



## Comparative Analysis of Machine Learning Based Approaches for Face Detection and Recognition

Ratnesh Kumar Shukla\* 

\*Corresponding Author, Ph.D. Candidate, Department of Computer Science & Engineering, Dr. APJ Abdul Kalam Technical University, Lucknow, Uttar Pradesh, India. E-mail: ratnesh.nittr@gmail.com

Arvind Kumar Tiwari

Associate Professor, Department of Computer Science & Engineering, Kamla Nehru Institute of Technology, Sultanpur, Uttar Pradesh, India. E-mail: arvind@knit.ac.in

### Abstract

This article discusses a device focused on images that enables users to recognise and detect many face-related features using the webcam. In this article, we are performing comprehensive and systemic studies to check the efficacy of these classic representation learning structures on class-imbalanced outcomes. We also show that deeper discrimination can be learned by creating a deep network that retains inter-cluster differences both and within groups. MobileNet, which provides both offline and real-time precision and speed to provide fast and consistent stable results, is the recently suggested Convolutional Neural network (CNN) model. The recently proposed Convolutional Neural Network (CNN) model is MobileNet, which has both offline and real-time accuracy and speed to provide fast and predictable real-time results. This also solved a problem related to the face that occurs in the identification and recognition of the face. This paper presents the different methods and models used by numerous researchers in literature to solve the issue of faces. They get a better result in using the highest number of layers. It is also noted that the combination of a machine learning approach with multiple image-based dataset increases the efficiency of the classifier to predict knowledge related to face detection and recognition.

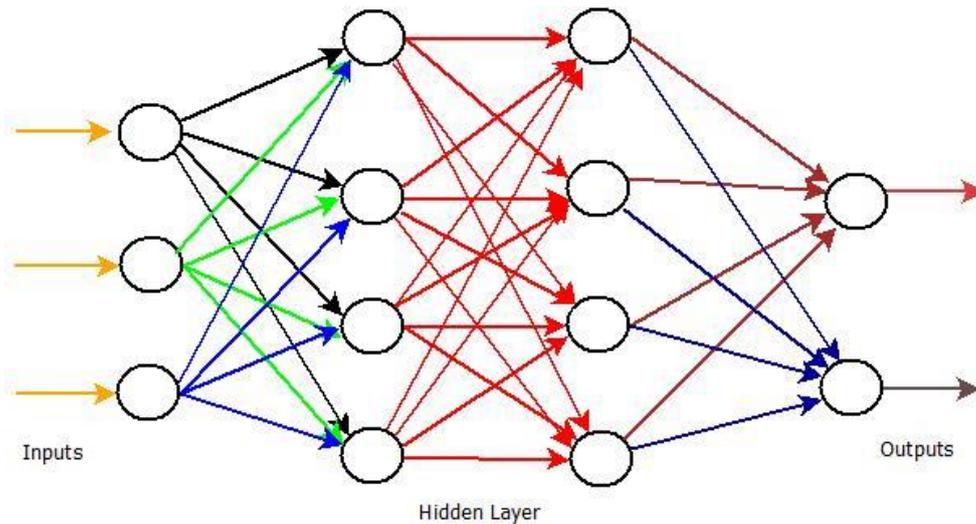
**Keywords:** Biometric machine, Convolution neural network, Deep neural network, Facial action unit, Random convolution neural network etc.

## Introduction

In current scenario face detection and recognition is very complicated and challenging problem in Pattern Analysis, Computer Vision, Neural Networks and Machine Learning. This problem is discussed in various learning communities such as controlled environment and uncontrolled environment. The facial applications are accepting 2D facial images and acquired different facial descriptors to use different learning techniques. In recent development deep learning has found a new approach for researchers in artificial intelligence in face detection and recognition. It is finding of the objective goals and it is very popular rapidly growing technique to use in face detection and recognition. It has developed to solve the complex problem of machine learning. It has worked on human and computers related problem. Deep learning is an advance idea and technique based concept to detect the faces for identification and verification. In deep learning the computer likes as learners and performs the classification task directly to the next video and images. It is solving the state of art problem and improves the accuracy and reliability of human. Deep learning has bounded and trained labeled data in neural network architecture. That contains many numbers of layers in convolution neural network. Hidden layer has worked as a bridge between input and output layers. They are working for simple and complex shape of images. It solves the identification and verification problems using maximum number of layers which are found in hidden layer. The first introduction of deep learning and theory is produced in 1980s. The following reasons they have become a powerful concept for feature extraction using in face detection and recognition.

## Convolutional Neural Network

It is the best concept of machine learning. This has been used in Pattern analysis, Computer vision and Machine learning. Convolution neural network is a best technique to analyse the previous knowledge or data for solved the future problem. It is depending on the transform an input image to convert in output image using multiple hidden layers. It is associated with machine learning and pattern analysis. The researchers are focusing on identification and verification problem of the images and videos. They have found a lot of solutions to solve the problem of face detection and recognition. In deep learning solution the term deep is refers to the concept of numbers of hidden layers present in max pooling networks. Generally they are two to three layers, but deep networks analysis hundreds of layers. While deep learning is working on face detection and recognition, it achieved the higher accuracy. They can help using their expectations. Deep learning is working as a robot because this is including artificial intelligence. It can learn their features with the help of feature in hidden layers.

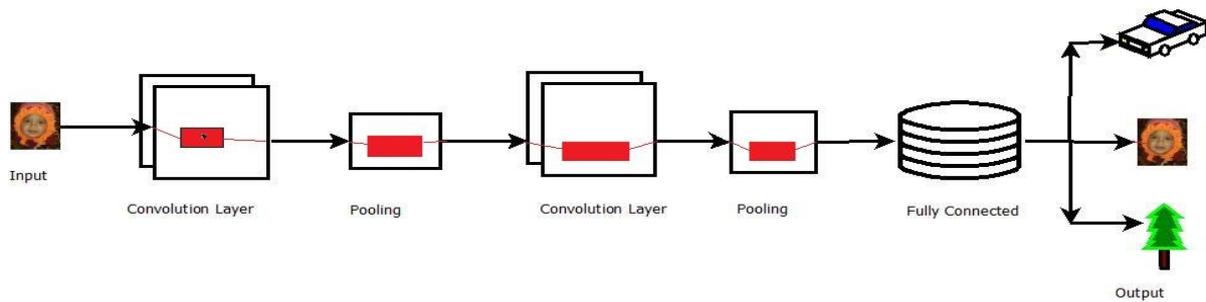


**Figure 1. Architecture of convolutional neural network from relationship between input, output and hidden layers**

In Figure 1 is showing architecture of convolution neural network from their relationship between input, output and hidden layers. Neural network architecture is organizing in the layers of consistency in a set of interconnected node. Which are known as hidden layers? They are classified the feature of the input images and mapped into the stored images. Hidden layer is occupied maximum number of layers regarding users. These have classified different-different features in simple and complex image in datasets. Figure 1 represents the architecture model of convolution neural network to including in the combination between input, hidden and output layers. These layers are the most important part of the convolution neural network, because they are reading all the features of the input images and classifying according their features in hidden layers. Then organize the output of the input images.

### **Machine Learning Methods in Face Detection and Recognition Algorithms**

It has been used for large number of labeled data present in datasets. That is matching pattern of features directly involved in datasets and using without manual feature extractions. A deep neural network is combination of non-linear processing layers using in machine learning and pattern analysis. They have interconnected with multiple layers like that input layer for input data, hidden layer for features and output layers for object. Deep learning is solving the problem of faces in the real time images and databases. They are working on the collection of images and normalized the networks. Deep learning has a latest idea and concept to improve the face related problem in real time environment.



**Figure 2. Processing flow of deep neural network**

## Alex Net

Alex Net is the first deep learning architecture which was introduced in 1980s by Geoffery Hinton and his colleagues. It is very popular and simple model for researchers to solve the problem in face detection and recognition. It has a simple architecture to combination of convolutional layer and max pooling layers. In deep neural network there is a combination of convolutional layer, multiple pooling and fully connected layer. They have collected bunch of information from the input and the filtered it in convolution layer including different type of hidden layers. So there are various categories of input object. In figure 2 deep neural network collects the input image and identifies or categories the object. They categorize the object and give the efficient outputs. Deep convolution neural network is working in different layers. Convolution layer contains the input image through a set of transform neural network. These transform learning and recognize the features of the images. Pooling layer is simplifying the output by performing non-linear reducing, down sampling the number of dimensions for using about the features of images. Fully connected layer is playing a role to recognize the object. After completing the feature detection in the deep convolution network they will go to next layer called fully connected layer.

## Feed forward Neural Network

It has most popular algorithm in deep learning concept. It will work on the feature of the faces in machine learning. It has worked to identify the features from input layers. They are not using manual feature extraction. So they have reduced the time for identifying features in. It has solved and classified the features in databases. Figure 3 Shows architecture model feed forward network containing both weight and bias features. They are also containing the three layer concept, so it has assigned multiple neural networks. In this model  $h$  is known has hidden layers shown in figure 3. They have performed  $vo_i$  (output layer),  $h$  has hidden layer and  $wo_i$  as bias layers. Bias has assigned weights in network model. Figure 3 architecture of feed forward network is found three layers. There are proving a good result to communicate between collections of layers. The input layers  $x_1$  to  $x_n$  have connected to the hidden layer  $h_1$  to  $h_n$  layers. Those have connected to output layers  $y_1$  to  $y_n$  and communicated and

interconnected with weight and bias in network model. They are flowing in reverse direction. Computational complexity is depending on the amount of hidden layers. They are performing high accuracy in network models. The bias is more effective in working with input and output layers.

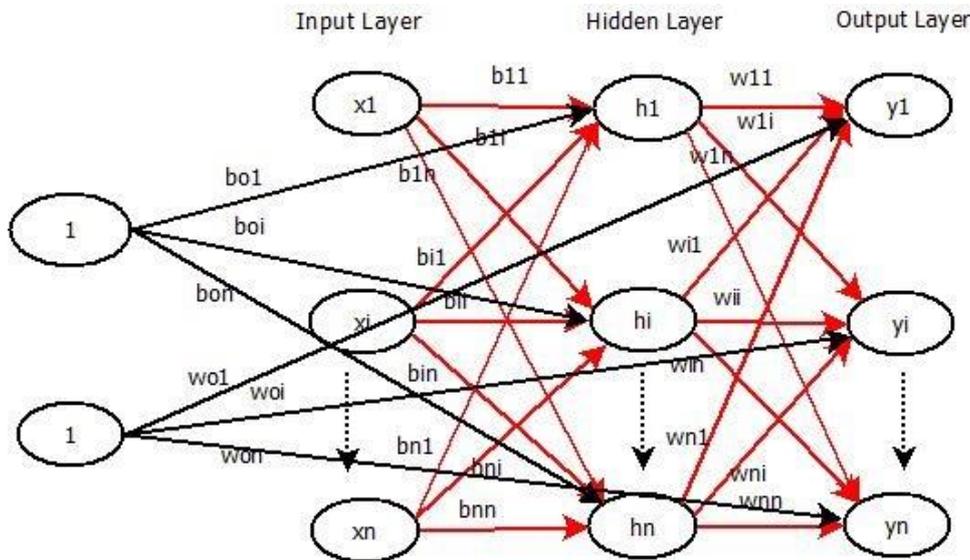


Figure 3. Architecture of feed forward network using Bias and weight

### Deep Convolution Neural Network

It has communicated non-linear model using maximum number layers. It has operated the faces and recognizes using machine learning. This network has assigning multiple hidden layers for used multiple feature of the faces. These layers have worked with neurons or nodes. DCNN has processed the set of objects and after processed, they have automatically recognized the object face to corresponding input images. In DCNN the labelled data worked has a training data in datasets. It has been used to understand the images and matched these features in matching categories in datasets. DCNN has transferred the previous data to collecting features in next layers of the architecture. DCNN has improved the complexity and accuracy of the object. They are working on the pattern from layer to layer in DCNN.

### K-nearest Neighbors

KNN algorithm is one of the best algorithms for introducing the feature of nearest neighbor or nodes. This is also known as lazy algorithms. K means the number of iterations. This algorithm is using simple to understand and communicate to each other's. K-nearest neighbor algorithm is working on nearest node of features of the object. So it has worked as a non-parametric algorithm. The term non-parametric is not making any assumptions to underlying

data distribution. This has been used real word databases system. Maximum practical data does not follow the typical theoretical analysis in feature learning process. KNN algorithm is assuming the features space in nearest feature in the database.

### **Principal Component Analysis**

It has a very popular and oldest technique in the image processing, and pattern analysis. This has developed by Pearson (1901) and improved by Hotelling (1933). It has worked on the Eigen value and Eigen vectors using matrix form. It has used different applications with different variety. The concept of PCA has to reduce the dimensionality features of databases. It is capable to large number of datasets. Their features have interrelated to remaining possible databases in the system.

### **Linear Discriminant Analysis**

This method has similar as fisher discriminant analysis. It is used to describe the images including local features. These features have worked in the form of pixel value. They have following defined as known shape feature, texture feature and color features. It has identified the features to use between linear separating vectors. And also have used similar feature in the image. These procedures have been used to maximize between class scatter and intra class variance in face detection and recognition.

### **Feature Extraction**

It has a very popular concept used to extract the features from images in face detection and recognition. It has been widely used in different approaches, such as Digital Image processing, Pattern Classification, Computer Vision, and Deep Learning. It has transformed the input materials or images in pixel. This pixel value has transformed a combination of features in database. Because these selected features are containing most appropriate information in the original data. It is very useful in biometrics applications and Machine Learning.

### **Geometric Based Methods in Feature Extraction**

Geometric features based approaches is computing a set of faces such as a mouth, eyes, ears and nose. In this geometric representation shows the position of eyes, mouth, ear and nose is a form of feature vectors. They are reliable to detect the automatic feature extraction and significant for face detection and recognition. Geometric feature are representing the shape, location and color of the facial components, which are extracting a feature component for recognizing the faces.

## **Holistic Based Methods in Feature Extraction**

It has most useful technique in face detection and recognition. They are using feature description methods based appearance in face detection and recognition. Holistic used feature extraction in for any local extraction method is a form of information of data and reduced a typical transformation that describes a large data from images in the database. Holistic based feature extraction is converting the image into a low- dimensional feature space with improving discriminant power of the faces.

### ***Related work of Machine Learning Approaches in Face Detection and Recognition***

In literature survey from various researchers, they have proposed the machine learning approaches. They have found several solutions and techniques for detection of faces and recognizing it. (Lee, Won, & Hong, 2020) proposed assembled ResNet-50 reveals increase in top-1 accuracy from 76.3 per cent to 82.78 per cent, mCE from 76.0 per cent to 48.9 per cent and mFR from 57.7 per cent to 32.3 per cent of ILSVRC2012 validation collection. (Bau, Zhu, Strobel, Lapedriza, Zhou, & Torralba, 2020) has studied a convolution neural network (CNN) that is specialized in the classification of scenes and discover units that fits a complex range of object concepts. (Liu, et al., 2020) launched a FinRL DRL library that will make it easier for beginners to invest in quantitative finance and develop their own stock trading strategies. (Carion, Massa, Kirillov, & Zagoruyko, 2020) proposed key components of the new system, called DETection TRansformer or DETR, are a set-based global loss that forces specific predictions by bipartite matching, and transforming encoder-decoder architecture. In view of Fixed limited collection of learned object questions, DETR explanations for object relations, and global image meaning for direct production.

(Agarap, 2019) proved that in the case as CNN-SVM achieved test accuracy of 90.72 per cent, while CNN-Softmax achieved test accuracy of 91.86 per cent. (Chopade, Edwards, Khan, & Pu, 2019, November) proposed explicitly interested in how teamwork skills and team interactions are demonstrated as verbal and non-verbal actions, and how these behaviours can be recorded and evaluated by passive data collection of F. (Trabelsi, Chaabane, & Ben-Hur, 2019) worked on deepRAM, an end-to-end deep learning platform that provides for the implementation of novel and previously proposed architectures; its fully automated model selection process allows for a rational and impartial evaluation of deep learning architectures. A work (Bali, Kumar, & Gangwar, 2019) proposed a speed of wind speed weather forecasting to calculate the wind speed to using deep learning technique. They are using different approaches to provide and solution of wind speed weather forecasting. (Howard, et al., 2019) created a model and generalized and extended to object recognition and semantic segmentation tasks. They suggest a new effective segmentation decoder Lite Reduced Atrous Spatial Pyramid Pooling for the role of semantic segmentation (or any dense pixel prediction) (LR-ASPP). (Shao, Zhang, & Fu, 2017) proposed a Sparse many to one

encoder (SMF) and collaborative random faces (RFs). They have worked on pose invariant face representation and detect the faces. Author is working on different paper using Multi PIE pose datasets; you tube datasets (YTF) and real world datasets. They have improved the performance from 7 to 14% in face detection and recognition. (Tsai, Li, Hsu, Qian, & Lin, 2018) proposed an unsupervised learning framework and joint optimization framework. This technique has enhancing co-segmentation mask for improving the co-saliency features. Unsupervised learning and joint optimization framework is exploring the concept of the objectness and saliency in different type of multiple images or datasets Cosal2015, iCoseg, Image Pair and MSRC datasets are providing high quality result on both co-saliency and co-segmentation.

Hu, Cho, Wang, & Yang (2014) proposed a Non-blind deconvolution method to suppress the ringing artifacts caused by light for face detection and recognition. Non-blind deconvolution method is detecting light stretches for corrupted images and incorporates it into an optimizing framework for face detection and recognition. The author has worked on png and jpeg images using low light environment. (Anwar, 2018) proposed to deblurring for class specific problem and class genetic blind deconvolution. This proposed method has used to overcome the limitation of existing method. When dealing with blurred image lacking high frequency. This is focusing only blurred images containing a single object and class specific training with using CMUPIE, CAR, FTHZ and INRIA datasets. (Wang, Ma, Chen, & You, 2017) proposed RegionNET or RexNet and Salient Object detection techniques to solve the problem in face detection and recognition. RexNet has provided saliency mapping between end to end with shape object boundaries for VGG, ImageNet, ContexNet, ECSSD, DOTOMRON and RGBD1000 datasets. A RexNet has worked on a detection and multi scale conceptual robustness technique in face detection and recognition. (Deng, Hu, & Guo, 2018) proposed an image filtering binarization and spatial histogram for face detection and recognition. The author has developed scattering Compressive Binary Pattern (SCBP) descriptor to improve the performance of face image. SCBP is using handcraft by 6RF Eigen filters to achieved accurate and robust performance. CBP is also used to improving the robustness of LBP. In this paper authors used to DFD and Cbfd for derived noise sensitive filtering adapting to fine grained structure for FERET, LPW and PaSC databases in Face detection and recognition. Koteswar Rao *et.al.* (2018) used co-saliency estimation method in different datasets for face detection and recognition. Co-saliency estimation method used simple scale estimation application as demonstrated on the large scale ImageNet, MSRC, iCoseg and Coseg-Rep datasets. This method is solving map problem with well separated background and foreground. This framework is able to achieve very competitive results. (Tulyakov, Jeni, Cohn, & Sebe, 2017) proposed a regression forest algorithm for detecting the problem of face detection and recognition. They have solved the problem in efficient manner with consistency. These have robust in 3D face rotation of MultiPIE, HBPD, BU-4DFE and

MultiPIE-VC datasets. These approaches are finding effective face pose estimation and viewpoint consistency on multiple measurements. This method is performing highly competitive score on a range of benchmarks. (Dong, Zheng, Ma, Yang, & Meng, 2018) proposed a Multi model and Self-Paced Learning Algorithms for Detection (MSPLD) and few example object detection (FEOD) for face detection and recognition. This model has used a large number of pools in unlabeled image and using a few labeled images per category. This is being used for discriminating knowledge for different detection model. This method is giving better result in PASCAL-VOC2007, PASCAL-VOC2012, MSCOCO2014, ILSURC2013 and ImageNet-COCO datasets.

(Mafi, Rajaei, Cabrerizo, & Adjouadi, 2018) proposed a Switching based Adaptive Median and Fixed based Weighted Mean Filter (SAMFWMF) for face detection and recognition. SAMFWMF has controlled the similar edge detection and sharpening in Lena (5128512), Cameraman (250\*250), Coins (300\*246) and Checkboard (256\*256) images. SAMFWMF is performing better structural metrics. They are solving better result in contract to other common thresholding method in detecting the faces and then recognizing it. (Tao, Guo, Li, & Gao, 2017) proposed a model of Tensor Rank Preserving Discriminant Analysis (TRPDA) technique to solve the problem of face detection and recognition. They have performed robust performance and produced the high rate in UMIST, ORL and CAS-PEAL-R1 datasets. TRPDA has extracted the feature with the rank information and elimination. They are usable manifold learning method in face detection and recognition. (Wang, Yan, Cui, Feng, Yan, & Sebe, 2018) proposed recurrent face aging (RFA) and RNN in face detection and recognition. RFA has improved 65.43% and the accuracy of bilayer has 61%. It shows RFA works slightly better than bilayer RNN. The author is using LFW and CACD datasets for better performance in face detection and recognition. RFA framework consists of triple layer GRU, There are giving better performance of the identity information in bilayer GRU for face detection and recognition. (Jeong, Lee, Kim, Kim, & Noh, 2017) used Markov random field energy optimization. This method has especially worked as wide base line multi view environment. This segmentation method has improved the performance with similar quality. That has produced in the current state of art models. They have captured features in various critical conditions. They have included maximum number of rotations, views and distance between captured by cameras. A sparse wide baseline has captured very efficiently. (Sagonas, Ververas, Panagakis, & Zafeiriou, 2017) proposed Joint and Individual Variance Explained (JIVF) and Robust JIVE (RJPVE) in face detection and recognition. It has improved the accuracy of faces. It has also identified information of the faces remain in the RJIVE based progression of FG-NET datasets. The accuracy is depending on pair of images converted to age difference. (Zhu, Liu, Lei, & Li, 2017) worked on a 3D Dense Alignment (3DDFA) and 3D Morphable model (3DMM) in face detection and recognition. Face alignment has covered with full pose range. It has achieved poses variation in face alignment

of ALFW, AFW, LFPW, HELEN, IBUG, 300W and AFLW 2003D datasets. Comparing performance the drop brought is replacing boundary poses. This method demonstrates best robustness initialization 3DDFA person in face detection and recognition. (Wang, Yan, Cui, Feng, Yan, & Sebe, 2018) used one to more different techniques. That has been used in face detection and recognition. (Ding, Xu, & Tao, 2015) proposed Controlled Face Feature (CPF) in face detection and recognition. Using CPF in faces of extensive experiments shows the demonstrate superiority in both learning representation and rotating frontal images. Face recognition experiment on MultiPIE database is providing more evidence that illustrate the task of strength in specific methods. (Qian, Deng, & Hu, 2018, May) proposed a facial action unit in face detection and recognition. Facial action unit has derived to solve the identification problem using comprehensive computer vision algorithms. These datasets are DISFA and AM-FED used for color features. This color can also be used to detect of action unit activation. (Cruz, Foi, Katkovnik, & Egiazarian, 2018) proposed a single image super resolution (SISR) and CNN for face detection and recognition. 1D wiener filtering is working on similarity domains. They are giving effective solution for specific problem of SISR. These results is sharper reconstruction and in SET5, SET14 and Urban datasets. This method works well only on image substantial self-similarity. (Wang, Ma, Chen, & You, 2017) proposed an algorithm to solve the cross age face verification for face problem in face detection and recognition. These have comparatively worked on effective balance between feature share and feature exclusion. (Malhotra, Bali, & Paliwal, 2017, January) proposed to solve the problem of cryptography and network security to using intrusion detection system. These techniques have solved and implement the problem of machine learning very frequently. (Duong, Quach, Luu, N., & Savvides, 2017, October) proposed a Temporal Non-Volume Preserving (TNVP) and Generative Adversarial Network (GAN) for face detection and recognition using FGNET, MORPH, CACD and AGFW datasets. TNVP has evaluated both terms of synthesizing progressed faces of ages and cross face verification age with consistency. TNVP has guaranteed attractable density function. They have extracted features information and inference the value of consecutive stage of faces in evaluation of embedded datasets. (He, Wu, Sun, & Tan, 2017, February) proposed visual verses infrared (VSS-NIR) and invariant deep representation (IDR) using CASIA, NIR-VIS2.0 and Large Scale VIS datasets for face detection and recognition. They are achieving 94% verification rate in large scale VIS data comparing with state of art. This is reducing the error rate of 58% only with a compact 64 D representation.

### ***Summary Sheet of Machine Learning Approaches for Face Detection and Face Recognition***

Table 1 shows the summary of the machine learning approaches for face detection and recognition in images and videos. Table 1 also presents all the methods of face detection and

recognition by the various technique and datasets with results. Table 1 also shows the merit and demerit of the technique by using various authors.

**Table 1. Summary sheet of technique, authors, datasets, results, merit and demerits of the machine learning approaches in face detection and recognition**

S. No.	Author's	Data Sets	Technique	Result	Merit	Demerit
1	(Lee, Won, & Hong, 2020)	ISLRVC2012, CVPR2019	CNN	It will improve accuracy 76.3% to 82.78%.	It will provide backbone of the network performance.	It will provide better result in CVPR2019.
2	(Carion, Massa, Kirillov, & Zagoruyko, 2020)	COCO object detection	Detection Transformer (DETR)	It has generally performed panoptic segmentation used in unified manner.	It is used very frequently and used other or extra special library.	It has worked on small database to provide better results.
3	(Bau, Zhu, Strobel, Lapedriza, Zhou, & Torralba, 2020)	PASCAL365	CNN and GAN	It has developed and improved the classifier in the networks.	To solve critical task to over large datasets.	It has not improved classification accuracy.
4	(Agarap, 2019)	Google TensorFlow, CNN-SVM	MNIST	It is improving good result of classification accuracy.	It will provide better accuracy of CNN based test models.	They are not involving data processing of the fashion MNIST.
5	(Chopade, Edwards, Khan, & Pu, 2019, November)	NLTK, CIS frame, few team members	Collaborative Problem Solving (CPS) and Multimodal framework	It is improving performance of collaborative scientific educational area.	It has been developed only performance of industrial and organizational SEL skills.	It is not working good result in low level feature extraction.
6	(Trabelsi, Chaabane, & Ben-Hur, 2019)	Bioinformatics online datasets, deepRAM	Hybrid CNN/RNN	It is helpful to improve the model accuracy using CNN.	To improving deep features of biological sequence data.	It is commonly used for prediction side of DNA and RNA
7	(Howard, et al., 2019)	MobileNet V3	NetAdapt algorithm	MobileNetV3-Large LRASPP is 34% faster than MobileNetV2 R-ASPP	These models are directly taking full-resolution inputs.	It will work very efficiently on new work generation algorithms.
8	(Shao, Zhang, & Fu, 2017)	Multi PIE, you Tube Database (YTF) and Real world database	Collaborative Random Faces(CRF) and Sparse Many-to-one Encoder (SME)	SME and RF model are improving 7% on MultiPIE and 14% You Tube Database(YTF)	RF Model is working on pose variant faces. SME works on comparing multiple image to single image	RF and SME are not giving positive results on constraint poses.

S. No.	Author's	Data Sets	Technique	Result	Merit	Demerit
9	(Tsai, Li, Hsu, Qian, & Lin, 2018)	Image Pair, iCoseg, Cosal2015, MSRC	Unsupervised learning and joint optimization framework has improved the performance of co-segmentation. They have also improved co-saliency priors.	This technique has explored the concept of saliency and objectness in different type of multiple images or datasets.	This method has worked on publically using datasets. These techniques have provided high quality result on both co-saliency and co-segmentation.	The object segmentation iteratively works out the region-wise. Adaptive saliency map fusion has transfer useful information to different task.
10	(Hu, Kan, Shan, Song, & Chen, 2017, May)	JPEG & PNG, 700*1000 resolution and 40 natural low-light images.	Non blind Deconvolution scheme has to suppress to the ringing artifacts covered by the light.	This algorithm has performed favorably against state-of-art deblurring for low-light images.	This method has detected light stretches in blurry images. And incorporates them into an optimization frame works.	Non blind Deconvolution method has not generated satisfactory results in present of drastic loss of information.
11	(Anwar, 2018)	CMU PIE, CAR dataset.	This technique has worked on deblurring with the class specific priors.	They have improved to result to comparatively previous existing methods.	This method has focus on deblurring of image.	They have not worked on blur method.
12	(Wang, Zhou, Kong, Currey, Li, & Zhou, 2017, May)	VGG, ECSSD, DUTOMRON, RGBD1000	CNN, R-CNN, RegionNET OR RexNET, Salient object detection.	RexNet is providing saliency mapping between end to end with sharp object boundaries.	RexNet has achieved clear detection boundary. And it has also achieved multi scale conceptual robustness.	RexNet is based on the segmentation of images.
13	(Deng, Hu, & Guo, 2018)	FERET, LPW, PaSC	Image filtering binarization and spatial histogram, Scattering Compressive Binary Pattern (SCBP)	SCBP descriptor is handcraft by 6RF Eigen filters. They are sufficient to achieve accurate and robust performance.	CBP is improving the robustness of LBP. DFD and Cbfd are derived noise sensitive filtering adapting to fine grained structure.	The major issue has solved by an optimized descriptors.
14	(Tulyakov, Jeni, Cohn, & Sebe, 2017)	MultiPIE, HBPD, BU-4DFE, MultiPIE-VC	Regression forest based algorithms	This has improved consistency and computationally in efficient manner.	They have found effective result in consecutively multiple related features in head pose and relative approaches.	This method has performed highly competitive accuracy.
15	(Dong, Zheng, Ma, Yang, & Meng, 2018)	PASCAL VOC2007, PASCAL VOC2012, MSCOCO 2014.	Multi Model Self-Paced Learning for Detection (MSPLD), Few example object detection (FEOD)	Object detection has used large scale unlabeled image. But also used few labeled image in some category.	It has used discriminative knowledge to solve the problem of images using for different image detection model.	They have not detected every complicated image in datasets.

S. No.	Author's	Data Sets	Technique	Result	Merit	Demerit
16	(Mehdipour Ghazi & Kemal Ekenel, 2016)	Lena (512*512), Camera Man (250*250)	Switching Adaptive Median and Fixed Weighted Mean Filter (SAMFWMF)	The similarity of edges has adopted the properties using median filters.	SAMFWMF is performed better structural metrics.	They have provided better result in high intensity impulse noise with edges.
17	(Tao, Guo, Li, & Gao, 2017)	UMIST, ORL, CAS-PEAL-R1	Tensor rank preserving discriminant analysis(TRPDA)	TRPDA has provided highest recognition rates and better performance.	They have extracted extract feature with the rank module. They have unstable manifold learning methods.	They have worked on two order tensor.
18	(Wang, Yan, Cui, Feng, Yan, & Sebe, 2018)	LFW, CACD	Recurrent face Aging (RFA) RNN	The accuracy has shown that RFA is worked better than RNN.	The triple layer RFA framework GRU gives the better identity information than bi layer GRU.	RFA framework does not work on integrate the age estimation.
19	(Jeong, Lee, Kim, Kim, & Noh, 2017)	Chair, Car Bike and Bus	Markov Random Field energy optimization.	This method has especially based on wide baseline multi view environment.	They have captured images in various conditions.	These systems are using sensitive camera parameters.
20	(Sagonas, Ververas, Panagakis, & Zafeiriou, 2017)	FG-NET	Joint and Individual Variance Explained (JPVE), Robust-JIVE (RJIVE)	They have improving accuracy and validate the identity of information.	They have improved accuracy when the differences of each pairs are maximum.	They have produces the problem of age in invariant faces.
21	(Zhu, Liu, Lei, & Li, 2017)	ALFW, AFW, LFPW, HELEN, IBUG, 300W.	3D Dense Face Alignment (3DDFA) and 3D Morphable Model (3DMM)	Face alignment has worked on pose range and also worked on 3D Morphable model.	This method has replaced the boundary boxes using 3DDFA.	They have produced large artifacts and invisible region filling.
22	(Ding, Xu, & Tao, 2015)	Multi PIE	Controlled Face Feature (CPF)	These methods have worked on superiority in oth learning representation.	Face recognition experiment on MultiPIE database provide more evidence and strength.	Unsupervised way is universal to all datasets.
23	(Benitez-Quiroz, Srinivasan, & Martinez, 2018)	DISFA, AM-FED	Facial Action Units(AUs)	They have provided to identification of AUs using color features in datasets.	They have used color model for detecting to AU activation.	They have not provided good efficiency in skin color.
24	(Cruz, Foi, Katkovnik, & Egiazarian, 2018)	SET S, SET 14, Urban100	Single Image Super Resolution, CNN	SISR have given good result in similar domain using 1D wiener filtering.	They have performed better result in self - similar object.	They have not leaded to training and relies input images.
25	(Wang, Ma, Chen, & You, 2017)	EER, MORPH, FG-Net	Cross-age face verification	These methods improved over the performance from	They have work on effectively balance feature sharing and	They have not provided good result in small

S. No.	Author's	Data Sets	Technique	Result	Merit	Demerit
				2.2% EER on MORPH 7.8% EER on FG-NET by more than 50% and 59.7%.	feature exclusion between the two tasks.	datasets and also a large number of datasets.
26	(Duong, Quach, Luu, N., & Savvides, 2017, October)	FG-NET, MORPH, CACD, AGFW	Temporal non-volume preserving (TNVP), Generative Adversarial Networks (GAN)	They have consecutively worked on synthesizing age progressed faces and cross age face verification.	They have guaranteed inference and evaluate the feature in consecutive stages.	This method has a big issue to solve large scale problem.
27	(He, Wu, Sun, & Tan, 2017, February)	CASIA, NIR-VIS2.0, Large Scale VIS data.	Visual versus near infrared (VIS-NIR), Invariant deep representation (IDR)	This technique achieves 94% verification rate large scale VIS data.	They have provided good result and reduce the error rate of 58% only with a compact 64D representation.	IDR performed very low observation of among three datasets.
28	(Hu, Kan, Shan, Song, & Chen, 2017, May)	Multi-PIE	Learning Displacement Field Network(LDF-NET)	LDF-NET achieved frontal image using useful information in original images.	They have provided good face recognition using Multi-PIE datasets.	They have perform low efficient better than 2D.
29	(Galea & Farrugia, 2017)	VGG Face, PRIP-HDC, MEDS-II.	Deep convolution neutral network (DCNN)	They have reduced the error rate of 80.7 to 32.5 in real world forensic images.	A face image recognize 3D Morphable model to improve facial features in new images	This algorithm is using primary the limited images.
30	(Ranjan, Sankaranarayanan, Castillo, & Chellappa, 2017, May)	ALF, CASIA, MORPH, CeLeb-A, IMDB+W IKI, PASCAL	Single deep convolution neutral network(SDCNN), Multi-task learning frame work (MTL)	This model has a better understand to the faces and achieved good result for most of these tasks.	This method performed better than HyperFace using MTL frameworks.	FDDB has failed to capture small faces in any region of the proposal.
31	(Trigeorgis, Bousmalis, Zafeiriou, & Schuller, 2016)	CMU-PIE, XMLVTS, CASIA	Semi-Non-Negative Matrix Factorization	This model has worked learn the two-dimensional representation.	They have not provided good result in datasets.	They are not able to solve the area of speech recognition.
32	(Masi, et al., 2018)	IARPA, LFW, CASIA, YTF, IJBA, PIPA	Pose aware Models (PAM), CNN	This model has to design for solving the regular problem and optimizes the point and loss minimization.	They have analysis of IJBA. They have evaluated landmarks and improve the accuracy of pose estimation.	These models have worked on only training PAMs with a single optimization framework.
33	(Li, Gong, Li, & Tao, 2016)	MORPH, FGNET, Album2	Hierarchical method based on two level learning.	These models have improved the accuracy of 94.2% using LPS.	These experiments perform better result in MORPH, Album2 datasets.	This method is not work better in low-level images.

S. No.	Author's	Data Sets	Technique	Result	Merit	Demerit
34	(Mehdipour Ghazi & Kemal Ekenel, 2016)	VGG, YTF, AR, CMU-PIE	VGG Framework, CNN	It have worked on pre-processed the face recognition and provide a powerful representation.	They have provided a multiple features in faces and evaluated a n under various circumstances.	They have find in limited data provided by the mismatched conditions
35	(Wandt, Ackermann, & Rosenhahn, 2016)	CMU MOCAP, KTH Football, Human Eva	A priorly trained base poses and Predefined Skeleton.	3D construction of human motion from monocular image sequence.	Proposed method performs well under occlusions, noise and real world data of the KTH datasets.	They are provide good result on high level noise of the reconstruction.
36	(Zhou, Lin, & Zhang, 2015)	YALEB, AR, PIE, UCF-50	Latest Low Rank Representation (LatLRR), PCA	Proposed method find better classification results other than representation based method.	On larger scale datasets by adopting the LI-filtering algorithms.	In the same spirit, we will try integrating other feature learning methods.
37	(Peng, Gao, Wang, Tao, Li, & Li, 2015)	CUHK, FERET, IIIT-D, FG-NET, LFW	Multiple Representation based Face Sketch Photo Synthesis (MrFSPS).	This approach has work superior performance in multiple datasets using existing method.	They have performed forensic sketch datasets using dependent style.	These datasets has work very less size so unfortunately.
38	(Dong, Loy, He, & Tang, 2016)	YCbCr, RGB Model	Single Image Super Resolution deep CNN (SRCNN)	Proposed SRCNN has capable to improve the reconstruction of images in natural corresponding channels.	They have learnt end to end communication using low level and high level resolutions.	The extra activity has explored more filters using other training strategies.
39	(Lei, Pietikäinen, & Li, 2013)	FERET, CAS-pEAL-R1, LFW, HFB	Gabor and Local Binary Patterns, discriminant face descriptor (DFD), Coupled-DFD(CDFD)	They have learnt to reduce image filters with heterogeneous gap in face images.	This has a good generalization and competitive descriptor in face recognition under various circumstances.	Proposed DFD does not work on video based analysis.
40	(Zhang, Shao, Wong, & Fu, 2013)	CMU, MULTI-PIE	High level feature learning scheme.	They have produces a novel technique.	They have reduces one to one and many to one encoder.	This method is working on high level feature learning.
41	(Ji, Xu, Yang, & Yu, 2012)	KTH, TRECVID	CNN, 3DCNN	They have extracts the features in both spatial and temporal dimensions.	These model have outperforms compared to TRECVID data.	3DCNN is working on a supervised learning training data.

### ***Observation and Recommendation of Machine Learning Approaches in Face Detection and Recognition***

Face detection and recognition applications are using on 2D dimensional face images. They need large number of feature matching in different techniques. Using learning applications, they have improved the accuracy of face images in datasets. The impact factor of pose, illumination and expression is the basic and complete information about the face images. Face identification and verification is very important factor of the unknown persons. The moving object is very challenging problem in face identification and also find another problem in aging and non-rigid motion of the object. Learning discriminant appearances in face representation they are depend on invariant poses in face detection and recognition. Face identification is a critical issue in face recognition systems. This problem is identifying to change of poses to the object in different angles. So it is comparing the database between test face image and registered face images.

In face detection and recognition are using collaborative random faces (RFs) guided encoder to recognize the facial appearances between the test faces and registered faces. Random features are matching the pattern of faces used in database. They are using three types of feature appearances learning techniques.

1. Supervised learning is the process of providing the series of sample faces and comparing with the resultant faces with the expected outcomes. In a face detection and recognition are recognizing in supervised learning.

2. Unsupervised learning is more complex and difficult to implement the database. Because after using unsupervised learning algorithms the target resultant faces are unknown. Unsupervised learning is also called self-organizing and self-learning. Unsupervised learning process extracts the statistical properties of the training image and group's similar vector into features.

Reinforcement learning, the output images are not recognizing after comparing parallel output images. In this technique teacher is present but they are confused about answers. For high level identity feature are used to supervise auto encoder. In this technique, Identify the similar feature value from output features value after extracting to identify the features, then they are improving the accuracy and reliability of the face using different-different datasets and algorithms.

Features of the faces and there level of alignment are managed to discriminative identity features and explore them. The face identification of the common structure is the same value of the pose, illumination and extraction but their identities are different. It means they have registered same faces on different identity so it is big problem to identify the true value of the

image. In this situation feature learning is done important role to find the true value of the images.

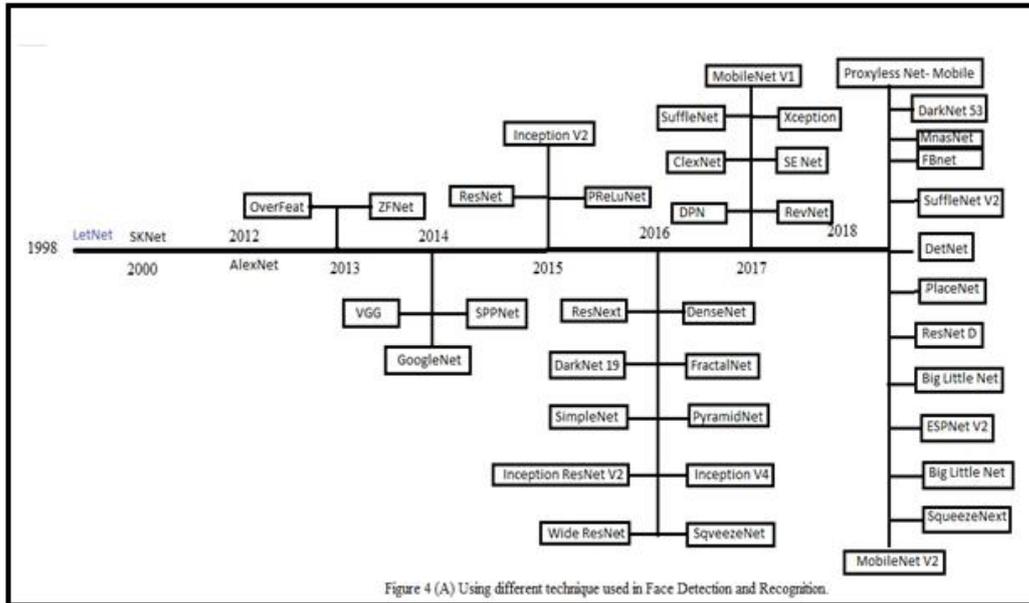


Figure 4. Different technology uses in different years

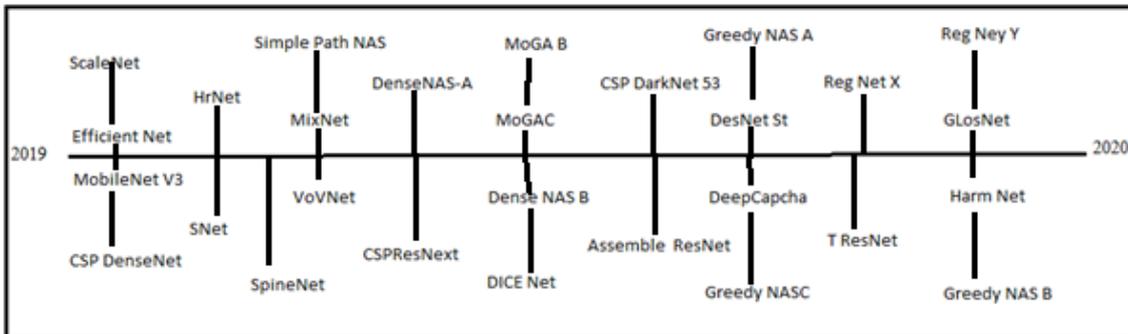


Figure 4 (B) Using Different technique used in Face Decetion and Recognition.

Figure 5. Different technology uses in different years

In figure 4 and Figure 5 is showing different technology used in to identify the deep feature of the faces. This concept is arising in 1998. In 1998 LetNet was very popular technique; they are basically depending on the neurons. But in 1998 to till 2020 there were very popular and many researchers used and develop different-different technologies to identify the features of the faces. In current scenario there are all most 96.76% is solved out. But still there are many problems are identified in our surrounding, so it is very important and useful to improve that technology as soon as possible. Machine learning based approaches work on the various problems related to pattern analysis, to detect the facial expression and face recognition.

## Conclusions

This article discusses a photos-based device that enables users to use the webcam to recognise and detect many face-related features. In this article, we are performing comprehensive and systemic study to check the efficacy of these classic representation learning structures on class-imbalanced outcomes. We also show that deeper discrimination can be learned by creating a deep network that retains inter-cluster differences both within and within groups. MobileNet, which provides both offline and real-time precision and speed to provide fast and consistent stable results, is the recently suggested Convolutional Neural network (CNN) model. This paper has evaluated the comprehensive survey of machine learning based approaches in face detection and recognition including various techniques and datasets. This paper also discussed the summary of related work presented by various researchers in face detection and recognition using different techniques. This paper also discussed the integration of machine learning based approach with multiple image related dataset and improving the performance of the faces using classifier.

## References

- Agarap, A. F. (2019). An architecture combining convolutional neural network (CNN) and support vector machine (SVM) for image classification. *arXiv preprint arXiv: 1712.03541*.
- Anwar, S. H. (2018). Image deblurring with a class-specific prior. *IEEE transactions on pattern analysis and machine intelligence*, 41(9), 2112-2130.
- Bali, V., Kumar, A., & Gangwar, S. (2019). Deep learning based wind speed forecasting-A review. *9th International Conference on Cloud Computing, Data Science & Engineering (Confluence)*, 426-431.
- Bau, D., Zhu, J. Y., Strobelt, H., Lapedriza, A., Zhou, B., & Torralba, A. (2020). Understanding the role of individual units in a deep neural network. *Proceedings of the National Academy of Sciences*, 117(48), 30071-30078.
- Benitez-Quiroz, C. F., Srinivasan, R., & Martinez, A. M. (2018). Discriminant functional learning of color features for the recognition of facial action units and their intensities. *IEEE transactions on pattern analysis and machine intelligence*, 41(2), 2835-2845.
- Carion, N., Massa, F. S., Kirillov, A., & Zagoruyko, S. (2020). End-to-End Object Detection with Transformers. *arXiv preprint arXiv:2005.12872*.
- Chopade, P., Edwards, D., Khan, S. M., & Pu, S. (2019, November). CPSX: Using AI-Machine Learning for Mapping Human-Human Interaction and Measurement of CPS Teamwork Skills. *In 2019 IEEE International Symposium on Technologies for Homeland Security(HST)*, 1-6.
- Cruz, C., Foi, A., Katkovnik, V., & Egiazarian, K. (2018). Nonlocality-reinforced convolutional neural networks for image denoising. *IEEE Signal Processing Letters*, 25(8), 1216-1220.
- Deng, W., Hu, J., & Guo, J. (2018). Compressive Binary Patterns: Designing a Robust Binary Face Descriptor with Random-Field Eigenfilters. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, 41(3), 758-767.

- Ding, C., Xu, C., & Tao, D. (2015). Multi-task pose-invariant face recognition. *IEEE Transactions on Image Processing*, 24(3), 980-993.
- Dong, C., Loy, C. C., He, K., & Tang, X. (2016). Image super-resolution using deep convolutional networks. *IEEE transactions on pattern analysis and machine intelligence*, 38(2), 295-307.
- Dong, X., Zheng, L., Ma, F., Yang, Y., & Meng, D. (2018). Few-example object detection with model communication. *IEEE transactions on pattern analysis and machine intelligence*, 41(7), 1641-1654.
- Duong, C. N., Quach, K. G., Luu, K. L., N., T. H., & Savvides, M. (2017, October). Temporal non-volume preserving approach to facial age-progression and age-invariant face recognition. *In Proceedings of the IEEE International Conference on Computer Vision*, 3737-3743.
- Galea, C., & Farrugia, R. A. (2017). Forensic face photo-sketch recognition using a deep learning-based architecture. *IEEE Signal Processing Letters*, 24(11), 1586-1590.
- He, R., Wu, X., Sun, Z., & Tan, T. (2017, February). Learning invariant deep representation for nir-vis face recognition. *In Proceedings of the AAAI Conference on Artificial Intelligence*, 31 (1).
- Howard, A., Sandler, M., Chu, G., Chen, L. C., Chen, B., Tan, M., et al. (2019). Searching for mobilenetv3. *In Proceedings of the IEEE International Conference on Computer Vision*, 1314-1324.
- Hu, L., Kan, M., Shan, S., Song, X., & Chen, X. (2017, May). LDF-Net: Learning a displacement field network for face recognition across pose. *In 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition*, 9-16.
- Hu, Z., Cho, S., Wang, J., & Yang, M. H. (2014). Deblurring low-light images with light streaks. *In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 3382-3389.
- Jeong, S., Lee, J., Kim, B., Kim, Y., & Noh, J. (2017). Object segmentation ensuring consistency across multi-viewpoint images. *IEEE transactions on pattern analysis and machine intelligence*, 40(10), 2455-2468.
- Ji, S., Xu, W., Yang, M., & Yu, K. (2012). 3D convolutional neural networks for human action recognition. *IEEE transactions on pattern analysis and machine intelligence*, 35(1), 221-231.
- Lee, J., Won, T., & Hong, K. (2020). Compounding the performance improvements of assembled techniques in a convolutional neural network . *arXiv preprint arXiv:2001.06268*.
- Lei, Z., Pietikäinen, M., & Li, S. Z. (2013). Learning discriminant face descriptor. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36(2), 289-302.
- Li, Z., Gong, D., Li, X., & Tao, D. (2016). Aging face recognition: a hierarchical learning model based on local patterns selection. *IEEE Transactions on Image Processing*, 25(5), 2146-2154.
- Liu, X. Y., Yang, H., Chen, Q., Zhang, R., Yang, L., Xiao, B., et al. (2020). FinRL: A Deep Reinforcement Learning Library for Automated Stock Trading in Quantitative Finance. *arXiv preprint arXiv:2011.09607*.
- Mafi, M., Rajaei, H., Cabrerizo, M., & Adjouadi, M. (2018). A robust edge detection approach in the presence of high impulse noise intensity through switching adaptive median and fixed weighted mean filtering. *IEEE Transactions on Image Processing*, 27(11), 5475-5490.

- Malhotra, S., Bali, V., & Paliwal, K. K. (2017, January). Genetic programming and K-nearest neighbour classifier based intrusion detection model. *In 2017 7th International Conference on Cloud Computing, Data Science & Engineering-Confluence*, 42-46.
- Masi, I., Chang, F. J., Choi, J., Harel, S., Kim, J., Kim, K. ..., et al. (2018). Learning pose-aware models for pose-invariant face recognition in the wild. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(2), 379-393.
- Mehdipour Ghazi, M., & Kemal Ekenel, H. (2016). A comprehensive analysis of deep learning based representation for face recognition. *In Proceedings of the IEEE conference on computer vision and pattern recognition workshops*, 34-41.
- Peng, C., Gao, X., Wang, N., Tao, D., Li, X., & Li, J. (2015). Multiple representations-based face sketch-photo synthesis. *IEEE transactions on neural networks and learning systems*, 27(11), 2201-2215.
- Qian, Y., Deng, W., & Hu, J. (2018, May). Task specific networks for identity and face variation. *In 2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)*, 271-277.
- Ranjan, R., Sankaranarayanan, S., Castillo, C. D., & Chellappa, R. (2017, May). An all-in-one convolutional neural network for face analysis. *In 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 17-24.
- Sagonas, C., Ververas, E., Panagakis, Y., & Zafeiriou, S. (2017). Recovering joint and individual components in facial data. *IEEE transactions on pattern analysis and machine intelligence*, 40(11), 2668-2681.
- Shao, M., Zhang, Y., & Fu, Y. (2017). Collaborative random faces-guided encoders for pose-invariant face representation learning. *IEEE transactions on neural networks and learning systems*, 29(4), 1019-1032.
- Tao, D., Guo, Y., Li, Y., & Gao, X. (2017). Tensor rank preserving discriminant analysis for facial recognition. *IEEE transactions on image processing*, 27(1), 325-334.
- Trabelsi, A., Chaabane, M., & Ben-Hur, A. (2019). Comprehensive evaluation of deep learning architectures for prediction of DNA/RNA sequence binding specificities. *Bioinformatics*, 35(14), i269-i277.
- Trigeorgis, G., Bousmalis, K., Zafeiriou, S., & Schuller, B. W. (2016). A deep matrix factorization method for learning attribute representations. *IEEE transactions on pattern analysis and machine intelligence*, 39(3), 417-429.
- Tsai, C. C., Li, W., Hsu, K. J., Qian, X., & Lin, Y. Y. (2018). Image co-saliency detection and co-segmentation via progressive joint optimization. *IEEE Transactions on Image Processing*, 28(1), 56-71.
- Tulyakov, S., Jeni, L. A., Cohn, J. F., & Sebe, N. (2017). consistent 3D face alignment. *Tulyakov, S., Jeni, L. A., Cohn, J. F., & Sebe, IEEE transactions on pattern analysis and machine intelligence*, 40(9), 2250-2264.
- Wandt, B., Ackermann, H., & Rosenhahn, B. (2016). 3d reconstruction of human motion from monocular image sequences. *IEEE transactions on pattern analysis and machine intelligence*, 38(8), 1505-1516.

- Wang, W., Yan, Y., Cui, Z., Feng, J., Yan, S., & Sebe, N. (2018). Recurrent face aging with hierarchical autoregressive memory. *IEEE transactions on pattern analysis and machine intelligence*, 41(3), 654-668.
- Wang, X., Ma, H., Chen, X., & You, S. (2017). Edge preserving and multi-scale contextual neural network for salient object detection. *IEEE Transactions on Image Processing*, 27(1), 121-134.
- Wang, X., Zhou, Y., Kong, D., Currey, J., Li, D., & Zhou, J. (2017, May). Unleash the black magic in age: a multi-task deep neural network approach for cross-age face verification. *In 2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)*, 596-603.
- Zhang, Y., Shao, M., Wong, E. K., & Fu, Y. (2013). Random faces guided sparse many-to-one encoder for pose-invariant face recognition. *In Proceedings of the IEEE International Conference on Computer Vision*, 2416-2423.
- Zhou, P., Lin, Z., & Zhang, C. (2015). Integrated low-rank-based discriminative feature learning for recognition. *IEEE transactions on neural networks and learning systems*, 27(5), 1080-1093.
- Zhu, X., Liu, X., Lei, Z., & Li, S. Z. (2017). Face alignment in full pose range: A 3d total solution. *IEEE transactions on pattern analysis and machine intelligence*, 41(1), 78-92.

---

**Bibliographic information of this paper for citing:**

Shukla, Ratnesh Kumar & Tiwari, Arvind Kumar (2021). Comparative Analysis of Machine Learning Based Approaches for Face Detection and Recognition. *Journal of Information Technology Management*, 13(1), 1-21.

---