Journal of Applied Information Science
10 (1) 2022, 01-09
http://www.publishingindia.com/jais

A Synopsis on Intelligent Face Discovery Frameworks

Nimisha Rajeev P.[1*], Anima M.[2], Heera Mohan K. V.[2], Pranav P. V.[2] and Swetha Pai[3]
[1]BTech Student, Computer Science and Engineering, Sree Narayana Guru College of Engineering and Technology, Payyanur, Kerala, India. Email: nimisharajeevan123@gmail.com
[2]BTech Student, Computer Science and Engineering, Sree Narayana Guru College of Engineering and Technology, Payyanur, Kerala, India.
[3]Assistant Professor, Computer Science and Engineering, Sree Narayana Guru College of Engineering and Technology, Payyanur, Kerala, India. Email: swethapai@sngcet.org
*Corresponding Author

Abstract: Image processing is a wide area which has attained attention over the last few decades. Multiple faces can be detected using Image Processing Techniques. Various algorithms are utilised to develop software and hardware that recognises the human face. The algorithm will compare the various pictures to pre-defined or learnt images, as well as real video images. Security and surveillance, authentication/access control systems, digital healthcare, photo retrieval, and other applications have all benefited from its use. This approaches needs maximum information, in certain conditions it is difficult to gain those informations such as small face detection, night person identification, partial face recognition, occlusion and so-forth. Opportunities and problems are inextricably linked. Growing business interest in face recognition is good, but it also proves to be a difficult undertaking when it comes to the difficulties that have plagued its quality of delivery. This paper gives a high-level overview of general solutions to these problems.

Keywords: Artificial Intelligence (AI), Image processing, IoT, Machine learning.

I. Introduction

There are many face acknowledgment frameworks present today. Surveillance cameras, when it comes to sending recordings at a high bit rate, chokes. It is additionally not functional to store recordings for long duration. The nature of the countenances that is recognized by the camera is additionally problematic. Face recognition has become a well known region of examination in computer vision and quite possibly the best utilizations of picture investigation and comprehension. Face detection and extraction are computer vision task, that have numerous applications and have direct significance to the face recognition issue. Human face recognition has a wide spectrum of applications such as mugshot identification, social networking, automated crowd surveillance both in public and restricted areas, design of human computer interface (HCI). Face acknowledgment frameworks use computer calculations to select explicit, particular insights concerning an individual's face. These details, like distance between the eyes or state of the jawline, are then changed over into a numerical portrayal and contrasted with information on different appearances gathered in a face acknowledgment data set.

Image processing can be characterized as the specialized examination of a picture by utilizing complex calculations. Here, image is utilized as the input, where the valuable data returns as the output. In everyday the (pre-) processing of a picture is frequently an underlying advance to later concentrate the highlights that would be utilized to prepare a machine learning classifier. Signal processing can be utilized to upgrade or wipe out properties of the picture that could improve the performance of the machine learning algorithm.

Personal re-identification has attracted increasingly more consideration the computer vision eld because of its significant applications in numerous genuine situations like video reconnaissance, advanced mechanics, robotized driving. For the most part, person re-identification alludes to coordinating with a given walker across non-covering cameras in an organization. By and by, an individual re-identification framework comprises of individual indicator, individual tracker and individual matcher. Since person recognition and person tracking are independent problems in computer vision, person re-identification as a rule centers around the third part and it is regularly viewed as a recovery issue.

Face detection is a central advance of many face related applications, for example, face alignment, face recognition, face verification, and face expression examination. Amazing face finders can precisely order and find faces from an image. Lately, deep learning techniques particularly convolutional neural organizations (CNN) have accomplished exceptional achievements in an assortment of computer vision errands, going from image classification, to protest identification, which additionally motivate face discovery. Dissimilar to customary strategies for hand-made highlights, CNN-based strategy can

remove face includes consequently. Anchor-based face locators assume a predominant part in CNN-based face indicators. They distinguish faces by ordering and relapsing a progression of pre-set anchors, which are created by consistently tiling an assortment of boxes with various scale on the pictures. So it is sensible to indicate attributes in face pictures which are sufficient to characterize an individual. The ideal attributes should take over intra-class variety at a similar time redesign between class varieties paying little mind to lighting, and outward appearance. Brief component extraction incorporates a portrayal of important highlights, its extraction as well as the inclination among a selection of highlights.

In a face recognition system, face recognition is the beginning interaction to begin with distinguishing faces. The face recognition step includes finding or recognizing a face if in the pictures caught. It distinguishes the face as well as finds the district of countenances in the picture. The subsequent stage is include extraction, it is an interaction where, it peruses the math of your face and key elements in a picture are distinguishes to get the facial milestones (Fiducial focuses). These pertinent and particular focuses are extricated as the descriptors from the caught picture. By utilizing this strategy a high dimensional facial picture space is limited to a low dimensional facial include space. A dimensionality reduction is a successful way to deal with scaling down information. In the caught pictures not all pixels or a piece of the picture will add the ID. Facial expression recognition extract the information about facial expression features from the original image and classify according to human emotions as aversion, happiness, neutrality and surprise [1] [2]. Numerous cctv observation frameworks send out footage in their own formats, which should be re-configuration or changed over to a reasonable arrangement that simpler for examination. In any case, this can frequently cause the lower of value also, data misfortune, which makes the assessment interaction difficult. The recording in advanced criminology is regularly utilized for relative investigation, including scientific examination, correlation of pictures of interrogated regarding know items like subjects, vehicles, dress, and weapons, with well-qualified assessment being giving on the findings. In numerous advanced CCTV frameworks, facial acknowledgment administrations are installed to recognize online lawbreakers or suspects. Other administrations, for example, movement detection, body and face acknowledgment, cross-present acknowledgment, step acknowledgment, are generally investigated in the previous few years. In some hard cases (poor seeing conditions), it is very difficult to distinguish people exploit face, body, still, and so forth Albeit many picture preparing methods have been created in the previous few decades, the vast majority of them don't exploit face, body, and so forth

II. Literature Survey

The Video Based Evidence Analysis and Extraction in Digital Forensic Investigation Jianyu Xiao *et al.* [3] explain the advanced forensic video analysis techniques to assist the forensic investigation [3]. It initially proposes a criminological video investigation structure that utilizes an effective video or picture upgrading calculation for the inferior quality of film examination. This method has four parts which are forensic object detection/identification in video/image, video enhancement techniques, hidden information extraction from corneal reflections in video and images and automated suspicious object detection in video.

In the first section the author addressed the problem of analyzing the visual data acquired in a crime scene using the object detection. Here the suspicious subject can be automatically detected and presented in the evidences. This strategy can separate some insight and build up connections that can assist an examination with relating diverse proof things or even crime locations. In the second segment video improvement methods center around methods that can be useful for recovering extra data from upgraded recordings [6]. The histogram balance based strategies can essentially improved the likelihood to reveal extra data from inferior quality video or pictures. AHE

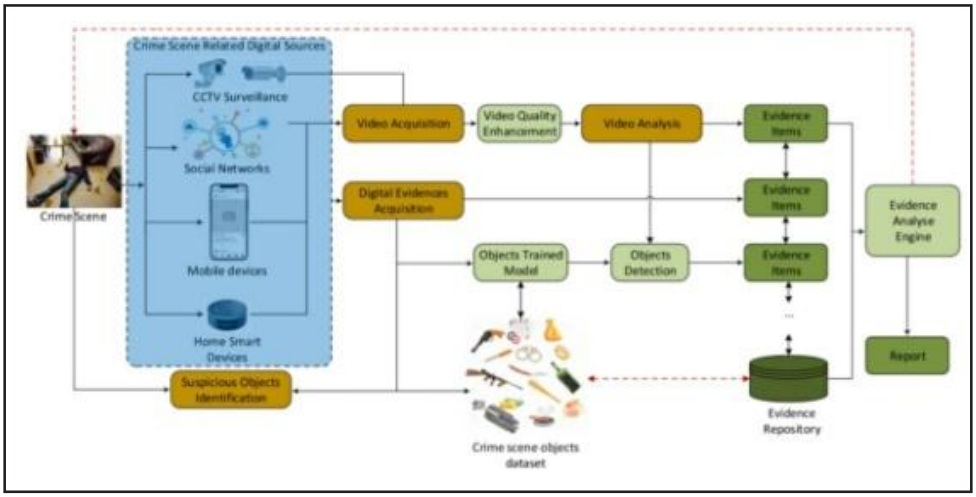

Fig. 1: Enhanced Forensic Video Analysis Framework [3]

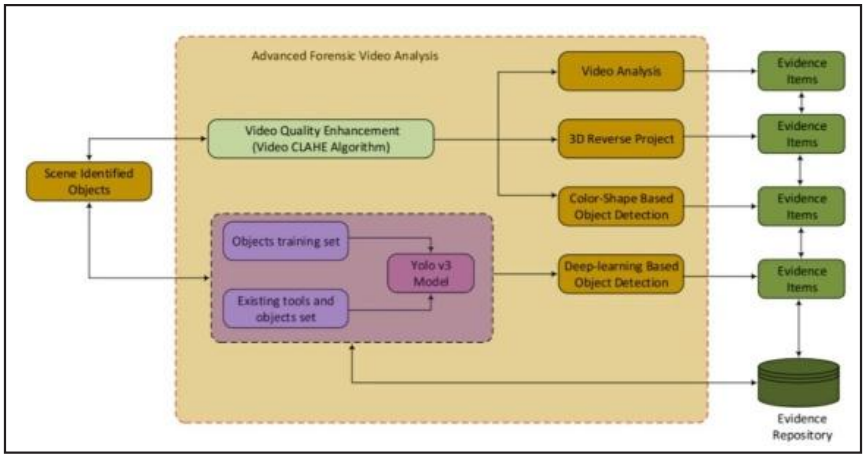

Fig. 2: Procedures for Advanced Forensics Video Analysis [3]

[4] and CLAHE [5] calculations are utilized for this strategy yet AHE calculation doesn't turn out viably for some CCTV recordings and CLAHE is more valuable for general casing/ pictures [7]. The next portion of the method comprises hidden information extraction from corneal reflections in video and images. Here it is feasible to mine concealed development and conceivable to remove the concealed scene data from the eye reflection and it is additionally conceivable to lead a face coordinating and which requires exceptionally high goal. The last segment portrays about robotized dubious article location in video. This segment intended to distinguish explicit dubious item that as of now recognize from proof and in criminological examination the dubious articles that is utilized in wrongdoing conduct can be arranged. OpenCV and GPU techniques are used for color and shape based object detection and tracking and deep learning techniques are used for automatic extraction of features.

The SFA: Small Faces Attention Face Detector Shi Luo *et al.* [8] the maker proposes a novel scale - invariant face marker named SFA. Here first present multi-branch face ID plan which zeros in nearer on face with restricted degree. This procedure in a general sense improves face area execution. The general designing of the procedure is, input an image with optional size is resized to outline an image variety [9]. Each image uses SFA to deliver acknowledgment result and union these results to get last area result, by then a part map mix is applied to merge feature maps. Finally by using a lot of multi-branch acknowledgment modules to bunch face/none-face and backslide the cutoff boxes. The multi-branch disclosure configuration explains that, for the extraction of features CNN-based face finders are used. It does convolution and pooling. The size of responsive fields creates. Here the appearances are perceived from 4 particular layers of VGG-16 using revelation modules, CNN and pool

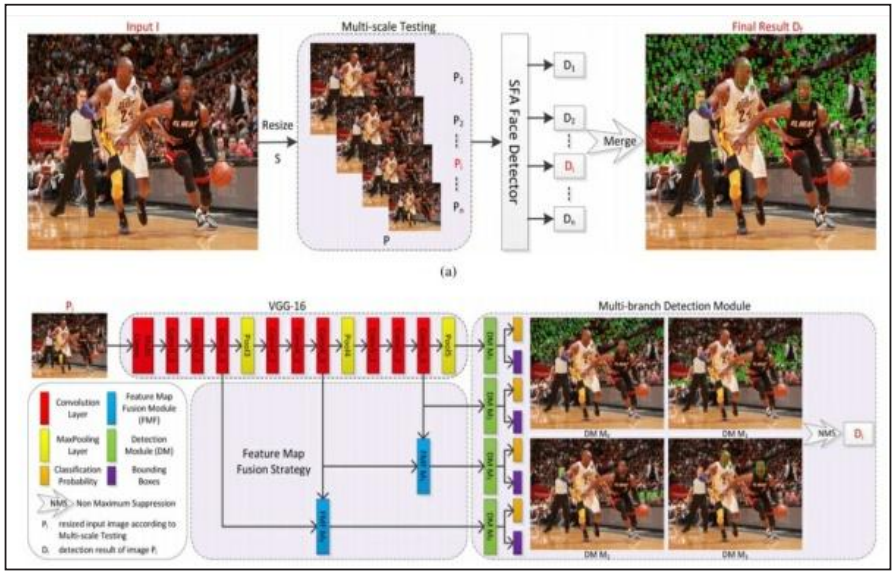

Fig. 3: (a) The Pipeline of Face Detection using SFA (b) The Network Architecture of SFA. It Consists of VGG-16, Feature Map Fusion, Multi-Branch Detection Module and Multi-Scale Testing [8]

5 are picked to interface the area modules autonomously [10]. Each acknowledgment module is set up to distinguish face from enormous degree. Division and defeat procedure is used to deal with unconstrained face acknowledgment in a lone locater. Finally the different branches are joined followed by NMS to shape the last area result. The third arrangement is little faces tricky anchor plan on a wide extent of size fluctuating from 4 to 512. Which guarantees that distinctive size of faces has enough features for ID. By then the anchors of size 4, 8, 16 and 32 are applied for faces with restricted degree. SFA can change the +ve and -ve gets on the establishment. Highlight map Fusion is the something else that is utilized to improve the limit of distinguishing hard faces with limited scope [11]. At that point significant level highlights into low level highlights are consolidated. After that meld the element guides of adjoining branches and apply the highlights. At last, multi-scale preparing and testing are performed [12]. Learn more highlights across a wide scope of scale, which makes the model more powerful. In the preparation stage, at first resize the most brief size of information picture and afterward scale it. In the testing stage construct a picture pyramid with a wide reach scale for each test picture. Each scale in the pyramid is autonomously tried and the identification results are converging to get eventual outcome. This technique accomplishes fulfilled location result on hard faces used to improve the limit of recognizing hard faces with restricted degree.

The Person re-identification through face detection from videos using Deep Learning Vimala Mathew *et al.* [13] talks about a technique for individual re-distinguishing proof from the recordings gathered from observation cameras [14]. The basic undertakings in this measure are face recognition from recordings and expectation of people utilizing Convolutional Neural Network models created utilizing the edited face pictures from the face location stage [15]. Distinguishing individuals in the recordings has numerous applications like taking care of security through reconnaissance cameras, crime recognition, and customized help for the poor, item buy advancement, representative observing, and so on. Exact individual re-identification from recordings has colossal potential that can reform the manner in which organizations work today.

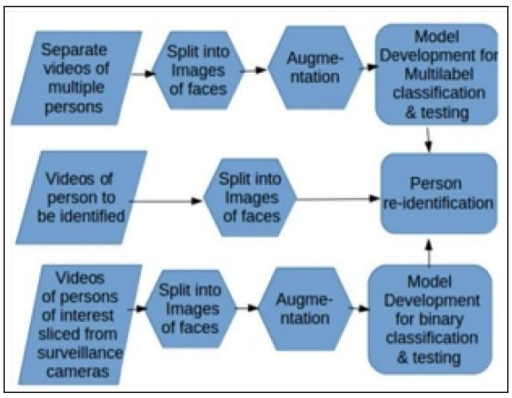

Fig. 4: Flowchart for Model Development and Person Identification [13]

Convolutional Neural Networks have got a trustworthy apparatus that can characterize pictures with accurate precision [16]. The video film gathered from the surveillance camera can be considered as a huge assortment of pictures that may incorporate at least one item. It talks about a technique for recognizing an individual of interest from observation recordings [17]. The principle commitment of this paper is the examination of two methodologies for face recognition to be specific Viola Jones calculation and Multitasking Convolutional Neural Networks. The difficult undertaking in the process is the exact preparing of the machine learning model so that there is no over fitting. Data pre-processing, Data augmentation and Model training are the three steps for training the machine. In data pre-processing it first collect videos of the persons of interest and then split the video into frames and then apply a face detection method. The two face detection used are Viola-Jones algorithm and Multi-tasking Convolutional Neural Network (MTCNN) [13]. In which MTCNN came about as a more exact face identifier since the face identification was finished with ideal edges of the appearances Fig. 5 [13]. The exactness of the deep learning models can be significantly improved by doing data augmentation. So prior to building up the face classifier, data augmentation methods are used. Model training is used for classify these images using labelled images. Two methods are used here. Feedforward Neural Network model and Convolutional Neural Network model (CNN). The execution of the CNN model was discovered better analyzed to the basic Feedforward model. The quality of the relied upon factors like the choice of the identification face location calculation, the nature of the recordings gathered from reconnaissance cameras and the sort of recognizable proof being made. Model exactness improves if the face pictures caught in these observation recordings are right posture and distance to the camera is least.

The Face and Gender Recognition System Based on Convolutional Neural Networks Yuxiang Zhou *et al.* [18] proposed Face and Gender Recognition System which is made out of two fundamental modules: face acknowledgment module and gender acknowledgment module [19] using public datasets. We use the public datasets Called Faces in the Wild (LFW),

Method	Accuracy percentage	
	Testing Accuracy	*Training Accuracy*
Viola Jones	98.5%	100%
MTCNN	99.5%	100%

Fig. 5: Comparison of Accuracy between the Two Models [13]

YouTube Face (YTF), and VGGFace2 to train CNN in the face recognition module, which improves precision. We use the public dataset Adience to train CNN in the gender recognition module, improving the best recognition accuracy from 91.80% to 93.22%. At the point when it is given a face picture, Face and Gender Recognition System will attempt to character data of the subject and its gender data independently, lastly yield the picture personality data and gender data all the while. As of

now, face acknowledgment is generally made out of four stages: identification, arrangement, portrayal and grouping [20]. Given a face portrayal with an intricate foundation, the framework can in any case anticipate character and sexual orientation by a grouping calculation.

Preprocessing and representation are the two main steps [21]. Firstly, the AdaBoost is utilized to recognize the face area in complex foundation, and afterward the face appearance picture is preprocessed. The picture size is first standardized to a size of 224 x 224 pixels and a 128- dimensional implanting is produced for each picture. Face location and arrangement handling are utilized as pre-handling to get ready for ensuing picture face acknowledgment. In representage stage, the strategies for face highlight extraction are summed up into two sorts: one is information portrayal techniques; the other is mathematical highlights or factual learning. Face and sex portrayal incredibly impacts the capacity of face and sex acknowledgment, which is the significant piece of the exploration. In this investigation, we construct the neural organizations of our framework dependent on the ResNet50 organizations to prepare face and sex datasets [22]. The neural organizations we constructed are a 50-layer networks that trains more than 3 million superstar pictures. In the neural organizations engineering, we use convolution to convolve highlights and utilize explicit layer capacities to get the attributes of each layer. Since the organizations construction of ResNet50 is unique in relation to other conventional convolutional networks structures, the forward proliferation of signs and the back spread of slopes become more muddled. To settle the forward engendering of the sign and the back proliferation of the inclination during preparing, we keep on utilizing Batch Normalization to tackle this issue. Before the last picture yield, we utilize the Global Average Pool (GAP) rather than the completely associated layer to diminish the size of organization model, along these lines improving preparing rate and preparing exactness of the neural organizations.

The Night Person Re-Identification and a Benchmark Jian'an Zhang *et al.* [23] the author focused on night person re-identification problems. For this purpose the author contributed a large and real-scenario person re-identification dataset for night scenario named KnightReid. The KnightReid dataset is the first to be checked for the realistic individual re-identification issue in a night scenario. The KnightReid dataset was gathered from a real-world video surveillance device on a campus. All three cameras are high definition (HD) with a resolution of 1920 x 1080 pixels. All three cameras were labelled in continuous night time for an average of three hours per camera [24]. In total, this dataset includes 937 identities captured by at least two cameras, as well as more than 31k bounding boxes. Proposed a new deep network model that combines a denoising network and a re-identification network is designed especially for the person re-identification in night scenario Fig. 3 [23]. Extensive experiments are conducted on the proposed dataset and benchmark results are provided including hand-designed feature methods, metric learning methods and deep learning methods.

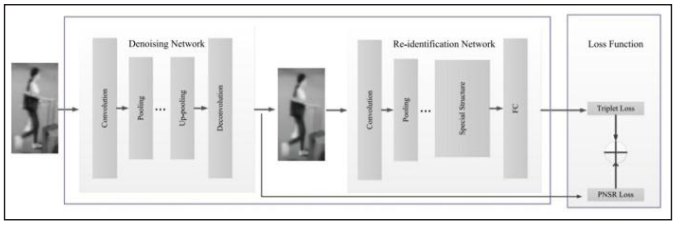

Fig. 6: The Proposed Pipeline Network for the Night Scenario Person Re-Identification [23]

The biggest difficulty of re-identification in a night scenario is that the pictures are low quality and noisy. As a result, it's only logical to conclude that certain techniques, such as image denoising, will help us enhance the quality of images. Both the image denoising and person re-identification can be modelled as deep neural network models in order to boost the quality of the image with a noising network before it feeds to a person re-identification network. Through this we can get rid of the noise and one the other hand we can get low- level features as a deep neural network has the ability [25]. Convolution and pooling modules are typically found in the lower layers of a denoising network, whereas up-pooling and deconvolution modules are found in the upper layers. Since they can't accommodate various image sizes, there are no completely linked layers. Aside from general convolution and pooling modules, individual re-identification networks often have a unique structure that is tailored to this problem, such as attention modules. In fact, it can be thought of as a basic pipeline that can be combined with any network. We use the N3 [26] network for image denoising and HACNN [27] for individual re-identification in this study since they demonstrate state-of-the-art efficiency in image denoising and person re- identification, respectively.

In reality, it can be thought of as a simple pipeline that can be combined with any network. In addition, when training a deep neural network model, it is important to use appropriate loss functions. PSNR (Peak Signal to Noise Ratio) is a commonly recognised evaluation indicator in the field of image denoising that can be used as a loss function. The triplet loss is the most commonly used in person re-identification. Finally, we compare and contrast various types of methods on the dataset, including hand-crafted feature methods, metric learn- ing methods, deep learning methods, and combined denoising methods, and provide benchmark results

The face emotion recognition method using convolutional neural network and edge computing Hongli Zhang *et al.* [28] proposes the technique for recognizing facial emotions using convolutional neural network and image edge detection. It extracts different features from our original input facial expression through image processing technology. Face expression recognition method will classify features according to different human emotions such as surprise, happiness, neutrality and aversion [1] [2]. It also plays an important role in research of emotional quantification. This method uses computer as assistant tool and combines with specific algorithm

to judge the inner emotions of human face expression. The developments of this technique is as follows, First we will locate the face in our image and cut it out, then we normalize our image to a specific size and then edge of each layer of our image will be extracted by convolution technique. The separated edge data is superimposed on each feature image to protect the edge structure data of surface picture. Dimensionality reduction of our extracted features is done by maximum pooling method. Atlast the expression of our test sample image is classified by softmax classifier. Using our method we can extract facial features effectively.

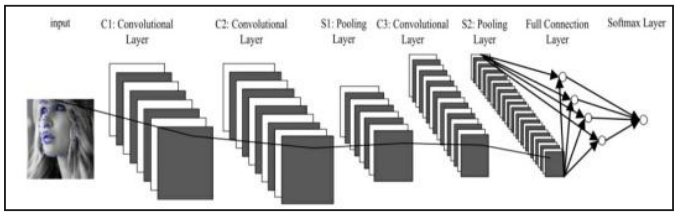

Fig. 7: CNN Structure for Facial Expression [28]

The Approaches on Partial Face Recognition: A Literature Review M R Reshma *et al.* [29] discuss about different methods for partial face recognition. For perceiving a discretionary picture caught in an unconstrained climate is a seriously difficult condition in video surveillance, handheld gadgets and so forth.

Various strategies have been concocted for unconstrained face recognition with amazing accomplishments. This paper gives an outline of various partial face recognition approaches proposed. The Face Recognition System (FRS) is relied upon to perceive a face with the assistance of assessing calculations, and by coordinating with an obscure face picture. It tends to be finished by contrasting the obscure face and the countenances put away in the data sets. When generally considering the Face acknowledgment framework it has three distinct stages, each one has a significant part in the last yield. The stages are face recognition, feature extraction (recognizes facial mile-stones) also, face recognition (coordinating with the data set of perceived faces). Brief feature extraction incorporates a portrayal of important highlights, its extraction as well as the inclination among a selection of highlights. A portion of the prominently utilized highlights are SIFT (Scale-invariant feature transform) [30] [31], LBP (Local binary patterns) [32] [33], Gabor and HOG (Histogram of oriented gradients) [34].

The customary utilization of face acknowledgment framework, particularly those utilized for access control, validation, also, observation, as a rule need greatest data of face to accomplish extraordinary execution in acknowledgment. At times it is hard to get access of a full face under certain confined circumstances. For instance, a non-agreeable face can prompt face impediment. In specific applications where pictures are caught by reconnaissance cameras or hand-held gadgets with less participation of human, the presentation of face acknowledgment frameworks are diminished because of the failure to access the face in the caught picture as it is impeded by certain articles like sunglass, scarf or different people which leaves just the fractional face on the pictures caught. The different approaches used for the partial face recognition are the keypoint-based approach, region-based approach, and CNN-based approach. The majority of the as of late proposed answers for partial face recognition depend on Keypoint based approach. A curiosity approach is proposed which is an arrangement free strategy for fractional face acknowledgment called as MKD- SRC [29], multiple keypoints descriptor- Sparse representation-based classification. In this technique numerous keypoints descriptor are inferred for facial attributes portrayal. At that point SRC is utilized for classification purpose. A system which is equipped for adjusting the test face picture to display pictures vigorously is proposed, it is called as Robust Point Set Matching (RPSM) [35]. This technique works in unconstrained circumstances like impediment, and outward appearances, and so forth. In Keypoint-based methodology for halfway face coordinating with they are utilizing SIFT descriptor, Speeded-Up Robust Features (SURF) descriptor [36] and LBP [37]. Aside from Keypoint based methodology for halfway face acknowledgment Region-based methodologies likewise offer arrangements. A Region-based strategy fundamentally relies upon the sub-locales which are the contribution to which the adjoining pixels inside one area have comparable esteem. The ordinary practice is to coordinate with pixels inside its neighbors. On the off chance that they are fulfilling the states of being comparative, at that point pixel can be set have a place with bunch as one or a greater amount of its neighbors. In this proposed a strategy with a wonderful presentation called Multi-Scale Region-based CNN (MR-CNN) [38] model for halfway face acknowledgment. In this technique at first, it deteriorates a picture with countenances or incomplete appearances into the changed district while it separates the highlights for every locale. At that point with the CNN, it empowers halfway face acknowledgment utilizing locale to area coordinating.

The development of deep Convolution Neural Organizations (CNN) acquired a sped up upgrade the execution of face recognition. The new methodology nearby for halfway face acknowledgment which goes under CNN-based is called Dynamic Feature Matching (DFM) [39]. It could oversee incomplete face pictures of irregular sizes without extra preprocessing with high veracity and processing exactness. In this methodology at first, the FCN is utilized to choose spatial highlights diagram of the slanted exhibition and the test faces. In the last stage, the SRC [40] is utilized to accomplish an arrangement free coordinating. SRC [40] offers a one-example per-class approach, for acquiring halfway face coordinating without arrangement. As to sum up the novel methodology of Dynamic Feature Matching, it mixes FCN with SRC [40], resolve progressed accomplishment in computational capacity and acknowledgment veracity. DFM works with halfway faces on a self-assertively estimated picture without arrangement. MKD-SRC in key point based approach and MR-CNN in region based approach are the best methods. CNN-based approach is better than other two and best method in CNN-based is DFM.

DFM is having a promising application in different video acknowledgment approaches in the future. The new exploration works are for the most part on the CNN-based approach.

The Multimodel Camera Based Gender Recognition using Human Body Image with Two Step Reconstruction Network Na Rae Baek *et al.* [41] propose a technique of gender detection for low resolution human full body pictures [26]. The existing gender recognition models uses high resolution facial images that are captured in short distance. In real life we know that images may be back-view or blurred images or with low resolution. Same gender people may wear different types of clothes or may have different hairstyle [42]. So existing method have different difficulties in gender recognition. To solve these problems in our method remote body shape based gender recognition was performed by two convolutional neural network which improved resolution of visible light pictures. This method improved performance of gender detection using deep CNN-based denoising and super resolution reconstruction. We used infra red camera for decreasing the background noise in the image. Then we combine visible light and IR image for describing the person's body shape [43]. The method has shown higher performance rate comparing to existing studies.

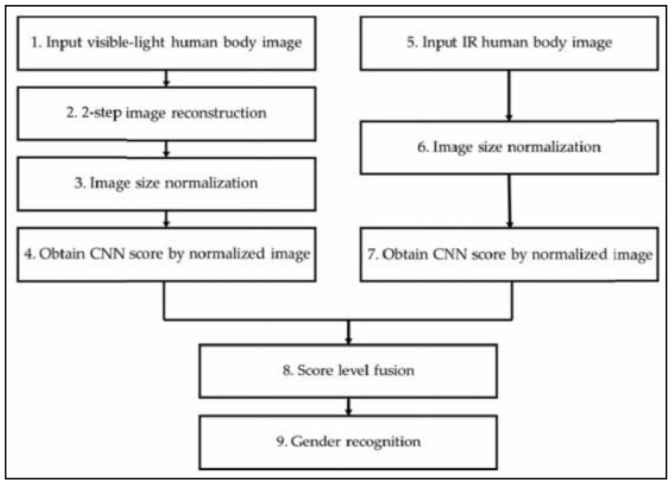

Fig. 8: Flowchart [41]

III. Conclusion and Future Work

Person re-identification proof from recordings utilizing face location was done in different settings and the elements influencing the exhibition of the models were distinguished. The quality of the recognizable proof relied upon factors like the determination of the face location calculation, the nature of the recordings gathered from surveillance cameras and the sort of recognizable proof being made. One of the method provides enhanced quality video with extraction of more evidence items. It is useful in intelligent investigation system with more efficiency. Face and Gender Recognition System integrates an image face recognition module with a gender recognition module to allow not only face but also gender recognition in complex backgrounds. KnightReid is a new dataset developed for use in night scenario person re-identification analysis. It is the first person re-identification dataset designed for a night scenario. To adapt to such a unique problem, the N3 denoising network was combined with the HACNN re-identification network after observing the properties of images in a night scenario. A new loss function that incorporates the triplet loss and the PSNR loss is proposed to jointly train the proposed pipeline in order to make the two types of networks compatible. A CNN model extracts facial features efficiently and is used to recognise facial expressions. Unlike conventional approaches, the proposed method will learn pattern features automatically and minimize incompleteness induced by artificial design features.

By utilizing one of the technique, ready to recognize faces with various scales i.e. particularly for little faces. Also, we can accomplish fulfilled recognition results on hard faces brought about by an average posture. The SFA strategy is quick, productive and strong to identify little faces. Also, it accomplishes promising execution on testing face recognition seat marks like more extensive face and FDDB datasets with competitive inference speed. The proposed approach uses training sample image data to directly input the image pixel value. By rising the highest achievable accuracy from 91.80% to 93.22%, we obtain the best results ever on the Adience dataset. Despite the fact that our work has enhanced the accuracy of face and gender recognition, there is still a lot of space for progress in gender recognition accuracy. Create a connection between current evidence and the item that was discovered.

References

[1] X. U. Feng, and J.-P. Zhang, "Facial microexpression recognition: A survey," *Acta Automatica Sinica*, vol. 43, no. 3, pp. 333-348, 2017.

[2] M. S. Zerdem, and H. Polat, "Emotion recognition based on EEG features in movie clips with channel selection," *Brain Inf.,* vol. 4, no. 4, pp. 241-252, 2017.

[3] J. Xiao, S. Li, and Q. Xu, "Video-based evidence analysis and extraction in digital forensic investigation," *IEEE Access*, vol. 7, pp. 55432-55442, 2019.

[4] S. Li, S. Li, K.-K. R. Choo, Q. Sun, W. J. Buchanan, and J. Cao, "IoT forensics: Amazon echo as a use case," *IEEE Internet of Things Journal,* vol. 6, no. 4, pp. 6487-6497, Aug. 2019.

[5] L. G. Mor, M. A. Brizuela, H. L. Ayala, D. P. Pinto-Roa, and J. L. V. Noguera, "Parameter tuning of CLAHE based on multi-objective optimization to achieve different contrast levels in medical images," in *Proc. IEEE Int. Conf. Image Process. (ICIP),* Sep. 2015, pp. 4644-4648.

[6] G. Gilboa, N. Sochen, and Y. Y. Zeevi, "Image enhancement and denoising by complex diffusion

processes," *IEEE Trans. Pattern Anal. Mach. Intell.,* vol. 26, no. 8, pp. 1020-1036, Aug. 2004.

[7] M. F. E. M. Senan, S. N. H. S. Abdullah, W. M. Kharudin, and N. A. M. Saupi, "CCTV quality assessment for forensics facial recognition analysis," in *Proc. 7th Int. Conf. Cloud Comput., Data Sci. Eng. Conuence*, Jan. 2017, pp. 649-655.

[8] S. Luo, X. Li, R. Zhu, and X. Zhang, "SFA: Small faces attention face detector," *IEEE Access*, vol. 7, 2019.

[9] X. Xiong, and F. D. la Torre, "Supervised descent method and its applications to face alignment," *Proc. CVPR*, Jun. 2013, pp. 532-539.

[10] A. Jourabloo, M. Ye, X. Liu, and L. Ren, "Pose-invariant face alignment with a single CNN," *Proc. ICCV*, Oct. 2017, pp. 3219-3228.

[11] F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: A unified embedding for face recognition and clustering," *Proc. CVPR,* Jun. 2015, pp. 815-823.

[12] Y. Wen, K. Zhang, Z. Li, and Y. Qiao, "A discriminative feature learning approach for deep face recognition," *Proc. Eur. Conf. Comput. Vis.*, Oct. 2016, pp. 499-515.

[13] V. Mathew, T. Toby, A. Chacko, and A. Udhayakumar, "Person re-identification through face detection from videos using Deep learning," *2019 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS)*, 2019.

[14] Muhd. U. Yaseen, A. Anjum, and O. Rana, "Deep learning hyper-parameter optimization for video analytics in clouds," *IEEE Transactions on Systems Man and Cybernetics: Systems*, vol. 49, no. 1, pp. 253-264, Jan. 2019.

[15] K. Zhang, Z. Zhang, Z. Li, and Y. Qiao, "Joint face detection and alignment using multitask cascaded convolutional networks," *IEEE Signal Processing Letters,* vol. 23, no. 10, pp. 1499-1503, Oct. 2016.

[16] R. Zhao, W. Oyang, and X. Wang, "Person re-identification by Saliency learning," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 39, no. 2, pp. 356-370, Feb. 2017.

[17] M. Cokun, O. Yldrm, and Y. Demir, "Face recognition based on convolutional neural network," *International Conference on Modern Electrical and Energy Systems (MEES)*, 2017, pp. 376-379.

[18] Y. Zhou, H. Ni, F. Ren, and X. Kang, "Face and gender recognition system based on convolutional neural networks," *2019 IEEE International Conference on Mechatronics and Automation (ICMA)*, vol. 41, 2007, pp. 59-72.

[19] Y. Sun, and D. Liang, X. Wang, and X. Tang, "Deepid 3: Face recognition with very deep neural networks," arXiv preprint arXiv:1502.00873, 2015.

[20] F. Ren, and Z. Huang, "Automatic facial expression learning method based on humanoid robot XIN-REN," *IEEE Trans Hum Mach Syst*, vol. 46, no. 6, pp. 810-821, 2016.

[21] Q. Cao, L. Shen, W. Xie, O. M. Parkhi, and A. Zisserman, "Vggface2: A dataset for recognising faces across pose and age," in *IEEE Conference on Automatic Face and Gesture Recognition (F&G),* 2018.

[22] O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," *British Machine Vision Conference (BMVC),* Sep. 2015, pp. 41.1-14.12.

[23] J. Zhang, Y. Yuan, and Q. Wang, "Night person re-identification and a benchmark," 18 Jul. 2019, pp. 95496-95504.

[24] L. Sun, Z. Jiang, H. Song, Q. Lu, and A. Men, "Semi-coupled dictionary learning with relaxation label space transformation for video-based person re-identication," *IEEE Access*, vol. 6, pp. 12587-12597, 2018.

[25] D. Gray, and H. Tao, "Viewpoint invariant pedestrian recognition with an ensemble of localized features," in *Proc. ECCV*, Munich, Germany, 2008, pp. 262-275.

[26] D. Liu, B. Wen, Y. Fan, C. C. Loy, and T. S. Huang, "Non-local recurrent network for image restoration," in *Proc. NIPS*, Paris, France, 2018, pp. 1673-1682.

[27] W. Li, X. Zhu, and S. Gong, "Harmonious attention network for person re-identication," in *Proc. CVPR*, Salt Lake City, UT, USA, Jun. 2018, pp. 2285-2294.

[28] H. Zhang, A. Jolfaei, and M. Alazab, "A face emotion recognition method using convolutional neural network and image edge computing," *IEEE Access*, pp. 159081-159089.

[29] M. R. Reshma, and B. Kannan, "Approaches on partial face recognition: A literature review," *2019 3rd International Conference on Trends in Electronics and Informatics (ICOEI)*, pp. 185-198.

[30] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," *Int. J. Comput. Vis.*, vol. 60, no. 2, pp. 91-110, 2004.

[31] D. G. Lowe, "Object recognition from local scale-invariant features," *IEEE International Conference on Computer Vision*, vol. 2, p. 1150, 1999.

[32] T. Ahonen, A. Hadid, and M. Pietikainen, "Face description with local binary patterns: Application to face recognition," *IEEE TPAMI*, vol. 28, no. 12, pp. 2037-2041, 2006.

[33] W. Zhang, S. Shan, W. Gao, X. Chen, and H. Zhang, "Local gabor binary pattern histogram sequence (LGBPHS): A novel non-statistical model for face representation and recognition," in *IEEE ICCV*, vol. 1, pp. 786-791, Oct. 2005.

[34] N. Dalal, and B. Triggs, "Histograms of oriented gradients for human detection," in *IEEE Computer*

Society Conference on Computer Vision and Pattern Recognition, 2005 (CVPR'2005)," vol. 1, pp. 886-893, Jun. 2005.

[35] R. Weng, J. Lu, and Y.-P. Tan, "Robust point set matching for partial face recognition," *IEEE Trans. Image Process.*, vol. 25, no. 3, pp. 1163-1176, Mar. 2016.

[36] H. Bay, T. Tuytelaars, and L. V. Gool, "Surf: Speeded up robust features," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2006, pp. 404-417.

[37] T. Ahonen, A. Hadid, and M. Pietikinen, "Face description with local binary patterns: Application to face recognition," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 28, no. 12, pp. 2037-2041, Dec. 2006.

[38] L. He, H. Li, Q. Zhang, Z. Sun, and Z. He, "Multiscale representation for partial face recognition under near infrared illumination," in *Proc. IEEE Int. Conf. Biometrics Theory, Appl. Syst. (BTAS)*, Sep. 2016, p. 17.

[39] L. He, H. Li, Q. Zhang, and Z. Sun, "Dynamic feature matching for partial face recognition," *IEEE Trans. on Image Processing*, vol. 28, no. 2, Feb. 2019.

[40] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210-227, Feb. 2009.

[41] N. R. Baek, S. W. Cho, J. H. Koo, N. Q. Truong, and K. R. Park, "Multimodal camera-based gender recognition using human-body image with two-step reconstruction network," *Label Information of Sun Yat-sen University Multiple Modality Re-ID (SYSU-MM01) Database and CNN Models,* May 2019.

[42] S. Baluja, and H. A. Rowley, "Boosting sex identification performance," *Int. J. Comput. Vis.*, vol. 71, no. 1, pp. 111-119, Jan. 2007.

[43] J. Mansanet, A. Albiol, and R. Paredes, "Local deep neural networks for gender recognition," *Pattern Recognit. Lett.*, vol. 70, pp. 80-86, Jan. 2016.