Vol.01/ Issue.02 Pages: 73-82

http://irojournals.com/aicn/

DOI: https://doi.org/10.36548/jaicn.2019.2.003

SURVEY ON EVOLVING DEEP LEARNING NEURAL NETWORK

**ARCHITECTURES** 

Dr. Abul Bashar,

Department of Computer Engineering, Prince Mohammad Bin Fahd University, Al Khobar, Kingdom of Saudi Arabia - 31952.

Email: abashar@pmu.edu.sa

Abstract: The deep learning being a subcategory of the machine learning follows the human instincts of learning by example to produce accurate results. The deep learning performs training to the computer frame work to directly classify the tasks from the documents available either in the form of the text, image, or the sound. Most often the deep learning utilizes the neural network to perform the accurate classification and is referred as the deep neural networks; one of the most common deep neural networks used in a broader range of applications is the convolution neural network that provides an automated way of feature extraction by learning the features directly from the images or the text unlike the machine learning that extracts the features manually. This enables the deep learning neural networks to have a state of art accuracy that mostly expels even the human performance. So the paper is to present the survey on the deep learning neural network architectures utilized in various applications for having an

accurate classification with an automated feature extraction.

Keywords: Deep Learning, Deep Neural Network, Machine Learning, Traditional Neural Network, Speech Recognition, Image

Recognition, NLP.

1. INTRODUCTION

The deep learning is a subset of the machine learning and based on the algorithms that are stimulated by the

functioning of the brain and the way they are structured. The deep learning educates the computer system to produce

results by training them on the previously available examples. The capability of the deep learning has enabled it to

accomplish results that were impossible once upon a time. The remain as the pivotal mechanics behind most of the

applications that are found nowadays such as the self-driving vehicles, voice control in consumer electronics such as

the phone, Television, smart machines etc.

ISSN: 2582-2012 (online)

73

Vol.01/ Issue.02 Pages: 73-82

http://irojournals.com/aicn/

DOI: https://doi.org/10.36548/jaicn.2019.2.003

The computer frame work in deep learning is trained to directly classify the tasks from the documents available in the form of the image or text or sound. The deep learning models ensure a state of art accuracy and a performance the excels the humans

The deep learning was conceived in 1980s, but has become predominant only in the recent years this is because of the substantial amount data set and the computing power that is required for the training. Recently the utilization of the cloud computing has reduced the training time of the deep learning to hours from weeks. Few areas where the deep learning is prominent are automated driving, aerospace and defense, medical research, industrial automation, consumer electronics etc. The fig.1 below shows the performance of the deep learning and the older learning algorithms under various areas.

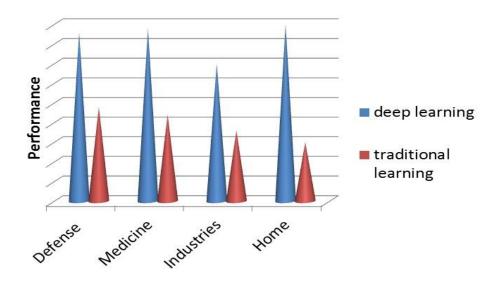


Fig.1 Performance of Deep Learning

The fig.1 present the performance of the deep learning from the huge amount of data set collected from different areas from defense to consumer devices at home. Many of the deep learning techniques utilizes the neural network architectures, to classify a data set acquired and are most probably referred as the deep neural networks. In deep neural networks the word deep points to the number of layers that are hidden, unlike traditional neural networks that can hold two or three hidden layers; the deep neural network can hold more than hundred and fifty hidden layers. This makes them capable of directly learning the features from the data itself eluding the necessity of the manual feature extraction and ensures higher accurate results compared to the traditional neural networks. The fig.2 below

Vol.01/ Issue.02 Pages: 73-82

http://irojournals.com/aicn/

DOI: https://doi.org/10.36548/jaicn.2019.2.003

shows the difference between the traditional learning and the deep learning in categorizing an object. Where the fig 2. (a) Is the classification using the traditional learning method and the fig 2. (b) Is the deep learning method that automatically learns from the data set. Fig shows how the learning algorithm identifies a desired object using the traditional learning and the deep learning.

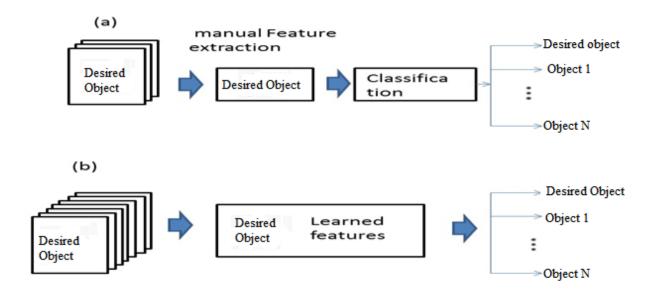


Fig. 2 (a) Traditional Learning, (b) Deep learning

The most popular DNN is the convolutional neural network. It convolutes the learned features and the input data and utilizes the two dimensional convolutional layers, it eludes the necessity for the manual feature extraction. The convolutional neural network works by extracting the features directly from the images. The necessary features are not pre-trained but learned in parallel during the training of the network on the collection of the images. This makes the DNN more compatible and suitable for the accurate classification of the objects.

The paper aims in presenting the survey of the evolving deep neural network architectures in various applications that are related to the speech recognition, image recognition, the natural language processing and the other applications.

Vol.01/ Issue.02 Pages: 73-82

http://irojournals.com/aicn/

DOI: https://doi.org/10.36548/jaicn.2019.2.003

The remaining paper proceeds with the literature survey presenting the evolving deep learning neural network, in

various applications in section 2 and the its uses in the other application in section 3 and the conclusion in the

section 4.

2. LITERATURE SURVEY

The neural network is fundamentally built to imitate the activity of the human brain. The experts reveal the deep

neural network as the frame work that is composed of three layers that is the input, output and the hidden layer that

is usually layered in between the input and the output layer. The deep neural network [1] is based on the concept of

deep learning that is the sub section of the machine learning and utilizes the facets of the artificial intelligence to

classify and order the data's, the following section provides the deep neural network architectures employed in the

various areas providing a state of art accuracy

2.1. SPEECH RECOGNITION

Deng, et al [2]. The author presents the review of the deep neural networks for the speech recognition by providing

an optimized neural architectures that provides a better optimization and the improved ways to enhance the huge

number of hyper parameters of the deep neural network with the accurate ways of preprocessing the speech for the

deep neural networks and the possible ways of extracting the many number of languages at ease using the deep

neural network than the Gaussian mixture model. So the author highlights the advantages of using the deep learning

in the speech recognition than the GMM Peddinti, et al [3] providing a time delay deep neural network in the

framing of the long temporal contexts utilizing the i-vector based network adaption. The time delay neural networks

developed are viewed as the forerunner of the CNN ensuring a long term temporal dependencies from the short term

features extracted from the speech, allowing a lower layer of the network to get updated with the gradient that are

accumulated over every time step of the input temporal context during the back-propagation and a speedy training

by the effective selection of the sub-sampling indices. Zhang et al [4], the convolutional neural network, which is the

most common type of deep neural network employed in broader range of applications, are more efficient as they

provide with the enhanced way of designing the spectral correlations in the acoustic features for an automated

speech recognition and the diminishing of the spectral variations, for having an computationally effective end to end

ISSN: 2582-2012 (online)

76

Vol.01/ Issue.02 Pages: 73-82

http://irojournals.com/aicn/

DOI: https://doi.org/10.36548/jaicn.2019.2.003

speech recognition the CNN are integrated with the connection temporal classification for sequence labeling. Jaitly et al [5], Palaz, et al [6], the deep belief network for the pre-trained ANN/HMM for a continuous speech recognition in the large vocabulary is proposed in the paper and shows much improvement in the speech recognition than the GMM/HMM Manohar, et al [7] overlapping speech recognition with the acoustic Modelling utilizing the data augmentation approaches integrated with the DNN architectures, front –end speech echoing, beam forming and robust i-vector extraction enables to have a considerable improvement in the speech recognition than the traditional method with an diminished error rate. (CNN+TDNN+LSTM). The table.1 below summarizes the deep neural network architecture in the speech recognition.

Deep Neural Network in Speech Recognition	Literature	Technology used	Features
	Deng, et al [2],	DNN	Extracts many languages at ease than GMM
	Peddinti, et al [3]	TDNN+ I vector based adaption	Provides a high Speed Training
	Zhang et al [4]	CNN+CTC	Computationally effective in end to end speech recognition
	Palaz, et al [6], ]	DBN+trained ANN/HMM	Continuous speech recognition for a large text
	Manohar, et al [7]	CNN+LSTM+TDNN	Improved speech recognition and diminished error rate

## 2.2. IMAGE RECOGNITION

Bischke et al [8], the author provides a multi tasks leaning by introducing the cascaded multi task loss for the preserving of the semantic segmentation of the foot print for the building along with the deep neural networks. This method of image segmentation combining the multitask learning with the deep neural network eluding the necessity for the post –processing ensures an enhanced performance than the conventional methods. Stricker, et al [9], the image recognition using the deep convolutional network ensures a lesser number of pre-processing techniques along with the increased accuracy in the "multi-font printed Amharic character image recognition" Fu et al [10], "the froth

Vol.01/ Issue.02 Pages: 73-82

http://irojournals.com/aicn/

DOI: https://doi.org/10.36548/jaicn.2019.2.003

image recognition" using the three different type of convolutional neural network. Alex net, VGG16 and the ResNet are done in the paper has identified that the Alex net outperforms the previous methods by a significant margin. Sun et al [11], the "DEEPID3" (GoogLeNet +VGG net) is utilized in recognizing the face by reconstructing the two architectures from the "stacked convolution and the inception layers" further the combined face identification and the verification signs are included to the intermediate as well as the final layers of the feature extraction during the process of training. Simonyan et al [12], the very deep convolution network with the depth of 16-19 weight layers are utilized in the system to recognize the images of large scale. Sainath et al [13], the author proposes the method of the "deep convolutional neural network" for the LCVSR tasks. Krizhevsky et a [14] the "Imagenet classification with the deep convolutional neural network" is proposed in the paper to reduce the over fitting in the fully connected layer. He et al [15] the "deep residual learning" to reduce the complexity in the training is utilized for the image recognition with a relative enhancement in the object detection, Imagenet detection and the localization.

## 2.3. NATURAL LANGUAGE PROCESSING

The recent advancements in the neural network along with the deep learning have led to high performance in the various tasks associated with the NLP. The conventional methods of the natural language processing along with the DNN have caused the NLP to have an increased accuracy. Young et al [16], the author has presented a review of the significant past, present and the feature deep learning architectures in the natural language processing and provides a pavement for their evolution. Zeng, et al [17], Hasan et al [18] present the solutions for the clinical natural language processing in the "adverse drug detection from the social media and the medical image caption generation" using the deep learning neural network architectures. The table 2 below provides the over view of the deep neural network in the Image recognition and the natural language processing.

Deep neural network in the Image recognition and the natural language processing	Literature	Technology used	Features
	Bischke et al [8],	DNN +Multi task learning	Eludes post processing.
	Stricker, et al [9],	DCNN	Eludes preprocessing and provides improved accuracy
	Fu et al [10],	Alex net , VGG16,	Improved accuracy

Vol.01/ Issue.02 Pages: 73-82

http://irojournals.com/aicn/

DOI: https://doi.org/10.36548/jaicn.2019.2.003

		ResNet	
	Sun et al [11],	GoogLeNet, VGG net	Reduced error rate
	Simonyan et al [12]	Very deep CNN	Better performance
	Sainath et al [13]	DCNN	Improved performance
	Krizhevsky et a [14]	Image net with DCNN	Reduces over fitting in fully connected layer
	He et al [15]	DNN+DRL	Reduces complexity in training
	Young et al [16], Hasan et al [18]	DNN +NLP	Increased accuracy

## 3. OTHER APPLICATIONS

In relation to the other applications the deep neural networks are employed in the detection of the bot Kudugunta, et al [19] has proposed a bot detection methodology by combining the LSTM and the deep neural network this proposed model of "deep neural network based on the contextual long short-term-memory architecture" ensures the detection of the bot in the social media such as the twitter Facebook etc. Nakazawa et al [20] proposes a "Anomaly detection and the segmentation of the wafer defect patterns utilizing the deep convolutional encoder-decoder neural network architectures in the manufacturing of the semiconductor" is done performing the training with the synthetic wafer maps with the basic patterns to make them capable of identifying the defects patterns from the real wafer maps. To detect the traffic in the web that is based on the number of people visiting the websites and the amount of the information transmitted to and fro the website Kim et al [21] employs the CLSTM and the DNN to have flawless anomaly detection with a heightened accuracy and recall. The deep neural networks finds its necessity in the medical field not only with the drug detection, but also in the identification of the tumors Mohsen et al [22] proposes a deep neural network in combination with the discrete wavelet transform, feature extraction tools and the principal component analysis for the brain tumor detection utilizing the dataset obtained from the MRI and categorizing them

Vol.01/ Issue.02 Pages: 73-82

http://irojournals.com/aicn/

DOI: https://doi.org/10.36548/jaicn.2019.2.003

into four types as the normal, sarcoma, metastatic-bct and glioblastoma. And Lee et al [23] utilizes the deep

Convolutional neural network architecture in the detection and diagnosis of dental caries further Nair et al [24]

utilizes the convolutional neural network in detection of the retinal alignment using the fundus images, providing an

accurate, cost-effective and a comfortable solution in the retinal alignment. The table3 below provides the summary

of the survey done based on the deep neural architectures in other applications.

The survey presented bring out the essential qualities of the deep neural network in the various application that are

associated with the speech recognition, image recognition, natural language processing, and the other application

such as the fraud detection, disease detection etc. Though the DNN has many advantages, they require a high

number of data sets for training along with the huge computational power. so the utilization of the cluster computing

or the cloud computing along with the deep neural network would be ensure the training time and the power

consumption by an considerable amount. The efficiency of the Deep neural networks along with the cluster /cloud

computing is to be continued as the future work.

4. CONCLUSION

The paper presenting the survey of the deep learning neural networks proceeds with the different areas the deep

learning neural network architectures are applied such as the in speech recognition, image recognition, natural

language processing and other applications such as the bot and fraud detection, and disease diagnosis. The review

presented details the deep learning neural network as the accurate way of classifying and predicting utilizing the

sorting and ordering in a feature hierarchy even for the data that are unlabeled and unstructured. The automated way

of feature extraction followed in the deep neural networks make them more compatible in the object detection and

other applications of the computer vision compared to the traditional neural networks. So in future the paper is to

proceed with the utilization of the deep neural network in the computer vision tasks such as the object classification

and later with the analysis of the deep neural network efficiencies when integrated with the cloud computing.

80

Vol.01/ Issue.02 Pages: 73-82

http://irojournals.com/aicn/

DOI: https://doi.org/10.36548/jaicn.2019.2.003

## References

1. Miikkulainen, Risto, Jason Liang, Elliot Meyerson, Aditya Rawal, Daniel Fink, Olivier Francon, Bala Raju et al. "Evolving deep neural networks." In *Artificial Intelligence in the Age of Neural Networks and Brain Computing*, pp. 293-312. Academic Press, 2019.

- 2. Deng, Li, Geoffrey Hinton, and Brian Kingsbury. "New types of deep neural network learning for speech recognition and related applications: An overview." In 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, pp. 8599-8603. IEEE, 2013.
- Peddinti, Vijayaditya, Daniel Povey, and Sanjeev Khudanpur. "A time delay neural network architecture for efficient modeling of long temporal contexts." In Sixteenth Annual Conference of the International Speech Communication Association. 2015.
- 4. Zhang, Ying, Mohammad Pezeshki, Philémon Brakel, Saizheng Zhang, Cesar Laurent Yoshua Bengio, and Aaron Courville. "Towards end-to-end speech recognition with deep convolutional neural networks." *arXiv* preprint arXiv:1701.02720 (2017).
- 5. Jaitly, Navdeep, Patrick Nguyen, Andrew Senior, and Vincent Vanhoucke. "Application of pretrained deep neural networks to large vocabulary speech recognition." In *Thirteenth Annual Conference of the International Speech Communication Association*. 2012.
- 6. Palaz, Dimitri, Mathew Magimai Doss, and Ronan Collobert. "Convolutional neural networks-based continuous speech recognition using raw speech signal." In 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 4295-4299. IEEE, 2015.
- Manohar, Vimal, Szu-Jui Chen, Zhiqi Wang, Yusuke Fujita, Shinji Watanabe, and Sanjeev Khudanpur.
  "Acoustic Modeling for Overlapping Speech Recognition: Jhu Chime-5 Challenge System." In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 6665-6669. IEEE, 2019.
- 8. Bischke, Benjamin, Patrick Helber, Joachim Folz, Damian Borth, and Andreas Dengel. "Multi-task learning for segmentation of building footprints with deep neural networks." In 2019 IEEE International Conference on Image Processing (ICIP), pp. 1480-1484. IEEE, 2019.
- 9. Stricker, Didier. "Multi-font Printed Amharic Character Image Recognition: Deep Learning Techniques." In Advances of Science and Technology: 6th EAI International Conference, ICAST 2018, Bahir Dar, Ethiopia, October 5-7, 2018, Proceedings, vol. 274, p. 322. Springer, 2019.
- 10. Fu, Y., and C. Aldrich. "Flotation froth image recognition with convolutional neural networks." *Minerals Engineering* 132 (2019): 183-190.
- 11. Sun, Yi, Ding Liang, Xiaogang Wang, and Xiaoou Tang. "Deepid3: Face recognition with very deep neural networks." *arXiv preprint arXiv:1502.00873* (2015).

Vol.01/ Issue.02 Pages: 73-82

http://irojournals.com/aicn/

DOI: https://doi.org/10.36548/jaicn.2019.2.003

- 12. Simonyan, Karen, and Andrew Zisserman. "Very deep convolutional networks for large-scale image recognition." *arXiv preprint arXiv:1409.1556* (2014).
- 13. Sainath, Tara N., Abdel-rahman Mohamed, Brian Kingsbury, and Bhuvana Ramabhadran. "Deep convolutional neural networks for LVCSR." In 2013 IEEE international conference on acoustics, speech and signal processing, pp. 8614-8618. IEEE, 2013.
- 14. Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "Imagenet classification with deep convolutional neural networks." In *Advances in neural information processing systems*, pp. 1097-1105. 2012.
- 15. He, Kaiming, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. "Deep residual learning for image recognition." In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp. 770-778. 2016.
- 16. Young, Tom, Devamanyu Hazarika, Soujanya Poria, and Erik Cambria. "Recent trends in deep learning based natural language processing." *ieee Computational intelligenCe magazine* 13, no. 3 (2018): 55-75.
- 17. Zeng, Daojian, Kang Liu, Siwei Lai, Guangyou Zhou, and Jun Zhao. "Relation classification via convolutional deep neural network." (2014).
- 18. Hasan, Sadid A., and Oladimeji Farri. "Clinical Natural Language Processing with Deep Learning." In *Data Science for Healthcare*, pp. 147-171. Springer, Cham, 2019.
- 19. Kudugunta, Sneha, and Emilio Ferrara. "Deep neural networks for bot detection." *Information Sciences* 467 (2018): 312-322.
- 20. Nakazawa, Takeshi, and Deepak V. Kulkarni. "Anomaly Detection and Segmentation for Wafer Defect Patterns Using Deep Convolutional Encoder–Decoder Neural Network Architectures in Semiconductor Manufacturing." *IEEE Transactions on Semiconductor Manufacturing* 32, no. 2 (2019): 250-256.
- 21. Kim, Tae-Young, and Sung-Bae Cho. "Web traffic anomaly detection using C-LSTM neural networks." *Expert Systems with Applications* 106 (2018): 66-76.
- 22. Mohsen, Heba, El-Sayed A. El-Dahshan, El-Sayed M. El-Horbaty, and Abdel-Badeeh M. Salem. "Classification using deep learning neural networks for brain tumors." *Future Computing and Informatics Journal* 3, no. 1 (2018): 68-71.
- 23. Lee, Jae-Hong, Do-Hyung Kim, Seong-Nyum Jeong, and Seong-Ho Choi. "Detection and diagnosis of dental caries using a deep learning-based convolutional neural network algorithm." *Journal of dentistry* 77 (2018): 106-111.
- 24. Nair, Lekha R. "RetoNet: a deep learning architecture for automated retinal ailment detection." *Multimedia Tools and Applications* (2019): 1-10.