International Journal of Advance Research in Science and Engineering
Volume No. 10, Issue No. 07, July 2021
www.ijarse.com
IJARSE
ISSN 2319 - 8354

A Survey on Deep Learning and Its Application

Anurag Sahani[1], Sapna Singh[2], Ranjeet Singh[3], Sristy Srivastav[4], Neeraj Tiwari[5], Akshay Kumar Maurya[6]

Department of Computer Science Engineering,[12456]

Buddha Institute of Technology, Gida, Gorakhpur, India[12456]

Asst. Professor, Dept. of C.S.E / I.T., Buddha Institute of technology, Gida, Gorakhpur, India[3]

Abstract

As deep learning gradually becomes the leader in this field of machine learning is in a golden age. Deep learning uses multiple layers to represent data abstraction to create computational models. Neural networks and model migration have completely changed our view of information processing; nevertheless, there is a revelation behind this extremely fast field, because it has never been presented in so many ways. These powerful technologies, such as black box machines, have fundamentally inhibited development. In addition, deep learning has repeatedly been regarded as a panacea for all obstacles to machine learning, which is far from the truth. The latest history and latest methods of text grammar, audio and image processing; social network analysis; and natural language processing, and then in-depth analysis of key and innovative developments in deep learning applications. It is also used to analyse deep learning challenges, such as unsupervised learning, black box models, and online learning.

Keywords: *Deep Learning, Machine Learning, Natural Language Processing, Neural Network, Big data*

1. INTRODUCTION

In current years, system getting to know has grown to be increasingly more famous in studies and has been integrated in a massive wide variety of programs, consisting of multimedia idea retrieval, photograph classification, video recommendation, social community analysis, textual content mining, and so forth. Among diverse system-getting to know algorithms, "deep getting to know," additionally referred to as illustration getting to know, is extensively utilized in those programs. The explosive increase and availability of facts and the perfect development in hardware technology have caused the emergence of recent research in disbursed and deep getting to know. [1] Deep getting to know, which has its roots from traditional neural networks, extensively outperforms its predecessors. It makes use of graph technology with modifications among neurons to expand many- layered getting to know models. Many of the trendy deep getting to know strategies were offered and feature established promising consequences of unique types of programs consisting of Natural Language Processing (NLP), visible facts processing, speech and audio processing, and lots of different famous programs. Traditionally, the effectiveness of machine learning algorithms largely depends on the representation quality of the input data. Compared with good data presentation, bad data presentation usually reduces performance. Therefore, the development of functions is an important research field in the machine field. Long-

IJARSE
ISSN 2319 - 8354

term education has focused on creating functions from raw data and has produced a lot of research. In addition, functional development is usually very domain-specific and requires a lot of manpower. For example, in the field of computer vision, they have been proposed and compared, including Histogram Oriented Gradient (HOG), Scale Invariant Transform (SIFT) and Bundle of Words (BOW). Once a new feature is proposed and works well, it will become a trend. Over the years,

similar situations have also a prepared in other fields, including speech recognition and NLP. [2]

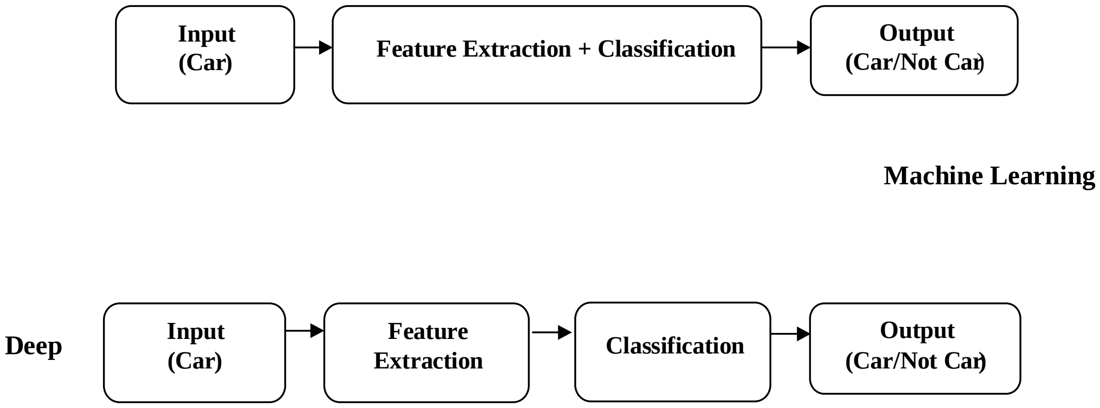

Fig. 1 Difference between machine learning and the deep learning

In contrast, deep learning algorithms perform feature extraction in an automated mode, enabling researchers to extract discriminative features with minimal domain knowledge and manpower. These algorithms include a hierarchical representation architecture in which high-level attributes can be extracted from the later layers of the network, and low-level attributes can be extracted from lower layers. These types of architectures were originally inspired by artificial intelligence (AI), which mimics the processing of key sensory areas in the human brain. Obtain data representing different scenarios. The input is the scene information received by the eyes, and the output is the classification object. This emphasizes the main advantage of deep learning, which is to imitate the work of the human brain. [3]

2. RELATED WORK

Although deep learning is regarded as a huge research field, this article aims to paint a larger picture and share research experience with peers. Although some previous studies only focused on deep learning [4, 5], the novelty of this article is that it focuses on all aspects of deep learning and outlines advanced articles, creative experience, and research and application progress in deep neural networks. The biggest challenge facing deep learning today is the huge data set available for training. As data sets become larger, more diverse, and more complex, deep learning has become an important tool for big data analysis. Our survey highlights the challenges and opportunities in key areas of deep learning that need to be focused on, including concurrency, scalability, performance, and optimization.

DL Networks	Descriptive Key Points	Papers
RvNN	Uses a tree-like structure Preferred for NLP	Goller et al. 1996 , Socher et al. 2011
RNN	Good for sequential information Preferred for NLP & speech processing	Cho et al. 2014 , Li et al. 2015
CNN	Originally for image recognition Extended for NLP, speech processing, and computer vision	LeCunn et al. 1995 , Krizhevsky et al.2012 , Kim2014 , Abdel-Hamid et al. 2014
DBN	Unsupervised learning Directed connections	Hinton 2009, Hinton et al. 2012
DBM	Unsupervised learning, Composite model of RBMs Undirected connections	Salakhutdinov et al. 2009, Salakhutdinov et al. 2012
GAN	Unsupervised learning Game- theoretical framework	Goodfellow et al. 2014, Radford et al. 2015
VAE	Unsupervised learning Probabilistic graphical model	Kingma et al. 2013

Table 1. Summary of the Deep Learning (DL) Networks

The types of deep networks are expressed in different fields, such as RNN for NLP and CNN for image processing. The article also introduces and compares popular deep learning tools such as Cafe, DeepLearning4j, Tensor Flow, Café and Torch, as well as the optimization techniques in each deep learning tool. In addition, various applications of deep learning are discussed to help other researchers expand their understanding of deep learning The rest of the article is structured as follows: Section 2 briefly introduces popular deep learning networks; Section 3 discusses various algorithms, technologies and frameworks of deep learning; since NLP deep learning is not only used in industrial applications, but also in speech and images Identification, Section 4 also introduces many deep learning applications, Section 5 describes future challenges and possible research directions.

2.1 DEEP LEARNING NETWORKS

This section discusses several popular deep learning networks, such as recurrent neural networks (RCNN), RNN, CNN, and deep generative models. However, due to the rapid development of deep learning, many new networks and new architectures will appear every few years. Months. This is beyond the scope of this article. Table 1 summarizes the deep learning networks introduced in this section, their key points, and the most representative articles. [6]

2.1.1 Recursive Neural Network (RCNN)

RCNN can make predictions in a hierarchical structure, and can also use combined vectors to classify the results. The development of RCNN is largely inspired by recursive self-associative memory (RAM), an architecture designed to deal with objects of arbitrary structure, such as trees or graphs. The method is to use a recursive variable-size data structure and create a distributed fixed-width representation

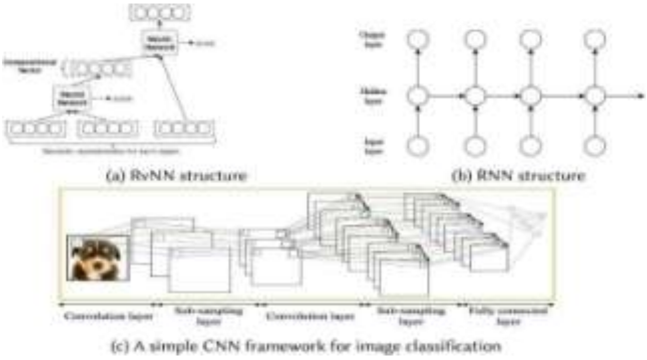

Fig. 1. RCNN, RNN, and CNN architectures.

A structured learning scheme (BTS) is introduced for network training. BTS uses a method similar to the standard back propagation algorithm, but also maintains a tree structure. The network reproduces the pattern from the input layer to the output layer through self-association learning. RNN is particularly successful in NLP. 2011 Soccer et al. proposed an RNN architecture that can handle inputs from various modalities. Shows two examples of using RNN to classify natural images and natural language sentences. When the image is broken down into various parts of interest, the sentence is broken down into words. RNN evaluates possible pairs to combine them and create a syntax tree. For each pair of units, RNN calculates an estimate of the likelihood of merging. The pair with the highest score is then combined into a combined vector. After each merge, RNN creates (1) a larger multi-element region, (2) a combination vector representing the region, and (3) a class name (e.g. If both units are two noun words, the class label of the new area is a noun phrase). The root of the RNN tree structure is the combined vector representation of the entire region. Figure 1 (c) shows an example of the RNN tree. [7]

2.1.2 Recurrent Neural Network (RNN):

In deep learning, especially in NLP and language processing, another widely used and popular algorithm is RNN. Compared with traditional neural networks, RNN uses sequential information in the network. This property is important in many applications, where the structure of the embedded data stream provides useful knowledge. For example, to understand the words in a sentence, you need to know the context in order to treat the RNN as a memory block. They include input layer x, hidden layer (state)s, and output layer y. Figure 1 (b) shows a typical RNN expanded view of the input sequence. In, there are three deep RNN methods, including entry to potential, potential to exit, and potential to potential. Based on these three solutions, a deep RNN is proposed, which not only uses a deeper RNN, but also reduces the complexity of learning in the deep network. RNN is particularly successful in NLP. 2011 Soccer et al. proposed an RNN architecture that can handle inputs from various modalities. shows two examples of using RNN to classify natural images and natural language sentences. When the image is broken down into various parts of interest, the sentence is broken down into words. RNN evaluates possible pairs to combine them and create a syntax tree. For each pair of units, RNN calculates an estimate of the likelihood of merging. The pair with the highest score is then combined into a

International Journal of Advance Research in Science and Engineering
Volume No. 10, Issue No. 07, July 2021
www.ijarse.com
IJARSE
ISSN 2319 - 8354

combined vector. After each merge, RNN creates (1) a larger multi-element region, (2) a combination vector representing the region, and (3) a class name (e.g. If both units are two noun words, the class label of the new area is a noun phrase). The root of the RNN tree structure is the combined vector representation of the entire region. Figure 1 (c) shows an example of the RNN tree.

$$h^k = f\,(W^k * x + b^k).$$

Then, each feature map is down-sampled in the down-sampling layer to reduce the parameters in the network, speed up the learning process, and control over-fitting. For all function mappings, grouping operations (such as average or maximum) are performed in the continuous range p × p (where p is the filter size). Finally, the final layer is usually fully connected, which is known in traditional neural networks. Low-level and intermediate-level functions and the generation of high-level abstractions from the last level (such as Soft max or SVM) can be used to generate classification scores. The probability that each score represents a particular class of a particular instance is. [8,9]

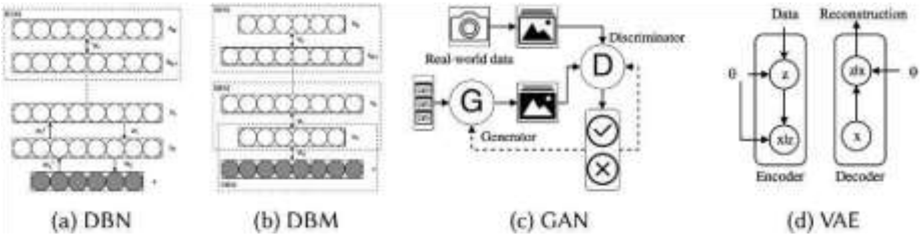

Fig. 2. The structure of generative models.

2.1.3 Convolutional Neural Network (CNN):

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics. The architecture of a ConvNet is analogous to that of the connectivity pattern of Neurons in the Human Brain and was inspired by the organization of the Visual Cortex. Individual neurons respond to stimuli only in a restricted region of the visual field known as the Receptive Field. A collection of such fields overlap to cover the entire visual area. [10]

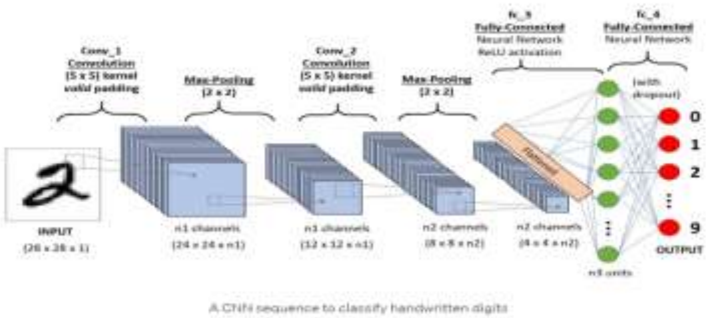

Fig. CNN workflow.

IJARSE
ISSN 2319 - 8354

3. ADVANCED ARCHITECTURES OF DEEP NEURAL NETWORK

Owing to many flexibilities provided by the neural network, deep neural network can be expressed by a diverse set of models. These architectures are called deep models and consist of:

3.1 AlexNet: AlexNet is the first deep architecture which was introduced by one of the pioneers in deep learning – Geoffrey Hinton and his colleagues. It is a simple yet powerful network architecture, which helped pave the way for ground-breaking research in Deep Learning as it is now.

3.2 VGG Net: The VGG Network was introduced by the researchers at Visual Graphics Group at Oxford (hence the name VGG). This network is specially characterized by its pyramidal shape, where the bottom layers which are closer to the image are wide, whereas the top layers are deep.

3.3 GoogleNet: GoogleNet (or Inception Network) is a class of architecture designed by researchers at Google. GoogleNet was the winner of ImageNet 2014, where it proved to be a powerful model. In this architecture, along with going deeper (it contains 22 layers in comparison to VGG which had 19 layers), the researchers also made a novel approach called the Inception module.

3.4 ResNet: ResNet is one of the monster architectures which truly define how deep a deep learning architecture can be. Residual Networks (ResNet in short) consists of multiple subsequent residual modules, which are the basic building block of ResNet architecture.

3.5 ResNeXt: ResNeXt is said to be the current state-of-the-art technique for object recognition. It builds upon the concepts of inception and resnet to bring about a new and improved architecture.

3.6 RCNN (region Based CNN): Region Based CNN architecture is said to be the most influential of all the deep learning architectures that have been applied to object detection problem. To solve detection problem, what RCNN does is to attempt to draw a bounding box over all the objects present in the image, and then recognize what object is in the image. It works as follows:

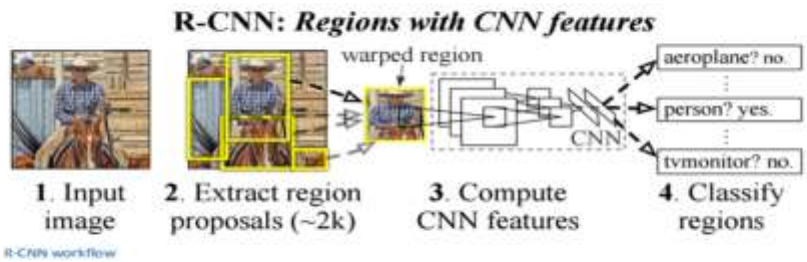

Fig. R-CNN workflow.

3.7 YOLO (You Only Look Once): YOLO is the current state-of-the-art real time system built on deep learning for solving image detection problems. As seen in the below given image, it first divides the image into defined bounding boxes, and then runs a recognition algorithm in parallel for all of these boxes to identify which object class do they belong to. After identifying this classes, it goes on to merging these boxes intelligently to form an optimal bounding box around the objects.

International Journal of Advance Research in Science and Engineering
Volume No. 10, Issue No. 07, July 2021
www.ijarse.com
IJARSE
ISSN 2319 - 8354

All of this is done in parallely, so it can run in real time; processing upto 40 images in a second.

3.8 SqueezeNet: The squeezeNet architecture is one more powerful architecture which is extremely useful in low bandwidth scenarios like mobile platforms. This architecture has occupies only 4.9MB of space, on the other hand, inception occupies ~100MB! This drastic change is brought up by a specialized structure called the fire module

3.9 SegNet: SegNet is a deep learning architecture applied to solve image segmentation problem. It consists of sequence of processing layers (encoders) followed by a corresponding set of decoders for a pixel wise classification. Below image summarizes the working of SegNet. [11].

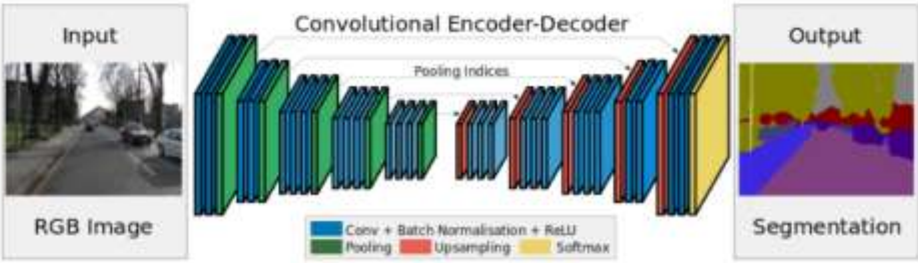

Fig. 2. Working of SegNet.

4. VARIOUS APPLICATIONS OF DEEP LEARNING

The target approach of deep learning is to resolve the sophisticated aspects of the input by using multiple levels of representation. This new approach to machine learning has already been doing wonders in the applications like face, speech, images, handwriting recognition system, natural language processing, medical sciences, and many more. Its latest researches involve revealing the optimization and fine tuning of the model by using gradient descent and evolutionary algorithms. Some major challenges that the deep learning technology is facing undoubtedly are the scaling of computations, optimization of the parameters of deep neural network, designing and learning approaches. A detailed investigation in various complex deep neural network models is also a big challenge to this potential research area. The combination of fuzzy logic with deep neural network is another provoking and demanding area which needs to be explored. [1]

A few years ago, we would've never imagined deep learning applications to bring us self-driving cars and virtual assistants like Alexa, Siri and Google Assistant. But today, these creations are part of our everyday life. Deep Learning continues to fascinate us with its endless possibilities such as fraud detection and pixel restoration. [12]

4.1 Application of Deep Learning Across Industries:

- Self-Driving Car
- News Aggregation and Fraud News Detection
- Natural language Processing
- Virtual Assistants
- Entertainment
- Visual Recognition
- Fraud Detection
- Healthcare
- Personalisation
- Detecting Developmental Delay in Children
- Colourisation of Black and White images
- Adding sounds to silent movies
- Automatic Machine Translation
- Automatic Handwriting Generation
- Automatic Game Playing
- Language Translations
- Pixel Restoration
- Photo Descriptions
- Demographic and Election Predictions
- Deep Dreaming

4.1.1 Self-Driving Car: Deep Learning is the force that is bringing autonomous driving to life. A million sets of data are fed to a system to build a model, to train the machines to learn, and then test the results in a safe environment. The Uber Artificial Intelligence Labs at Pittsburg is not only working on making driverless cars humdrum but also integrating several smart features such as food delivery options with the use of driverless cars. The major concern for autonomous car developers is handling unprecedented scenarios. A regular cycle of testing and implementation typical to deep learning algorithms is ensuring safe driving with more and more exposure to millions of scenarios. Data from cameras, sensors, geo-mapping is helping create succinct and sophisticated models to navigate through traffic, identify paths, signage, pedestrian-only routes, and real-time elements like traffic volume and road blockages.

4.1.2 News Aggregation and Fraud News Detection: There is now a way to filter out all the bad and ugly news from your news feed. Extensive use of deep learning in news aggregation is bolstering efforts to customize news as per readers. While this may not seem new, newer levels of sophistication to define reader personas are being met to filter out news as per geographical, social, economic parameters along with the individual preferences of a reader. Fraud news detection, on the other hand, is an important asset

International Journal of Advance Research in Science and Engineering
Volume No. 10, Issue No. 07, July 2021
www.ijarse.com
IJARSE
ISSN 2319 - 8354

in today's world where the internet has become the primary source of all genuine and fake information. It becomes extremely hard to distinguish fake news as bots replicate it across channels automatically.

4.1.3 Virtual Assistants: The most popular application of deep learning is virtual assistants ranging from Alexa to Siri to Google Assistant. Each interaction with these assistants provides them with an opportunity to learn more about your voice and accent, thereby providing you a secondary human interaction experience. Virtual assistants use deep learning to know more about their subjects ranging from yours dine-out preferences to your most visited spots or your favourite songs. They *learn* to understand your commands by evaluating natural human language to execute them.

4.1.4 Entertainment (Netflix, Film Making etc.): Wimbledon 2018 used IBM Watson to analyse player emotions and expressions through hundreds of hours of footage to auto-generate highlights for telecast. This saved them a ton of effort and cost. Thanks to Deep Learning, they were able to factor in audience response and match or player popularity to come up with a more accurate model (otherwise it would just have highlights of the most expressive or aggressive players). Netflix and Amazon are enhancing their deep learning capabilities to provide a personalized experience to its viewers by creating their personas factoring in show preferences, time of access, history, etc. to recommend shows that are of liking to a particular viewer.

4.1.5 Visual Recognition: Imagine yourself going through a plethora of old images taking you down the nostalgia lane. You decide to get a few of them framed but first, you would like to sort them out. Putting in manual effort was the only way to accomplish this in the absence of metadata. The maximum you could do was sort them out based on dates but downloaded images lack that metadata sometimes. In comes, Deep Learning and now images can be sorted based on locations detected in photographs, faces, a combination of people, or according to events, dates, etc. Searching for a particular photo from a library (let's say a dataset as large as Google's picture library) requires state-of-the-art visual recognition systems consisting of several layers from basic to advance to recognize elements. Large-scale image Visual recognition through deep neural networks is boosting growth in this segment of digital media management by using convolutional neural networks, Tensor flow, and Python extensively.

4.1.6 Deep Dreaming: In 2015, Google researchers found a method that used Deep Learning Networks to enhance features in images on computers. While this technique is used in different ways today, one of the Deep Learning applications essentially involves the concept of Deep Dreaming. This technique, as the name suggests, allows the computer to hallucinate on top of an existing photo – thereby generating a reassembled dream. The hallucination tends to vary depending upon the type of neural network and what it was exposed to. This deep dreaming technique has been used by a group of researchers from the university of Sussex, to create a hallucination Machine which allows users to experience psycho-pathological conditions or psychoactive substances through a virtual reality. This successful experiment further opens up possibilities of using deep neural network algorithms for more induced dreaming experiences. [11,12]

5. CHALLENGES OF DEEP LEARNING

Although deep learning techniques are proving its best and has been solving various complicated applications with multiple layers and high level of abstraction. It is surely accepted that the accuracy, acuteness, receptiveness and precision of deep learning systems are almost equal or may sometimes surpass human experts. To feel the exhilaration of victory, in today's scenario, the technology has to accept many challenges. So, here is the list of challenges which deep learning has to overcome is:

- Deep learning algorithms have to continuously manage the input data. [1]
- Algorithms need to ensure the transparency of the conclusion.
- Resource is demanding technology like high performance GPUs, storage requirements.
- Improved methods for big data analytics. Deep networks are called black box networks.
- Presence of hyper parameters and complex design.
- Need very high computation power.
- Suffer from local minima.
- Computationally intractable.
- Need a large amount of data.
- Expensive for the complex problems and computations.
- No strong theoretical foundation.
- Difficult to find the topology, training parameters for the deep learning.
- Deep learning provides new tools and infrastructures for the computation of the data and enables computers to learn objects and representations.

6. CONCLUSION AND FUTURE ASPECTS

Deep learning is indeed a fast growing application of machine learning. The rapid use of the algorithms of deep learning in different fields really shows its success and versatility. Achievements and improved accuracy rates with the deep learning clearly exhibits the relevance of this technology, clearly emphasize the growth of deep learning and the tendency for the future advancement and research. Additionally, it is very important to highlight that hierarchy of layers and the supervision in learning are the major key factors to develop a successful application with deep learning. The reason behind is that the hierarchy is essential for appropriate data classification and the supervision believes the importance of maintaining database. Deep learning relies on the optimization of existing applications in machine learning and its innovativeness on hierarchical layer processing. Deep learning can deliver effective results for the various applications such as digital image processing and speech recognition. During the current era and in coming future, deep learning can be executed as a useful security tool due to the facial recognition and speech recognition combined. Besides this, digital image processing is a kind of research field that can be applied in multiple areas. For proving it to be a true optimization, deep learning is a contemporary and exciting subject in artificial intelligence. At last, we conclude here, that if we follow the wave of success, we will find that with the increased availability of data and computational resources, the use of deep learning in many applications is genuinely taking off towards the acceptance. The technology is really ephebic, young and specific and in the next few years, it is expected that the rapid advancement of deep learning in more and more applications with a great boom and prosperity e.g. natural language processing, remote sensing and healthcare will certainly achieve targets and height of triumph and satisfaction. Future Aspects of deep learning includes: [1,5]

- Working of deep networks with the sophisticated and non-static noisy scenario and with multiple noise types?
- Raising the performance of the deep networks by improving the diversity between the features?
- Compatibility of deep neural networks in the unsupervised learning online environment.
- Deep reinforcement kind of learning will be the future direction.
- With deep networks, in coming future the factors like inferences, efficiency and the accuracies will be desirable.
- Maintenance of wide repository of data.
- To develop deep generative models with superior and advanced temporal modelling abilities for the parametric speech recognition system.
- Automatically assess ECG signal through deep learning.
- Use of deep neural network in object detection and tracking in videos.
- Deep neural network with the fully autonomous driving. It is submitted that the deep learning methodologies have got great interest in each and every area where conventional machine learning techniques were applicable. It is finally to say that deep learning is the most effective, supervised and stimulating machine learning approach. It can provide researchers a quick evaluation of the hidden and incredible issues associated with the application for producing better and accurate results. [1,6]

7. REFERENCES

[1] A Survey of Deep Learning and Its Applications: A New Paradigm to Machine Learning Shaveta Dargan, Manish 1071-1092(2020) Published: 01 June 2019.

[2] Rami Al-Rfou, Guillaume Alain, Amjad Almahairi, Christof Angermueller, Dzmitry Bahdanau, Nicolas Ballas, Frédéric Bastien, Justin Bayer, Anatoly Belikov, Alexander Belopolsky, Yoshua Bengio, Arnaud Bergeron, James Bergstra.

[3] Johannes Abel and Tim Fingscheidt. 2017. A DNN regression approach to speech enhancement by artificial bandwidth extension. In *IEEE Workshop on Applications of Signal Processing to Audio and Acoustics*. IEEE, 219– 223.

[4] Sami Abu-El-Haija, Nisarg Kothari, Joonseok Lee, Paul Natsev, George Toderici, Balakrishnan Varadarajan, and Sudheendra Vijayanarasimhan. 2016. YouTube-8M: A large-scale video classification benchmark. *CoRR* abs/1609.08675 (2016). Retrieved from http://arxiv.org/abs/1609.08675.

[5] Ting Chen and Christophe Chefd'hotel. 2014. Deep learning based automatic immune cell detection for immunohistochemistry images. In *International Workshop on Machine Learning in Medical Imaging*. Springer, 17– 24.

[6] Nicolas Ballas, Li Yao, Chris Pal, and Aaron C. Courville. 2015. Delving deeper into convolutional networks for learning video representations. *CoRR* abs/1511.06432 (2015). Retrieved from http://arxiv.org/abs/1511.06432

[7] Min Chee Choy, Dipti Srinivasan, and Ruey Long Cheu. 2006. Neural networks for continuous online learning and control. *IEEE Transactions on Neural Networks* 17, 6 (2006), 1511–1531.

[8] Christof Angermueller, Tanel Pärnamaa, Leopold Parts, and Oliver Stegle. 2016. Deep learning for computational biology. *Molecular Systems Biology* 12, 7 (2016), 878.

[9] Jonathan Berant, Andrew Chou, Roy Frostig, and Percy Liang. 2013. Semantic parsing on freebase from questionanswer pairs. In *Empirical Methods in Natural Language Processing*, Vol. 2. Association for Computational Linguistics.

[10] https://towardsdatascience.com/a-comprehensive-guide-to-convolutional-neural-networks-the-eli5-way-3bd2b1164a53

[11] https://www.analyticsvidhya.com/blog/2017/08/10-advanced-deep-learning-architectures-data-scientists/

[12] https://www.mygreatlearning.com/blog/deep-learning-applications/