A ADVANCED JOURNAL IN ARTIFICIAL INTELLIGENCE AND MACHINE LEARNING TRENDS IN BIOLOGICAL SCIENCES
Open Journal

Review

Deep Learning: An Overview and Practical Guide to the Emerging Paradigm in Machine Learning

Shaveta Dargan*; Munish Kumar; Maruthi Rohit Ayyagari; Gulshan Kumar

Kyushu Institute of Technology, Japan

***Corresponding author**
Shaveta Dargan
Kyushu Institute of Technology, Japan

Article information
Received: March 31st, 2023; **Revised:** May 30th, 2023; **Accepted:** July 29th, 2023; **Published:** September 24th, 2023

Cite this article
Dargan S, Kumar M, Ayyagari MR, Kumar G. Deep learning: An overview and practical guide to the emerging paradigm in machine learning. 2023; 2(2).
doi: https://doi.org/10.70705/ppp.bioai.2023.v02.i02.pp65-73

ABSTRACT

Everyday social and commercial activities are aided by artificial intelligence (AI), a crucial piece of technology. It helps alleviate a lot of societal issues while also making a significant contribution to Japan's economy's long-term prosperity. Artificial intelligence (AI) has been a hot topic in the last few years, with both developed and emerging nations seeing it as a potential economic driver. New forms of artificial intelligence (AI) in ICT and RT have received the bulk of the research and development efforts. New artificial intelligence systems do a great job at identifying patterns, but they still have a long way to go. In addition to being extremely complex and reliant on massive data, the majority of ICT models do not have a self-idea function. This paper's focus is on creating a new notion of general-purpose intelligence cognition technology, dubbed "Beyond AI," as opposed to only creating AI for the future generation. Our focus is on creating a "Brain Intelligence (BI)" model of intelligent learning that can use an imagine function to conjure up novel interpretations of occurrences even when its developers lack first-hand knowledge with them. Additionally, we will showcase the created BI intelligence learning model using autonomous vehicles, precise healthcare, and industrial robots.

Keywords
Deep learning; Artificial intelligence; Machine learning.

INTRODUCTION

A subfield of AI, machine learning enables systems to gain information and understanding automatically, without human intervention or programming. Preparing for data characteristics and patterns, leading to improved future findings and decisions, begins with observations, such as first-hand experiences. Machine learning techniques that model data at a high level using several nonlinear regressions are the backbone of deep learning. transformations. The ANN system is the foundation of deep learning technology. The effectiveness of training processes may be enhanced by continually expanding the quantities of data, and these ANNs constantly adopt learning methods. Greater data quantities are necessary for efficiency. As the neural network's level grows during training, the process is referred to as deep learning. There are two main parts of deep learning, and they are the training and inferring phases, without which the process would not function. Both the training and inferring phases involve labeling and determining the matching properties of massive volumes of data.issues with drawing judgments and classifying novel, undiscovered information based on what they already know. One method that allows systems to comprehend complicated perceptual tasks with utmost precision is deep-learning. Hierarchical learning, deep structured learning, and multi-layer deep learning all use nonlinear processing units to convert data and extract features.

layer receives as input the output of the preceding layer. Using different levels of representation and diverse stages of abstraction, learning may occur in either a supervised or unsupervised manner. An essential computational unit in deep learning or a deep neural network is the neuron, which receives input from a variety of signals. It transfers the combined signals across the nonlinear tasks to generate outputs, after linearly integrating them with the weight.
The word "deep" refers to the many levels of processing that are used in the "deep learning" approach to data transformation. A key component of these systems is the depth of the credit assignment path (CAP), which denotes the number of steps involved in converting inputs to outputs and stands for the impulsive link between the input and output layers. Keep in mind that deep learning and representational learning are not the same thing. A collection of techniques that aid machines in taking raw data as input and determining

representations for detection and classification purposes is known as representational learning. Purely speaking, deep learning approaches are a form of learning method that operates at a higher abstract level and employs several layers of representation. The distinctions between deep learning and machine learning are shown in Figure 1.

In order to make advantage of massive datasets, deep learning methods use nonlinear transformations and abstract models at a high level. Additionally, it explains how a machine takes in the representations and abstractions from one layer and uses them to change its internal attributes, which are needed to list the descriptions in each tier. Adaptive testing, big data, cancer detection, data flow, healthcare, object detection, object classification, speech recognition, image classification, natural language processing, and voice activity detection are some of the many areas that make extensive use of this innovative learning approach.

In order to discover a wide range of applications, the deep learning paradigm first employs large amounts of ground truth-designed data to identify unique features and combinations of characteristics. It next builds an integrated model for feature extraction and classification. One important feature of deep learning is that it may information that does not need the involvement of human engineers and makes use of general purpose procedures with a variety of substantial characteristics. For the purpose of cleaning out spam communications and classifying massive amounts of data, Facebook has also developed Deep Text.

The fundamental principles of deep learning technique are:
Nonlinear processing in multiple layers or stages.
- Supervised or unsupervised learning.

Nonlinear processing in multiple layers refers to a hierar- chical method in which the present layer accepts the results from the previous layer and passes its output as input to the next layer. Hierarchy is established among layers so as to organize the importance of the data. Here supervised and unsupervised learning are linked to the class target label. Its availability means a supervised system and absence indi- cates an unsupervised system. Soniya et al. [56] presented current trends, models, architecture and the limitations of deep learning. They explored some of the characteristics like learning techniques, optimization methods and tuning of these models. They also focused on the use of large datasets for the deep learning. They also discussed the challenges for the deep learning.

2 Basic Architectures of Deep Neural Network (DNN)

Different names for deep learning architectures embrace deep belief networks, recurrent neural networks and deep neural networks. DNN can be constructed by adding mul- tiple layers which are hidden layers in between the input layers and the output layers of Artificial Neural Network with various topologies. The deep neural network can model convoluted and non-linear relationships and generates mod- els in which the object is treated as a layered configuration of primitives. These are such feed forward networks which have no looping and the flow of data is from the input layer to the output layer. There are wide varieties of architectures and algorithms that are helpful in implementing the concept

Table 1 Years with the usage of architectures of deep learning

Year	Architecture of deep learning
1990–1995	Recurrent neural network
1995–2000	Long short term memory, convolutional neural network
2000–2005	Long short term memory, convolutional neural network
2005–2010	Deep belief network
2010–2017	Deep stacked network, gated recurrent unit of deep learning.

Table 1 depicts the year wise distribution in the architecture of deep learning.

Here, we will discuss six basic types of the deep learning architectures and these are:-

- Auto-Encoder (AE)
- Convolutional Neural Network (CNN)
- Restricted Boltzmann Machine (RBM)
- Deep Stacking Network (DSN)
- Long Short Term Memory (LSTM)/Gated Recurrent Unit (GRU) Network
- Recurrent Neural Network (RNN)

Out of these, LSTM and CNN are two of the fundamental and the most commonly used approaches.

2.1 Auto-Encoder (AE)

An Auto-encoder (AE) is a type of neural network which is based on unsupervised learning technique and uses the back propagation algorithm. The network first sets the target result values to be equal to the input values. The network tries to understand an approximation which is equivalent to the identity function. Its architecture consists of three lay- ers which are an input, a hidden called encoding layer, and a decoding layer. The network tries to reconstruct its input, which forces the hidden layer to learn the best representa- tions of the input. The hidden layer is used to describe a code which helps to represent the input. Auto-encoders are neural networks, but they are also closely related to PCA (Principal Component Analysis).

Some Key Facts about the Auto-encoder are:-

- Auto-encoders are neural network.
- Auto-encoders are based on the unsupervised machine learning algorithm.
- These are closely resembled with the Principal Component Analysis (PCA).
- It is more flexible than the PCA.
- It minimizes the same objective function as PCA
- The neural network's target output is its input

Although Auto-encoders are same as PCA, but the flex- ibility of auto-encoder is quite high. Auto-encoders allow the representation in both linear and non-linear way in the encoding whereas linear transformation is possible in PCA. Due to the network representation, Auto-encoders can be stacked and layered to produce a deep learning network.

Following are the types of Auto-encoders:

1. De-noising Auto-encoder: It is an advanced version of basic auto-encoders. To addresses the identity functions, these encoders corrupt the input and afterwards, recon- struct them. It is also called the stochastic version of the auto-encoders.
2. Sparse Auto-encoder: These auto-encoders have the learning methods that automatically extract the features from the unlabeled data. Here the word sparse indicates that hidden units are allowed to fire only for the certain type of inputs and not too frequently.
3. Variational Auto-Encoder (VAE): The variational auto- encoder consists of an encoder, decoder and a loss func- tion. They are used for the designing of the complex models of the data that too with large datasets. It is also known as high resolution network.
4. Contractive Auto-encoder (CAE): These are robust networks as de-noising auto-encoders but the difference is that the contractive auto-encoders generate robust- ness in the networks through encoder function whereas de-noising auto-encoders work with the reconstruction process.

Auto-encoders are used to operate with high dimensional data and explains the representation of a set of data via dimensionality reduction. Auto-encoder (AE) uses mainly two structures, called, De-noising Auto-encoder and Sparse Auto-encoder. For De-noising Auto-encoder, it uses data from noise to experience the network weight and for Sparse Auto-encoder, they bound the activation state of hidden units. Working of an Auto-encoder considers the input and afterwards maps it to an inherent transformation with the help of nonlinear mapping.

2.2 Convolutional Neural Network (CNN)

CNN is a neural network with multiple layers and is based on the animal visual cortex. The first CNN was developed by LeCun et al. [27]. Application areas of CNN include mainly image-processing and handwritten character recognition e.g. postal code interpretation. Considering the architecture, ear- lier layers are used for identifying the features such as edges and the later layers are used for the recombination of features to form high level attributes of the input followed by the classification. Then pooling will be done, which mitigates the dimensionality of the extracted features.

The next step is to perform convolution and then again pooling, that is fed into a perfectly linked multilayer percep- tron. Responsibility of the concluding layer called an output layer is to recognize the features of the image by using back- propagation algorithms. In CNN, the advantage of deep lay- ers of processing, convolutional layer, pooling, and a fully connected classification layer reveals various applications such as speech recognition, medical applications, video rec- ognition and various tasks of natural language processing. CNN produces better accuracy and improves the perfor- mance of the system due to its exclusive features such as local connectivity and shared weights. It works much better than any other deep learning methods. It is the most com- monly used architecture as compared to others. Figure 2 depicts the working of CNN with the flow of data from the inputs, convolutional, pooling layers, hidden layers and the outputs.

2.3 Restricted Boltzmann Machines and Deep Belief Network

Restricted Boltzmann Machine (RBM) is such an undirected graphical and modeled representation of the hidden layer, a visible layer and the symmetric connection between the layers. In RBM, there is no connection in between an input and the hidden layer. The deep belief network represents multilayer network architecture that incorporates a novel training method with many hidden layers. Here every pair of connected layers is a RBM and is also known as a stack of restricted Boltzmann machines. The input layer constitutes the basic sensory input, and the hidden layer characterizing the abstract description of this input. The job of the output layer is to only perform the network classification.

The training part is done in two stages: Unsupervised pre training and supervised fine-tuning. In unsupervised pre training, from the first hidden layer, RBM is skilled to recon- struct its input. The next RBM is qualified similar to the first one, and the first hidden layer is taken as the input and visible layer, and the RBM is worked by taking the outputs from the first hidden layer. Hence, every layer is pre skilled or pre-trained. Now when the pre training is completed, steps of supervised fine-tuning start. In this step, the nodes representing the output are marked with the values or labels so that they can help in the learning process and later on full network training is done with the gradient descent learning or back-propagation algorithm.

2.4 Deep Stacking Networks

Deep Stacking Networks (DSN) is also acknowledged as deep convex networks. DSN is different from other tradi- tional deep learning structures. It is called deep because of the fact that it contains a large number of deep individual networks where each network has its own hidden layers.

The DSN believes that training is not a particular and isolated problem, but it holds the combination of individual training problems. The DSN is made up of a combination of modules which are part of the network and present in the architecture. There are three modules that work for the DSN. Here every module in the model is having an input zone, a single hidden zone and an output zone. Subroutines are placed one over the top of another with the input to the The second gate called forgets port controls and is used when an existing piece of information is forgotten and helps the cell to memorize the new data.

The job of the output gate is again to control the informa- tion that is present in the cell and is used as the output of the cell.

The weight of the cell can be used for the controlling purpose. There is a need for the training method which is commonly called as Back propagation through time (BPTT) that enhances the weight. The method requires network out- put error for the optimization.

The Gated Recurrent Unit (GRU) includes two gates called as an update gate and a reset gate. Responsibility of an update gate is to tell the requirement of the contents of the previous cell for the maintenance. The reset gate describes the carrying process of previous cell contents with the new input. The GRU represents a standard RNN by initializing the reset gate to 1 and update gate to 0. Working capability of GRU model is simple as compared to the LSTM. It can be skilled in a short time and it is considered to be more efficient in terms of its execution.

2.6 Recurrent Neural Network

every module is taken as the outputs of the preceding layer and the authentic input vector. Figure 3 depicts the process of working of the layers that helps to resolve the complex classifications. In DSN, every module is trained in isolation so as to make it productive and competent with the ability to work in coordination. The process of supervised method of training is practiced as the back-propagation for each mod- ule and not for the entire network. DSNs works superior than typical DBNs making it suitable and accepted network architecture.

2.5 LSTM/GRU Network

The Long Short Term Memory (LSTM) was designed with the efforts of Hochreiter and Schimdhuber, and used for many applications. IBM selected LSTMs mainly for speech recognition. The LSTM uses a memory unit called a cell which can hold its value for a sufficient time and treats it as a function of its input. This helps the unit to memorize the last calculated value.

The memory unit or a cell is made up of three ports called gates, which control the movement of information in the unit, i.e. into the cell and out of the cell.

• The input port or the gate manages flow of new information into the memory.

RNN consists of a rich set of architecture and is the basic network architecture. The important characteristic of a recurrent network is that the recurrent network has a connec- tion that can be given as feedback into prior layers as com- pared to the complete feed-forward connections. It takes the previous memory of input and models the problems within time. These networks can be upgraded, skilled and expanded with standard back-propagation called as back-propagation through time (BPTT). Table 2, describes the various application areas of different architecture of deep neural networks.

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:

• AlexNet The net is named for the researchers. It was the earliest deep learning architecture and was developed by Alex Krizhevsky, Geoffrey Hinton and his colleagues, who gave ground breaking research in deep learning. The architecture consists of the convolutional layers and the pooling layers which are stacked on one another and then followed by completely interlinked layers on the top. The benefits and superiority lie in the fact that the scalability

Table 2 Architectures of deep neural network and their major application areas

Architecture	Major application areas
Auto-encoder	Natural language processing, understanding compact representation of data
Convolutional neural networks	Document analysis, face recognition, image recognition, natural language processing, video analysis
Deep belief networks	Failure prediction, information retrieval, image recognition, natural language understanding
Deep stacking networks	Continuous speech recognition, Information retrieval
LSTM/GRU networks	Gesture recognition, handwriting recognition, image captioning, natural language text compression, speech recognition
Recurrent neural networks	Handwriting and speech recognition
Restricted Boltzmann machine	Collaborative filtering, classification, dimensionality reduction, feature learning, regression, and topic modeling

and the use of GPU are incomparable. AlexNet has high speed of processing and training because of the use of GPU.

• Visual Graphic Group Net This net was developed by the technicians at the Visual Graphics Group from the Oxford and is in pyramid shape. The model consists of the bottom layers which are wide and the top layers are deep. VGG accommodates successive convolutional lay- ers and then the pooling layers to make the layers narrow.

• GoogleNet The architecture was introduced by the researchers at Google and hence the name of the Net. It involves 22 layers whereas VGG had 19 layers. Google Net is based on the novel technique which is known as the inception module. Here single layer carries multiple kinds of the feature extractors that help the network to perform better. When multiple of these inception mod- ules are stacked one over the other, it becomes the final one. The model converges faster because of the joint and the parallel training. Training of GoogleNet is faster than VGG with small size of the pre-trained GoogleNet.

• ResNet Residual Network incorporates numerous succes- sive residual modules also called as the basic building block of ResNet. The residual modules are placed on one over the other and form a successful and complete node to node network. The main benefit of ResNet is that many residual layers are capable of forming a trained network.

• ResNeXt It is constructed based on the concepts of ResNet with novel and enhanced architecture with improved performance.

• RCNN (Regions with Convolutional Neural Network) It depends upon designing a bounding box over the objects in the image and identifies the object given in the image.

• YoLo (You only look once) This architecture solves image detection problems. To identify the class of the object, the image is divided into bounding box and then a rec- ognition algorithm is executed which is common for all the boxes. After identification of the classes, the boxes are merged very carefully to make a best bounding box around the objects. It is used in real time for handling day-to-day problems.

• SqueezeNet The SqueezeNet architecture is the most powerful architecture to select with the low bandwidth. This network architecture takes space of 4.9 MB and the inception process will take 100 MB. A fire module is used for handling the drastic change.

• SegNet SegNet is considered as the best model for the image segmentation problems. SegNet is a deep neural network, which is used to solve the image segmentation complexities. It is made up of an arrangement of pro- cessing layers which are called encoders, and interre- lated set of decoders for pixel wise classification. The important feature of SegNet is the ability to retain very high frequency details in the segmented image. Herein the encoder network and the decoder network, pooling indices are connected. The flow of information is also straight.

- GAN Generative Adversarial Networks is a unique network architecture, which creates an entirely novel and different images, which are not already present in the available training dataset.

4 Characteristics of Deep Learning

Deep learning is a broad term used for the machine learning and for the artificial intelligence. Because of the following mentioned characteristics, deep learning techniques have achieved the heights of success in the variety of applica- tion areas. For example, new areas such as decision fusion, on-board mobile devices, transfer learning, class imbal- ance problems and human activity recognition have gained improvement in the performance and the accuracy.

So, here are the following characteristics of deep learning:

- Extensively powerful tool in many fields.
- It is purely based on neural networks with the addition of more than two layers and so called deep.
- Have strong learning ability.
- Can make use of datasets more effectively.
- Learn feature extraction methods from the data.
- Surpass human ability to solve highly computational tasks.
- Very little engineering by hand is required in deep learning.
- Optimized results.
- Deep learning networks depend upon the nature of the network structure, activation function and data represen- tation.
- Describe highly variant features in a few parameters.
- Prediction performance can be greatly improved.
- Solve highly computational tasks.
- Capability to extract features from high dimensional sensory inputs.
- Secure and robust generalization capability and with less requirement of training data.
- Fuse the benefits of multiple features for voice activity detection.
- Stronger than machine learning model in feature representation.
- Covariance estimation can be improved for the prediction applications.
- Deep learning networks do not rely on prior data and knowledge.
- DNN has a unique representation and having innovative methods to understand the representations even with large-scale and unlabeled data.
- With high-level abstraction, these networks can extract complicated features.
- Good recognition ability approaches in the big data era.

5 Motivation to Use Deep Learning

Deep learning technology has a conception that there is nothing inherently challenging the applications to enhance the performance, e.g. handwriting recognition of the machines achieves human level of performance, same as for face recognition and the object recognition metrics. It is to be admitted that the deep learning begins from the hand- writing recognition. Its architecture called CNN was cre- ated successfully in order to read handwritten postal codes. The motivation for the use of deep learning occurs from the many facts as listed below:-

- Undoubtedly, deep learning will definitely drive AI adoption into the enterprise also.
- Deep Learning is the main driver and the most essential approach to AI.
- Deep learning is a collection of methods and techniques based on artificial neural networks with multiple layers and increased functionality.
- Deep learning perceives tremendous growth because it has deep layered neural networks and the support of graphical processing units to improve the execution.
- Deep neural networks mainly include feed-forward networks with convolution and pooling layers.
- There is no sequence and inputs and outputs are independent.
- Deep neural networks achieved eminence, 4–5 years back when deep models started replacing the traditional approaches, especially in handwriting recognition, healthcare, image classification, speech recognition and natural language processing.
- Deep neural networks can be disciplined and analyzed by many researchers and academia.
- Deep learning techniques and methodologies are more accurate when skilled with large amount of data.
- NVIDIA will influence the space in 2017 because they are having the affluent deep Learning ecosystem. Intel Xeon Phi solutions are buried on influx with respect to deep learning.
- Designers will depend on meta-learning.
- Reinforcement learning will only become more crea- tive.
- Adversarial and cooperative learning will be the king.

6 Deep Learning vs. Machine Learning

Deep learning architecture is constructed from many hidden layers and multiple neurons per layer. The multilayer archi- tecture facilitates with the mapping of the input to higher level representation. Here we discuss the major differences that are found between two learning techniques:-

- Deep learning constructs algorithms in various layers to make an artificial neural network, which can learn and take intelligent decisions on its own, whereas machine learning needs algorithms to interpret data, learn from that data and then synthesized informed decisions.
- Deep learning takes a large amount of data while machine learning needs a small amount of data to work and arrive at a conclusion.
- Deep learning requires hardware with very high performance.
- Deep learning creates new features by its own processes

and techniques, whereas in case of machine learning, features are accurately and precisely recognized by the users.

- Deep learning solves the problem on end to end basis, whereas machine learning solves it by decomposing a bigger task into smaller tasks and then combining the results.
- Deep networks are black box networks and their working is very difficult to understand because of hyper parameters and complex network design.
- Time requirement to train is much more in deep learn- ing than in machine learning.
- Transparency is shown by machine learning methods rather than the deep learning methods.
- Accuracy rate achieved by deep learning is very satis- factory as compared to machine learning.
- Challenging and complex feature engineering phase is eliminated in the deep learning which is present in the machine learning.
- Deep networks need high-end graphical processing units which are very expensive and are skilled in suf- ficient time with big data.

7 Deep Learning vs. Conventional Learning

The major differences that are present between the deep learning methodologies and the conventional learning are as described below:-

- Extraction of features and their Representation

- From the raw sensor, deep learning methods can learn features and finds the most suitable pattern for improving the recognition accuracy.
- Conventional learning worked on the feature vec- tors which are manually produced and applications dependent. These features are difficult to model in complexities.

- Generalization and Diversity

- It is possible to extract spatial, scale invariant and temporal features from the unlabeled raw data in deep learning.
- Conventional learning used labeled sensor data. And also focus on feature selection with dimen- sionality reduction methods.

- Data preparations

- In deep learning, pre-processing of the data and normalization are not mandatory.
- Conventional learning extracts features by using sensor appearance and within the active windows.
- Temporal and Spatial changes in Activities

- Use of hierarchical features and translational invari- ant features can solve the complexities present in intra-class variability's in handcrafted features.
- Handcrafted features are not suitable and inefficient in solving the inter-class variability's and inter-class relations in the conventional learning.

- Model Training and Execution time

- To avoid over fitting, deep learning requires large amounts of sensor dataset. It is also used for reduc- ing high computations. Graphical Processing Unit (GPU) is used to speed up the training.
- Less training data is required with less time for com- putation and memory utilization is also less in con- ventional training.

8 Reported Work on Various Applications of Deep Learning

The target approach of deep learning is to resolve the sophis- ticated aspects of the input by using multiple levels of rep- resentation. 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 evolu- tionary algorithms. Some major challenges that the deep learning technology is facing undoubtedly are the scaling of computations, optimization of the parameters of deep neu- ral network, designing and learning approaches. A detailed investigation in various complex deep neural network mod- els 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. Numerous applications of deep learning are depicted in Fig. 4.

- Acoustic Modeling

Mohamed et al. [37] proposed deep learning network, which contains multiple layers of features with many parameters for the phone recognition. They replaced Gaussian mixture models and used TIMITT dataset. They trained deep learning networks as a multilayered genera- tive model. After designing features of pre-trained deep network, the next step was to perform discriminative fine tuning with the back propagation, distribution so as to adjust the features for the better prediction of probability distribution. They worked on such applications of acous- tic modeling where multiple layers of features were pre- trained. They explicitly exemplary the covariance structure of the input features. They were trying to reveal alterna- tive representations of the input that helps deep neural networks to gather the relevant information in the sound- wave. They also explored various ways of using recurrent neural networks for increasing the amount of past detailed information that helps in the interpretation of the future.

Ling et al. [29] presented in a very systematically way the review of the speech generation approaches. They created interest in the mind of readers to learn the existing paramet- ric speech generation methods and also stimulated for the generation of developing new methods. They concluded in their findings that for parametric speech recognition, RBM and DBN which are called deep joint models and CRBM and DNN are better to represent the complicated and nonlinear relations. Santana et al. [54] presented a unique method for the acoustic modeling. In the presence of noise, they devel- oped the system for the speech recognition by using deep neural network. This is the big challenge for the researchers for developing speech recognition system with the presence of noise speech signals. For their experiment, they used CNN and the recurrent architecture. CNN was used for the acoustic modeling and recurrent method with connection- ist and temporary classification was used for the sequential modeling. Their method worked well as compared to the classical model such as HMM with the BioChaves datasets.

9 Conclusion and Future Aspects

Indeed, one of the most rapidly expanding areas of machine learning is deep learning. The success and versatility of deep learning algorithms are shown by their fast adoption in several disciplines. The advancements and higher accuracy rates achieved via deep learning demonstrate the technology's importance, highlight its progress, and point to its future potential for study and improvement. Furthermore, it is crucial to emphasize that the primary critical success criteria in developing a deep learning application are the layer hierarchy and supervision in learning. Because proper data categorization requires a hierarchical structure, and because supervision places a premium on database maintenance, this is the rationale behind it.

The foundation of deep learning is its new use of hierarchical layer processing and its improvement of preexisting machine learning applications. Deep learning is capable of producing efficient outcomes for a wide range of uses, including voice recognition and digital picture processing. The combination of deep learning with face recognition and voice recognition makes it a promising tool for security both now and in the future. Aside from this, digital image processing is a multi-disciplinary topic of study. Deep learning is an interesting and cutting-edge area of AI research since it has the potential to prove to be an optimization. Finally, we can say that if we ride the success wave, we will see that deep learning is really taking off in many applications, thanks to the growing availability of data and processing power. It is believed that in the coming years, the fast development of deep learning in an increasing number of applications—including healthcare, remote sensing, and natural language processing—will undoubtedly achieve goals and reach new heights of success and happiness. This technology is really ephebic, young, and specific.

What is ahead for deep learning is:

• How well do deep networks perform in complex, dynamic, and multi-type noisy environments?

• Enhancing deep networks' performance with a more diverse set of features?

• Deep neural networks' compatibility with online learning without supervision.

• Moving forward, deep reinforcement learning is the way to go.

Eventually, inferences, efficiency, and accuracy will be desirable with deep networks.

• Keeping a large database up to date.

The goal of this project is to build generative models for the parametric voice recognition system that are deep and have excellent temporal modeling capabilities.

• Use deep learning to automatically analyze ECG signals.

The use of a deep neural network for video object tracking and detection.

• An autonomous vehicle that uses a deep neural network.

We argue that deep learning approaches have garnered a lot of attention in every domain where traditional machine learning methods have been useful. Finally, deep learning has shown to be the most exciting, effective, and supervised method of machine learning. In order to get better and more accurate findings, it may help researchers quickly assess the application's hidden and unbelievable difficulties.

REFERENCES

1. Paul B. Abadi, Jianmin C. Jianmin, Andy D. Jeffrey, Matthieu D. (2016) An infrastructure for massively parallel machine learning: Tensorflow. Volume 16, pages 265-283, published in: OSDI'16, the 12th USENIX conference on operating systems design and implementation.

Ibrahim MEA, Jaffar MA, and Abbas Q. (2018) A thorough analysis of the most current developments in deep vision systems. "Artif Intell Rev." https://doi.org/10.1007/s10462-018-9633-3.

Thirdly, in 2017, Affonso, Rossi, Vieria, and Carvalho published a study. Biomedical picture categorization using deep learning. Reference: Expert Syst Appl 85:114-122

Fourth, in 2016, Alwzwazy HA, Albehadili HA, Alwan YS, and Islam NE used convolutional neural networks to recognize handwritten digits. Volume 4, Issue 2, Pages 1101–1106 of the Proceedings of the International Journal of Innovative Research in Computer and Communication Engineering

Fifthly, in 2017, Amato et al. included Gennaro, Falchi, Carrara, Meghini, and Vairo. Deep learning for decentralized parking lot occupancy detection. Applied System Expertise 72:327-334

In 2017, Araque, Corcuera-Platas, Sánchez-Rada, and Iglesias published a study. Improving social application deep learning sentiment analysis with ensemble methods. Advanced Systems Analysis 77: 236–246

Seventh, Ashiquzzaman and Tushar (2017) Using deep learning neural networks to recognize Arabic numerals written by hand. In: Images, vision, and pattern recognition: proceedings of the IEEE international conference, pages 1–4. URL: https://doi.org/10.1109/ICIVPR.2017.78908 Table of Contents

According to Azar and Hamey (2017), Summarizing text using deep learning without supervision. Advanced Systems Analysis 68:93-105

Deep learning for big data: problems and solutions (Chen XW, Lin X, 2014). The URL for the article is https://doi.org/10.1109/ACCES S.2014.2325029 and the publication number is IEEE 2:514-525.

In 2016, Chen, Lee, and Lu published a paper. An architecture for deep learning on the go and its implementation in smart vehicle cameras. Section A, Pages 14–25, International Conference on the Internet of Vehicles Book of Abstracts. Accessed at https://doi.org/10.1007/978-3-319-51969-22.

Chapter 11: The Authors Chen, Papandreou, Kokkinos, Murphy, and Yuille (2018) Segmenting semantic images using DeepLab's convolutional neural networks, arousal convolution, and fully linked convolutional neural networks (CRFs). (40(4):834-848) IEEE Transactions on Pattern Analysis and Artificial Intelligence

This is the 12th publication of Cheng et al. (2018). Person re-identification using deep feature learning with structured graph Laplacian embedding. Journal of Pattern Recognition, 82:94-104

Thirteen. Chong E, Han C, and Park FC (2017) Methodology, data formats, and case studies of deep learning networks for analysis and prediction of the stock market. The article "Expert Syst Appl 83:187-205"

In 2014, Chu and Srihari found 14. Using a deep neural network to identify writers. Pp. 1–7 in: Computer Vision, Graphics, and Image Processing 2014, Indian Conference Proceedings.

15. A framework for deep inference learning in healthcare (Dai Y, Wang G, 2018). Recognition of Patterns Letter. See the link at https://doi.org/10.1016/j.patre c.2018.02.009 for more information.

16. Research by Dhieb, Ouarda, Boubaker, and Alilmi (2016) A deep neural network trained on the beta-elliptic model for the purpose of identifying online writers. Volume 18, Issue 4, Pages 1863–1870, International Joint Conference on Neural Networks

In 2017, Falcini, Lami, and Costanza published a study. Advanced algorithms for use in automotive software. "IEEE Software, Volume 34, Issue 3, Pages 56–63, 2017" (https://doi.org/10.1109/MS.2017.79).

In 2017, Gheisari, Wang, and Bhuiyan published a paper. Deep learning with large data: a comprehensive overview. Pages 1–8 of the IEEE International Conference on Embedded and Pervasive Computing (EUC) proceedings

9. Ghosh MMA, Maghari AY (2017) A neural network-based comparative study on hand-writing digit identification. Pages 77–81 in: Proc. of the International Conference on Potential Electronic Technologies (ICPET).

20. Fink GA, Sudholt S, and Gurjar N (2018)." Training with little guidance to learn deep representations for word detection. Pages 7–12 of the 13th International Association for Pattern Recognition (IAPR) Workshop on Document Analysis Systems (DAS).

Rapid and efficient speaker adaption of models for voice recognition based on convolutional neural networks (Hamid OA, Jiang H, 2013). Pages 1248–1252 in: INTERSPEECH edition.

22. Ignatov A. (2018) Convolutional neural networks for real-time human activity identification from accelerometer data. Science of Computer Applications 62: 915-922

approach for deep learning-based image recognition (Jia X, 2017). The 29th IEEE Chinese control and decision conference (CCDC), with references 4730–4735, is published in the volume.

the year 2015 by Kannan and Subramanian A survey on an adaptive strategy to tamil character recognition using deep learning and large data. Expert Systems with Computers: 557-567

Kaushali, Khehra, and Sharma (2018) conducted a state-of-the-art assessment on object identification and tracking methods based on soft computing. Proceedings of the Association for Computing Machinery, 70: 423-464.

In 2018, Krishnan, Dutta, and Jawahar published a study. Word spotting and rec- ognition utilizing deep embedding. Included in: The 13th International Workshop on Document Analysis Systems (DAS) by the Institute of Advanced Paper Research (IAPR). http://doi.org/10.1109/das.2018.70 has been cited.

27. "Deep learning" by LeCun, Bengio, and Hinton (2015) in Nature 521: 1–10.

Section 28: Lee SH, Chan CS, Mayo SJ, and Remagnino P (2017); The process by which deep learning identifies plant species based on leaf characteristics. Finding patterns in 71:1-13

In 2015, Ling et al. (1915) conducted research on a variety of topics. An analysis of current and future developments in deep learning for acoustic modeling in parametric speech production. 32(3):35-52 IEEE Signal Processing Magazine

In 2018, Ling et al. (30) conducted research. Taguchi method-based spectrum prediction using long short-term memory in deep learning... The article is published in IEEE Access and has the DOI number 6(1): 45923-45933.

Thirty-one. Deep learning and its use in general picture categorization (Liu PH, Su SF, Chen MC, Hsiao CC, 2015). One page to one page long, published in the proceedings of the international conference on cybernetics and informatics for computational social systems.

Thirdly, in 2017, Looks, Herreshoff, Hutchins, and Norvig published a... Dynamic computation graphs for deep learning. Page 1 through 12 of the international conference on learning representation's proceedings

Thirty-three. Lopez, Rivas, and Gualdron (2017) Cognitive radio wireless networks' principal user characterization via a deep learning neural system. Analysis of Artificial Intelligence Review: 1-27

35. Luckow A, Cook M, Ashcraft N, Weill E, Djerekarov E, Vorster B (2017) Tools and applications of deep learning in the automobile sector. Includes pages 3759–3768 from the IEEE International Conference on Big Data.

In 2016, Makhmudov and Abdukarimov published a study. Deep learning algorithms for speech recognition. Volume 10, Issues, Methods, and Technologies in Informatics, Proceedings of the International Conference on, pages 10-15.

The authors of the 2018 study are Markovnikov, Kipyatkova, Karpov, and Filchenkov. Russian voice recognition using deep neural networks. Journal of Artificial Intelligence, Natural Language Processing, and Computer Science 789:54-67. http://doi.org/10.1007/978-3-319-71746-3_5

Thirdly, in 2009, Mohamed A. Dahl and Geoffrey H. Deep belief networks for phone recognition. Pp. 1–9 in: Deep Learning for Speech Recognition and Related Applications: Proceedings of the NIPS Workshop

Mussa AM, El-Makky NM, and Ghanem N. (2017) in 38. Using deep learning to identify authors. Presented at the 15th Annual IEEE