RESEARCH ARTICLE OPEN ACCESS

Recent Breakthroughs and Advances in Deep Learning

Anam Bansal
Lovely Professional University
Punjab

ABSTRACT
Deep learning (DL) is an emerging branch of learning algorithms, particularly machine learning which include algorithms driven by artificial neural networks. It includes the neural networks that are rebranded to contain multiple layers. Recent breakthroughs in this domain have renewed interests of researchers as it can model the high-level abstractions and classify the non-linear data, the task which is difficult for shallow neural networks. The parameters are learned from data and prediction is made by deep learning models. Various models of deep learning such as Convolutional Neural Networks, Recurrent Neural Networks etc. have been employed in various applications. This paper presents the recent advancements in the field of deep learning and its applications in various domains like digital image processing, speech recognition, text recognition etc..
Keywords :— Convolutional Neural Network, Recurrent Neural Network, hierarchical feature learning

I. INTRODUCTION

Deep learning is inspired by the structure and functioning of the human brains. Recent popularity of deep learning is due to advancement in techniques, hardware, and data [1]. Earlier the required hardware was not available and computers were quite slow. Today, availability of high GPU computing has supported the growth of the field of deep learning. Similarly, earlier in techniques like backpropagation [2], the methods used for initializing weights was not correct. In deep learning, better weight initialization techniques are provided through supervised learning. Dataset available earlier was not enough and small in size. Today, large data set is available and hence deep learning excelled due to this.

The field of deep neural networks has surpassed various other neural networks because of its unmatched capabilities like scalability, hierarchical feature learning and unbeaten performance in analog domain [1]. In previously used models of machine learning, the performance will at some point get stabled despite feeding them with more data. On contrary, the per- formance of deep learning will keep on increasing as more and more data is fed (Figure 1). Then, deep learning models perform automatic feature extraction from the raw data. They learn hierarchies of feature i.e. learn higher layer features from the composition of lower layer features. Thus, the system can learn the complex functions which map input to output directly from data and need not rely on handcrafted features. In addition, it can give good results in the domains where there are analog inputs(and even output). The inputs can be documents of text data, images of pixel data or files of audio data.

There is a recent move toward deep learning models because the machine learning algorithms rely heavily on the representation of the data [3]. So, there is a lot of computational effort required for preprocessing and transforming data into the particular form for feeding them to machine learning models. The features extracted are fed to the classifier in case of machine learning algorithms. The feature engineering entails huge dependency and poses the weakness of machine learning algorithms. As described in Figure 2, the deep learning models perform feature extraction and classification in one step without requiring particular representation of features for feeding them to classifiers.

II. BACKGROUND AND RELATED WORK

McCulloch and Pitts were the first to develop the neural network model in 1943 [4]. In 1949, Hebb’s rule was proposed by Donald Hebb. Hebb’s rule described the way to update the weights of connections in the neural network. Then, the perceptron was created by Frank Rosenblatt in 1958 [5]. The cells in Visual Cortex were elaborated by Hubel and Wiesel in 1959 [6]. Then in 1975, Paul J. Werbos developed the backpropagation algorithm [7]. The hierarchical artificial neural network named Neocognitron was brought into the picture in 1980 [8]. Finally, in 1990, Convolutional Neural Networks were developed [9].In 2006, the notion of deep learning came into existence as a subfield of machine learning. The concept of deep learning includes hierarchical feature learning [10], also called non-linear learning. Since there are many layers, the output from the lower layer is given as input to the layers above it [11]. Also, the deep learning models can be supervised or unsupervised [12]. Supervised models work with labeled training data but the unsupervised models work with unlabelled training data, forming clusters.

Researchers have proposed several deep learning neural networks for various tasks (Fig- ure 3). Commonly used deep learning architectures are deep belief networks [13], Deep Boltzmann Machines [14], GoogLeNet (Convolutional Neural

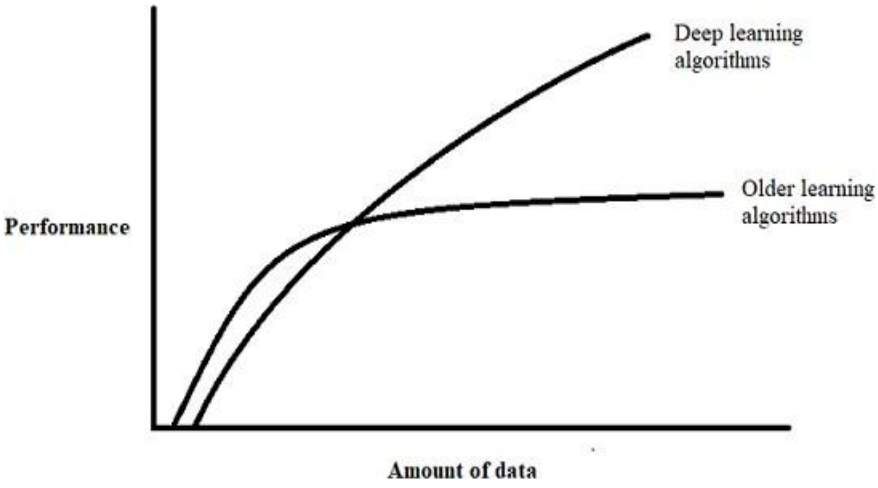

Figure 1: Effect of more data on deep learning algorithms

Network) [15], AlexNet (Convolutional Neural Network) [16], Deep Autoencoder / Stacked Autoencoder [17], LeNet-5 (Convolutional Neural Network) [18], Network In Network (Convolutional Neural Network) [19], Recurrent Neural Networks [20], Long Short-Term Memory Neural networks [21].

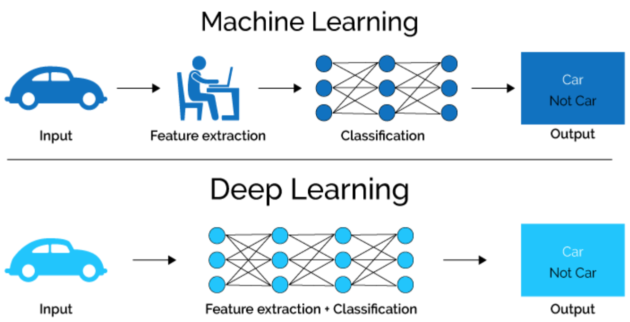

Figure 2: Machine Learning vs Deep Learning

III. APPLICATIONS

Deep learning neural networks have been employed in number of fields (Figure 4). They have widespread applications in Digital Image Processing [22] [23], Speech Recognition [24] [25] [26], Healthcare [27] [28], Natural Language Processing [29] [30], Object Recognition[31] [32], Customer Relationship Management [33], Recommender Systems [34], Biometrics[35] [36] [37] etc. The applications of deep neural networks in various domains is illustrated below:

A. Digital Image Processing

Deep neural networks can process images and predict the class. Handwritten digit recognition through the convolutional neural network is a good example of the deep learning based image recognition [38]. The Convolutional Neural Network was also employed for iris recognition [39]. Facial recognition is another domain of application of DL models [40]. Gender can also be determined through these models [41] [42]. Even, the images of the humans' faces can be processed to determine the age [41] and emotions of the people [43].

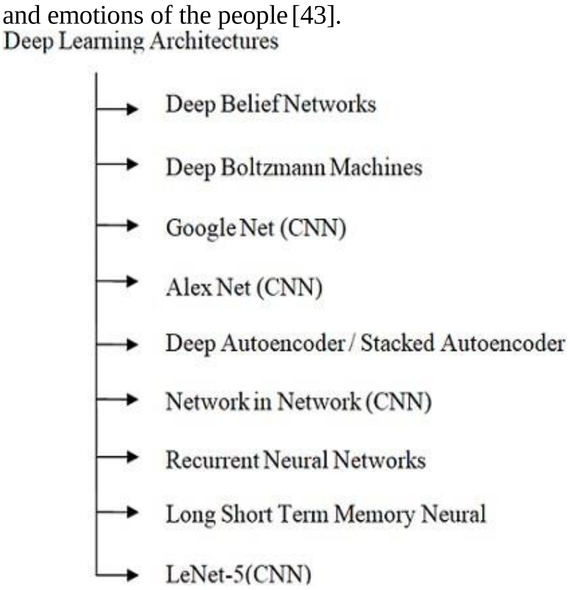

Figure 3: Different architectures of Deep Learning

B. Automatic Speech Recognition

Deep learning models are employed for automatic speech recognition. Deep learning model, Recurrent Neural Network is used for transcribing the speech directly to text without intermediate phonetic representation [44]. These deep models are used for recognizing several different languages such as English, Mandarin Chinese [45] etc. Speech recognition has helped to recognize emotions of a person [46]. This is through the use of deep learning only.

C. Healthcare

Deep learning is applied extensively in healthcare domain. It has been used for disease prediction through sounds[47] orimages[48]. Ophthalmology is an area in which these models are used universally [49]. Deep learning has also been used for tumor detection as the features learned from the medical images are used for prediction by DL algorithms [50]. The severity of Diabetic Retinopathy can be detected by these algorithms [51]. Magnetic Resonance Imaging (MRI) was used to predict the Alzheimer disease through deep neural networks [52]. These models are also employed for Optical Coherence Tomography(OCT) [53]. So, the hospitals are now using deep learning methodologies widely to treat and detect various diseases.

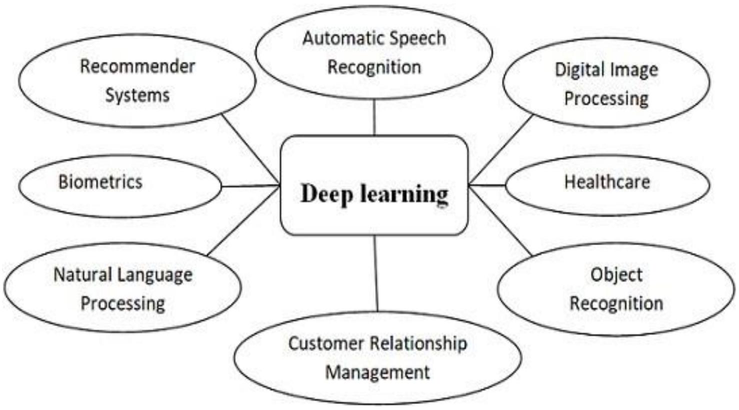

Figure 4: Applications of Deep Learning

D. Natural Language Processing

Deep learning is universally applied for natural language processing. The convolutional neural network with 29 convolutional layers was utilized for text processing [54]. This was the first time, the deep neural network was applied for text processing. Videos can be translated directly to sentences by unifying convolutional neural network and recurrent neural network [55].

E. Object Recognition

Recognising objects through deep neural networks has surpassed the shallow neural net- works. The recurrent Convolutional neural network gave good performance for object recog- nition [56]. Convolutional feature learning is used for object recognition [57]. A novel architecture, RGB-D was proposed for object recognition [32].

F. Customer Relationship Management

Customer relationship management (CRM) is a method to retain the potential customers. The business companies analyze the past records of customers' deals with the company and exploit that to enhance competitiveness at e-commerce [58]. The automatic control of the CRM is achieved through deep reinforcement learning [33]. Recurrent Neural Networks and Reinforcement learning can be combined for managing customer relations with the companies [59].

G. Recommender Systems

The recommender system is used by several big companies to recommend products to the customers based on their preferences. Companies like Amazon analyze the surfing behavior of the customer and advertise the same products to the customers. Collaborative deep learning is used effectively for the recommender systems [60] [61]. Youtube videos can be recommended based on the user's past preferences analyzed based on deep learning [62]. Even the recommendations can be based on small sessions through the Recurrent neural network [63]. Deep learning can be used for recommending an appropriate developer for fixing bugs in the reports [64].

IV. CONCLUSIONS

Deep learning is growing at very fast pace. The versatility of deep learning in various applications demonstrate the wide acceptance and reliability of deep neural networks. These have been used effectively for security systems such as face recognition and other biometric systems. Various deep neural network architectures can be modified internally for various other applications. These networks can be integrated end to end to achieve higher accuracy in various classification systems. In nutshell, though deep neural networks are complex systems but they are very reliable as they combine memory, reasoning, andlearning.

REFERENCES

[1] Goodfellow, I., Bengio, Y., Courville, A. & Bengio, Y. *Deep learning* (MIT press Cambridge, 2016).

[2] Hecht-Nielsen, R. in *Neural networks for perception* 65–93 (Elsevier, 1992).

[3] Mohan, R. & Sree, P. K. An Extensive Survey on Deep Learning Applications (2017).

[4] McCulloch, W. S. & Pitts, W. A logical calculus of the ideas immanent in nervous activity. *The bulletin of mathematical biophysics* **5,** 115–133 (1943).

[5] Schmidhuber, J. Deep learning in neural networks: An overview. *Neural networks* **61,** 85–117 (2015).

[6] Hubel, D. H. & Wiesel, T. N. Receptive fields, binocular interaction and functional architecture in the cat's visual cortex. *The Journal of physiology* **160,** 106–154 (1962).

[7] Werbos, P. J. The roots of backpropagation: from ordered derivatives to neural networks and political forecasting (John Wiley & Sons, 1994).

[8] Fukushima, K. & Miyake, S. Neocognitron: A new algorithm for pattern recognition tolerant of deformations and shifts in position. *Pattern recognition* **15,** 455–469 (1982).

[9] Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks in Advances in neural information processing systems (2012), 1097–1105.

[10] Vargas, R., Mosavi, A. & Ruiz, L. DEEP LEARNING: A REVIEW.

[11] Paul, S., Singh, L., et al. A review on advances in deep learning in Computational Intelligence: Theories, Applications and Future Directions (WCI), 2015 IEEE Workshop on (2015), 1–6.

[12] Lee, H., Grosse, R., Ranganath, R. & Ng, A. Y. Convolutional deep belief networks for scalable unsuper- vised learning of hierarchical representations in Proceedings of the 26th annual international conference on machine learning (2009), 609–616.

[13] Hinton, G. E. Deep belief networks. *Scholarpedia* **4,** 5947 (2009).

[14] Salakhutdinov, R. & Larochelle, H. Efficient learning of deep Boltzmann machines in Proceedings of the thirteenth international conference on artificial intelligence and statistics (2010), 693–700.

[15] Zhong, Z., Jin, L. & Xie, Z. High performance offline handwritten Chinese character recognition using GoogLeNet and directional feature maps in Document Analysis and Recognition (ICDAR), 2015 13th International Conference on (2015), 846–850.

[16] Ballester, P. & de Araújo, R. M. On the Performance of GoogLeNet and AlexNet Applied to Sketches.in AAAI (2016), 1124–1128.

[17] Vincent, P., Larochelle, H., Lajoie, I., Bengio, Y. & Manzagol, P.-A. Stacked denoising autoencoders: Learning useful representations in a deep network with a local denoising criterion. *Journal of Machine Learning Research* **11,** 3371–3408 (2010).

[18] LeCun, Y. *et al.* LeNet-5, convolutional neural networks. *URL: http://yann. lecun. com/exdb/lenet*,20 (2015).

[19] Lin, M., Chen, Q. & Yan, S. Network in network. *arXiv preprint arXiv:1312.4400* (2013).

[20] Medsker, L. & Jain, L. Recurrent neural networks. *Design and Applications* **5** (2001).

[21] Hochreiter, S. & Schmidhuber, J. Long short-term memory. *Neural computation* **9,** 1735–1780 (1997).

[22] Tapia, J. & Aravena, C. in *Deep Learning for Biometrics* 219–239 (Springer, 2017).

[23] Chellappa, R., Castillo, C. D., Patel, V. M., Ranjan, R. & Chen, J.-C. in *Deep Learning in Biometrics* 33–64 (CRC Press, 2018).

[24] Yu, D. & Deng, L. *AUTOMATIC SPEECH RECOGNITION*. (Springer, 2016).

[25] Chan, W., Jaitly, N., Le, Q. & Vinyals, O. *Listen, attend and spell: A neural network for large vocabulary conversational speech recognition* in *Acoustics, Speech and Signal Processing (ICASSP), 2016 IEEE International Conference on* (2016), 4960–4964.

[26] Noda, K., Yamaguchi, Y., Nakadai, K., Okuno, H. G. & Ogata, T. Audio-visual speech recognition using deep learning. *Applied Intelligence* **42,** 722–737 (2015).

[27] Tamilselvan, P. & Wang, P. Failure diagnosis using deep belief learning based health state classification. *Reliability Engineering & System Safety* **115,** 124–135 (2013).

[28] Rav`ı, D. *et al.* Deep learning for health informatics. *IEEE journal of biomedical and health informatics* **21,** 4–21 (2017).

[29] Hirschberg, J. & Manning, C. D. Advances in natural language processing. *Science* **349,** 261–266 (2015).

[30] Sarikaya, R., Hinton, G. E. & Deoras, A. Application of deep belief networks for natural language understanding. *IEEE/ACM Transactions on Audio, Speech, and Language Processing* **22,** 778–784 (2014).

[31] Maturana, D. & Scherer, S. *Voxnet: A 3d convolutional neural network for real-time object recognition* in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on* (2015), 922– 928.

[32] Eitel, A., Springenberg, J. T., Spinello, L., Riedmiller, M. & Burgard, W. *Multimodal deep learn- ing for robust RGB-D object recognition* in *Intelligent Robots and Systems (IROS), 2015 IEEE/RSJ International Conference on* (2015), 681–687.

[33] Tkachenko, Y. Autonomous CRM control via CLV approximation with deep reinforcement learning in discrete and continuous action space. *arXiv preprint arXiv:1504.01840* (2015).

[34] Zuo, Y., Zeng, J., Gong, M. & Jiao, L. Tag-aware recommender systems based on deep neural networks.*Neurocomputing* **204,** 51–60 (2016).

[35] Vatsa, M., Singh, R. & Majumdar, A. *Deep Learning in Biometrics* (CRC Press, 2018).

[36] Bhanu, B. & Kumar, A. *Deep learning for biometrics* (Springer, 2017).

[37] Di, X. & Patel, V. M. in *Deep Learning for Biometrics* 241–256 (Springer, 2017).

[38] Ciresan, D. C., Meier, U., Gambardella, L. M. & Schmidhuber, J. *Convolutional neural network committees for handwritten character classification* in *Document Analysis and Recognition (ICDAR), 2011 International Conference on* (2011), 1135–1139.

[39] Gaxiola, F., Melin, P., Valdez, F. & Castro, J. R. in *Fuzzy Logic Augmentation of Neural and Optimization Algorithms: Theoretical Aspects and Real Applications* 69–80 (Springer, 2018).

[40] Hu, G. *et al. When face recognition meets with deep learning: an evaluation of convolutional neural networks for face recognition* in *Proceedings of the IEEE international conference on computer vision workshops* (2015), 142–150.

[41] Levi, G. & Hassner, T. *Age and gender classification using convolutional neural networks* in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (2015), 34–42.

[42] Zhang, K., Tan, L., Li, Z. & Qiao, Y. *Gender and smile classification using deep convolutional neu- ral networks* in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops* (2016), 34–38.

[43] Fan, Y., Lu, X., Li, D. & Liu, Y. *Video-based emotion recognition using CNN-RNN and C3D hybrid networks* in *Proceedings of the 18th ACM International Conference on Multimodal Interaction* (2016), 445–450.

[44] Graves, A. & Jaitly, N. *Towards end-to-end speech recognition with recurrent neural networks* in *International Conference on Machine Learning* (2014), 1764–1772.

[45] Amodei, D. *et al. Deep speech 2: End-to-end speech recognition in english and mandarin* in *International Conference on Machine Learning* (2016), 173–182.

[46] Han, K., Yu, D. & Tashev, I. *Speech emotion recognition using deep neural network and extreme learn- ing machine* in *Fifteenth Annual Conference of the International Speech Communication Association* (2014).

[47] Low, J. X. & Choo, K. W. Automatic Classification of Periodic Heart Sounds Using Convolutional Neural Network. *World Academy of Science, Engineering and Technology, International Journal of Electrical and Computer Engineering* **5** (2018).

[48] Kermany, D. S. *et al.* Identifying medical diagnoses and treatable diseases by image-based deep learning.*Cell* **172,** 1122–1131 (2018).

[49] Rahimy, E. Deep learning applications in ophthalmology. *Current opinion in ophthalmology* **29,** 254– 260 (2018).

[50] Shi, J. *et al.* Stacked deep polynomial network based representation learning for tumor classification with small ultrasound image dataset. *Neurocomputing* **194,** 87–94 (2016).

[51] Gulshan, V. *et al.* Development and validation of a deep learning algorithm for detection of diabetic retinopathy in retinal fundus photographs. *Jama* **316,** 2402–2410 (2016).

[52] Suk, H.-I., Lee, S.-W., Shen, D., Initiative, A. D. N., *et al.* Hierarchical feature representation and multimodal fusion with deep learning for AD/MCI diagnosis. *NeuroImage* **101,** 569–582 (2014).

[53] Devalla, S. K. *et al.* A Deep Learning Approach to Digitally Stain Optical Coherence Tomography Images of the Optic Nerve Head. *Investigative ophthalmology & visual science* **59,** 63–74 (2018).

[54] Conneau, A., Schwenk, H., Barrault, L. & Lecun, Y. Very deep convolutional networks for natural language processing. *arXiv preprint arXiv:1606.01781* (2016).

[55] Venugopalan, S. *et al.* Translating videos to natural language using deep recurrent neural networks. *arXiv preprint arXiv:1412.4729* (2014).

[56] Liang, M. & Hu, X. *Recurrent convolutional neural network for object recognition* in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (2015), 3367–3375.

[57] Agrawal, P., Girshick, R. & Malik, J. *Analyzing the performance of multilayer neural networks for object recognition* in *European conference on computer vision* (2014), 329–344.

[58] Navimipour, N. J. & Soltani, Z. The impact of cost, technology acceptance and employees' satisfaction on the effectiveness of the electronic customer relationship management systems. *Computers in Human Behavior* **55,** 1052–1066 (2016).

[59] Li, X. *et al.* Recurrent reinforcement learning: a hybrid approach. *arXiv preprint arXiv:1509.03044*