Convolutional Neural Networks: A Comprehensive Survey

[1]Rahul Haridas, [2]Jyothi R L

[1] *College of Engineering Chengannur, Kerala, India.*

Abstract

Convolutional Neural Networks(CNNs) are drifting territory in Deep Learning. These days CNNs are utilized in the greater part of the Object Recognition errands. It is used in different application areas like Speech Recognition, Pattern Recognition, Computer Vision, Object Detection and other image processing applications. CNN orders the information based on a likelihood esteem. In this paper, a detailed analysis of CNN structure and applications are demonstrated. A comparative study of different types of CNN are also depicted in this work.

Keywords: Convolutional Neural Network, Network Architectures, VGGNet, ResNet, DenseNet.

1. INTRODUCTION

Deep Learning is a broader class of machine learning which will bring intelligence to the machines. Nowadays Deep Learning is widely used for Object Recognition, Pattern Recognition, Natural Language Processing and other image processing tasks. But mostly it is utilized in Object Recognition tasks such as Driver assistance systems, Autonomous Driving system, Target Detection[1] etc. in real-life. But it has also some other interesting applications like Colorization of Black and White Images, Adding Sounds To Silent Movies, Automatic Machine Translation, Object Classification in Photographs, Automatic Handwriting Generation, Character Text Generation, Image Caption Generation and Automatic Game Playing.

 For the most part, deep Learning can be performed with the help of different Neural Network Architectures[2]. These Networks are essentially used for feature extraction and processing. There are 4 Major Neural Networks used in Deep Learning. They're Unsupervised Pretrained Networks (UPNs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks, Recursive Neural Networks. Unsupervised Pretrained networks are classified into Autoencoders, Deep Belief Networks (DBNs), Generative Adversarial Networks (GANs). DBNs are special type of network which will automatically extracts higher level features. GANs are popularly used in the areas of Sound, Video and Text extraction. CNNs are the trending network architecture in Deep Learning. It will extract info by repeated convolutions. Mostly it is used in object recognition tasks. We will explain about CNN in detail in Chapter-2. Recurrent Neural Networks are in the same class of Feed-Forward Neural Networks.

Generally these types of networks are very hard to train. But the recent advances in Deep Learning will improve it. It can be used for Time-series prediction, Video analysis and Music information retrieval. Recursive Neural Networks have the ability to deal with hierarchical structures in the dataset. But here important problem is the recursive nature of the network. It will make some confusions in the training time.

 Nowadays Deep Learning has particularly significance, all things considered, since it tends to be utilized in assortment of fields of research. In this audit, we means to introduce a nitty gritty synopsis of various Convolutional Neural Networks which are utilized in Deep Learning. This work incorporates structure of CNN, distinctive models of CNN and its applications.

2. COVOLUTIONAL NEURAL NETWORK

This section includes Structure of CNN and its working. Before managing CNN we need to examine about Artificial Neural Networks.

A. Artificial Neural Network

Artificial Neural Network [20] is a figuring framework made up of various straightforward, very interconnected processing elements, which process data by their dynamic state reaction to outer information sources.

1) Basic Structure of ANN: The possibility of ANNs depends on the conviction that working of human cerebrum by influencing the correct associations, to can be imitated utilizing silicon and wires as living neurons and dendrites. The human brain is made out of 86 billion nerve cells called neurons. They are associated with other thousand cells by Axons. Upgrades from outer condition or contributions from tactile organs are acknowledged by dendrites. These sources of info make electric driving forces, which rapidly travel through the neural system. A neuron would then be able to send the message to other neuron to deal with the issue or does not send it forward. The human neuron system is illustrated in Fig.1[20]. ANNs are made out of various nodes, which copy natural neurons of human brain. The neurons are associated by connections and they interface with one another. The nodes can take input information and perform basic tasks on the information. The aftereffect of these activities is passed to different neurons. The yield at every node is called its activation or node value.
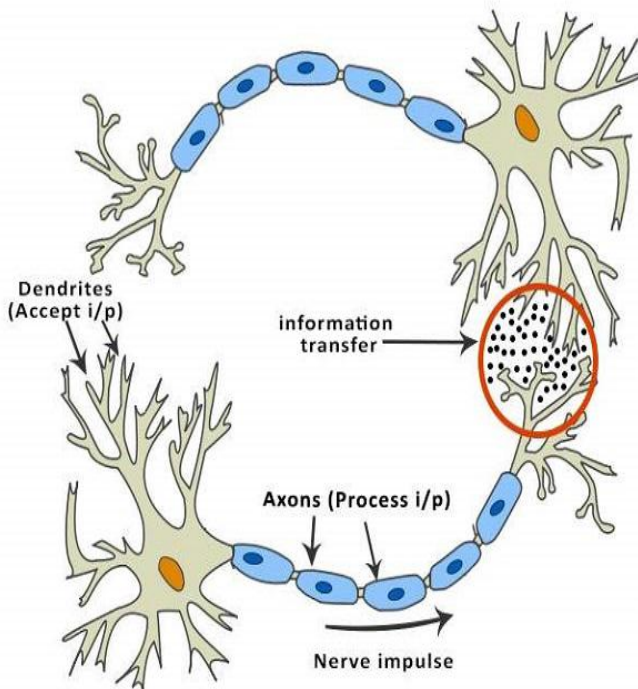
Fig. 1. Human Neuron System

Each connection is related with weight. ANNs are equipped for learning, which happens by changing weight esteems. Fig.2 demonstrates a straightforward ANN [20]
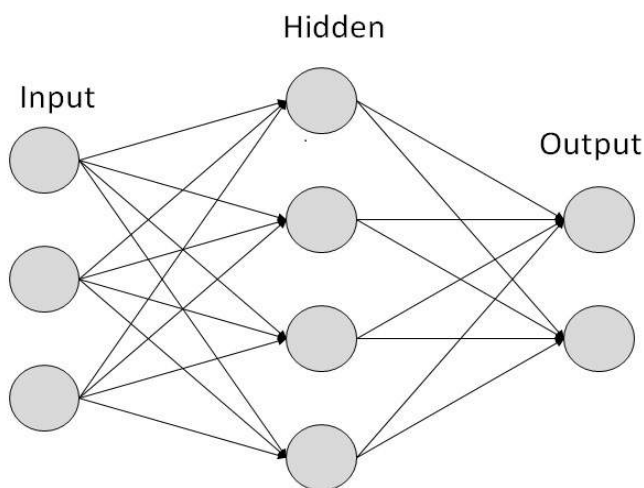

Fig. 2. Structure of ANN

2) Types of Artificial Neural Networks: There are two Artificial Neural Network topologies FeedForward and Feedback. In FeedForward ANN, the data stream is unidirectional. A unit sends data to other unit from which it doesn't get any data. There are no feedback loops. They are utilized in pattern generation/recognition/classification. They have settled sources of input and ouputs. It is Illustrated[20] in Fig. 3:
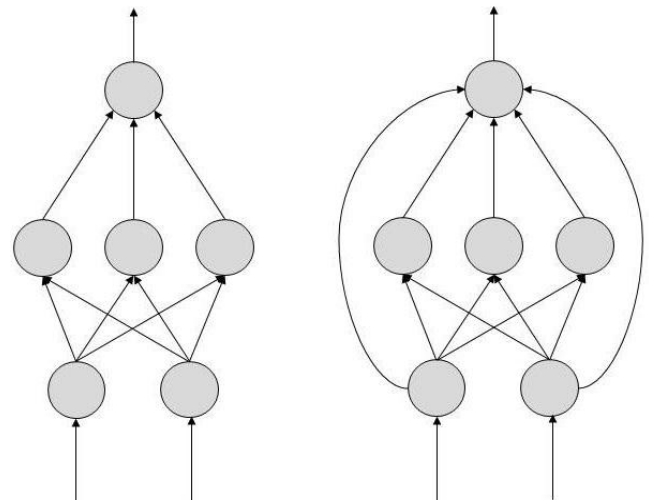

Fig. 3. FeedForward ANN

In FeedBack ANN, feedback loops are permitted. They are utilized in substance addressable recollections. It is represented in Fig.4[20]
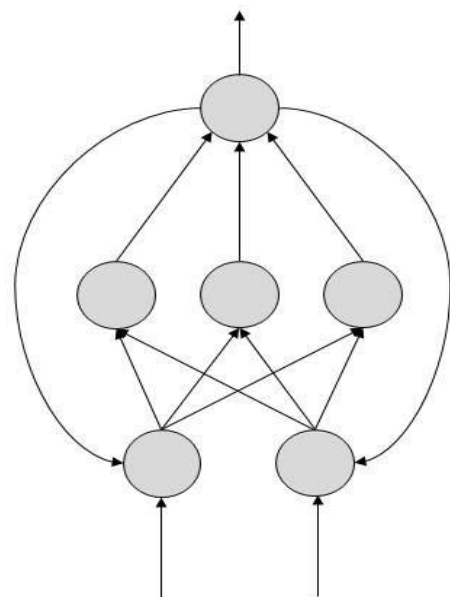

Fig. 4. FeedBack ANN

3) Machine Learning in ANNs: Learning in ANNs can be generally classified into 3. Supervised Learning, Unsupervised Learning and Reinforcement Learning.

In Supervised Learning, it involves a training algorithm to train the data. For example, the training algorithm will train ANN with some solutions. The network then guesses with training algorithm's correct solution and making necessary arrangements.

Unsupervised Learning doesn't provide a training algorithm for the data. Here data are classified based on similarity. Clustering and grouping are well-known examples of Unsupervised Learning.

In Reinforcement Learning, the network adjust the solution by observation. When the observation is negative, the network adjust its weight to make decisions.

4) Applications of ANN: Important Applications of ANN are in the fields of Aerospace, Automotive, Military, Electronics, Financial, Industrial, Medical, Telecommunications, Transportation, Anomaly Detection etc.

B. Basics of Convolutional Neural Networks

Convolutional Neural Networks(CNNs)[21] are artificial neural networks that are utilized principally to group images (for example name what they see), bunch them by comparability (photograph look), and perform object acknowledgment inside scenes. They are calculations that can distinguish faces, people, road signs, tumors, platypuses and numerous different parts of visual information.

Convolutional networks perform optical character recognition (OCR) to digitize content and make NLP possible on simple and manually written reports, where the pictures are symbols to be interpreted. CNNs can likewise be connected to sound when it is spoken to outwardly as a spectrogram. All the more as of late, convolutional networks have been connected specifically to content examination just as chart information with diagram convolutional networks.

The viability of convolutional nets (ConvNets or CNNs) in image acknowledgment is one of the primary reasons why the world has woken up to the adequacy of deep learning. They are controlling significant advances in Computer Vision (CV), which has evident applications for self-driving autos, apply robotics, drones, security, restorative conclusions, and medicines for the outwardly weakened.

Convolutional neural networks ingest and process pictures as tensors, and tensors are frameworks of numbers with extra dimensions.

They can be difficult to envision, so we should approach them by similarity. A scalar is only a number, for example, 7; a vector is a rundown of numbers (e.g., [7,8,9]); and a matrix is a rectangular grid of numbers involving a few lines and sections like a spreadsheet. Geometrically, in the event that a scalar is a zero-dimensional point, a vector is a one dimensional line, a matrix is a two-dimensional plane, a heap of frameworks is a three-dimensional 3Dcube,and when every component of those grids has a pile of feature maps atttached to it, you enter the fourth measurement. For reference, here's a 2 x 2 network:

[1, 2]

[5, 8]

A tensor includes the measurements past that 2-D plane. You can without much of a stretch image a three-dimensional tensor, with the variety of numbers orchestrated in a cube. Here's a 2 x 3 x 2 tensor[21] exhibited straight:

$$\left(\begin{pmatrix} 2 \\ 3 \end{pmatrix} \begin{pmatrix} 3 \\ 5 \end{pmatrix} \begin{pmatrix} 4 \\ 7 \end{pmatrix} \right)$$
$$\left(\begin{pmatrix} 3 \\ 4 \end{pmatrix} \begin{pmatrix} 4 \\ 6 \end{pmatrix} \begin{pmatrix} 5 \\ 8 \end{pmatrix} \right)$$

Fig. 5. Representation of a Tensor

C. Structure of CNN

We know that CNN are very effective for Image Recognition and Classification. So it can be used for identifying faces, objects and traffic signs apart from powering vision in robots and self driving cars. The Structure of CNN is Illustrated in Fig. 6[21] Convolutional Neural Networks are composed of several layers. They're, 1) Convolution Layer 2) Non Linearity (ReLU) 3) Pooling or Sub Sampling Layer 4) Fully Connected Layer
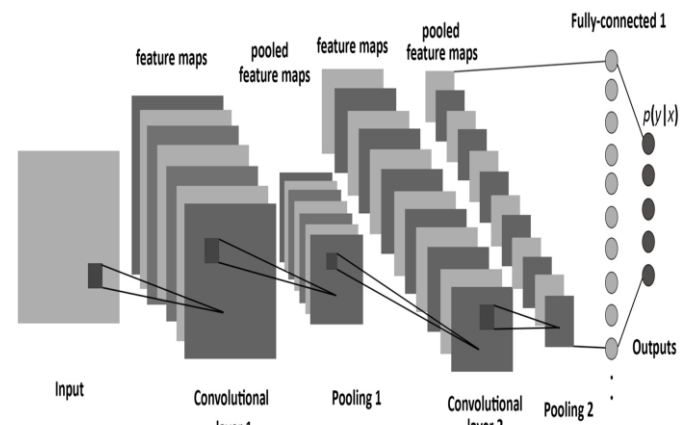

Fig. 6. A sample CNN

1) Convolution Layer: The main role of Convolution if there should arise an occurrence of a CNN is to extricate features from the input image. Convolution saves the spatial connection between pixels by learning image features utilizing little squares of input information. We won't delve into the numerical subtleties of Convolution here, however will endeavor to see how it functions over images. As we mentioned earlier an Image can be represented as matrix. Here we are considering a binary image in which matrix consists of values 0 or 1. Consider the 5 x 5 input image in Fig. 7[21],

Fig. 7. Input Image

Also, consider another 3 x 3 matrix as shown in Fig. 8[21]:

Fig. 8. Convolution Mask

Here the mask will convolove over each cell position of the input image and finally we will get a Convoluted Image. The Convolution of the 5 x 5 image and the 3 x 3 matrix can be computed as shown in Fig. 9[21] below:

Fig. 9. Convolution Output

The mask used for convolution is called as 'Filter' and the output of convolution is known as 'Feature Map'.

2) Rectified Linear Unit(ReLu): ReLu is a Non-Linear operation which replaces all negative pixel values in the feature map by zero. The motivation behind ReLU is to present nonlinearity in the CNN, since the greater part of this real world information we would need our CNN to learn would be nonlinear. ReLu operation can be represented[21] as,
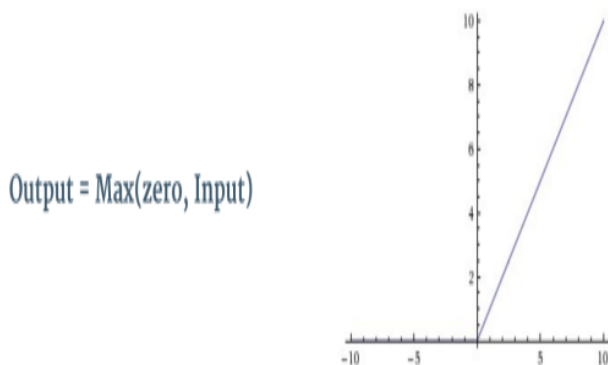

Fig. 10. ReLu Operation

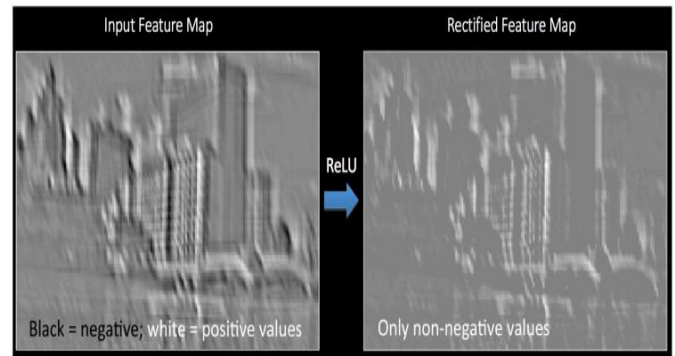ReLu operation can be applied on images as illustrated[21] below Fig. 11.

Fig. 11. ReLu Operation Example

3) Maxpooling: It is also known as Subsampling or Downsampling or Spatial-Pooling. It diminishes the dimensionality of each feature map but holds the most critical data. Maxpooling can be of different types like sum, average and max. If there should arise an occurrence of Max Pooling, we characterize a spatial neighborhood (for instance, a 2 x 2 window) and take the biggest component from the rectified feature map in that window. Rather than taking the biggest component we could likewise take the normal (Average Pooling) or total of all components in that window. By and by, max Pooling has been appeared to work better. Maxpooling is represented in Fig. 12[21].
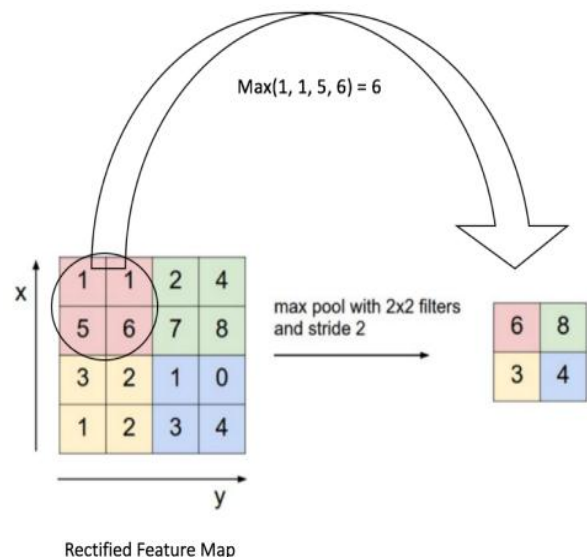

Fig. 12. Maxpooling

Here a 2 x 2 map is moved over the entire feature map and highest value from each region is selected then new

subsampled feature map is generated. pooling applied to rectified feature map can be visualized[21] in Fig. 13,
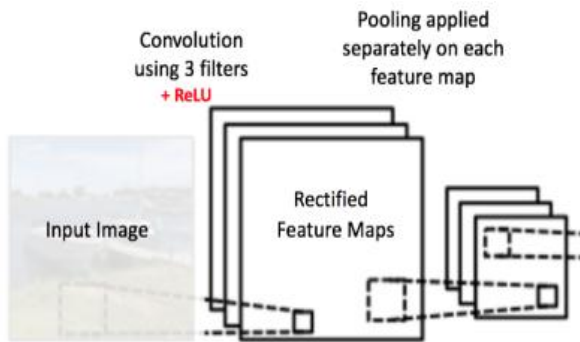

Fig. 13. Maxpooling

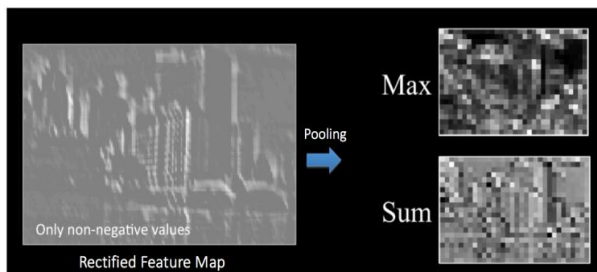The max and sum feature maps of pooling illustrated in Fig. 14[21].

Fig. 14. Types of Pooling

Generally Maxpooling are used for makes the input representations smaller and more manageable, reduces the number of parameters and computations in the network, therefore, controlling overfitting.

4) Fully Connected Layer: Fully Connected Layer is used for classification tasks. It uses softmax activation function for classification. The expression "Fully Connected" suggests that each neuron in the past layer is associated with each neuron on the following layer. The output from previous layers represent higher level features. The primary concern of this layer is to classify the input image based on the higher-level features. For example, image classification task of classifying dog, cat, boat and bird is represented[21] below Fig. 15.
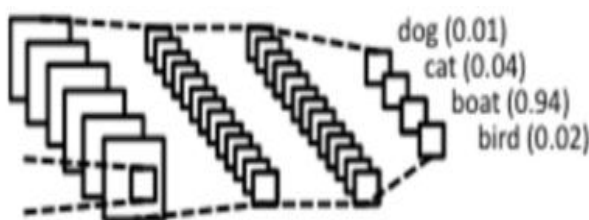

Fig. 15. Fully Connected Layer

3. CNN ARCHITECTURES

In this section, we will discuss about various Convolutional Neural Network architectures and the research works which used these architecture. Convolutional Neural Networks are evolved over time to improve their performance. Each architecture will differ from no. of layers used, no. of filters used for convolution and its complexity. Here we will analyze each architecture on the basis of winning ImageNet Large Scale Visual Recognition Competition(ILSVRC)[22].

A. LeNet-5

LeNet-5, a spearheading 7-level convolutional network by LeCun et al in 1998, that groups digits, was connected by a few banks to perceive transcribed numbers on (checks) digitized in 32 x 32 pixel greyscale input images. It was the most common architecture used by researchers for object detection. The basic architecture of LeNet-5 illustrated in Fig.16[22].
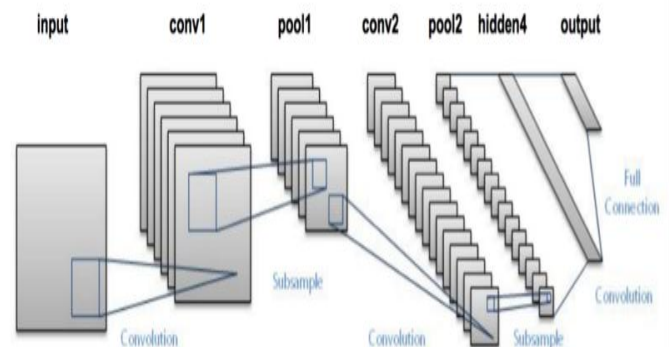

Fig. 16. Architecture of LeNet-5

From the architecture, It is clear that the general thought behind the plan is to play out various convolutions with max pooling between two activities and interfacing the last convolutional layer by means of completely associated layers to the output layer. Since all convolutional and pooling layers utilize a walk of one and no padding, the size of each measurement is diminished by 4 amid every convolution what's more, is split by each pooling activity. LeNet-5 architecture was used for extracting the features of Brain-Computer interface[3] by considering the common spatial patterns (CSPs) algorithm. But the main problem occured here was the mean accuracy obtained by the modified network is lower than the original accuracy.

B. AlexNet

The system had a fundamentally the same as design as LeNet by Yann LeCun et al. however was more deep, with more channels per layer, and with stacked convolutional layers. AlexNet was introduced in 2012 by Alex Krizhevsky, Geoffrey Hinton and Ilya Sutskever. Its architecture represented in Fig. 17[22].
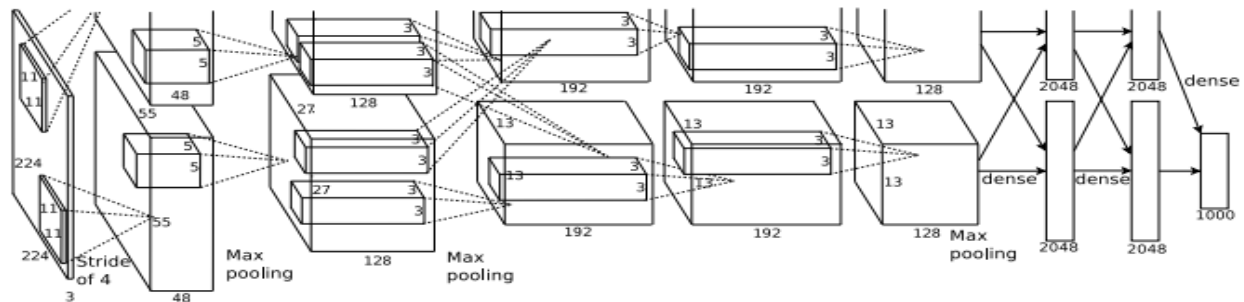
Fig. 17. Architecture of AlexNet

AlexNet is possibly the most powerful execution of CNNs up to date. It overcame traditional approaches to object detection. As we see it is similar to LeNet-5, but differs in using various parameters. It comprised 11x11, 5x5,3x3, convolutions, max pooling, dropout, data augmentation, ReLU actuations, SGD with force. It appended ReLU enactments after each convolutional and fully- connected layer. AlexNet was designed by the Super Vision group, consisting of Alex Krizhevsky, Geoffrey Hinton, and Ilya Sutskever.

AlexNet was implemented for detecting the lateral distance between vehicles[4]. This work has been successfully done using AlexNet beacause for efficiency, image blocks were searched in the database, thus the synthesis time was not affected by similarity values. As a result, it was possible to synthesise a large number of high-quality images in a short time. Instead of using a particular dataset, a real-time dataset was used in this scenario.

C. ZFNet and OverFeat

ZFNet won the ILSVRC one year after AlexNet in 2013. It was for the most part an accomplishment by tweaking the hyper-parameters of AlexNet while keeping up indistinguishable structure with extra Deep Learning components from talked about before in this work. ZFNet developed a method of visualizing convolutional neural networks. It is known as De-Convolutional Neural Network. It just opposite of the CNN which makes mapping features to pixels. Its architecture represented[22] in Fig. 18.

OverFeat won ILSVRC 2013 which is similar to AlexNet. But the difference is that AlexNet utilizes bounding boxes for classification while OverFeat used confine a given item at numerous areas and scales and combines these outputs to get the last outcomes.
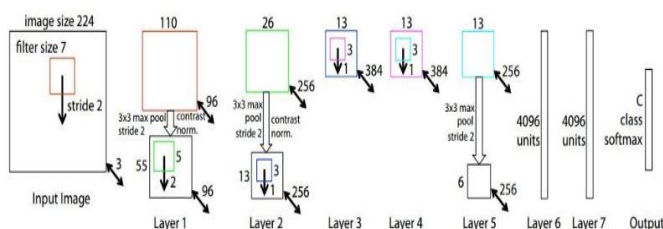

Fig. 18. Architecture of ZFNet

D. GoogleNet/Inception

The victor of the ILSVRC 2014 challenge was GoogleNet(a.k.a. Inception V1) from Google. It accomplished a main 5 blunder rate of 6.67%. This was near human level execution which the coordinators of the test were currently compelled to assess. This system was an enhancement of the AlexNet that was a lot further as well as diminished the quantity of parameters. It introduced Inception module which enable a network to recognize patterns of different sizes within the same layer. The commencement module plays out a few convolutions with various responsive fields in parallel and consolidates the outcomes by combining the depth slices of the diverse channels into one single layer. Architecture illustrated[22] in Fig. 19.
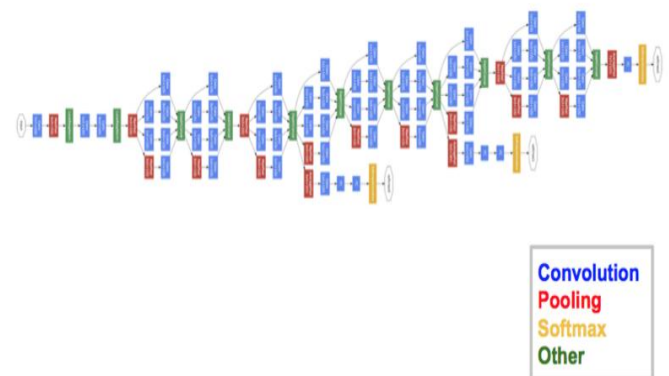

Fig. 19. Architecture of GoogleNet

The last GoogleNet comprised of a few such commencement modules stacked over one another with incidental pooling layers in the middle of, a couple of extra convolutional layers in the start of the system and a couple completely associated layers directly before the output layer.

 The GoogleNet likewise contained extra output layers closer to the center of the system and their outputs were joined with the output of the last layer of the system to get the complete expectation. This had some minor effect on the general outcome yet was primarily proposed to quicken the preparing of prior layers.

E. VGGNet

The sprinter up at the ILSVRC 2014 challenge is named VGGNet by the network and was created by Simonyan and Zisserman. VGGNet comprises of 16 convolutional layers and is extremely engaging a result of its exceptionally uniform architecture. In recent years, VGGnet is the most popular architecture used in the field of Deep Learning. It is similar to AlexNet having 3 x 3 convolutions but having lots of filters. It is at present the most favored decision in the network for extricating features from images. Its architecture represented in Fig. 20[22].
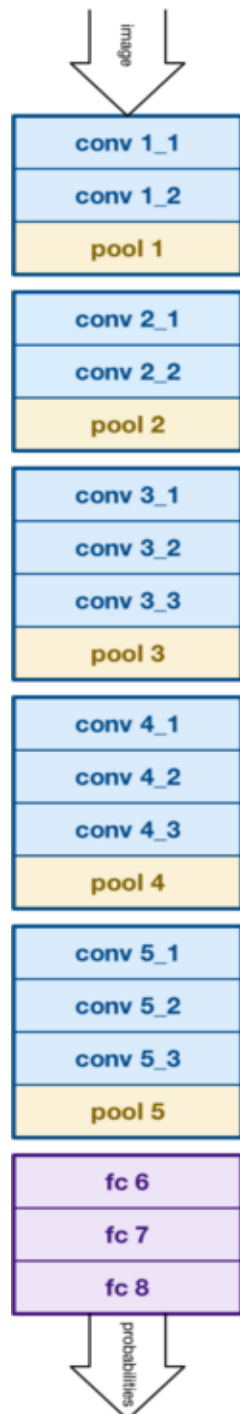
Fig. 20. Architecture of VGGNet

The load setup of the VGGNet is freely accessible and has been utilized in numerous different applications and difficulties as a benchmark include extractor. Be that as it may, VGGNet comprises of 138 million parameters, which can be somewhat testing to deal with.

VGGNet utilized in so many deep learning applications. Blind-Image Quality prediction[5] is a best work in which it outperformed all traditional methods. But here the issue was suffers from the overfitting problem leading to a degradation of performance.

Finger Vein based Biometric Authentication[6] was an excellent idea which achieve greater than 95% results in all datasets. Crack Detection[7] was another important application implemented using VGGNet in which it can't able to deal with noisy cracks. Main limitation of VGGNet was the usage of large no. of parameters as we discussed earlier. [8] is an example for this limitation.

 We know that VGGNet most commonly used for Object Detection. Object Tracking[9] is a sub-field of Object Detection which effectively utilized this architecture in a good manner. Nowadays so many research works are done in the field of Intelligent Transportation Systems. Self Driving Car Navigation[10] is an impotant application of it. It efficiently made use of VGGNet but the problem was camera latency.

F. ResNet

Residual Neural Network was introduced by Kaiming He et al with "skipconnections" and features heavy batch normalization in 2015. Such skip associations are otherwise called gated units or gated recurrent units and have a solid comparability to later fruitful components connected in RNNs. Because of this system they could prepare a NN with 152 layers while as yet having lower unpredictability than VGGNet. It accomplishes a best 5 blunder rate of 3.57% which beats human-level execution on this dataset. Its architecture illustrated in Fig.21. ResNets are started using widely for several research works like[11]. In this it introduces skip-layer connections to achieve both faster convergence in training and better performance. [12] proposes a refeelction removal system by using this architecture. In that case network is precise and effective for reflection removal.
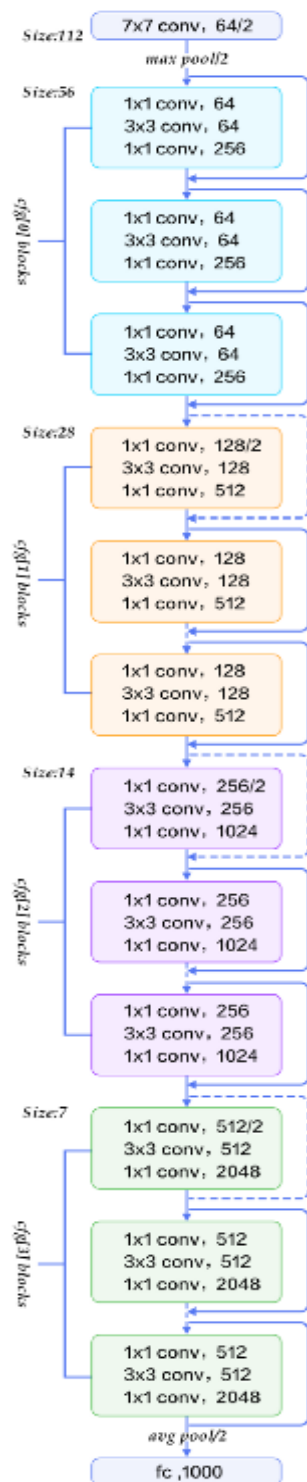
Inception-ResNet introduced as a combination of Inception network and Residual Network. by combining the inception module and residual connection as shown in Fig. 22[23].
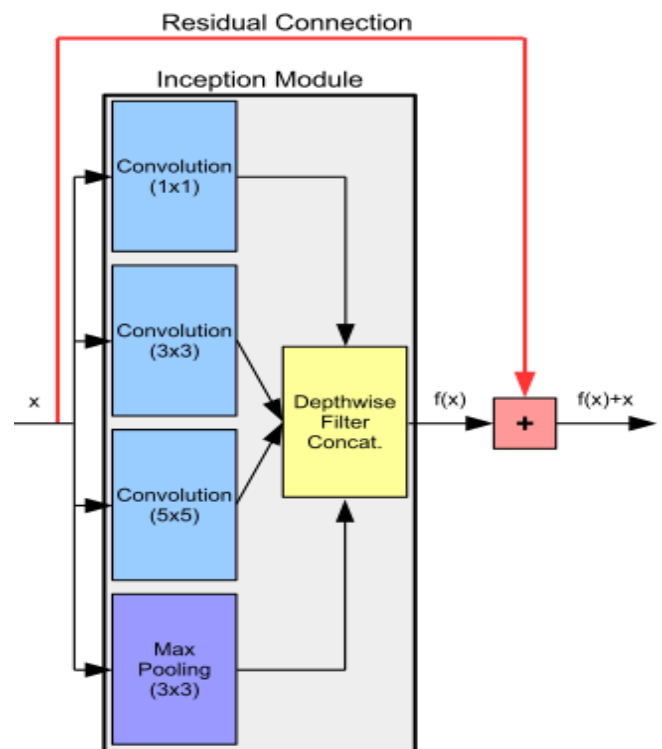

Fig. 22. Architecture of Inception-ResNet

This network is more efficient and time saving than Inception Network.

The second place in the 2016 ILSVRC classification challenge was achieved by ResNeXt. It joins the remaining forming blocks and associations of ResNets with the parallelization methodology of inception structures. It is Illustrated in Fig. 23[23]

Fig. 23. Architecture of ResNeXt

Fig. 21. Architecture of ResNet

F. Inception-v4, Inception-ResNet-v1 and ResNeXt

Inception-v4 was a modification of GoogleNet. It added several layers to the network. It used three different kinds of inception modules. In rent times this architecture is successfully used in[13] which has higher accuracy than VGGNet and Inception-v3. Also used in [14] to find out nutritional value in food.
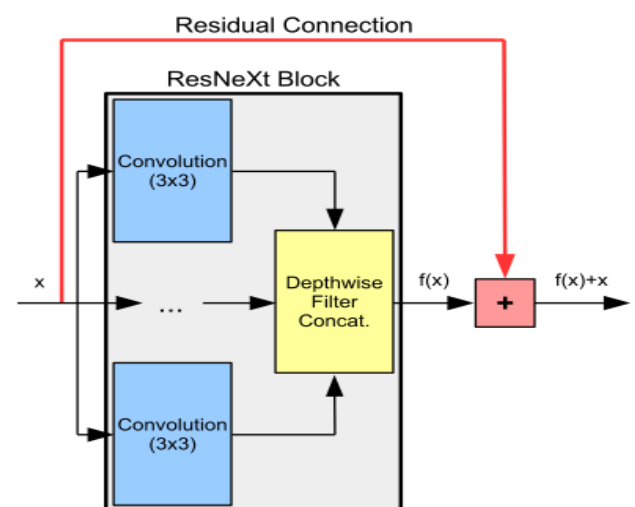
G. Densenet, DPN and MobileNets

DenseNets join the two outputs by depthwise channel link, as performed in inception modules. Moreover, DenseNets are including one such association from each layer to every resulting one with coordinating input sizes. Thusly, the educated features of a layer can be reused by any of the accompanying layers. Therefore, later layers need to create many less component maps, bringing about less unpredictable structures with less parameters.

 Since the width and height of layers in CNNs are step by step diminishing, interfacing every single perfect layer is partitioning the organize into Dense Blocks. Between these squares, pooling layers are utilized to change the sizes in like manner. These layers are alluded to as Transition Layers. [15] utilized DenseNet for recognizing 3-D actions. DenseNet is represented in following Fig. 24[23]:
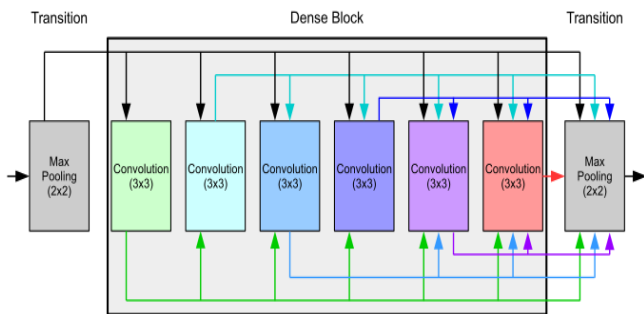

Fig. 24. Architecture of DenseNet

Dual Path Network (DPN) won first place in the 2017 ILSVRC localization challenge and finished top three in both classification and detection. So as to join the systems, the output of a layer is part and one section is joined with a remaining association, though the other is sent to every single ensuing layer, as performed in DenseNets. It is Illustrated in Fig. 25[23].
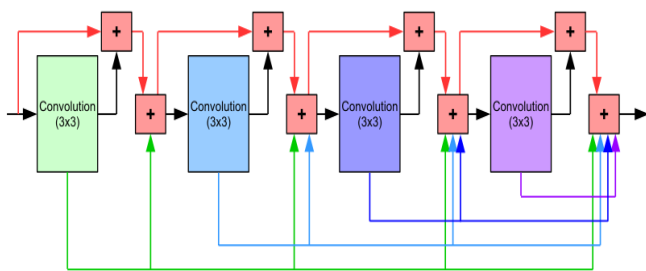

Fig. 25. Architecture of DPN

MobileNets are worked to be as time productive as could be expected under the circumstances by supplanting standard convolutions by Depthwise Separable Convolutions. Depthwise distinct convolutions split this procedure into two stages: A Depthwise Convolution and a Pointwise Convolution. The depthwise convolution goes about as a channel by just considering the square measured locales inside

a solitary depth slice. The point insightful convolution at that point plays out a 1 x 1 convolution to consolidate the data over the entire depth. Also, two new hyperparameters are utilized for MobileNets, which empower manifestations of considerably quicker structures by exchanging off exactness furthermore, execution time. MobileNet is represented[23] in following:
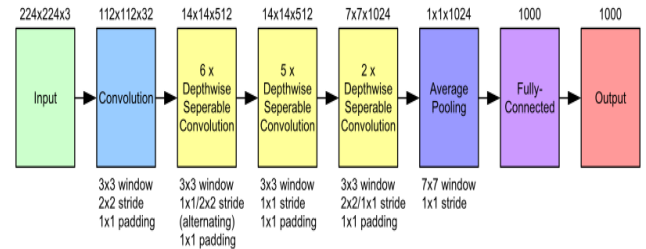

Fig. 26. Architecture of MobileNet

1) Comparison of CNN Architectures: Here we will perform a comparison between different generations of CNN architectures upto ResNet. The comparisons are based on ImageNet Large Scale Visual Recognition Challenge (ILSVRC) competition. ImageNet is an image database organized according to the WordNet hierarchy, in which each node of the hierarchy is depicted by hundreds and thousands of images. The comparison is illustrated in Table.1

4. CONCLUSION

 In this paper, we have discussed about various types of Convolutional Neural Networks. The different types are Le-Net-5, AlexNet, ZFNet and OverFeat, GoogleNet or Inception, VGGNet, ResNet, Inception-v4, ResNext, DenseNet, MobileNet etc. Functioning and evolution of different types of CNN are depicted. A comparative analysis of different networks is carried out and found that VGGNet is the most utilized type of CNN. But it was analyzed that CNNs like ResNet and DenseNet are outperformed VGGNet based on ImageNet Large Scale Visual Recognition Challenge (ILSVRC).

Table I Comparison of CNN Architectures

Year	CNN	Developed By	Error Rates	No. of Parameters	Dataset
1998	LeNet	Yann LeCun		60 Thousand	
2012	AlexNet	Alex Krizhevsky, Geoffrey Hinton and Ilya Sutskever	15.3 %	60 Million	ImageNet
2013	ZFNet	Matthew Zeiler, Rob Fergus	14.8 %		
2014	GoogleNet	Google	6.67 %	4 Million	
2014	VGGNet	Simonyan, Zisserman	7.3 %	138 Million	
2015	ResNet	Kaiming He	3.6 %		

REFERENCES

[1] Srinivas S. S. Kruthiventi , Kumar Ayush , R. Venkatesh Babu "DeepFix: A Fully Convolutional Neural Network for Predicting Human Eye Fixations",IEEE Transactions on Image Processing ,Vol. 26(9), pp. 4446 - 4456, Sept. 2017

[2] Zhang Bo ; Zhang Ling, "The analysis and improvement of artificial neural network models,",in 1997 IEEE International Conference on Intelligent Processing Systems (Cat. No.97TH8335), Oct. 1997.

[3] Siavash Sakhavi, Cuntai Guan , Shuicheng Yan, "Learning Temporal Information for Brain-Computer Interface Using Convolutional Neural Networks,", IEEE Transactions on Neural Networks and Learning Systems , vol. 29(11), pp. 5619 - 5629, March 2018.

[4] Xiang Zhang , Wei Yang , Xiaolin Tang , Yun Wang,"Lateral distance detection model based on convolutional neural network,",IET Intelligent Transport Systems, vol. 13(1), pp. 31 - 39, Jan 2019.

[5] Jongyoo Kim , Anh-Duc Nguyen , Sanghoon Lee,"Deep CNN-Based Blind Image Quality Predictor",IEEE Transactions on Neural Networks and Learning Systems, vol. 30(1), pp. 11 - 24, June 2019.

[6] Rig Das , Emanuela Piciucco , Emanuele Maiorana , Patrizio Campisi, "Convolutional Neural Network for Finger-Vein-Based Biometric Identification,", IEEE Transactions on Information Forensics and Security, vol. 14(2), pp. 360 - 373, Feb. 2019.

[7] Qin Zou , Zheng Zhang , Qingquan Li , Xianbiao Qi , Qian Wang , Song Wang, "DeepCrack: Learning Hierarchical Convolutional Features for Crack Detection", IEEE Transactions on Image Processing, vol. 28(3), pp. 1498 - 1512, Oct. 2018.

[8] Zhen Dong , Zheng Zhou , Zefan Li , Chen Liu , Peng Huang , Lifeng Liu , Xiaoyan Liu, "Convolutional Neural Networks Based on RRAM Devices for Image Recognition and Online Learning Tasks,",IEEE Transactions on Electron Devices, vol.66(1), pp.793 - 801, Dec.2018.

[9] Xiaopeng Hu , Jingting Li , Yan Yang , Fan Wang,"Reliability verification-based convolutional neural networks for object tracking,",IET Image Processing, vol. 13(1), pp. 175 - 185, Dec. 2018.

[10] Truong-Dong Do , Minh-Thien Duong , Quoc-Vu Dang , My-Ha Le, "Real-Time Self-Driving Car Navigation Using Deep Neural Network,", in 2018 4th International Conference on Green Technology and Sustainable Development (GTSD), 23-24 Nov. 2018.

[11] Yue Li , Dong Liu , Houqiang Li , Li Li , Zhu Li , Feng Wu, "Learning a Convolutional Neural Network for Image Compact-Resolution,",IEEE Transactions on Image Processing , vol. 28(3), pp. 1092 - 1107, sept. 2018.

[12] Yakun Chang , Cheolkon Jung "Single Image Reflection Removal Using Convolutional Neural Networks,",IEEE Transactions on Image Processing, vol. 28(4), pp. 1954 - 1966, Nov. 2018.

[13] Omer Deperlıoˇglu , Utku Kˇose "Diagnosis of Diabetic Retinopathy by Using Image Processing and Convolutional Neural Network,",in 2018 2nd International Symposium on Multidisciplinary Studies and Innovative Technologies (ISMSIT), 19-21 Oct. 2018

[14] Raza Yunus ; Omar Arif ; Hammad Afzal ; Muhammad Faisal Amjad ; Haider Abbas ; Hira Noor Bokhari ; Syeda Tazeen Haider ; Nauman Zafar ; Raheel Nawaz, " A Framework to Estimate the Nutritional Value of Food in Real Time Using Deep Learning Techniques ",in IEEE Access, vol.7, pp. 2643 - 2652, Dec. 2018.

[15] Chuankun Li , Yonghong Hou , Pichao Wang , Wanqing Li, "MultiviewBased 3-D Action Recognition Using Deep Networks ",IEEE Transactions on Human-Machine Systems, vol. 49(1), pp. 95 - 104, Dec. 2018.

[16] Armin Kappeler , Seunghwan Yoo , Qiqin Dai , Aggelos K. Katsaggelos "Video Super-Resolution With Convolutional Neural Networks,",IEEE Transactions on Computational Imaging , vol. 2(2) ,pp. 109-122, June 2016.

[17] Chenhao Lin , Ajay Kumar "A CNN-Based Framework for Comparison of Contactless to Contact-Based Fingerprints",IEEE Transactions on Information Forensics and Security, vol. 14(3), pp.662-676, July 2018.

[18] GuangYang,HaiboHe,QianChen"Emotion-Semantic-EnhancedNeural Network",IEEE/ACM Transactions on Audio, Speech, and Language Processing , vol. 27(3), pp. 531 - 543, March 2019.

[19] Michael G. Bechtel , Elise Mcellhiney , Minje Kim , Heechul Yun, "DeepPicar: A Low-Cost Deep Neural Network-Based Autonomous Car",2018 IEEE 24th International Conference on Embedded and RealTime Computing Systems and Applications (RTCSA), 28-31 Aug. 2018.

[20] https://www.tutorialspoint.com/artificial_intelligence /artificial intelligence neural networks.htm

[21] https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

[22] https://medium.com/@sidereal/cnns-architectures-lenet-alexnet-vgg-googlenet-resnet-and-more-666091488df5

[23] Felix Altenberger, Claus Lenz,"A Non-Technical Survey on Deep Convolutional Neural Network Architectures",in arXiv.org , 6 Mar 2018