Deep Learning Through Convolutional Neural Networks

Pooja Kalange, Megha Mutalikdesai

Abstract: Deep learning which is associated with the basics of Machine Learning has become popular over the years because of its fast paced adaptability and ability to handle complex problems. Prior to this technology breakthrough traditional methods of machine learning were used in applications of Image processing and pattern recognition, and analytics .With the advent of CNNs it has become easy to combat complex learning problems using the property of specificity and accuracy in CNN architectures and methodologies. This paper gives an introductory insights in CNNs like the feed-forward propagation networks and Back propagation Networks. The paper explains steps followed by CNNs for classifying the input and generating a predefined output. It also explains evolution of multiple Image CNN architectures which find applications in multiple domains of Computer Science like Image Processing & Segmentation, Pattern Recognition & Predictive Analytics, Text Analytics to name a few .

Index Terms: Keywords: Deep Learning, Convolutional Neural Network (CNN),CNN Architectures

I. INTRODUCTION

Deep learning are machine learning algorithms used to deal with abstractions of data at high level by usage multiple non linear transformations in the model architectures. Deep learning is a method of Machine learning which trains and builds neural networks ,which ultimately help to solve complex problems. [1-3]. Deep algorithms used the basics of non linear transformations where as shallow algorithm use only predefined levels of data abstraction. There are 4 Major Neural Networks used in Deep Learning. Namely Unsupervised Pre-trained Networks (UPNs), Convolutional Neural Networks (CNNs), Recurrent Neural Networks, Recursive Neural Networks[3]. Out of these Convolutional Neural Networks (CNNs) are used in variety of Applications like Image Processing & Segmentation, Big Data Analytics, Speech Processing, Pattern Recognition, Textual and Predictive Analytics[1-7]. Hence they are most booming Deep learning architecture in these days and acquired lot of attention in Research as well. Here we will be considering various Image-Net CNNs for understanding.[1-3][17].

II. CNN STRUCTURE

A CNN architecture basically consists of layers that are fed in forward manner and comprises of i)convolutional filters ii)pooling layers ii)activation [1-8][17] .After the last pooling layer CNN adopts several fully- connected layers that work on converting the 2D feature mapping of the previous layers into 1D vector for classification. Some necessary functions include preprocessing, segmentation and classification also[7-11]. It basically adopts a feed forward propagation or Backward propagation methodology [fig. 1] . The basic neural network consists of single layer perceptrons .It consists of output nodes which are single layer and the inputs use series of weights and bias which are given to the outputs so that there will be unidirectional flow of data. As multiple perceptron's come together they form a multilayered neural network, which follow the feed forward way. In a neural network, having back propagation the weights and biases of the neurons are updated on the basis of the error at the output. Hence Back propagation calculates the minimum value error function in weight space a rule called delta rule which is also known as gradient descent[1,13]. The weights with minimized error function can help then in finding solution for a problem.

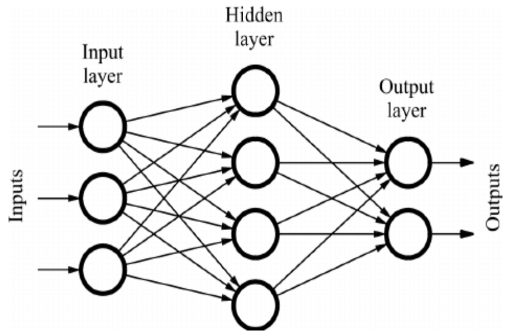

Fig. 1 Feed Forward Propagation in CNNs

A. *Convolution Layer*

This is the first layer in CNN it consists of an actual image or may be a matrix or any input from say Tensor Flow which would be having height, width and depth as parameter for consideration which is called kernel size. Typical Kernel size is 3x3 or 5x 5 of depth 1 for gray scale images and depth of 3 for color(RGB) images. The convolutional layers help in feature extraction as they slide over convolution kernel for the input images[17,18,20]. The operations mainly performed are taking the dot product of the input image and kernel and summation of both so as to reduce the size and map feature from the elements of image. The pixel selection and movement is known as stride . Multiple such mapping help in generation of compressed feature maps which fed into next layers[1-3][17].

Revised Manuscript Received on September 15, 2019
Ms. Pooja Kalange, Computer Department, Institute of Industrial & Computer Management & Research, Pune, India.
Ms. Megha Desai, Computer Department, Institute of Industrial & Computer Management & Research, Pune, India.

B. *Activation Layer*

Neural network are composed of neurons and in deep learning they are also called as filters[17-22]. They work in line with weight, bias and their respective activation function. Non linear activation function is used in deep neural networks to process more advanced, non-linear data since convolutional layers do not possess the property of non-linearity. It's the Activation function which decides the activation of the neuron which eventually introduces non-linearity into the output of a neuron.They compute element-wise multiplications between a filter matrix and a matrix that contains a part of an image . The activation functions variants are linear function, sigmoid function, tanh, ReLU (Rectified Linear Unit),Leaky ReLU, Softmax Function. The most significantly used activation function in CNNs is the ReLu and has gained its usage in variety of applications[1-10].

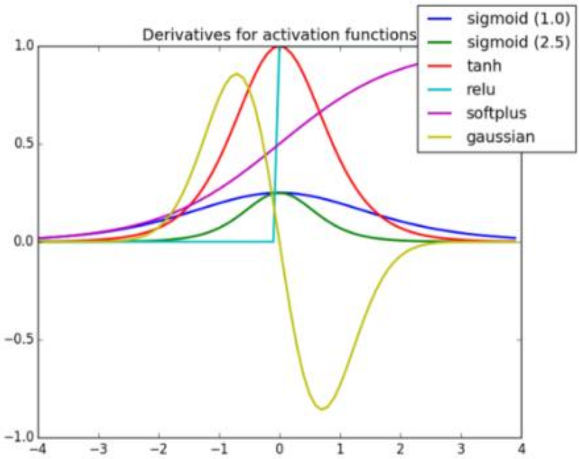

Fig. 2 Derivatives of Activation Functions

C. *Max Pooling Layer*

Multiple CNN convolutions result in reduction of output. However a small filter say may lead to smaller output generation [1-10][17-20]. But images may have a larger size than 32 pixels hence there arises the concept of pooling which is also called as Sub-sampling or Down- pooling. Therefore multiple layers of pooling help in abstraction , generalization and reducing the dimensionalities but keeping the critical information intact. Pooling is normally performed after non-linear activation, where the pooling layer helps reduce the number of parameters and avoids over fitting, and it also serves as a smoothing measure to eliminate unwanted noise. In addition we can use PCA(Principal Component Analysis)in pooling[4-10][17-22] .

D. *Fully Connected Layers*

Non-linear layers and pooling layers present in convolution layers allowed us to identify patterns in images hence comes the use of classification and regression. All neurons in the previous layer are connected to each neuron in next layer. Fully-connected layers do not preserve spatial information[1,17]. The fully-connected layer is followed by an output layer. We use softmax for classification and regression tasks as generates a well-performed probability distribution of the outputs. SVM(Support Vector Machines),DT(Decision Tree),KNN(k-Nearest neighbor),ANN(Artificial neural networks) are other classifiers which are used[1-17]. This process is followed by calculation of loss function i.e. difference between estimated and actual output values. Loss function is inversely proportional to accuracy and robustness. Lesser loss function eventually helps in greater accurate ,improved & optimized model could be generated.

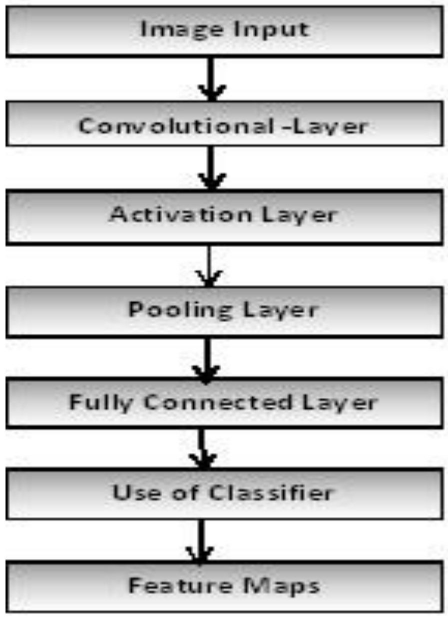

Fig. 3 Processes & Steps in CNN from I/p to O/p

III. CNN ARCHITECTURE APPLICATIONS

The CNN dataset considered in this paper is ImageNet, which basically finds enormous use in image or Text processing, pattern recognition and Natural Language processing. There are various CNN architectures like LeNet, AlexNet, GoogleNet ,ZFNet, VGGNet and ResNet.. LeNet is typical neural network architecture used in handwritten digit recognizes patterns as one of the applications.it basically works on Gray scale images. AlexNet was introduced in 2012 which uses ReLU as an activation function.It uses dropout instead of regularization to handle over fitting. Pytorch has one the implementation of AlexNet .VGGNet is a simple architecture which uses fewer number of parameters.GoogleNet is called as Inception Module ,it has less error rates and greater number of parameters.ZFNet has more fine- tuned parameters as compared to others. ResNet is also called residual neural network by Kaiming .It calculates transformation using ReLU activation function.Table[] illustrates comparative analysis of all CNN ImageNet architectures[1-17].

Table.1 Comparative Study of CNN Architectures.

CNN	Year	Developer Name	Error Rate (%)	No of Layers	Parameters	External data Usage
LeNet	1998	Yancunet Al	-	7 Layers	60 thousand	Yes
Alex Net	2012	Alex Krizevskey ,Geoffery Hinton,Ilya Sutskever	15.3	8 layers	60 million	no
ZFNet	2013	Mathew Zealer,Rob Furgus	14.8	-	-	no
GoogleNet	2014	Google	6.67	22 Layers	4 million	no
VGG Net	2014	Simonyan, Zisserman	7.3	19 layers	138 million	no
ResNet	2015	Kiamang he	3.6	152 layers	-	Yes

IV. CONCLUSION

The CNNs as part of Deep Learning has brought revolutionary changes in Technology. If the CNNs are trained appropriately they can find solutions and applications in variety of fields. As per the comparative study of different CNN Architectures as told in the paper they can be used in classification and pattern recognition. Of course if CNNs are combine with break through methodologies may be supervised or unsupervised learning they can be used immensely in biometrics, Internet of things (IOT),radio-genomics, geo-spatial detections, medical analysis which will help in identifying real time problems and applied solutions.

REFERENCES

1. "Application Of Deep Convolution Neural Network", Jiudong Yang, Jianping Li, 978-1-5386-1010-7/17/$31.00 ©2017 Ieee
2. "Simple Convolutional Neural Network on Image Classification", Tianmei Guo, Jiwen Dong ,Henjian Li˒Yunxing Gao, 2017 IEEE 2nd International Conference on Big Data Analysis.
3. "Convolutional Neural Networks: A Comprehensive Survey", 1Rahul Haridas, 2Jyothi R L, International Journal of Applied Engineering Research ISSN 0973-4562 Volume 14, Number 3 (2019) pp. 780-789 © Research India Publications. http://www.ripublication.com
4. "Classification using deep learning neural networks for brain tumors", Heba Mohsen a, El-Sayed A. El-Dahshan b,c , El-Sayed M. El-Horbaty d , Abdel-Badeeh M. Salem, Future Computing and Informatics.Journal-3.(2018)/68e71http://www.journals.elsevier.com/future-computing-and-informatics-journal
5. "Classifiers in Image processing", Rama Gaur ,Dr. V.S. Chouhan ,International Journal on Future Revolution in Computer Science & Communication Engineering ISSN: 2454-4248 Volume: 3 Issue: 6.
6. "Brain Tumor Segmentation Using Convolutional Neural Networks in MRI Images", Sérgio Pereira*, Adriano Pinto, Victor Alves, and Carlos A. Silva, IEEE TRANSACTIONS ON MEDICAL IMAGING, VOL. 35, NO. 5, MAY 2016.
7. "Speech Recognition Based on Convolutional Neural Networks", Du Guiming1 , Wang Xia1 , Wang Guangyan2 , Zhang Yan1 , Li Dan1, 2016 IEEE International Conference on Signal and Image Processing.
8. "Classification of Tumors in Human Brain MRI using Wavelet and Support Vector Machine", IOSR Journal of Computer Engineering (IOSRJCE) ISSN: 2278-0661, ISBN: 2278-8727Volume 8, Issue 2 (Nov. - Dec. 2012), PP 25-31, Mubashir Ahmad 1 , Mahmood ul-Hassan2, Imran Shafi 3 , Abdelrahman Osman4.
9. "Asynchronous Multiplex Communication Channels in 2-D Neural Network With Fluctuating Characteristics", Shinichi Tamura , Life Fellow, IEEE, Yoshi Nishitani, Chie Hosokawa, and Yuko Mizuno-Matsumoto, IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS.
10. K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in ICLR, 2015.
11. Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives[J]. Pattern Analysis and Machine Intelligence, IEEE Transactions on, 2013, 35(8): 1798-1828.
12. Ciregan,D.,U.Meier, and J. Schmidhuber. Multi-column deep neural networks for image classification. in Computer Vision and Pattern Recognition (CVPR), 2012 IEEE Conference on. 2012. IEEE.
13. M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," in ECCV, 2014.
14. Cun Y L, Boser B, Denker J S, et al. Handwritten digit recognition with a back-propagation network[C]// Advances in Neural Information Processing Systems. Morgan Kaufmann Publishers Inc. 1990:465.
15. Arslan B, Yorulmaz E, Akca B, Sagiroglu S (2016) Security perspective of biometric recognition and machine learning techniques. In: 15th IEEE Int. Conf. on Machine Learning and Application (ICMLA), Anaheim, California, USA, pp. 492-497.
16. D. E. Rumelhart, G. E. Hinton, R. J. Williams, et al., "Learning representations by back propagating errors," Cognitive modeling, vol. 5, no. 3, p. 1, 1988.
17. https://medium.com/@RaghavPrabhu/cnn-architectures-lenet-alexnet-vgg-googlenet-and-resnet-7c81c017b848.
18. https://www.machinecurve.com
19. https://hub.packtpub.com/convolutional-neural-networks-reinforcement-learning/
20. https://medium.com/@RaghavPrabhu/understanding-of-convolutional-neural-network-cnn-deep-learning-99760835f148
21. http://deeplearning.net/tutorial/lenet.html
22. https://www.coursera.org/lecture/convolutional-neural-networks/transfer-learning-4THzO

AUTHORS PROFILE

Ms.Pooja Kalange obtained UG Degree from Pune University and PG Degree from JNTU University .She is presently working as Asst.Professor in Institute of Industrial & Computer Management & Research, Pune.She has 9+ years of Teaching Experience.Her areas of interest include Data Warehousing & Data mining,Big data analytics,Data visualization & Reporting,Artificial Intelligence.

Ms.Megha Mutalikdesai obtained UG Degree from VTU University and PG Degree from JNTU University .She is presently working as Asst.Professor in Institute of Industrial & Computer Management & Research, Pune.She has 8+ years of Teaching Experience. Her areas of interest include IOT,Python,Image Processing,Data Science.