Proceedings of the 13th INDIACom; INDIACom-2019; IEEE Conference ID: 46181
2019 6th International Conference on "Computing for Sustainable Global Development", 13th - 15thMarch, 2019
Bharati Vidyapeeth's Institute of Computer Applications and Management (BVICAM), New Delhi (INDIA)

Impact of Deep Learning Approaches in Computer Vision System

Preeti Saini
[1]Department of Computer Science & Engineering.,
AIACTR, New Delhi (INDIA)

Bharti Nagpal
[2]Department of Computer Science & Engineering.,
AIACTR, New Delhi (INDIA)

Abstract- In Computer Vision Systems the task of information analysis has been grown in the most recent decade at a great rate with enormous ingress of multimodality data. A strategy with its establishment in artificial neural system, Deep Learning is rising as an adequate archetype for machine learning and it is assuring to alter the dictum in the field of artificial intelligence. It results into growing the activity in the origination of data driven model & analytical models on the basis of machine learning in Computer Vision Systems. Additionally, fast data storage, parallelization and Quick changes in computational power have added to the agile take-up of the innovation because of its prescient power and capacity to produce consequently improved abnormal state highlights & linguistic translation from the data. This paper mostly targets on keyareas of "Deep learning" technique in an arena of Face Recognition, Intelligent transportation systems, Action recognition, Biometrics and human computer interaction. Basically, this article subject to a evaluative inquiry in regard to merit, & likely demerits of various deep learning techniques as well as its forthcoming outlook & at present provide a thorough analysis of research conducting around deep learning technique in Computer Vision Systems.

Keywords-Deep learning; RNN, CNN Computer Vision

I. INTRODUCTION

A variant approach of machine learning is deep learning that favor computer systems to learn with experience and facts. This technique emerges as a novice pattern in machine learning and it conquers the disadvantage of signal processing's feature-extraction phase which is very time consuming & most often the extracted key points are not well enough to conveys the relation between patterns. Due to this advanced feature extraction techniques have been grown. Deep Learning Methodologies have proven to attain improved results than any other techniques in the context of image analysis which are based on human-elected features. The stochastic formulation of deep learning technique are very much established in the traditional neural network system (NN) research, although deep learning approach is distinctive to more traditional Neural network by virtue of it represents the utilization of many hidden neurons and layers usually in excess of two which go about as a plan advantage here related with new training criterions than in Neural network systems. In the meantime resorting to numerous neurons permits a noteworthy inclusion of the fundamental information within reach, a course of nonlinear mix of their yields the layer by layer a lower dimension projections of the information field.

Several models emerge in fame among the different methodological variations of of deep learning approach. Especially, CNN's (Convolutional Neural Networks) within the area of computer recognition systems had the tremendous impact. A high-level abstract feature emerges at each layer in the network. This CNN architecture is a biologically-inspired architecture that simulates a scheme to grasp imaged data in the manner of responsive field by a visual cortex. A CNN blueprint is explained as follows an interlaced set of feedforward layer's that equipped the convolutional filters succeeded via rectification, pooling layers or reduction process. On the Other hand, plausible methods of given technique insert which are established in design of RBMs(restricted Boltzmann machine's) for example, stack Auto encoders worked as deep Auto encoders, deep belief networks (DBNs) extending artificial intelligence. So far, alot of trial undertaking have been actualized different deep learning approaches for Computer Vision Systems, many of them providing comparable execution or as a rule better that of other substitute strategies. Despite the fact that, the usage of deep learning out how to acknowledgment System raises various issues which required to be fathomed. For Instance, deep learning approaches desires extensive computational assets without that training can become unreasonably long delayed. In additional, a large volume of labeled data during training is required by deep architecture, which is very difficult to attain in the field of Computer Vision. Sometimes a deep learning approaches can be afflicted via over fitting & convergence issues as a result auxiliary training methods are required to point out these concerns [8].

The framework of the given research is described as follows : II section of this paper examine various Deep learning methods to explain its potential pitfalls and relative pros. Section III outlines the literature survey on various applications of DNN and lastly, conclusions are condensed in Section IV.

II. DEEP LEARNING APPROACHES

One of the prior artificial neural networks (ANN) is "Perceptron" for supervised learning of binary classification. It scientifically defines structure how a biological neuron operates.

To resolve many complicated issues, neural networks have minimum one of hidden layers of Perceptron's been conferred. To make these neural network systems, many stages are commonly executed in which a new input label is provided to

 ISSN 0973-7529; ISBN 978-93-80544-32-8

the network, particularly the weights of neuron are changed on the basis of a learning procedure known as delta rule.. In particular, with no earlier learning arbitrary esteems are given to the network weights. A delta rule is generally executed by using back propagation technique & utilized by normal class of supervised neural network techniques amid the preparation To restrict the variance between the system output and the desired yield these network weights are altered.

A deep architecture is to be developed by appending many hidden layers to the network which can convey more complicated premises as the hidden layer represents nonlinear connection, these type of neural networks are referred as deep neural network. The concept of Deep learning has given new refined ways to deal with prepare DNN structures. In usual, DNNs can be prepared either with supervised learning and unsupervised algorithms. In the course of supervised methodology, training data are employ to train the deep neural network & then take in the weights that can decrement the mistake to conjecture a target an incentive for reason for classification. During unsupervised learning, training data is not required for training to be executed; this learning is commonly utilized for clustering, dimensionality reduction, and feature extraction.

Until now, a lot of deep neural network architectures have been made known in research which are discussed here.

A. *Recurrent Neural Network*

During 1980's **recurrent neural network** (**RNN**) were introduced [1] in which relation between the nodes forms a directed cycle to reveal dynamic temporal response. In RNN to present the yield of the data it utilizes two origins of input i.e. the present and the recent past. Consequently, it is well known that memory is available for RNN. Besides that, RNN is basic and capable layout it likewise endures from the vanishing slope and detonating inclination issues [2]. A Variation of recurrent neural network known as LSTM's (long short term memory units), was essentially best in class to tackle a purpose of the vanishing slant made via large info successions.

A RNN (recurrent neural network) is unlike to distinct sorts of deep neural network that utilizes varied weights on every layer, a LSTM or a RNN has same weights at over all the step as shown in Fig.1. This results in incredibly diminishing the added number of specifications that the system required to memorize. A recurrent neural network have been indicated extraordinary results in various natural language processing task for example in the arena of language modeling, bioinformatics, and speech recognition.

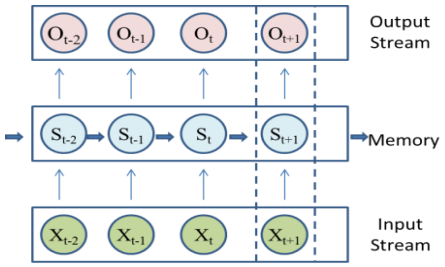

Fig. 1. Recurrent Neural Network (RNN)

B. *Restricted Boltzmann machine (RBM)*

A RBM (Restricted Boltzmann machine) was initially suggested in 1986 by Paul Smolensky and it is alternative to Boltzmann machine. These systems are shown by resort to stochastic units by an individual dissemination (for instance Gaussian) [3]. A RBMs learning method includes a few stages known as Gibbs sampling, that basically used to repeatedly change the weights to restrict the recreation mistake. These type of neural network are valuable in the event where it is demanding to display probabilistic connections between the factors. Restricted Boltzmann machine have numerous applications in classification, reduction of dimensions, learning of features.

An advantageous key point regarding RBM is binding appropriation upon hidden units that impacts on a given units. Applying RBM's as learning modules there are two important deep learning techniques has been utilized: a Deep Boltzmann machine & a Deep belief network.

1) *Deep Belief Network*

DBNs can be seen as an organization of straightforward, unsupervised systems, for example, Restricted Boltzmann machines (RBMs) or auto encoders in which every sub system's hidden layer fill as the visible layer for the succeeding next layer as mentioned in given below Fig. 2.

Fig. 2. Deep Belief Network's

2) *Deep Boltzmann Machines*

A DBM (Deep Boltzmann Machine's) is another deep neural network alternative which is based on the Boltzmann background proposed in [3]. The primary contrast of DBM with DBN is that the previous has undirected associations between all layers of the system. Therefore, to train a Deep Boltzmann Machines a stochastic maximum likeliness based approach is normally utilized to widen the lower bound of the probability. The principle inconvenience of Deep Boltzmann Machines is time unpredictability. In addition this perform improvement of the variables that is not useful for large intstruction set.

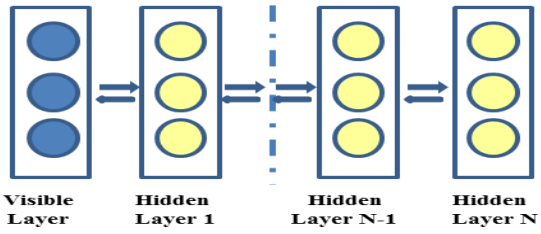

Fig. 3. Deep Boltzmann Machines

C. *Convolutional Neural Networks*

All the Deep neural networks which are discussed so far can't be extent efficiently via multidimensional information which have local coordinate information such as in a picture. The elementary worry is the quantity of parameters and the

quantity of nodes that it require for make it tremendous, which can't be deal with these deep neural networks. CNNs architecture as shown in Fig. 4 have been proposed to investigate visual imagery. A key contrast between fully connected artificial neural network and a convolutional example as a multilayer perceptron having neighborhood availability and shared the weights among a neurons. Within a layer in which every neuron is associated with other neuron in the information pattern. Due to the usage of convolution operator which are easy to conduct complex operations, the name of this network encounters from here. The process is illustrated as follows:

1) Many small filters are used to convolute the input image.
2) Sub sampling of the output produced at step1
3) The subsampling and Convolution forms are repeated as long as main features get extracted by taking into account the output produced at step2 as the new input.

The Fully convolutional network (FCNs) were introduced as a modification of CNN model which were beforehand intended for the purpose of classification of images. FCNs expand prior models, for example, AlexNet, GoogLeNet and VGG net by replaced the finally associated layers towards the ending of mentioned systems via convolutional layers.

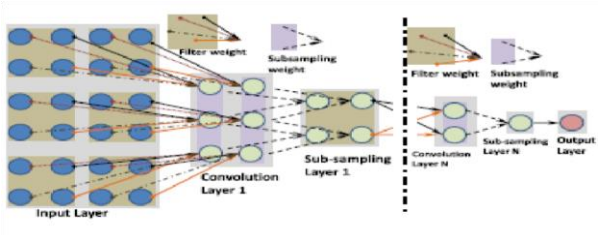

Fig. 4. Convolutional Neural Networks [8]

III. Hardware & Software Implementations

Table I records the most prominent programming packages that permit usage of altered deep learning systems in view of the methodologies portrayed up until now. Table II compiles the numerous applications of Deep neural network in the areas of computer vision system considered in this paper. In addition to the developing pattern of new deep learning structures are replaced into open source extends a few organizations for example, Nervana Systems [19] and Wolfram Mathematic [18] scientists to speed up the training procedure have chosen to give a cloud based administrations.

Framework	Developed by	License	Platform	Algorithms Supported	Core Language
Café [30]	Berkeley Vision & Learning Center	BSD 2-Clause license	Linux, Win, OSX, Android	RNN,CNN	C++,Python

IV. Conclusions

In the recent years the concept of deep learning has picked up a focal position in a field of pattern recognition and in machine learning. This paper focus key points that reduce the measure of human intercession in this procedure. This approach is useful for many concern in recognition systems and provides an incredible solution for unstructured information for example, those emerging from remote sensing, Biometrics etc. The notion of deep learning have been utilized mostly in applications where the datasets were balanced, other concern is that this technique dominatingly relies upon a lot of training information. Such prerequisites make basic established passage boundaries of machine learning i.e., information accessibility and protection of data. It is important to note that a great degree of empowerment in this field is given by the powerful IT organizations e.g. Microsoft, Google, Baidu and Facebook etc. they hold, are forceful worked by information gathering and tremendous storage facilities. To apply deep learning to computer vision domain analysts have been urged by providing free software packages for assist this research. Taking a vision at it from the superb side which refined it an intriguing incline and supported the desires of what machine learning techniques realizing could accomplish alone. Eventually, we ought not to consider deep learning as a silver shot for each and each confornt set for computer vision arena. From implementation point of view, by considering another learning algorithms that can provide same ability with lesser resources, less parameterization, tuning, and higher interpretability it is however uncertain whether the huge amount of training particulars & computational measures are required to implement deep learning at full performance is either valuable. Therefore , this paper infer that deep learning has given a good improvement to neural networks & connectionism from the real combination of new advancement in area of parallel processing also accredited via coprocessors.

V. References

[1] R. J. Williams, D. Zipser(1989), "A learning algorithm for continually running fully recurrent neural networks," Neural Comput., Vol. 1, No. 2, pp. 270–280.

[2] Y. Bengio, P. Simard, P. Frasconi (1994), "Learning long-term Dependencies with gradient descent is difficult," IEEE Transaction Neural Netork, Vol. 5, No. 2, pp. 157–166.

[3] R. Salakhutdinov , G. E. Hinton (2009), "Deep boltzmann machines." In Proc. Int. Conf. Artif. Intell. Stat., Vol. 1, Art. No. 3.

[4] Y. LeCun, L. Bottou, Y. Bengio, P. Haffner(1998), "Gradient-based learning databases," Pattern Recognition, Vol. 44, No. 3, pp.572–587[5] J. Long, E. Shelhamer , T. Darrell (2015), "Fully convolutional networks for semantic segmentation," in Proceedings of IEEE Conference Comput. Vision Pattern Recognition., pp. 3431–3440.

[5] Yuting Zhang, Kui Jia, Jiwen Lu Shenghua Gao, and Yingying Zhang Member, IEEE ,"Single Sample Face Recognition via Learning Deep Supervised Autoencoders".

[6] Esen, E., & Collison, J. (2005). Employee development survey report. Society For Human Resource Management, V(A) 4.

[7] Y. LeCun, Y. Bengio, G. Hinton(2015), "Deep learning," Nature, Vol. 521, No. 7553, pp. 436–444, 2015.

[8] D. Lee and Y. Park(2009), "Vision-based remote control system by motion detection and open finger counting," IEEE Trans. Consum. Electron., Vol.5, No. 4, pp. 2308-2313.

[9] Y. Wang, and Y. Ruoyu(2013), "Real-time hand posture recognition based on hand dominant line using Kinect," in Proc. ICME, San Jose, CA, pp. 1-4.

[10] P. Bao et al.: Tiny Hand Gesture Recognition without Localization via a Deep Convolutional Network

[11] M. El Ayadi, M. S. Kamel, and F. Karray(2011), "Survey on speech emrecognition: features, classification schemes, and databases," Pattern Recognition, Vol. 44, No. 3, pp. 572–587.

[12] C.-N. Anagnostopoulos, T. Iliou, and I. Giannoukos(2015), Artificial Intelligence Review: a survey from 2000 to 2011, "Features and classifiers for emotion recognition from speech", vol. 43, no. 2, pp. 155–177.

[13] Eyben, K. R. Scherer, B. W. Schuller, J. Sundberg, E. Andr´e, C. Busso, L. Y. Devillers, J. Epps, P. Laukka, S. S. Narayanan et al. (2016), "The Geneva minimalistic acoustic parameter set (GeMAPS) for voice research and affective computing," IEEE Transactions on Affective Computing, vol. 7, no. 2, pp. 190–202.

[14] E.Schroff, D. Kalenichenko, and J. Philbin(Jun. 2015) "FaceNet: A unified embedding for face recognition and clustering," in Proc. IEEE Conf. Comput.Vis. Pattern Recognition., pp. 815–823.

[15] W. Huang, G. Song, H. Hong, and K. Xie,(Oct. 2014) "Deep architecture for traffic flow prediction: Deep belief networks with multitask learning," IEEE Trans. Intell. Transp. Syst., vol. 15, no. 5, pp. 2191–2201.

[16] StrategyYue Huang, Ruiwen Wu, Ye Sun, Wei Wang, and Xinghao Ding (AUGUST 2015)," Vehicle Logo Recognition System Based on Convolutional Neural Networks With a Pretraining", IEEE TRANSACTIONS ON INTELLIGENT TRANSPORTATION SYSTEMS, VOL. 16, NO. 4.

[17] Shiqing Zhang, Shiliang Zhang, Member, IEEE, Tiejun Huang, Senior Member(IEEE) and Wen Gao(MARCH 2016)," Speech Emotion Recognition Using Deep Convolutional Neural Network and Discriminant Temporal Pyramid Matchingn", JOURNAL OF LATEX CLASS FILES, VOL. XX, NO. XX1.

[18] Arief Koesdwiady, Student Member, IEEE, Ridha Soua, and Fakhreddine Karray, Senior Member(IEEE) DECEMBER 2016," Improving Traffic Flow Prediction With Weather Information in Connected Cars:A Deep Learning Approach", IEEE TRANSACTIONS ON VEHICULAR TECHNOLOGY, VOL. 65, NO. 12.

[19] Furkan Isikdogan, Alan C. Bovik, Fellow, IEEE, and Paola Passalacqua," Surface Water Mapping by Deep Learning" , IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING.

[20] B. Pan, Z. Shi, and X. Xu(May 2017.), "R-vcanet: A new deep-learning-based hyperspectral image classification method," IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 10, no. 5, pp. 1975–1986.

VI. TABLE I: WELL-KNOWN SOFTWARE PACKAGES THAT PROVIDE DNN IMPLEMENTATION

Theano [36]	University of Montreal	BSD	Cross-platform	RNN,CNN,DBN	Python
NVIDIA CUDA	NVIDIA	Freeware	Windows XP and later, macOS, Linux	RNN,CNN	C, C++, and Fortran
TensorFlow	Google Brain Team	Apache 2.0 open source license	Linux, macOS, Windows, Android	RNN,CNN,DBN	Python, C++
Torch	Ronan Collobert	BSD License	Linux, Android, Mac OS X, iOS	RNN,CNN,DBN	C, CUDA, C++
MXNet	Distributed Machine Learning Community	Apache 2.0	Windows, Linux	RNN,CNN,DBN	C++, Python, R, Julia, JavaScript, Scala, Go, Perl
Keras	François Chollet	MIT	Cross-platform	RNN,CNN,DBN	Python
Eclipse Deeplearning4j	Apache 2.0	Adam Gibson, Chris Nicholson, Josh Patterson	Linux, macOS, Windows, Android	RNN,CNN,DBN	Java, Scala, CUDA, C, C++, Python, Clojure
SINGA	Apache Software Foundation	Apache License 2.0	Linux, Mac OS, Windows	RNN,CNN,DBN	C++, Python, Java

TABLE I. DIFFERENT DEEP LEARNING METOD VIA AREAS AND APPLICATION

Areas	Application	Input Data	Base Method	References
Intelligent transportation systems	vehicle logo recognition	Images are first captured from the monitoring framework and the vehicle tag is then recognized for each picture	Convolutional neural networks (CNNs)	Yue Huang et. al
Forensic	Face Photo-Sketch Recognition	facial attributes	DCNN	Christian Galea et.al
Biometrics	Face Recognition	shape and position of facial highlights	Learning Deep Supervised Autoencoders	Shenghua Gao et.al
Remote sensing	classifying water bodies	Landsat imagery		Furkan Isikdogan et. al
Human computer interaction (HCI)	Action recognition	sparse spatio-temporal interest	Deep convolutional neural	Xin Chen et.al, L. Wang et. al

