

These days, one of the biggest computer vision technologies is facial recognition. Face identification in computer vision, lighting position, and facial expression is always an extremely challenging issue. In real-time video pictures captured by a video camera, face recognition tracks specific objects. Put simply, it is a system tool that uses a still picture or video frame to automatically identify a person. In this research paper we use different different algorithm for face recognition like Viola-Jomes algorithm, PCA algorithm, KLT algorithm and Emgu CV. We can carry through the aforementioned procedures using a variety of strategies. The present work here depicts the details of various technologies which have been researched and discovered. Let us take an example of designing color filters that we use in cameras to make them more colorimetric.

Keywords: Face recognition, PCA, viola-JOMES, Dual mode QR codes, video anomaly detection

INTRODUCTION

Human faces are always essential in applications like security systems, debit and credit card verification, and monitoring on criminal activity in public areas. The primary goals of the system are to develop a facial recognition system that can someday eclipse human capabilities. The frontal faces of humans are the system's primary focus. There are several facial recognition algorithms available, and each offers advantages of its own. When we look at a face, most of the time we can instantly recognize it if we are already acquainted with it. The Viola-Jones cascade classifier and the KLT algorithm were utilized, however the camera continually catches things. A significant portion of machine learning research is devoted to enhancing the current methods in the face identification and identification domain. Real-time face detection and identification is one of the numerous features of this program. Utilizing a programmable system cuts prices and time. Face picture matching is done automatically or semi-automatically. Face recognition is the process of precisely recognizing a previously spotted item as a known or unknown face, utilizing a database of faces to verify the input face. Face detection is the process of locating and detecting a face's object in an input the image.

***Author for Correspondence**
Saurabh Shandilya
E-mail: saurabh.shandilya@poornima.org

[1]Professor, Department of Computer Engineering Poornima College of Engineering Jaipur, Rajasthan, India
[2]Student, Department of Computer Engineering Poornima College of Engineering Jaipur, Rajasthan, India
[3]Student, Department of Computer Engineering Poornima College of Engineering Jaipur, Rajasthan, India
[4]Student, Department of Computer Engineering Poornima College of Jaipur, Rajasthan, India
[5] Student, Department of Computer Engineering Poornima College of Jaipur, Rajasthan, India

Received Date: May 30, 2024
Accepted Date: August 06, 2024
Published Date: August 30, 2024

Citation: Saurabh Shandilya, Shubham Prajapati, Vishal Agrawal, Rahul Arora, Sohaib Nasir. AI-Powered Face Detection and Recognition Using Machine Learning. Trends in Machine Design. 2024; 11(2): 26–35p.

RELATED WORK

Face Detection

Real-time facial object detection and continuous tracking of the same item are the goals of this technique. Here, we harness the training sample photographs of various items that you might choose for the training classifier to understand and track. A facial recognition system includes face tracking. Here, we can identify particular, distinguishing features on a person's face using a system algorithm. Facial recognition is continuously improved by an

adequate and effective facial detection algorithm. Numerous facial algorithms include machine learning and face-to-face geometry-based techniques. Viola and Jones suggested a framework that is quick and provides an outstanding rate of detection out of these approaches [1].

The Viola-Jones detection process is reliable and quick. Thus, the Viola-Jones face identification algorithm—which makes use of Integral Image—was our choice. We've seen that in a range of lighting circumstances, our approach produces better results. A programmable system cuts expenses and time. Face depict matching is done automatically or semi-automatically. Face recognition is the process of determining if an item has already been detected, classifying it as a known or unknown face, identifying the precise face it is, and utilizing a database of faces to verify the input face. Face detection involves locating and recognizing the face's object inside the input the image [2].

Most of The frontal faces of persons are the system's primary focus. There are several facial recognition algorithms offered, and each offers advantages of its own. The majority of the instant we identify a face when we glance at it, especially if we are previously familiar with it. If feasible, this innate talent can be justified and applied in practical ways. There are several face detection techniques at that time. The Viola-Jones cascade classifier and the KLT algorithm were employed, however the camera is continually detecting things. A large number of research issues in the field of machine learning are centered on enhancing the current methods for face detection and identification. This application has many features which have many advantages of face detection and recognition in real time. Using a programmable system reduces the time and cost (Figure 1 and Figure 2).

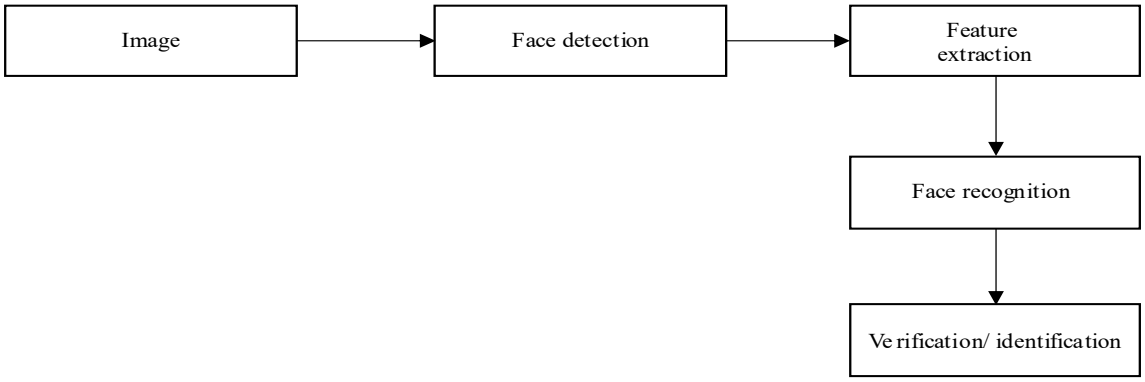

Figure 1. Process of face detection.

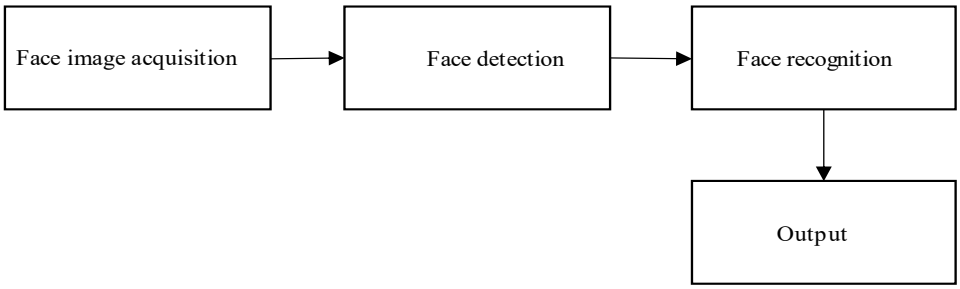

Figure 2. Process of real time face detection system approach.

Viola-jones algorithm

An astounding face recognition system that can recognize faces in frontal view in real time was introduced by Viola and Jones. The first object in real-time that Paul Viola and Michael Jones proposed in 2001 is the Viola-Jones object detection methods. The technique was developed as a solution to the face identification problem. It may be taught photos to recognize a range of object types. Open CV uses this technique to recognize objects employing CV Haar. Features that Jones and Viola used. It is possible for calculating hundreds of characteristics speedily by

Verification/ Identificatio n The two rectangle characteristics are used by three types: two, three, and four-rectangle, viola, and Jones.

The integral image at location x, y contains the sum of the pixels above and to the left of x, y, inclusive:

$$ii(x,y) = \sum_{x1=y1sy} i(x1,y1)$$

Proposed Algorithm

1. Capture the person's image.
2. Apply face detection algorithm to detect face.
3. Use Viola Jones and KLT algorithm.
4. Do feature Extraction based on the extraction process using Eigenvector selection.
5. Save that captured face with the name into the database.

$$s(x,y) = s(x,y-1) + i\,(x,y)$$
$$ii(x,y) = ii\,(x-1,y) + s(x,y)$$

Value: oriented gradients of pixel intensities. The four features matched by this algorithm are then sought in the image of a face. Rectangle features. Value=Σ (pixels in black area) – Σ (pixels in white area)

Three types: two, three, four-rectangle, viola & jones used two rectangle features. The integral image at location x, y contains the sum of the pixels above and to the left of x, y, inclusive:

Face recognition is the identifying an already detected object as a known or unknown face, and telling exactly who face it is and using for a database of faces in order to validate [6].

IMAGE RECOGNITION

Face image

Face recognizers examine shots of faces and identify a number of crucial features, including the lips, nose, eyes, eyebrows, and corners of the mouth. There are 66 of these spots whose coordinates are referred to as facial-feature dots. In this way and identify to face recognition and different-different technique are used (Figure 3).

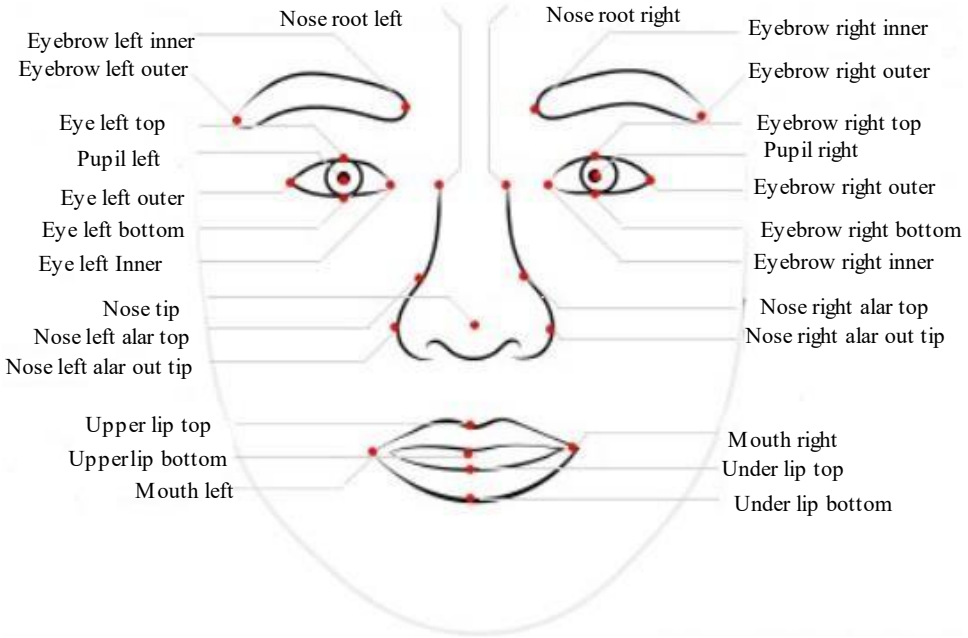

Figure 3. Face image.

Dual module qr codes

DMQR, Dual mode QR codes is the safe way through which users can transmit or interface private information in public places or in front of the people without any second thoughts. QR is a 2-dimensional barcode type. QR stands for quick response. It is used to have information about some product or service on spot. We can even make transactions by just scanning the QR codes. Every associated product has their own unique QR codes just to avoid any cross transaction or any confusion. The same module design, together with syncing and alignment patterns, are employed by DMQR codes to convey discrete data pieces. Each module has primary data coupled by intensity modulation that is compatible with traditional QR code honors [7].

Dual module QR codes provide remarkably better contiguous privacy for secondary data as compared to classic QR codes. Dual module QR codes are flexible and can work in any shape, not restricted to any particular pattern or colour. The direct privacy of dual module QR codes can be reconstructed to meet the requirements of the applications by congruous the eccentricity parameter that handles the acclimatisation modification for the secondary data (Figure 4).

In early times most of the work was done using traditional features such as Dense Trajectories (DT), Improved Dense

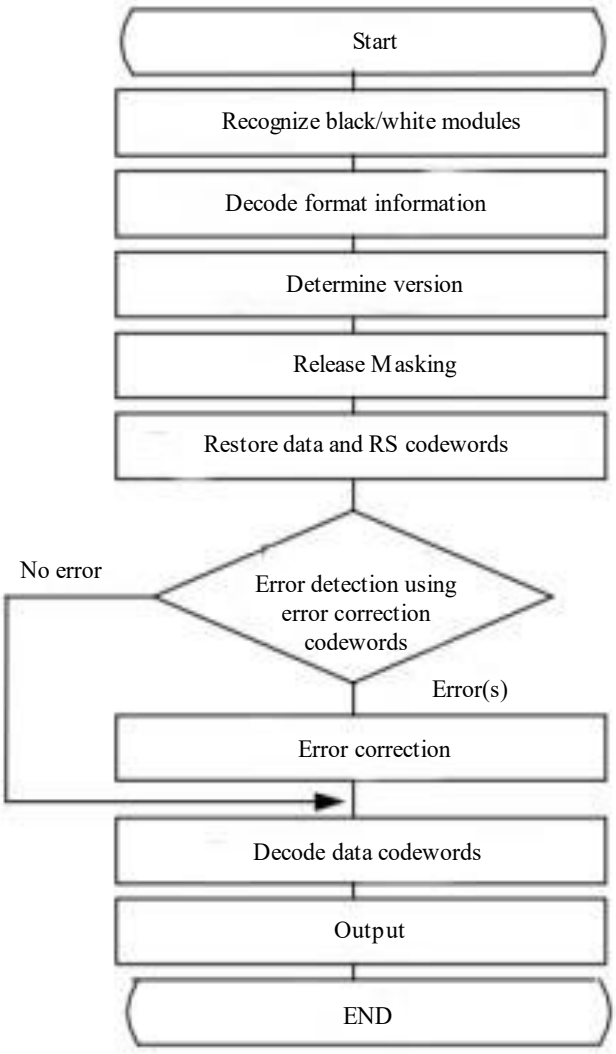

Figure 4. Working of QR codes.

COLOUR MATCHING FUNCTIONS

Colour matching functions basically provides a measurable link between the physical light stimuli and the colours received by the human vision system. The given figure shows a setup for the colour matching experiment. Viewed through the lens, the person watching sees a bipartite field, or two distinct sections, with one side illuminated by a test light and the other by light mixes of the three primaries, or monochromatic red, green, and blue lights. The observer varies the intensity of the three principal lights provided to achieve a visual match, meaning that if the observer cannot discern any difference, the two stimuli on either side of the field match. There might not be every time a match [9].

In this instance, the test field of lights has to include one more main light. We can depict it as though we were deducting a portion of the main light. We may estimate the Color Matching Functions by sequentially configuring the test light to be a unit power monochromatic light at sample locations across the visible spectrum. In other words, we identify the RGB mix that provides a match for each wavelength. Given the colour matching functions, we can calculate the match for any test light as colour matching is linear (the total of two test spectral lights is matched by the sum of their distinguish matches). Each relevant product has its own distinct QR code to prevent misunderstanding or cross-selling. The same module design, which is together with synchronization and alignment patterns, are employed by DMQR codes to convey discrete data fragments. Primary data is embedded through intensity modulation within each module [10].

The main objective is to use the technique of the appearance and motion deep network to generate discriminative features and then use a one-class SVM to predict anomaly scores of each test feature. Coincidentally, Wu et al introduced a one- stage deep one-class classifier for anomaly detection; the core of this model is that it can jointly learn discriminative features and train a one-class classifier (Figure 5).

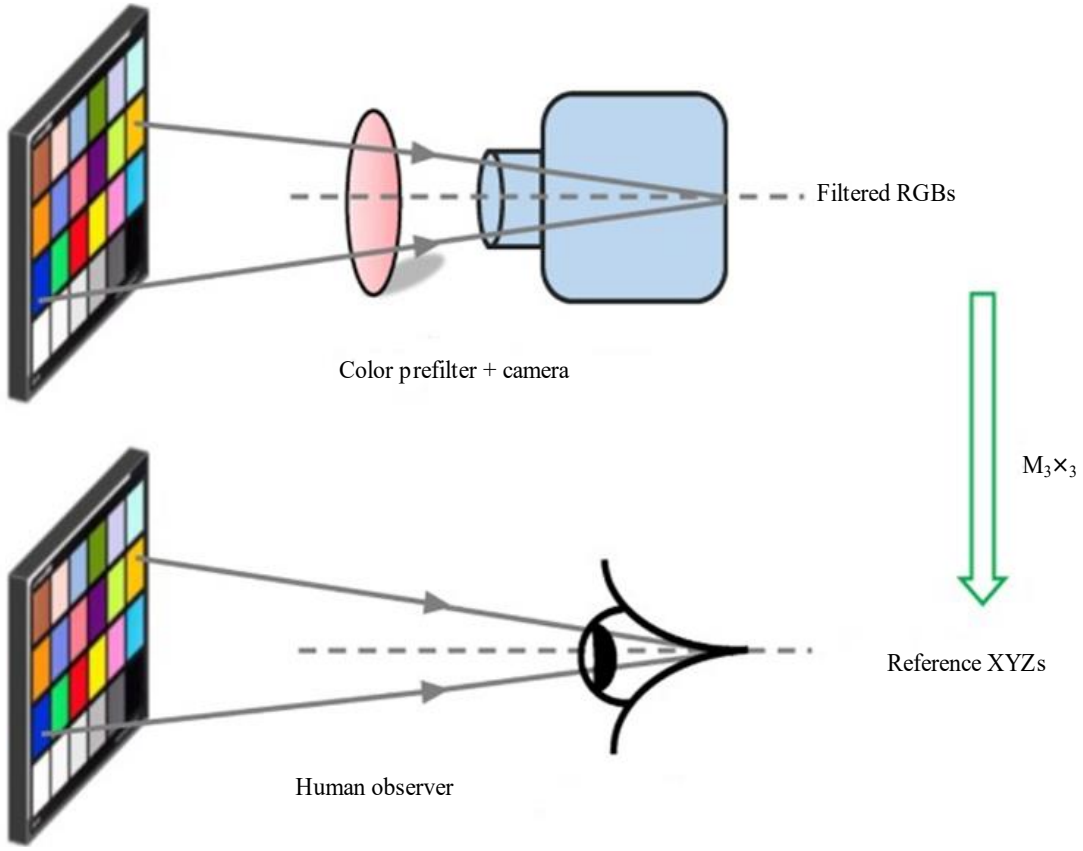

Figure 5. A typical setup that can be used for the colour matching experiment.

Video Anomaly Detection

As we have discussed earlier, video anomaly detection comprises mainly two categories, namely, semi-supervision and weak supervision. Common semi-supervised solutions are generally based on reconstruction of errors or one-class classifiers, e.g., the sparse reconstruction costs over the normal dictionary to detect anomalies. It was discovered through research that a stacked recurrent neural network based on reconstruction errors is used to classify whether a test event belongs to normal or abnormal activities. The main objective is to use the technique of the appearance and motion deep network to generate discriminative features and then use a one-class SVM to predict anomaly scores of each test feature [11,12]. Coincidentally, Wu et al introduced a one-stage deep one-class classifier for anomaly detection; the core of this model is that it can jointly learn discriminative features and train a one-class classifier (Figure 6).

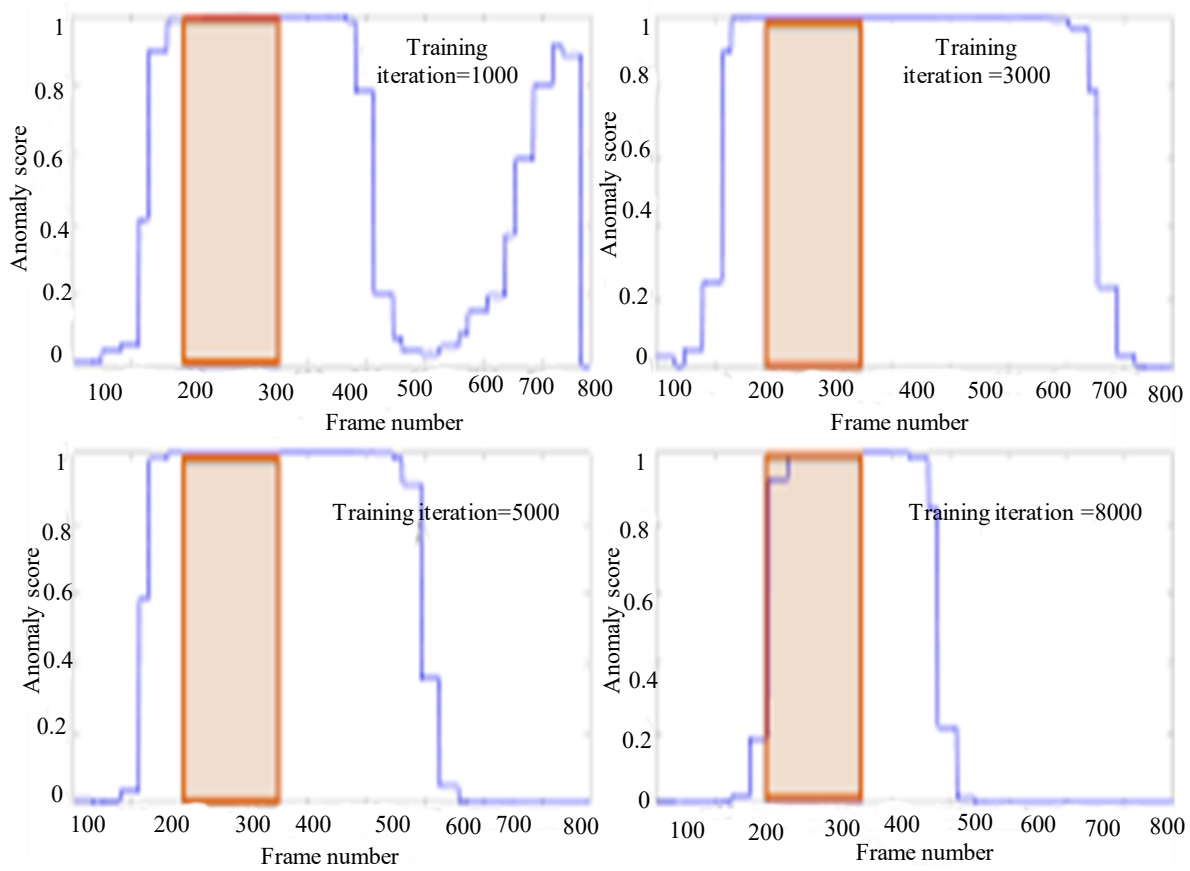

Figure 6. Video anomaly detection.

E Together Recognizing, Localising and Summarising Actions in Egocentric Videos Egocentric discernment is a renowned research area in computer vision and multimedia communities, because of its various advantages in human action object event recognition [13].

ACTION RECOGNITION

We can find many action-recognition methods in the last 2 decades, but we focus on action and activity recognition methods mainly in egocentric videos. Many object-centric representations have been developed for recognizing daily activities from first camera views. The research paper has focused on the information regarding the work that discussed how different object models can be learned from egocentric video of household activities.

In early times most of the work was done using traditional features such as Dense Trajectories (DT), Improved Dense Trajectories (IDT) and pyramid histogram of oriented gradient (PHOG). Later on,

recognition and localization of single as well as multiple actions using super pixel level solution was achieved. While in recent times, deep features have shown notable performance development for self-centred action recognition.

ACTION RECOGNITION FROM EGOCENTRIC VIDEOS

Different models have been proposed, some of are following: -

Centre Surround Model

In this research paper, a centre -surrounded model (CSM) has been developed to distinguish centre and surrounding regions in each frame of egocentric video. The frame is said to be a union of a centre QT=Q(t) region, and a surrounding region RT=R(t). In the works, super-pixels have been used for precise localization and for accurate recognition of single and multiple actions. SLIC has a very simple algorithm, so it has been crossed over to other sophisticated super-pixel extraction techniques available.

video represent graph

This section focuses on the details of construction of a weighted sparse spatiotemporal video representation graph (VRG) which is provided in deep space.

Random Walk Based Action Labelling

It is used for 0btaining the labels of actions of the various nodes in the VRG, used for extraction of VRG in proper order. Probability with which walker can move to a marked

vertex vm ∈ VM is given by:

$$p(vn, vm) = w(e(vn, vm))\deg(vn) \text{ is the degree of the vertex vn}$$

Time complexity

Computational complexity of the proposal algorithm is presented, further it is calculated that the time complexity of evaluating the centre-surround cost for all n frames is 0(n) (Figure 7).

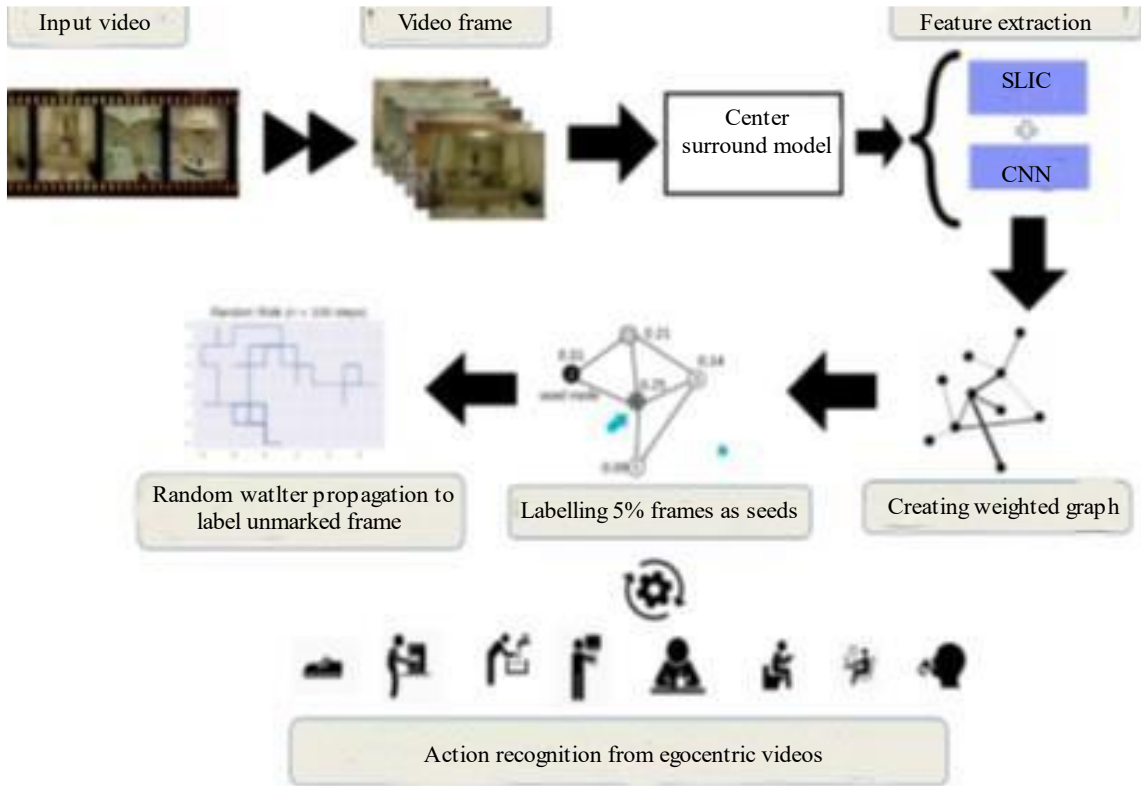

Figure 7. Flowchart showing the proposed framework.

Blur Removal Via Blurred Noisy Image Pair

One of the most common examples of using picture DE blurring techniques to recover excellent images from blurry images is photographing moving things in a night scene, where one can capture extraordinarily clear and bright photos.

Single image blind deblurring

Blind DE blurring is a technique for calculating an accurate estimate of an unknown blur kernel, which is then used to perform DE convolution to recover the sharp image. DE blurring images can be accomplished using a variety of techniques, including maximum a posteriori (MAP), vibrational Bayes, edge prediction, and sparse coding. Multiple image DE blurring attempts have been undertaken. The superiority of multiple image DE blurring is due to the complimentary information offered in the photos, whereas Depending on the components that are lacking owing to blur, many blurred images can give the appropriate frequency [14]. It, on the other hand, has a hard time dealing with object movement. In fact, none of the aforementioned methods can deal with sophisticated blur. To provide better solutions in solving non-uniform blur, other methods have been proposed, in which they give leverage to different forms of multiple observations, for e.g., a blurred-noisy pair, a blurred-flash pair and multiple blurred images. The optical flow approach is used to locate the corresponding patches between the blurred and noisy images. DE blurring of images consists of two stages: DE blurring and adding detail layer, where detail layer contains a post-processing or refinement. Decomposition of blurred image I1 into a collection of overlapping square patches C is equal to c1,....., ci,....., cP , in which ci RM

M = s1 s1 s1 s1 s1 s1 s1 s1 s1 P is the number of patches, and s1 is the first patch.

M is the number of pixels, and signifies the patch size in I1.

In every patch the collection of pixel intensities in a given image.

X (X RM) represents the patch.

Deep Face Rectification For 360 Dual Fisheye Cameras

In simple words, we want to develop a face recognition system that works for both fisheye and rectilinear images. In addition, the system we want to be agnostic of the camera parameters so that it can function without any knowledge of the fish eye cameras [15].

Our approach is to accomplish face recognition on the fish eye images. Various algorithms for this approach can be divided into two parts. The first part will create a new dataset of artificial images generated from an existing set of rectilinear images And then use it to rescale the face recognition model. The other part modifies the architecture of the underlying neural network and captures the complete fish eye images rather than a cropped part of the image as input.

3D face from x; learning face shape from diverse sources

We propose a way for weakly supervised anomaly detection. First we proposed a causal temporal relation (CTR) module to capture temporal cues among video snippets. here should be no information leakage in the anomaly detection task from future to past.

Our CTR module is local which only captures temporal relations between the current snippets and historical neighbours within a local range (Figure 8).

Light Field Image Super-Resolution Using Deformable Convolution

We examine the important works in single image SR (SISR), LF image SR, and deformable convolution in this part. SISR's job is to take a blurry LR image and turn it into a clean HR image. SISR is a powerful tool since an input

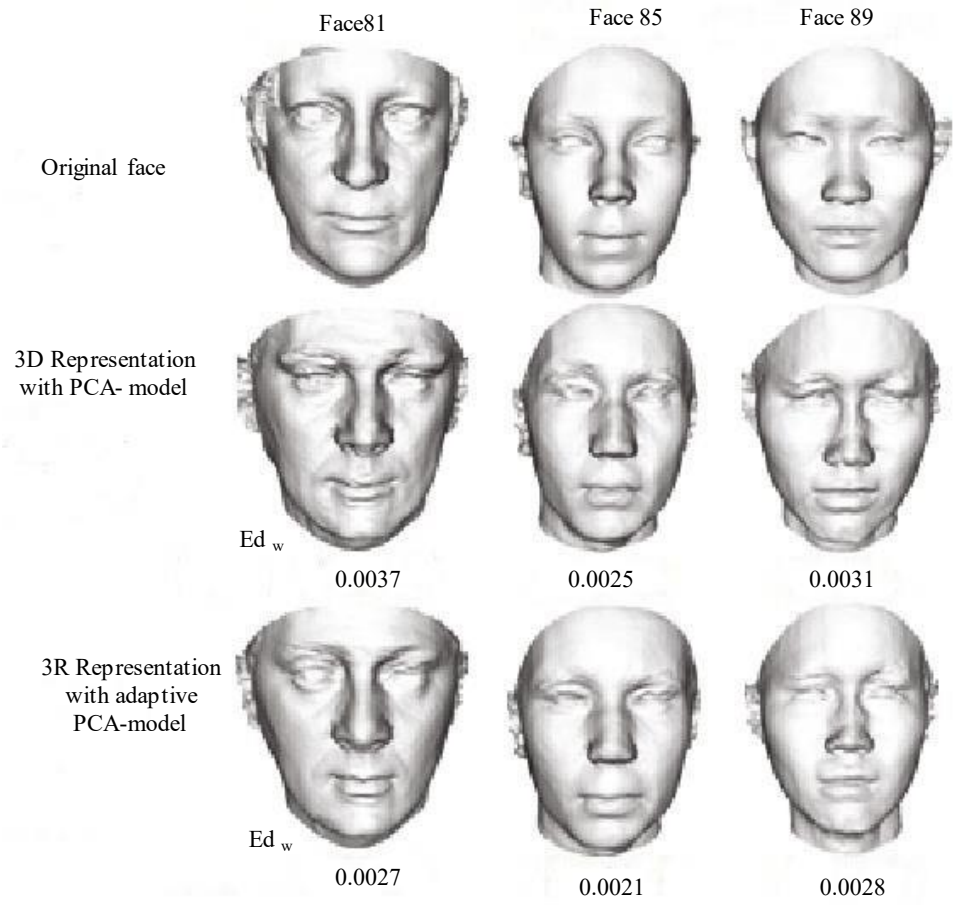

Figure 8. Visual comparison of represented 3D face

LR image can be linked to several HR outputs. A full review of SISR approaches has been published. We will just discuss a few literary classics in this article. Since Dong et al. the seminal work has been proposed based on CNN the SISR technique is a method that is used to calculate deep learning- based SRCNN. Because of their outstanding performance, techniques have dominated this field. In terms of precision and efficiency, performance is excellent. Without a doubt, various networks have been proposed to continuously improve the SISR performance. Kim et al. proposed a very deep SR network (i.e., VDSR) and achieved a significant performance improvement over SRCNN. Lim et al [33]. proposed an enhanced deep SR network (i.e., EDSR). With the combination of local and global residual connections, EDSR won the NTIRE 2017 SISR challenge. In this research, we present a deformable convolution network (LF- DFnet) to handle the disparity problem for LF image SR, inspired by the success of deformable convolution in video SR [16].

Weighted Feature Histogram of Multi-Scale Local Patch Using Multi-Bit Binary Descriptor for Face Recognition

The three main procedures in the WFH test phase are MSLPG, multi-bit local binary descriptor extraction, and global face representation construction. Patches that are locally multi-scaled the initial data is generated using MSLPG. PDVs are extracted first, followed by local binary descriptor extraction, during multi-bit processing. The MBLBDs were then projected with the learned mapping. Finally, global face representations are generated using feature histograms. After learning the codebook using the concatenation of the weighted feature histograms, MBLBDs are utilised to extract all local patches using the final face representation. We learned about weights. The proposed WFH can improve the robustness of face recognition for the following reasons. For starters, the MSLPG can extract multi-scale face data for WFH. On a multi-scale, confront local patches. Second, MBLBDL can extract binary descriptors with a large number of bits without increasing the mapping quantization loss, which can be

decreased with the use of matrix parameters. The performance of WFH has to be improved. RWL also learns a set of weights for various local patches in order to improve the functions of critical.

CONCLUSION

Face recognition is an essential component of the face recognition system. Unknown people can be identified with this approach. The facial expression recognition system provides a robust face identification model in real-time settings by mapping behavioral traits with physiological biometric parameters. The physiological features of the human face that are relevant to different facial emotions, such pleasure, pain, fear, rage, surprise, and disgust, are linked to geometrical structures that serve as the recognition system's basic matching template. And In this work, we have offered a systematic analysis of proximal privacy for DM QR codes via smartphone screens as well as we have also discussed, how to robust image DE blurring method? We have introduced a joint framework for recognition, localization and summarization of actions in egocentric videos, the incongruity problem has been clearly handled by performing feature alignment using the designed angular deformable alignment module (ADAM). JCS technique has been widely used during the COVID that helped patients by providing explainable reports. Image recognition algorithms based on different technologies have been evaluated and compared. Every approach has its own set of benefits and drawbacks, but progress is being made at breakneck pace, indicating the scope and dependability of image recognition in the near future.

REFERENCES

1. Barron I, Hsin Jui Yeh, Karthik Dinesh, Sharma G. Dual Modulated QR Codes for Proximal Privacy and Security. IEEE Transactions on Image Processing . 2021;30:657–69.
2. Graham D. Finlayson, Member, IEEE, and Yuteng Zhu. Designing Color Filters That Make Cameras More Colorimetric. IEEE TRANSACTIONS ON IMAGE PROCESSING, VOL. 30, 2021. 853-867.
3. Wu P, Liu J. Learning Causal Temporal Relation and Feature Discrimination for Anomaly Detection. IEEE Transactions on Image Processing . 2021;30:3513–27.
4. Sahu A, Chowdhury AS. Together Recognizing, Localizing and Summarizing Actions in Egocentric Videos. IEEE Transactions on Image Processing .2021;30:4330–40.
5. Gu C, Lu X, He Y, Zhang C. Blur Removal Via Blurred-Noisy Image Pair. IEEE Transactions on Image Processing . 2021;30:345–59.
6. Li YH, Lo I-Chan, Chen HH. Deep Face Rectification for 360° Dual-Fisheye Cameras. IEEE Transactions on Image Processing. 2021 ;30:264–76.
7. Guo Y, Cai L, Zhang J. 3D Face From X: Learning Face Shape From Diverse Sources. IEEE Transactions on Image Processing . 2021 ;30:3815–27.
8. Kasar MM, Bhattacharyya D, Kim T. Face Recognition Using Neural Network: A Review. International Journal of Security and Its Applications. 2016;10(3):81–100.
9. Liu Z, Wang Y. Face detection and tracking in video using dynamic programming. 2002 ; Vancouver, BC, Canada. DOI: 10.1109/ICIP.2000.900890.
10. Ritesh Boda, M Jasmine Pemeena Priyadarsini, J Pemeena. Face detection and tracking using KLT and Viola Jones .ARPN journal of Engineering and Applied Sciences. 2016; 11(23):13472-13476
11. Turk M, Pentland A. Eigenfaces for Recognition. Journal of Cognitive Neuroscience. 1991;3(1):71–86.
12. Beymer D, Poggio T. Face recognition from one example view. Proceedings of IEEE International Conference on Computer Vision. 2002; Cambridge, MA, USA. DOI: 10.1109/ICCV.1995.466898.
13. Wu YH, Gao SH, Mei J, Xu J, Fan DP, Zhang RG, et al. JCS: An Explainable COVID-19 Diagnosis System by Joint Classification and Segmentation. IEEE Transactions on Image Processing. 2021 ;30:3113–26.
14. Wang Y, Yang J, Wang L, Ying X, Wu T, An W, et al. Light Field Image Super-Resolution Using Deformable Convolution. IEEE Transactions on Image Processing . 2021;30:1057–71.
15. Yang H, Gong C, Huang K, Song K, Yin Z. Weighted Feature Histogram of Multi-Scale Local Patch Using Multi-Bit Binary Descriptor for Face Recognition. IEEE Transactions on Image Processing. 2021;30:3858–71.

