A Comprehensive Study of Deep Learning Architectures, Applications and Tools

Nilay Ganatra

IJCSE International Journal of Computer Sciences and Engineering Open Access
Review Paper Vol.-6, Issue-12, Dec 2018 E-ISSN: 2347-2693

A Comprehensive Study of Deep Learning Architectures, Applications and Tools

Nilay Ganatra[1*], Atul Patel[2]

[1,2]Dept. of Computer Science and Applications, CHARUSAT, Changa, Gujarat, India

[*]*Corresponding Author: nilayganatra.mca@charusat.ac.in Tel.: +91 – 99240 48124*

Available online at: www.ijcseonline.org

Accepted: 25/Dec/2018, Published: 31/Dec/2018

***Abstract*—** The Deep learning architectures fall into the widespread family of machine learning algorithms that are based on the model of artificial neural network. Rapid advancements in the technological field during the last decade have provided many new possibilities to collect and maintain large amount of data. Deep Learning is considered as prominent field to process, analyze and generate patterns from such large amount of data that used in various domains including medical diagnosis, precision agriculture, education, market analysis, natural language processing, recommendation systems and several others. Without any human intervention, deep learning models are capable to produce appropriate results, which are equivalent, sometime even more superior even than human. This paper discusses the background of deep learning and its architectures, deep learning applications developed or proposed by various researchers pertaining to different domains and various deep learning tools.

***Keywords*—** deep learning, architectures, applications, tools

I. INTRODUCTION

Deep Learning is considered as the subset of machine learning which is intern subset of Artificial Intelligence, a prominent field of computer science over the past decade. Artificial Intelligence makes machines to think intelligently without minimal human intervention. Machine learning comprises with various algorithms that are capable to model high level abstractions from input data. Deep Learning provides a more adaptive way using deep neural network that learns feature itself from the given input data and make machine capable for taking decision. Unlike task specific algorithms of machine learning, deep learning is a method based on learning data representations. Learning can be supervised, semi-supervised or unsupervised. Deep learning provides set of algorithms and approaches that learns features and tasks directly from data. Data can be of any type, structured or unstructured, including images, text or sound. Deep learning is often referred as end-to-end learning because it learns directly from data. Moreover, Deep learning techniques works without human mediation and sometime capable of producing more accurate result than human being itself. Nowadays, deep learning is widely used in the areas like computer vision, natural language processing, pattern recognition, object detection. Representative learning methods of deep learning provides multiple level of representation, generated by linking simple but non-linear modules that transmute the representation at one level into a representation at next higher layer, slightly in more abstract way.

II. DEEP LEARNING: OVERVIEW AND BACKGROUND

Machine Learning provides the vast collection of algorithms including singular algorithms, together with statistical approach like Bayesian networks, function approximation such as linear and logistic regression, or decision trees. These algorithms are indeed powerful, but there are few limitations of these algorithms when to learn for enormously complex data representations. Deep learning is emerged from cognitive and information theories and human neurons learning process along with strong interconnection structure between neurons is looking for to imitate. One of the key feature of computing neurons and the neural network model is, it can be able to apply generic neurons to any type of data and learn comprehensively [1]. Deep learning is considered as a promising avenue of research as it is capable of automatically identifying the complex features at a high level of abstraction. It is about learning multilevel representation and abstraction, which is useful for the data, such image, text and sound. One of the exceptional Deep learning characteristics is its capability of using unlabeled data during the training process [2].

According to the definition of Deep learning, it is considered as the application of multi-layer neural network with multiple neurons at each layer to perform the desired tasks like classification, regression, clustering and others. Fundamentally, each neuron with activation function is considered as the single logistic node, which connected to the

input in the next layer of it, loss function is calculated to modify the weights at each neuron and optimize to make it suitable for input data. Each layer of neural network layer multiple neurons initiated with dissimilar weights and try to learn on the input data concurrently. Thus, in multiple layers with multiple nodes, each node learns from the output of the previous layers, and gradually decreases the approximation of the real input data to provide accurate output representation set [3]. This leads to lot of complexity between multiple interconnected neurons.

The term "deep learning" was initially used by Igor Aizenberg and colleagues in or around 2000 while talking about Artificial Neural Network (ANN). However, the first mathematical model of neural network was introduced in 1943 by Walter Pitts and Warren McCulloch and published in the seminar "A Logical Calculus of Ideas Immanent in Nervous Activity". In 1965 Alexey Ivakhnenko and V.G. Lapa created first working deep learning network by applying theories and idea to the point. After that, Kunihiko Fukushima had introduced "Neocognitron"- and artificial neural network that learned how to recognize visual patterns in 1979-80. Likewise, many authors contributed in subsequent years. However, in 2006 Hinton claimed that he knew how the brain works, and highlighted the idea of unsupervised pretraining and deep belief nets. Then, in 2009 Fei-Fei Li launched ImageNet, in 2011 Alex Krizhevsky created AlexNet built upong LeNet (by Yann LeCun years earlier). In 2012, convolutional neural network (CNN) is used by Geoff Hinton and his team. They found that it is capable to learne its own features and the error rate decreases to 18.9% [4]. In 2014, GoogLeNet - a Google's own deep learning algorithm considered as the one of the base in the research field of deep learning as it is able to down the error rate about 6.7%. After this, deep learning has successfully applied to one of the promising application, speech recognition. Deep learning architectures have produced impressive results in all most all natural language processing tasks starting from including sentiment analysis [5] to vocal language understanding [6] including information retrieval, machine translation, contextual entity linking, and many more. One of the notable application of deep learning algorithms are product recommendations that are implemented by all e-commerce websites. Other Deep learning models are also used in the fields of object tagging, face recognition, drug innovation and toxicology, weather forecasting, financial assistance and many more [7].

III. DEEP LEARNING ARCHITECTURES

The advancement in the field of deep learning architectures in last two decades offers significant opportunities for applying it in various fields and its applications. This paper presents three most popular deep learning architectures - Deep Belief Network, Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs).

A. Deep Belief Network (DBN)

Deep belief network (DBN) is a typical unsupervised network architecture that provides a unique training algorithm. The DBN is multilayer network contains many hidden layers in which each pair of connected layer is made up with a restricted Boltzmann machine(RBM) [8] [9]. As it contains multiple layers, a DBN is also represented as a stack of more than one RBMs. The symmetric connections exist between top two layers, forms an associative memory. Moreover, top-down connection is formed by the lower layer to the layer above [10]. The training process of DBN starts with learning a layer of features from the visible units. For that, it uses an algorithm namely Contrastive Divergence (CD). Then after, activations of previously trained features as visible units and learns about the features of features in a second hidden layer. The entire DBN training is completed after the learning of the final hidden layer. The following figure 1 displays the general architecture of DBN [11].

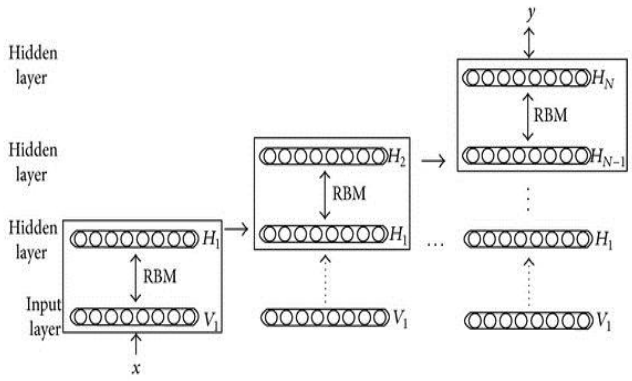

Figure 1. A General Architecture of Deep Belief Network (DBN)

B. Recurrent Neural Networks (RNN)

Recurrent neural network (RNN) is one of the first algorithm that remembers its input, as it consist of internal memory, and therefore most suitable for solving machine learning problems which involve sequential data, such as speech and language. For the given input sequence, RNNs process one element at a time and retained in hidden units as it is having an internal memory [12]. A state vector implicitly contains past elements information as history of the given sequence. Training of RNNs are like typical neural networks and uses backpropagation learning algorithm. Although, training of RNNs are seems to be problematic as the characteristics of backpropagation gradients. They may raise or shrink at every timestamp and over many steps likely to be exploded or vanished [13] [14]. The following figure 2 demonstrates a conventional architecture of recurrent neural network [S53]. The figure 2 shows the architecture of RNN [15].

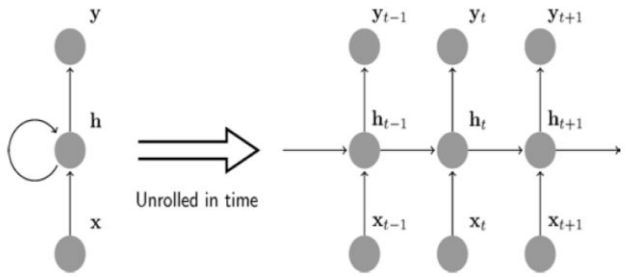

Figure 2. An Architecture of Recurrent Neural Network (RNN)

C. Convolutional Neural Networks (CNN)

A Convolutional Neural Network (CNN) is derived from the neurobiological process of visual cortex, which defines convolutional layer(s) followed by fully connected layer(s). Between these layers, it may exist subsampling tasks. It is considered as variant of the multilayer perceptron. Generally, CNN are most appropriate for the data sets that requires large number of nodes and parameters to be processed. A Convolutional Neural Network architecture contains neurons in three dimensions (width, height, depth), as presented in one of the layers as shown in figure 3 [write reference]. The deep architecture of the network provides hierarchical feature extraction e.g. the filters of the first layer cable of identifying edges or color blobs, second layer some curves of the shapes, next layer filters start learning object parts and final layer filters are capable of identifying the object. [16]. Some prominent CNN architectures are AlexNet [17], VGGNet [18], and GoogLeNet [19]. Figure 3 demonstrate the architecture of Convolutional Neural Network which contains Convolutional layer, pooling layers and fully connected layers [20].

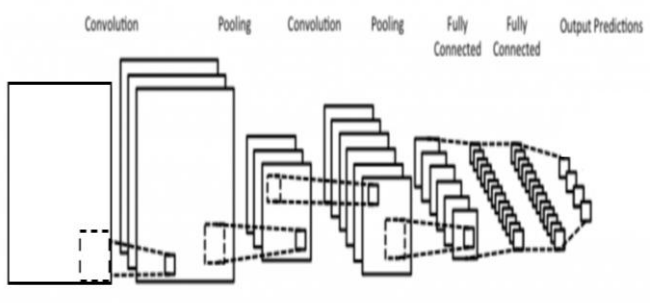

Figure 3. An Architecture of Convolutional Neural Network (CNN)

IV. APPLICATIONS OF DEEP LEARNING

In this section of paper, the major applications of deep learning have been reviewed. The emergence of applications that have made up using deep learning has grown steadily since the last decade. The major areas, where deep learning has successfully applied are medical diagnosis, precision agriculture, education, market analysis, natural language processing, recommendation systems and several others. . Moreover, advancements in the field of big data and cloud computing are enable machine learning to grow and expand. It also makes possible to execute deep models [21].

Shine et al. [write reference] proposed a deep learning architecture with a stacked sparse autoencoder. It is successfully applied to identify the organ from the complex dataset, for which it is very difficult to detect the labels of the sample in given dataset. In this paper, they have generated sparse encoder model by stacking number of unsupervised feature learning layers and trained the using greedy methods.

Tai Sing Lee and David Mumford (2003) [22] proposed a hierarchical Bayesian inference framework to be used in the visual cortex. In this work, they represented use of recurrent feedforward/feedback loops in visual cortex. They have used it to integrate top-down contextual priors and bottom-up observations and implemented concurrent probabilistic interface with the visual hierarchy. They have suggested the algorithms of particle filtering and Bayesian-belief propagation.

Geoffrey E. Hinton et al. (2006) [23] demonstrated the use of "complementary priors" to overcome the difficulty of explaining-away effects that make inference difficult in belief nets with many hidden layers. They have derived greedy algorithm which is fast and fine-tunes the weights using a contrastive version of the wake-sleep algorithm with slower learning procedure. After proper training a model gives better classification for handwritten digit images and their labels than the best discriminative learning algorithms.

A language modeling (LM) is a capacity that catches the remarkable salient statistical characteristics of the successions of words in natural language. Minh and Hinton (2007) [24] used temporally factored RBM for language modeling. The factored RBM considered distributed representation also for the words being predicted other than the context words. This methodology is summed up to more profound structures as detailed in (Mnih and Hinton, 2008).

For speech recognition, neural network has been given very promising results since many years. On the contrary, its performance was observed as lower compared to the Hidden Markov Models (HMMs). Mohamed et al. (2009) [25] presented the work in which they have used a five layer DBN as a replacement of the Gaussian mixture component of the GMM-HMM .Here in this work, as a modeling unit, the monophone state was used. This novel DBN integrated HMM approach with monophones was seemed to realize higher telephone recognition precision than the triphone GMM-HMM frameworks.

A modified DBN is introduced by Nair and Hinton (2009) [26] where they have shown the utilization of a third-order Boltzmann machine by top-layer. . This kind of DBN is connected to the NORB database- a 3-dimensional object recognition task. The error rate is reported very close to best result published on this task and proved the better accuracy of DBN over other models like SVMs.

For the robustness of DBN, two approaches are suggested by Tang and Eliasmith (2010) [27]. In the first layer of DBN, they have introduced sparse connection and standardized the model. Other than this, probabilistic denoising algorithms are also developed in the second approach.. For recognition of noisy images, both approaches are proven feasible as they have improved occlusion and arbitrary noise.

In the area of speech recognition, Abdel-Hamid Ossama et al. (2012) [28] suggested a hybrid NN-HMM model that is based on CNN . For regularization of speaker adjustment, they suggested of using local filtering and max pooling approach in frequency domain. During experiments, results have shown that there is a 10% relative error reduction as they used CNN model.

To provide ease for the training of deeper networks, Kaiming He et al. (2015) [29] presented a framework based on a residual learning. Despite of learning unreferenced functions, the layer is reformulated as learning residual function in reference to the layer inputs. They have got very promising results while evaluating residual nets on the ImageNet dataset with a depth of up to 152 layers that achieves 3.57% error on the ImageNet test set.

David Silver et al. (2016) [30] have introduced the new approach to computer GO which uses 'value networks' to assess the positions on the board and 'policy network' to decide moves. These profound neural networks are trained using combination of supervised learning form human expert, and reinforcement learning from game of self-play. By introducing new search algorithm with combination of Monte Carlo simulation with value and policy network their program named AlphaGo accomplished a 99.8% wining chances compare with Go programs. A novel approach using CNN For classification of carotid plaques, Yuxi Dong et al. (2017) [31] proposed a novel approach using CNN. Carotid plaques may cause strokes. For accommodation of the multi-contrast MRI images, they have used models that do pixel wise predictions.

V. DEEP LEARNING LIBRARIES

There are various tools and libraries are available that are used to implement real time applications that are based on deep learning models.

A. Theano

It is Python based numerical library for fast mathematical computations that can be run on the both CPU or GPU. It is tightly integrated with NumPy [32]. This library makes writing deep learning models easy, and provide the option of training models and a GPU for better performance. Neural network and data is represented as matrices and operation on data are defined as matrix calculations. Vectored code run quickly because of parallel execution. This library was designed to deal with large neural network computations used in Deep Learning. It was one of the pioneer industry standard library of its kind for Deep Learning research and development.

B. Deeplearning4j

It is open source deep learning framework developed by machine learning group and supported by Skymind team. Deeplearning4j is written in Java and it is compatible with any JVM based language, like Scala, Kotlin or Clojour. The primary computations are written in C, C++ and CUDA language. DL4J takes the benefits of various distributed computing frameworks like Apache Spark and Hadoop to speed up the training. Performance of DL4J on multiple GPU is equivalent to Caffe. It offers support for all deep nets like neural tensor network, deep autoencoder, deep belief nets, restricted Boltzmann machine, etc.

C. Caffe

Berkeley AI Research (BAIR) and community contributors have introduced this open source deep learning toolkit It is written in C++, with a Paython interface. This deep learning framework is made with expression, speed and modularity in mind [33]. It supports Models and optimization by configuration without hard-coding. It supports NVIDIA cuDNN and Intel MKL, GPU and CPU based acceleration computational kernel libraries.It supports many types of deep learning architecture for image classification and segmentation. It mainly supports CNN, LSTM, RNN and deep neural network. It provides very limited abstraction, so it is extremely easy to do unconventional, hard-core modifications.

D. Torch

It is open-source scientific framework which provides huge support for machine learning algorithms. It is scripting language based on Lua programming language. It also provides support for wide range of deep learning algorithms, and uses LuaJIT scripting language and using it provide interface to C. The core package torch supports flexible N-dimensional array or Tensor, that provide routines for indexing, transposing, slicing, resizing and cloning. Torch provide optimized libraries to simplify the use of popular neural network. You can develop arbitrary graphs of neural network, and run them parallel on CPUs and GPUs in effective manner [34].

E. TensorFlow

It is open source software library for high performance mathematical computations. Flexible architecture of TensorFlow allows easy deployment of computations on variety of platforms like CPUs, GPUs and TPUs, and from single machine to clusters of server to hand-held and edge device. It was invented with the intention to express a numeric computation as a graph; G(e,n) where n stands for node and e denote for edge, mathematical operations of the graph are represented as nodes, while edge of the graph represents the tensor(multidimensional data) transferred between them.

VI. CONCLUSION

Deep learning models possesses a capability to produce appropriate results without human intervention and it makes them very promising for solving real time problems. In this paper, the introduction, background and usage of deep learning architectures are explained. Also, real time applications based on deep learning architectures are also discussed. Moreover, the tools and libraries currently available to realize deep learning applications are discussed to provide insight for the researchers working in the similar area.

REFERENCES

[1]M. Nielsen, "Neural networks and deep learning," 2017. [Online]. Available: http: //neuralnetworksanddeeplearning.com/.

[2]U. V. SurajitChaudhuri, " Anoverview of business intelligence technology," Communications of the ACM, vol. 54, no. 8, p. 88–98, 2011.

[3]A. P. Sanskruti Patel, "Deep Leaning Architectures and its Applications A Survey," INTERNATIONAL JOURNAL OF COMPUTER SCIENCES AND ENGINEERING , vol. 6, no. 6, pp. 1177-1183, 2018.

[4]I. S. a. G. E. H. A. Krizhevsky, "Imagenet classification with deep convolutional neural networks," Advances in neural information processing systems, p. 1097–1105, 2012.

[5]A. P. J. Y. W. J. C. C. D. M. A. Y. N. a. C. P. R. Socher, "Recursive deep models for semantic composition- ality over a sentiment treebank," in in Proceedings of the conference on empirical methods in natural language processing, Citeseer, 2013.

[6]Y. D. K. Y. Y. B. L. D. D. H.-. T. X. H. L. H. G. T. D. Y. a. G. Z. G. Mesnil, "Using recurrent neural networks for slot filling in spoken language understanding," IEEE/ACM Transactions on Audio, Speech, and Language Processing, vol. 23, no. 3, pp. 530-539, March, 2015.

[7]Q. V. Le, "Building high-level features using large scale unsupervised learning," in IEEE International Conference on Acoustics, Speech and Signal Processing, May-2013.

[8]G. E. H. a. T. J. S. S. E. F. G. E. H. a. T. J. S. S. E. Fahlman, "Massively parallel architectures for AI: Netl, thistle, and Boltzmann machines," Washington, DC, USA, 1983.

[9]T. J. S. a. D. H. A. G. E. Hinton, "Boltzmann machines: Constraint satisfaction networks that learn," Technical Report CMU-CS-84-119, CarnegieMellon University, Pittsburgh, PA, USA, 1984.

[10]D. b. n. S. 4. Geoffrey E. Hinton (2009), "Deep belief networks, Scholarpedia," 2009.

[11]H. &. C. Y. &. C. L. Wang, "A Vehicle Detection Algorithm Based on Deep Belief Network," The Scientific World Journal, vol. 10.1155/2014/647380, 2014.

[12]J. Schmidhuber, "Deep Learning," Scholarpedia.

[13]S. U. z. d. n. N. Hochreiter, " Diploma thesis," in German, 1991.

[14]Y. S. P. &. F. P. Bengio, " Learning long-term dependencies with gradient descent is difficult," IEEE Trans, vol. Neural Networks 5, p. 157–166 , 1994.

[15]S. &. L. B. &. Y. S. Min, "Deep Learning in Bioinformatics. Briefings in Bioinformatics.," 18. 10.1093/bib/bbw068, 2016.

[16]N. A. a. G. M, "A Review on Deep Convolutional Neural Networks," IEEE Xplore, pp. 0588-0592, 2017.

[17]I. S. a. G. E. H. A. Krizhevsky, "ImageNet classification with deep convolutional neural networks," in in Proc. NIPS, 2012.

[18]K. S. a. A. Zisserman, "Very deep convolutional networks for large-scale image recognition," CoRR, vol. vol. abs/1409.1556, Apr. 2015.

[19]C. S. e. al, "Going deeper with convolutions," in in Proc. CVPR, 2015.

[20]"4. Convolutional Neural Networks Tutorial in TensorFlow, retrieved from http://adventuresinmachinelearning.com/convolutional-neural-networks-tutorial-tensorflow/," May 2018. [Online].

[21]H. Z. W. G. H. J. N. a. W. Y. Z. Chen, "A streaming-based network monitoring and threat detection system," in IEEE 14th International Conference on Software EngineeringResearch, Management and Applications (SERA), 2016.

[22]T. S. L. a. D. Mumford, "Hierarchical Bayesian inference in the visual cortex," J. Opt. Soc. Am. A, vol. Vol. 20, July 2003.

[23]S. O. Y.-W. T. Geoffrey E. Hinton, "A Fast Learning Algorithm for Deep Belief Nets," Neural Computation, pp. Pages 1527-1554, July 2006.

[24]A. M. a. G. Hinton, "Three new graphical models for statistical language modelling," ICML '07 Proceedings of the 24th international conference on Machine learning, pp. 641-648, 20 - 24 June 2007.

[25]G. D. a. G. H. Abdel-rahman Mohamed, "Deep Belief Networks for phone recognition," in Proceedings of the NIPS Workshop on Deep Learning for Speech Recognition and Related Applications, 2009.

[26]V. N. a. G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," ICML'10 Proceedings of the 27th International Conference on International Conference on Machine Learning, pp. 807-814 , June 2010.

[27]Y. T. a. C. Eliasmith, "Deep networks for robust visual recognition," in Proceedings of the 27th International Conference on Machine Learning (ICML-10), Haifa, Israel, 2010.

[28]A.-r. M. H. J. G. P. Ossama Abdel-Hamid, "Applying Convolutional Neural Networks concepts to hybrid NN-HMM model for speech recognition," in 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2012.

[29]K. H. X. Z. S. R. J. Sun, "Deep Residual Learning for Image Recognition," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2015.

[30]A. H. a. m. David Silver, "Mastering the game of Go with deep neural networks and tree search," in Nature 529(7587):484-489, 2016.

[31]Y. P. J. Z. a. W. X. Y. Dong, "Learning to Read Chest X-Ray Images from 16000+ Examples Using CNN," in 2017 IEEE/ACM International Conference on Connected Health: Applications , Systems and Engineering Technologies (CHASE), Philadelphia, 2017.

[32]O. B. F. B. P. L. R. P. G. D.-. j. J. T. D. W.-F. a. Y. B. J. Bergstra, "Theano: A cpu and gpu math compiler in python," in 9th Python in Science Conf, 2010.

[33]E. S. J. D. S. K. J. L. R. G. S. G. a. T. D. Y. Jia, "Caffe: Convolutional architecture for fast feature embedding," in 22nd ACM interna- tional conference on Multimedia. ACM, 2014.

[34]K. K. a. C. F. R. Collobert, "Torch7: A matlab-like environment for machine learning," in in BigLearn, NIPS Workshop, no. EPFL-CONF-192376, 2011.

Authors Profile

Mr. Nilay Ganatra is currently working as an Assistant Professor in Faculty of Computer Science and Applications, Charotar University of Science and Technology, Changa. He is currently pursuing his research in the field of deep learning and machine learning. He is having 8 years of experince inclunding acadamics and research. He has published many research papers in various national/international journals.

.

Dr. Atul Patel is currently working as a Dean and Professor in Faculty of Computer Science and Applications, Charotar University of Science and Technology, Changa. He is having more than 20 years of experince inclunding acadamics and research. He has published more than 40 research papers in various reputed international journals. His area of intrest is Computer Network, Data Analytics, and Machine Learning.

