Journal of Data Acquisition and Processing
ISSN: 1004-9037
https://sjcjycl.cn/
DOI: 10.5281/zenodo.777463

A SURVEY ON CONVOLUTION NEURAL NETWORK FOR FACE RECOGNITION

[1] Mr. Zubin C. Bhaidasna, [2] Dr. Priya R. Swaminarayan

[1]Research Scholar,[2]Dean of Faculty of IT & CS.

[1]Parul University, Vadodara, India.[2]Parul University, Vadodara, India.

Abstract: Thinking about the presence of an exceptionally huge measure of accessible information archives and distinguished the data from a facial ID and creative framework have come a long way in the last several years. One of the key areas for identifying an individual's proof is face recognition by the offices of law organization in the latest things of forensic science. Many studies focus on applying deep learning techniques to benefit from machine learning models for highlight extraction and classification to recognize the unique person.The first half of this review study primarily focuses on deep learning methods for face identification and matching, which increased the accuracy of face recognition algorithms by training them with large volumes of data.In this case, the paper incorporates review of various strategies used to distinguish a face from pictures additionally talked about difficulties for recognition.

***Index Terms* –** Deep Learning, CNN, Face Recognition, LetNet, AlexNet, ZFNet, Google Net, Res Net, R CNN

I. INTRODUCTION

The recognition system using face will lead as future technology in Computer Science that will directly identify a face from an image or the video and its uses will span a variety of industries, including ATM, healthcare, driver's license, reservations for trains, and surveillance operations.However, face picture identification will always be a difficult problem for a big database.Looking in today's era of technology a person can be recognized by theirfingerprints, reading their palms, geometry of their hands, scanning their eye iris, recognizing their voice, or by other biometric trait.The objective in developing such biometric applications is the aim in supporting the theme of smart cities and globally many researchers and scientist engineers are working on developing more robust and accurate algorithms and methods, which can be used in daily life. The most commonly used recognition is password and the security systems must protect all these personal information and data. The main problem with authentication that lies in most of the system is data acquisition with finger print, speech recognition, iris scan, etc. as the person should properly keep his bio impression say for example his thumb, face or eye in proper place in order to read this biometrics. However, in case of face images, the method of acquiring is non-intrusive where the user is unaware about that as a subject he is being captured in the system. Being a universal feature of human being as face, it becomes important in the field of research applications and can be a problem solving tool in case of many applications of object recognition.

Face recognition system have mainly two aspects for a face that are captured from an image or the video as:

1. Verification of Face also known as authentication.
2. Identification of Face also known as recognition.

With reference to copy the human brain as the human brain network, Deep learning and machine learning are potential solutions to the aforementioned issue since they are both branches of artificial networks that may be used to replicate the human brain as the human brain network. For obtaining a better result, we can use deep learning concepts as a tool. The role of deep learning as a technology plays a vital role in surveillance system and on social media applications like Facebook in tagging people. But right now, identifying and recognizing the same individual is the most challenging issue if the person has undergone some changes like, grown beard, face mask, ageing, luminance, etc. where the demand is to design a more robust algorithm in the area of deep learning.

II. RELATED WORK

Since more than a decade, Face recognition plays a vital role in area of research.Face recognition research involves a variety of domains, such as machine learning, neural networks, and image processing and computer vision and pattern recognition.To recognize faces in videos, several methods and strategies have been put forth.Here, a few facial recognition algorithms and strategies have been described.

KangGeon Kim, Zhenheng Yang, IacopoMasi, RamakantNevatia Gerard Medioni[1]over her the authorshave proposed new algorithm with three steps like, Face detection using YOLO cascade, Body Pose estimation and data association. Face detection using Off the Shelf by retrained YOLO detector, OpenPse for body pose estimator that helps to estimate joints where body part is occluded and filter out false position in estimating pose. Final data association that uses score function from face and body detection to keep track if subject has movements.

Nate Cross white, Jeffrey Byrne, Chris Stauffer, Omkar Parkhi, Aiong Cao and Andrew Zisserman [2], author has studied the problem related to template adaption, which is a type of education that transfer set template media. A new kind of learning by transfer known as "template adaption" combines deep convolution network features that are learnt on a source domain of labelled faces with template-specific linear SVM that are taught on a target domain using media in a template.When used on IJB-A datasets, template adaptation, which has been demonstrated to be a straightforward and efficient technique for face verification and identification, has produced positive results.

Carolina TodedoFerraz and Jose Hiroki[3],the authors have proposed a new Local Binary Convolution Neural network (LBCNN),with relation to studies on convolution neural networks in various deep learning architectures; this has shown a decrease in the number of parameters that required to be trained on. Chokepoint data was used in the experiment, which led to the use of Laplacian steps and fewer LBC modules, making it feasible to train the LBCNN easily and more accurately.

Sushil Bhattacharjee, Amir Mohammadi and Sebastien Marcel [4] authors have gone through the vulnerability of Convolution Neural Network (CNN) for detecting faces from the masks that are made from silicon as presentation attacks (PA). The authors have demonstrated the susceptibility of each such face recognition system, by using a dataset of six silicon masks, we get a true match rate equivalent to 10 times the false match rate.For a straightforward presentation assault, they utilized a cheap thermal camera. PA are of two types: obfuscation, in which the attacker tries to disguise himself so that he cannot be identified, and impersonation, in which the attacker attempts to pass for the victim.They employed CNN's FaceNet implementation and its allied models. While verifying the CNN, a template is generated for probe face image and comparing it with the enrollment templates to identify the claim by using Cosine similar measure.

Jun-Cheng Chen, Rajeev Ranjan, Swami Sankaranarayanan, Amit Kumar. Ching-Hui Chen, Vishal M. Patel, Carlos D. Castillo, Rama Chellappa[5],for face alignment, face verification, and face recognition the authors have developed a single deep learning system architecture.This one is challengingsince it handles errors better and enhances the effectiveness of automatic face identification, face verification, and facial landmark detection.The performance of Deep Convolutional Neural Networks is good.

Ran He, Xiang Wu, Zhenan Sun and Tieniu Tan[8],A unique Wasserstein convolutional neural network approach is suggested for learning traits that are constant between images of faces taken in the near infrared (NIR) and visible (VIS) spectrums.The WCNN's low-level layers are trained using readily accessible VIS face pictures, NIR, VIS, and NIR-VIS shared layers make up the high-level layer, which are divided into three portions.This technique use a single network to project NIR and VIS images into a constrained Euclidean environment.The modality invariant identity information and the modality variant light spectrum information are separated into two orthogonal subspaces in the top-level layer of the CNN. The WCNN seamlessly incorporates invariant feature extraction and subspace learning into a CNN.

IacopoMasi, Feng-ju Chan, Jongmoo Choi, Shai Harel, Jungyeon Kim, KangGeon Kim, JatupornLeksut, Stephen Rawls, Yue Wu, Tal Hassner, Wael AbdAlmageed, Gerard Medioni, Louis-Philippe Morency, Prem Natarajan, Ram Nevatia[10], a novel technique has been put out that is intended for unrestricted face identification in the field, with a focus on severe out-of-plane position changes.The current approaches do this by either normalizing pictures that align in a single front posture or learning pose invariance by training the enormous volumes of data. The proposed method is able to tackle variations in pose where in it processes image by using different pose specific CNN networks. Additionally, the synthesis of different facial poses from input photographs using 3D rendering is employed to train models while offering resilience to position changes during testing.

Zhen Dong, Chenchen Jing, Mingtao Pei, Yunde Jia [11],the authors propose a deep convolutional neural network to train a discriminative binary hash video for face identification. A unified framework for face feature extraction and hash function is created by integrating in the network, making these two components compatible.Face frames are the input to the network, and they produce corresponding binary hash frames as the output.A face video clip's frame representation is hard voted together to produce a binary hash video representation.The entire process is carried out as: first, a deep CNN is trained using many face photos, next a face movie is used as a collection of face framesfor which a face movie is utilized to compile a series of face frames.Finally, by calculating the distances between the binary representations of the query and those found in the database, the retrieval procedure is completed. In the third stage, all the frames are combined to create a single binary representation for the facial movie.Here in below table we can see the comparative study of the different researcher.

Table 1 Comparative Study of the different methods

PAPER	JOURNAL & YEAR	METHOD USED	DATA SET	REMARKS
Face And Body Association For Video-Based Face Recognition. [1]	Winter Conference On Applications Of Computer Vision. IEEE. 2018.	Retrained Yolo Detector Using Resnet 50 Architecture	JANUS CS-3	Does not work on all poses of the face. Needs to improve on other factors like occlusion, side view, etc.
Template Adaptation For Face Verification And Identification. [2]	12th International Conference On Automatic Face & Gesture Recognition. IEEE. 2017.	Template Adaptation of DCNN	IJB-A	Template based media dependent. Performance decreases in case of single media.
A Comprehensive Analysis Of Local Binary Convolution Neural Network For Fast Face Recognition In Surveillance Video [3]	ACM. 2018	Local Binary CNN	CHOKE POINT Data Set Collecting Images From Real Video Surveillance	Extremely noise-sensitive and additional noisy photos must be added to get better results.
Unconstrained Still/Video-Based Face Verification	Springer. 2017.	Proposed DCNN	IJB-A, JANUS CS-3 & LABELED	Reliance on large training dataset. More time

With Deep Convolutional Neural Networks [5]			FACES IN WILD	consuming. Invariance in translation.
Large Margin Cosine Loss For Deep Face Recognition. Conference On Computer Vision And Pattern Recognition [7]	IEEE. 2018	Large Margin Cosine Loss function	Mega Face Challenge, YouTube Faces & LFW	In addition to outperforming other loss functions, it also outperforms A Softmax's performance with feature normalization.
Learning Invariant Features For Nir-Vis Face Recognition [8]	IEEE. 2017	Wasserstein Convolution Neural Network	Oulu-CASIA, BUAA-NIS-VIR, and CASIA NIR-VIS	Needs to improve distribution measure to accurately match the NIR and VIS distribution.
Adversarial Embedding And Variational Aggregation For Video Face Recognition. [9]	IEEE. 2018	Adversarial Embedding and Variational Aggregation	IJB-A, Celebrity 1000, YouTube Face, and YouTube Celebrities	The adversarial embedding learning approach can enhance the fundamental feature model.
Learning Pose-Aware Models For Pose-Invariant Face Recognition In The Wild. Transaction Pattern Analysis Machine Intelligence. [10]	IEEE. 2017	Pose Aware Models using several pose specific DCNN	IJB-A & PIPA	The method may be enhanced where a multi-branch network is trained with various convolution filters for faces in various locations.
Deep CNN Based Binary Hash Video Representations	Pattern Recognition. Elsevier 2018	Feature Extractor, and Fine Tuning with Learning	TV Series Datasets	Hash learning and feature integration are combined to

For Face Retrieval [11]		Hash Functions, Learning Face		optimize the framework.
Cross Euclidean To Riemannian Metric Learning With Application To Face Recognition From Video [12]	IEEE 2017.	CERML	Four Challenging DataSets	A novel manifold-to-manifold approach will be integrated with an existing method to improve the effectiveness of the suggested framework.

III. FACE RECOGNITION USING CONVOLUTIONAL DEEP LEARNING

Deep learning utilizes artificial neural networks to do computations on enormous amounts of data.Artificial intelligence known as "deep learning" is based on the structure and operation of the human brain.The three main categories of deep learning algorithms are reinforcement learning, unsupervised learning, and supervised learning.Neural networks, which are arranged similarly to the human brain, are composed of artificial neurons, commonly referred to as nodes.These nodes are stacked on top of each other in three levels: the input layer, any hidden layers, and the output layer. Deep belief networks, long short-term memory networks, multilayer perceptron, generative adversarial networks, convolution neural networks, and recurrent neural networks, etc. are only a few examples of the various types of neural networks that are accessible. The basic steps for the face recognition using deep learning are as shown below figure.

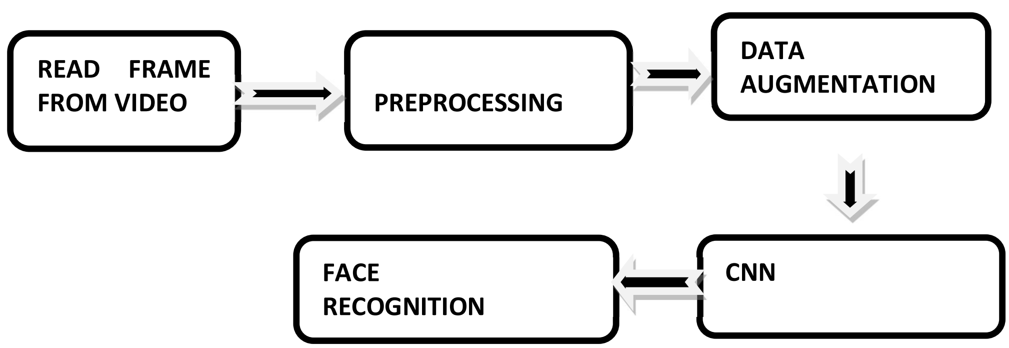

Figure 1 Basics steps for Face Recognition

As shown in above figure first frame read from the video and gathered the data.For diverse uses, such as frontal face photographs, posed images, and facial expression images, there are several databases online.After acquiring the data, preprocessing operations including resizing, filtering, and background removal will be used.To generate sufficient training examples, data augmentation is carried out following pre-processing.The process of creating synthetic samples from the original photos is known as data augmentation. After then, CNN will receive the final data.

One of the most often utilized models nowadays is CNN. This neural network computational model has one or more convolutional layers and a multilayer perceptron variant. It is frequently used in cases for classification. Convolution, pooling, and fully linked layer are the three basic processes.

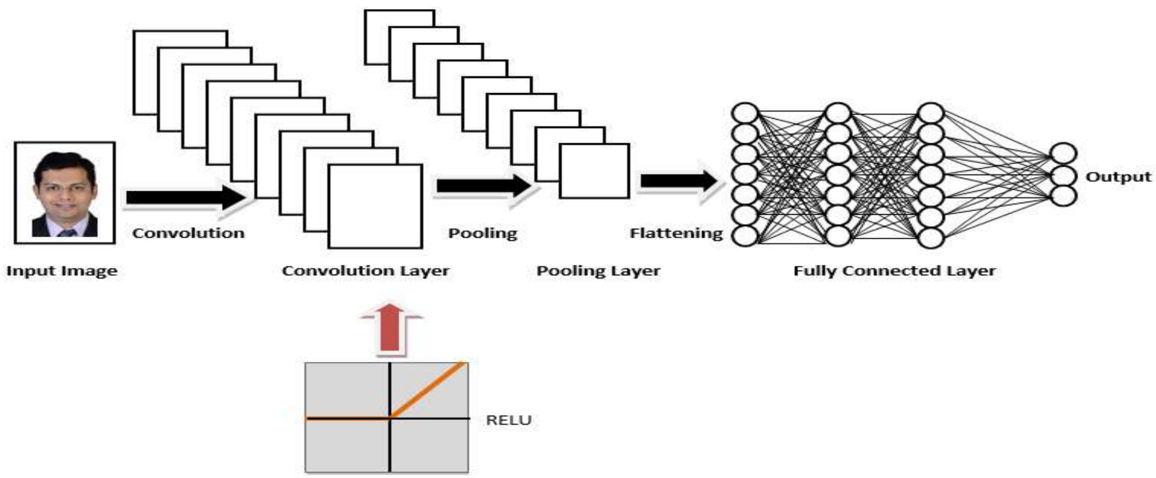

Figure 2 CNN Architecture

A convolution layer on CNN uses a number of filters to carry out the convolution process. To extract the characteristics from the image, various types of filters will be used in this layer.Convolution processes basically entail computing the dot product between the weights and the little region to which they are associated across the input matrix. Consider one input image (I) with 6*6 pixels and filter (f) 3*3 filters, then resultant matrix P

$$P = I * f \qquad (1)$$

I=

3	1	1	2	1	2
4	8	5	9	6	5
2	9	2	5	3	1
1	4	8	3	2	4
7	4	3	2	9	8
6	1	5	9	7	4

f =

1	0	-1
1	0	-1
1	0	-1

Let's start with Convolution operation

$$\begin{bmatrix} 3 & 1 & 1 & 2 & 1 & 2 \\ 4 & 8 & 5 & 9 & 6 & 5 \\ 2 & 9 & 2 & 5 & 3 & 1 \\ 1 & 4 & 8 & 3 & 2 & 4 \\ 7 & 4 & 3 & 2 & 9 & 8 \\ 6 & 1 & 5 & 9 & 7 & 4 \end{bmatrix} \times \begin{bmatrix} 1 & 0 & -1 \\ 1 & 0 & -1 \\ 1 & 0 & -1 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 2 & 8 \\ 8 & 4 & 4 & 7 \\ 3 & 7 & 1 & -3 \\ -2 & -5 & -2 & -2 \end{bmatrix}$$

Any activation function must be used for the non-linearity once features have been extracted. In general, RELU activation function is preferable for the convolution layer. Rectified Linear Unit (ReLU) layers are used in CNNs to manipulate elements. A corrected feature map is the result. The pooling layer receives the corrected feature map after that.

$$P1= RELU (P) \tag{3}$$

Downsampling is done along the breadth by the pooling layer, which reduces the dimensions.The two types of pooling are average pooling and maximum pooling. Reducing spatial dimensions is the primary goal of pooling. Downsampling technique called "pooling" reduces the size of the feature map. The pooling layer's improved output results may reduce the likelihood of over-fitting in the neural network.The pooling layer then flattens the resulting two-dimensional arrays from the pooled feature map into a single, protracted continuous linear vector.

$$\begin{bmatrix} 1 & 3 & 2 & 7 \\ 2 & 7 & 8 & 9 \\ 3 & 4 & 0 & 4 \\ 1 & 5 & 3 & 1 \end{bmatrix} \longrightarrow \begin{bmatrix} 7 & 9 \\ 5 & 4 \end{bmatrix}$$

$$P2=MaxPool(P1) \tag{4}$$

Pooling is followed by matrix conversion into one dimension and transmission to fully linked layer. Multiple neurons are connected to one another in a layer with all connections, which lays the groundwork for deep neural networks. Each neurons with different weight and bias values. Lastly non-linear activation function gets involved.

$$P3=W * P2 + b \tag{5}$$

A network that is fully linked will use any activation function, such as sigmoid, ReLU, etc.This neural network is straightforward feed forward. This layer's input is the result of the previous pooling or convolution layer.The fully connected neural network is given the outputs of the pooling layer and convolution layer architecture for final decision on classification.

IV. CONVOLUTION NEURAL NETWORK AND ITS TYPES

One of the most well liked Deep Learning methods is CNN. Particularly in applications connected to image processing and computer vision.Multiple-layer Convolutional Neural Networks (CNNs), commonly referred to as ConvNets, are used mostly for object detection, image classification, facial recognition, etc. Convolution and pooling layers alternate with one or more completely connected layers at the conclusion of a CNN design, in general.A global average-pooling layer may occasionally be used in place of a fully linked layer. To improve CNN performance, additional regulatory units such batch normalization and dropout are incorporated along with different mapping functions.

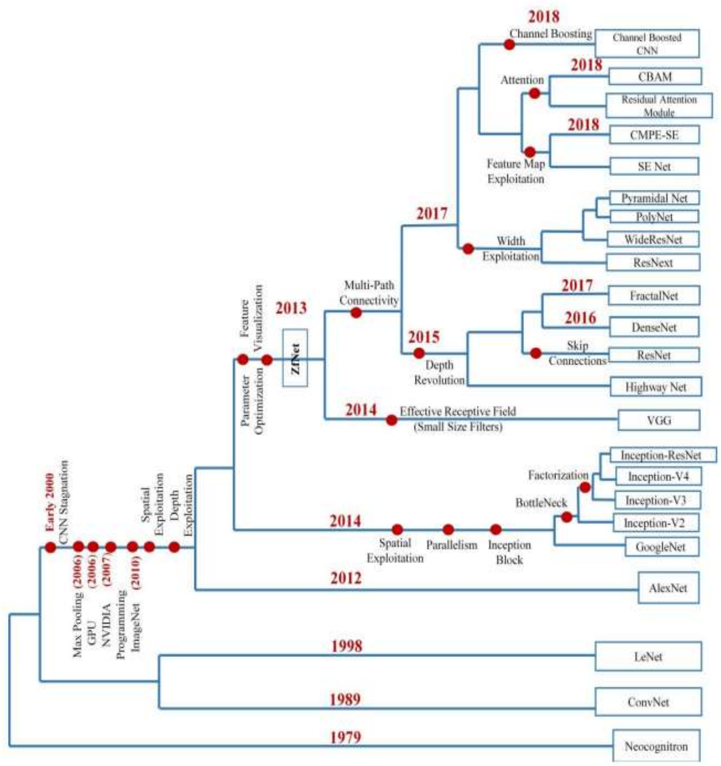

Figure 3 Evaluation of CNN [18]

4.1 LetNet

When it was still known as LeNet in 1988, the initial CNN was made by Yann LeCun.The LetNet architecture is the most often used CNN architecture. Several banks have used LeNet-5, a cutting-edge digit classification using a 7-level convolutional network, to identify handwritten numbers on cheque. This method is restricted by the availability of computational resources since processing power is needed to process pictures at increasing resolutions, which calls for larger and more convolutional layers. [20]. LeNet was the first CNN design that could automatically learn features from raw pixels additionally to cutting down on the quantity of parameters.

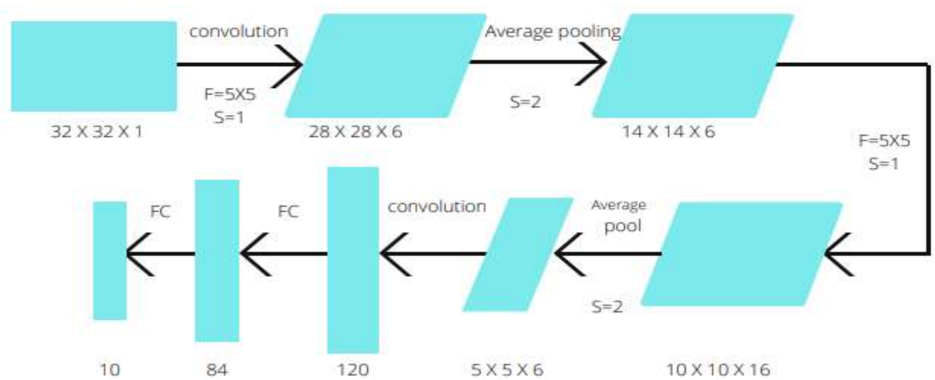

Figure 4 LetNet's Basic Architecture

LetNet's ability to use spatial correlation to cut down on computation and the number of parameters is one of its strengths.Additionally, prior to LetNet, conventional multilayered fully connected NN was employed which adds to the computational load and takes more time. In LetNet will take the advantage of Automatic learning of feature hierarchies. It give better result in compare to traditional NN model. The limitation of this LetNet model is poor scaling in various picture classes, large-size filters, and the extraction of low-level characteristics [18]. The fact that it was the first CNN to show off cutting-edge performance on hand digit identification tasks makes it famous for its historical relevance.

4.2 AlexNet

LeNetfounded in 1995 it started the era of deep CNN, was only capable of performing well on hand digit identification tasks at the time but gave poor results for other classes of images. In 2012, AlexNet was developed with more layers and features to overcome the problem of LetNet. AlexNet is the name of the first deep CNN architecture that produced groundbreaking results in image identification and classification.

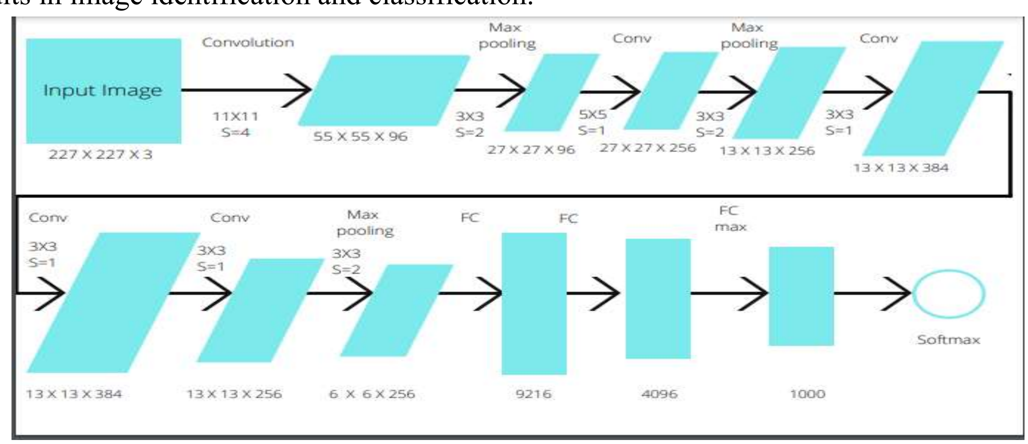

Figure 5 Basic Architecture of AlexNet

The design of the network was similar to LeNet's, except it was deeper, had layered convolutional layers and additional filters per layer.Convolutions, dropout, max pooling, ReLU activations, data augmentation and SGD with momentum were all used, as well as 11x5, 3x3,5x5 and 11x11. After each fully connected and convolutional layer, ReLU activations were added. Because of AlexNet's effective learning methodology, a new stage of the research into the architectural advancements of CNNs has started, and the next generation of CNNs will be greatly influenced by AlexNet.

4.3 ZFNet

Before, CNN's learning process relied heavily on trial-and-error, without a clear understanding of the cause of the improvement.Deep convolution neural network performance on complicated pictures is diminished as a result of this information gap. A fascinating multilayer Deconvolutional NN (DeconvNet) was proposed by Zeiler and Fergus in 2013 and became known as ZfNetby the names of the authors Zeiler and Fergus [18]. To quantitatively visualize network performance, it was developed. The network activity visualization's goal was to understand neuron activation in order to track CNN performance. Contrast to the current CNN architecture, in ZFNet reduced the size of filter and strides to improve the accuracy of the architecture. But same time for visualization purpose, architecture required to process more information.

4.3 GoogleNet

In the scientific article "Going Deeper with Convolutions," published in 2014, Google researchers suggested Google Net, commonly known as Inception-V1. When it came to the 2014 ILSVRC image classification competition, this architecture was the winner. Compared to CNN's earlier architecture, it has offered a significantly lower mistake rate. The GoogleNet architecture's primary goal was to achieve great accuracy at a low computing cost. GoogleNet's entire design is 22 layers thick; including 27 pooling levels, in this architecture researcher added 1 *1 convolution layer and average pooling. There was a risk of overfitting because the network in this case had extremely deep layers, therefore GoogleNet architecture was constructed with the notion of having filters of various sizes that could function on the same level. The network actually gets broader with this concept rather than deeper. Heterogeneous topology of GoogleNet, it needs modification from module to module, was the major issue. A representation bottleneck was yet another issue with GoogleNet, It greatly decreased the feature space in the following layer and occasionally resulted in the loss of crucial data.

4.4 ResNet

The foundation of deep CNN designs is the idea that as the network's depth increases, using a range of nonlinear maps and deeper feature hierarchies; it will be better able to approximate the objective function. Finally, Kaiming presented his so-called Residual Neural Network (ResNet) at the ILSVRC 2015. Based on the concept of "skip-connections," this employs a significant amount of batch normalization, which helps it train over thousands of layers efficiently without suffering long-term performance degradation.

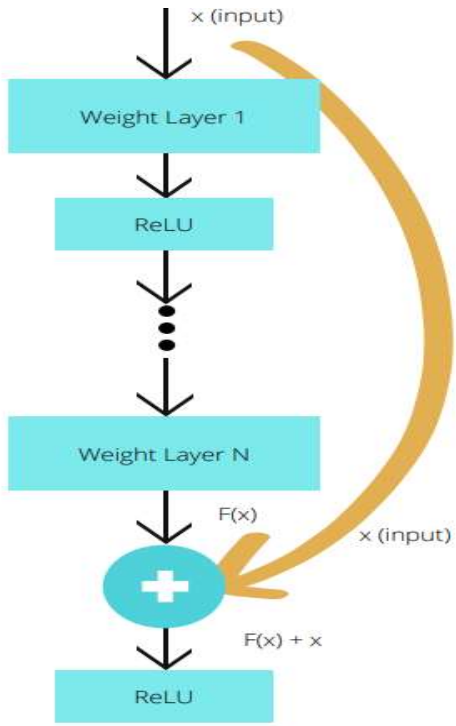

Figure 6 Basic Architecture of Residual

This sort of skip connection has the advantage that regularization will skip any layers that negatively affect architectural performance. When a gradient is back-propagated, the "vanishing gradient" problem occurs because repeated multiplication reduces the gradient to an endlessly tiny value. Performance suffers as a result of this. The ResNet algorithm's advantage is that it introduces discusses the issue of vanishing gradient and the idea of residual learning. However, the ResNet design is a bit convoluted, and it also degrades the feature map's feed forwarding information. The design has an extremely high computational expense.

4.5 R-CNN

A solution was proposed in 2014 byRoss Girshick et al, R-CNN's to address the issue of effective object localization in object identification. Since CNN directly extracts the characteristics from the data, the process of identifying an item of interest takes too long. The main problem with a traditional convolutional network followed by a fully connected layer is that the size of the output layer is variable rather than constant, which results in a picture with an unpredictable number of instances of various objects.

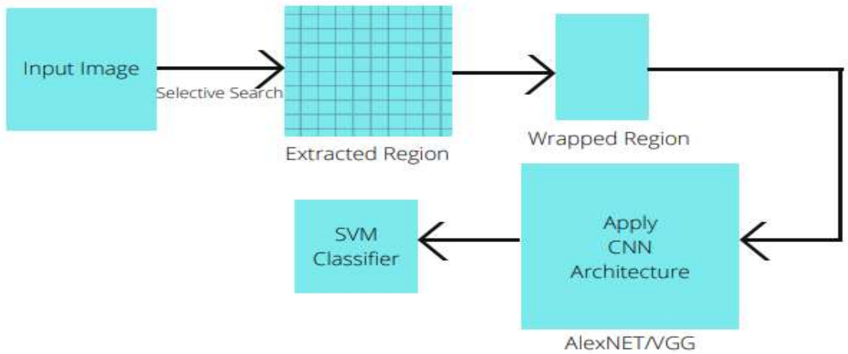

Figure 7 Basic Structure of R-CNN

Using a CNN to classify the presence of the object within several areas of interest from the picture would be a straightforward way to solve this problem. RCNN employs three distinct steps: In order to determine the collection of interesting point detections existing in an image, it first creates category-independent region recommendations. A deep convolutional neural network (AlexNet), which is the second module, extracts a feature vector from particular regions. After that, an SVM classifier was used to classify the information. Since we must partition a picture into 2000 regions on a recurring basis, it does not produce good results for real-time applications.

Here, we covered the fundamentals of CNN. CNN is a reliable deep learning technique for image processing tasks including face and image categorization. Object detection. Face recognition etc. The main advantage of the CNN is automatic detection of features without human interaction. However, there are some disadvantages of CNN are like;CNN does not encode the object's location or orientation. It often means that when the same object changes slightly in orientation or position, it may not activate the neurons that recognize the object. The training procedure will take a lengthy time if the CNN has several layers and the computer's GPU is not up to par.The other disadvantage is CNN required many data for training due to this a Convolutional neural network is significantly slower. The Pooling layer overlooks the relationship between the portion and the whole, which results in significant information loss. From video to identify the face will required data dependency. CNN is not give good result in time series problem. The quantity of data that convolutional network architecture offers is one of its key drawbacks. Expect CNNs to perform poorly if we provide them with insufficient data. CNNs have millions of parameters, thus using a limited dataset will cause them problems since they require many data to satisfy their thirst. You thereby disclose many data, CNNs are stronger and more willing to give you better performance, and you are giving fewer data. Therefore, to improve the result we can combine CNN algorithm with Recurrent Neural Network, LSTM or any other neural network to increase the efficiency of the CNN algorithm.

V. APPLICATION OF FACE RECOGNITION

There are numerous of applications exists in today's world where face recognition systems plays a vital role as a biometric system among which few are described below:

1. Surveillance System:

The surveillance system plays a vital role today on various fields like ATM machines, Railway stations, Airports, on public gatherings, on traffic signal cross roads, etc. The entire video is recorded about all the activities that are taking place at those places and the face of the person is captured and identified during the time of problems.

2. Indexing of videos:

The surveillance system used in police station as a tool to identify a criminal and to know about his past criminal record if any from the database by just inputting his image as face and identifying the records if exists if any. It is also used to label faces during biometric punch system when people use biometric as a device to register their in and out punch.

3. Investigation of Image Dataset:

The data set is used to find the missing children during the criminal like cases. It is also used in driving license system dataset in order to maintain the records of the users.

4. Security

Security and safety are now more crucial than ever for every enterprise, hospital, institute, etc. Face recognition technology has been used in many places for security systems like airport, hospital, access control in electronic device etc. We can also use face recognition system for mobile access, computer access and also even some industry entry time they used face authentication. Users are continuously verified, confirming that the user using the computer or in front of the screen is the same authorised person who signed in, to prevent someone else from changing files or conducting transactions while the authorised person momentarily leaves the computer terminal.

5. Biometrics

In many industries, institutes, many places face recognition system used as biometrics form for attendance monitoring system. Like example in education institute biometric system used for students attendance.

VI. Challenges In Face Recognition

At the time of Face Recognition, many problems or we can say challenges occur in deep learning algorithms like, the face may not be clearly visible due to low luminance or light, the face changes as per the age of the person, there is a grown beard on the face, etc. which challenges us to make a robust algorithm to detect and recognize a face.

3.1 Profile versus Frontal:

The cameras of a surveillance system are not constantly facing the subjects. The faces are seen from a certain angle. Face recognition ability is significantly impacted by the camera angle at which the person's photo was shot.

3.2 Occlusions:

The faces are quite hard to discern when they are partially obscured. In real-world applications, the camera view of people's faces may be partially obscured by the usage of objects like sunglasses, scarf, hands on the face, objects that people are carrying, and other sources.

3.3 Ageing

The texture of face changes with change of time and age of the person that is reflected on face. The face that we owe in the childhood is different when we turn 18 years and again it changes when we turn above 30 years. The change is mainly in the facial expressions, wrinkles, changes due to change in hormones in human body with respect to age. To overcome this, the dataset has to prepare of persons with different age groups over a period. The only characteristics that can be taken from the face, such as the brows, wrinkles, hairstyles, etc., will be used to identify a face.

3.4 Pose Angel

The face recognition system is highly sensitive to pose variation as the pose of the person changes constantly with respect to time. The pose remains constant only during capturing a photo but while recording a video the pose keeps on changing. This leads to change in pose of the person and the face may not be captured properly as frontal view of the face may not be possible throughout the video. In such cases, the face recognition system makes a challenge for us to detect and recognize face up to certain angle.

3.5 Illuminations

The meaning of the word illumination means variation in light. The slight change in lighting condition may yield to significant challenge to automated systems to identify the face. Now if the illumination may tend to vary then the face may be recorded with the same camera having identical facial expression and pose may appear somewhat different and because of this illumination, the face appearance may change drastically.

VII. CONCLUSION

In this review paper, the different Deep Learning approaches employed for the facial recognition system are summarized. An extensive examination of the literature revealed that Deep Learning Techniques have made considerable advancements in the field of facial recognition already. Numerous study publications have even suggested and put into practice numerous works that take into account various variations of face recognition, such as multiple expressions, time-invariant, weight variation, illumination variation, etc. Deep learning methods for face recognition systems have been the focus of a comparatively modest number of academic articles. utilizing modified CNN; it becomes abundantly obvious after compiling multiple evaluations. There is thus still a commendable amount of room for doing ongoing research using Deep Learning techniques for facial recognition.Discoveries show that relatively few articles have employed transfer-learning strategy for the facial recognition system after identifying various deep learning approaches used for the system. Therefore, in the future, research may be focused on face recognition utilizing deep learning techniques combined with a transfer learning strategy, which could result in a unique contribution. Upon closer examination, we discovered a wide variety of data sets that were utilized for the study.

References

1. Kanggeon Kim, Zhenheng Yang, IacopoMasi, RamakantNevatia Gerard Medioni., "Face And Body Association For Video-Based Face Recognition. Winter Conference On Applications Of Computer Vision". IEEE. 2018.

2. Nate Crosswhite, Jeffrey Byrne, Chris Stauffer, Omkar Parkhi, Aiong Cao And Andrew Zisserman, "Template Adaptation For Face Verification And Identification. 12th International Conference On Automatic Face & Gesture Recognition",. IEEE. 2017.

3. Carolina TodedoFerraz And Jose Hiroki. , "A Comprehensive Analysis Of Local Binary Convolution Neural Network For Fast Face Recognition In Surveillance Video." ACM. 2018

4. Sushil Bhattacharjee, Amir Mohammadi And Sebastien Marcel, " Spoofing Deep Face Recognition With Custom Silicone Masks" . IEEE. 2018.

5. Jun-Cheng Chen, Rajeev Ranjan, Swami Sankaranarayanan, Amit Kumar. Ching-Hui Chen, Vishal M. Patel, Carlos D. Castillo, Rama Chellappa." Unconstrained Still/Video-Based Face Verification With Deep Convolutional Neural Networks", Springer. 2017.

6. OtkristGuptaa, Dan Ravivb, Ramesh Raskar. " Illumination Invariants In Deep Video Expression Recognition.", Elsevier. 2017.

7. Hao Wang, Yitong Wang, Zheng Zhou, Xing Ji, Dihong Gong, Jingchao Zhou, Zhifeng Li And Wei Liu, " Cosface: Large Margin Cosine Loss For Deep Face Recognition. Conference On Computer Vision And Pattern Recognition." , IEEE. 2018.

8. Ran He, Xiang Wu, Zhenan Sun And Tieniu Tan. Wasserstein" Cnn: Learning Invariant Features For Nir-Vis Face Recognition.", IEEE. 2017.

9. Yibo Ju, Lingxiao Song, Bing Yu, Ran He, Zhenan Sun. "Adversarial Embedding And Variational Aggregation For Video Face Recognition.", IEEE. 2018.

10. IacopoMasi, Feng-Ju Chan, Jongmoo Choi, Shai Harel, Jungyeon Kim, Kanggeon Kim, JatupornLeksut, Stephen Rawls, Yue Wu, Tal Hassner, Wael Abdalmageed, Gerard Medioni, Louis-Philippe Morency, Prem Natarajan, Ram Nevatia, " Learning Pose-Aware Models For Pose-Invariant Face Recognition In The Wild. Transaction Pattern Analysis Machine Intelligence." IEEE. 2017.

11. Zhen Dong, Chenchen Jing, Mingtao Pei, Yunde Jia. ,"DeepCnn Based Binary Hash Video Representations For Face Retrieval. Pattern Recognition.", Elsevier 2018.

12. Zhiwu Huang, Ruiping Wang, Shiguang Shan, Luc Van Gool And Xilin Chen.," Cross Euclidean To Riemannian Metric Learning With Application To Face Recognition From Video.", IEEE 2017.

13. Ahmed Jawad A. AlBdairi, "Identifying Ethnics of People through Face Recognition: A Deep CNN Approach",Hindawi 2020.

14. Steve Lawrence , , C. Lee Giles , Ah Chung Tsoi "Face Recognition: A Convolutional Neural-Network Approach", IEEE Transactions On Neural Networks, JANUARY 1997

15. Peng Lu, Baoye Song, Lin Xu" Human face recognition based on convolutional neural network and augmented dataset, Systems Science & Control Engineering, 2020.

16. Shilpi Singh, S.V.A.V.Prasad, “Techniques and Challenges of Face Recognition: A Critical Review” ,
17. Procedia Computer Science, ScienceDirect- 2018.
18. Khan, Asifullah et al. “A survey of the recent architectures of deep convolutional neural networks.” Artificial Intelligence Review (2020): 1 - 62.
19. Stan Z. Li, “Face Recognition: Technical Challenges and Research Directions” , Advances in Biometric Person Authentication, Springer-2004.
20. Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, "Gradient-based learning applied to document recognition," in Proceedings of the IEEE, vol. 86, no. 11, pp. 2278-2324, Nov. 1998, doi: 10.1109/5.726791.
21. Batmaz, Zeynep et al. “A review on deep learning for recommender systems: challenges and remedies.” Artificial Intelligence Review 52 (2018): 1-37.
22. Gu, Jiuxiang et al. “Recent advances in convolutional neural networks.” Pattern Recognit. 77 (2018): 354-377.
23. Elngar, Ahmed A. et al. “Image Classification Based On CNN: A Survey.” Journal of Cybersecurity and Information Management (2021).
24. Indolia, Sakshi et al. “Conceptual Understanding of Convolutional Neural Network- A Deep Learning Approach.” Procedia Computer Science 132 (2018): 679-688.
25. Srivastava, Shrey et al. “Comparative analysis of deep learning image detection algorithms.” Journal of Big Data 8 (2021): 1-27.