A Comparative Study on Different Training Model in Machine Learning

Priyanka S Jigalur[1], Dr. B. G. Prasad[2]

[1]Department of Computer Science & Engineering, B.M.S. College of Engineering, Bengaluru, India
priyanka.jigalut[at]gmail.com

[2]Department of Computer Science & Engineering, B.M.S. College of Engineering, Bengaluru, India
bgprasad.cse[at]bmsce.ac.in

Abstract: ***Image reconstruction is the process of regeneration of an image from its distorted version. The image reconstruction problem is part of a larger field of image processing that is used in a variety of applications, including Forensic Analysis, Medical Science, Astronomy, Entertainment industry, Legal Services, etc. Designing a neural network that memorizes a given pattern or image and reconstructing the image requires resources, which is a tedious task and time-consuming. Many studies have been conducted in order to reconstruct images that have been distorted. This paper conducts a review of the current research to reconstruct the image from the distorted image using different Machine Learning methods. The domain of machine learning has led to the tremendous development of deep neural network techniques. Neural Networks are used to model complex patterns and predictive problems. This survey has compared different approaches for solving the problem and has proposed the neuromodulation technique as an efficient technique.***

Keywords: Image Reconstruction, Machine Learning, Forensic Analysis, Astronomy, Neural networks, Neuromodulation

1. Introduction

Machine learning is an artificial intelligence branch that allows systems to learn and build on prior knowledge without being explicitly programmed. Machine learning focuses on the creation of computer models that can process data and learn from the data being fed to the model. Pattern recognition is a field that can also use Machine Learning techniques for the identification of patterns. Pattern recognition is used to classify data based on knowledge previously acquired. It can also use statistical information extracted from patterns and their representation.

The domain of machine learning has led to the immense development of deep neural networks. Recurrent Neural Networks (RNNs) have evolved throughout the years to capture both long-term and short-term relationships. The added capacity is utilized to store the relationship information alongside the rest of the network's components.[1] Recurrent neural networks (RNNs) are a diverse family of deep neural networks. These recurrent associations enable RNNs to have an internal state that can memorize background information, allowing them to catch temporal patterns in sequential data sets.

Long-Short Term Memory network (LSTM) is the most widely used popular recurrent neural network structure in the deep learning field. They are designed to ignore long-term dependencies and are more focused on short-term information or dependencies. Other methods such as temporal convolutions or memory networks exist, which are well suited for handling time-based or temporal information.

Many of the current machine Learning techniques to know successes include learning one complex challenge thoroughly by extensive testing over millions of training samples. The agent's information is stored and unchanged after the end of training. In order to perform a particular function, the agent has to be re-trained again if the model is not trained before. They again involved many new training instances. The question arises of how appropriate propagation of learning signals to non-output neurons can be performed in a biological fashion in deep learning neural networks.

To resemble the biological neurons using back propagation, a new learning paradigm was introduced. This paradigm took inspiration from cellular neuromodulation, which was designed specifically to learn adaptive behaviors and create a new deep neural network architecture. The idea of Plasticity is derived from humanity's understanding of the functioning of the human brain, which is known to continuously form new connections between neurons based on the usage of the neuron. Also, it discards connections between them if they go unused or aren't strengthened. However, the automatic modification of synaptic weights is to be set as a function of ongoing activity. This concept of neurons strengthening and weakening their connections with other neurons has been thoroughly studied in the academic sphere.

In this way, the advancement of neural plasticity, where the memory access mechanisms in neural memory models are made plastic, gives changing attention levels to the stored data and activity.

This paper presents the comparison of various techniques of image reconstruction. Later discuss the efficient approach of the same.

2. Literature Survey

This section describes the different studies which the researchers on image reconstruction carried out. The advancement in Artificial Intelligence in computer science tries to work like the human brain in learning, then reasoning it and later adopting the model based on what it has learned. When these characteristics are put together in real-world testing, these techniques become inefficient and highly specialized to the task. The mathematician John McCarthy, who coined Artificial Intelligence in the mid-1950s, is recognized as the father of AI, used to describe computational machines that we can replicate the logical thinking and reasoning performed by a human, such as problem-solving abilities [8].

Machine Learning can be thought of as a black-box model that continuously learns from the input being fed. The learning is done through means such as back-propagation, etc. This learning process is modeled on the functioning human brain. The Machine Learning algorithms are trained to learn the input data's different features instead of memorizing the data itself. This input data is then further processed to generate learned weights for the model described and make predictions based on these learned weights. To understand how the predictions are made in machine learning algorithms, many intensive algorithmic approaches were mapped, and the tasks to implement the logical inference. Several models were designed to infer these algorithmic mechanisms to train the models for future insights.

There are many different approaches that show that ML algorithms can learn by themselves. Machine Learning was developed to make models that work similarly to the human brain, which can be used in various applications for prediction. Dartmouth proposed a Machine Learning hypothesis that “the intelligence of Humans can be so precisely described that a machine can be made to simulate it" [6]. The Dartmouth theory led to the advancement of a Neural Processing Network using biological mechanisms to make predictions of their actions. The biological neural network consists of numerous neurons, which exactly behave like human nerves. Each network is formed by connecting a neuron with all other different neurons connected.

The development of different ML algorithms led to the advancement of new computing paradigms. One of those computing models is Artificial Neural Network (ANN), coined by the neurophysiologist Warren McCulloch and logician Walter Pitts in 1943, which is considered the first neural architecture model. The Artificial neural network consists of numerous connected nodes called artificial neurons. These neurons are triggered based on a mathematical function called the activation function. Also, a bias value is attached to the neuron, which acts like an error parameter that tells how well the neuron's influence should propagate to the next neuron. Each neuron transmits the signal to another neuron, which receives the signal, and the same process is repeated. The neurons have weights assigned that can increase or decrease the signal strength. ANN utilizes simple logical concepts for processing the artificial neurons, resulting in the exhibition of complex behavior. The activity of the neurons and their connection between the different neurons can be calculated by propositional logic[7]. D. O. Hebb created the learning hypothesis based on the neural plasticity mechanism, known as Hebbian learning. He also explained the neuron connection in his book, The Organization of Behavior, which explains the synaptic plasticity in neurons, which can be weakened or strengthened based on the model's activity [9]. With the development of ANN over the years, Frank Rosenblatt, in 1958, modeled Perceptron, a new learning algorithm. The neural network, which is a single-layer perceptron, is used to classify the input data as belonging to one of two classes at the output layer. Collecting data or images to train the model and an output layer consisting of the desired output [10][11].

The perceptron can be used for making complex decisions, but there are some limitations. Marvin Minsky and Seymour's Paper explains the boundaries of the perceptron. The perceptron can be affected by the bias or weights, which results in changing the output range between 0 to 1 [12]. In the year 1975, Werbos Paul J introduced the concept of back-propagation to train multilayer perceptrons, and this resulted in the growth of new research of multilayer neural networks [13].

In 1982, the modern Neural Network's age began with the Hopfield network and continued its journey in 1985 with the Boltzmann machine's studies from Sejnowski. This resulted in the new change to Neural Networks and was renamed as the generative model from the passive associative machine [15] [14]. In 1997 Deep learning architecture was enhanced by Schmidhuber & Hochreiter proposed a recurrent neural network framework, Long Short-Term Memory (LSTM) [17]. These theories of RNNs were impelled to keep track of the input field, and these neural networks propagated as feed-backward connections that focused on long-term dependencies.

The major drawback of addressing the LSTM is that it takes a considerable time to train the model and have the Vanishing Gradient Problem, which resulted in difficulty remembering many previous connections. Also, LSTM has the problem of long-term dependencies[18]. Hence there is a need for a better neural network that can remember the previous connection for a longer time. In 2016 Thomas Miconi explained differentiable neural networks. . To train Differentiable Plasticity using backpropagation, the plastic rule has to be applied. Further, it is introduced over Hebbian learning to build long-term dependencies[1][2]. The connection between two neurons has both a plastic component and a fixed component. The fixed components act like the traditional neural network approach, where all the parameters such as bias and learning rate are fixed. In contrast, the plastic component is stored in Hebbian trace[1]. This results in the neural network for the validation of the long-term dependencies.

Comparative study for Machine Learning Algorithm

This section explains the comparative study of different neural networks such as Convolutional Neural Network (CNN), Artificial Neural Network(ANN), Recurrent Neural Network(RNN), and Differentiable Neural Network(DNN).

Table 1: Comparative study of different neural networks

	ANN	CNN	RNN	DNN
Data	Tabular data	Image data	Sequence data	Image/ Sequence data
Recurrent connection	No	No	Yes	Yes
Parameter sharing	No	Yes	Yes	Yes
Spatial Relationship	No	Yes	No	No
Vanishing and Exploding gradient	Yes	Yes	Yes	No

Table 1 explains the comparative study of different neural networks. The idea of adding plastic rules plays a vital role in learning completely naturally [1], which makes it interesting in modern neural networks. Also, using Differentiable plasticity models can prevent gradient explosion and vanishing decent problems. This study shows that Differentiable Neural Network strongly improves the performance compared to other different neural networks and shows that meta-learning can outperform the other alternative neural network approaches.

3. Conclusion and Future Scope

In this paper, various neural networks such as Convolutional Neural Network(CNN), Artificial Neural Network(ANN), Recurrent Neural Network(RNN), and Differentiable Neural Network(DNN) with respect to pattern recognition. By comparing different techniques, Neural networks found that Differentiable Plasticity has more efficiency than other techniques. It is an active research area, and various researchers work to improve the algorithm's efficiency by developing more efficient algorithms. There are several application domains of implementing neural networks on high-dimensional data to gain insights into the problems.

References

[1] Thomas Miconi, Jeff Clune, and Kenneth O. Stanley. Differentiable plasticity: training plastic networks with gradient descent. In Proceedings of the 35th International Conference on Machine Learning, 2018.

[2] Thomas Miconi, Aditya Rawal, Jeff Clune & Kenneth O. Stanley. Backpropomine: Training self-modifying neural networks with differentiable neuromodulated plasticity. Published as a conference paper at ICLR 2019.

[3] Thomas Miconi.Neural networks with differentiable structures. The Neurosciences Institute La Jolla, CA, USA, Aug 2016.

[4] https://eng.uber.com/differentiable-plasticity/

[5] B.Yegnanarayana, Artificial Neural Network, Indian Institute of Technology Madras, 2005.

[6] ”Philosophy of artificial intelligence,” February 2020, Wikipedia.

[7] W. McCulloch and W. Pitts “A Logical Calculus of Ideas Immanent in Nervous Activity,” Bulletin of mathematical biophysics, 1943.

[8] Jef Akst, “A Primer: Artificial Intelligence Versus Neural Networks,” the-scientisit.com, para. 1, April. 30, 2019. [Online].

[9] Hebb, D.O. “The Organization of Behavior: a neuropsychological theory,” 1949 [Online].

[10] Rosenblatt, F. “The Perceptron: A probabilistic model for information storage and organization in the brain.” Psychological Review, 65(6):386–408, 1958.

[11] Rosenblatt Frank, "The Perceptron—a perceiving and recognizing automaton.” Report 85-460-1, Cornell Aeronautical Laboratory, 1957.

[12] Marvin Minsky and Seymour Papert. “Perceptrons. An Introduction to Computational Geometry”. M.I.T. Press, Cambridge, Mass., 1969.

[13] Werbos, Paul J. (1994). The Roots of Backpropagation: From Ordered Derivatives to Neural Networks and Political Forecasting. New York: John Wiley & Sons. ISBN 0-471-59897-6.

[14] Hinton, GE, Sejnowski, TJ, and Ackley, DH. Boltzmann machines: Constraint satisfaction networks that learn. Technical report, Department of Computer Science, Carnegie-Mellon University, 1984.

[15] Hopfield, JJ. Neural networks and physical systems with emergent collective computational abilities. Proc. Nat. Acad. Sci. USA, 79(8):2554–2558, 1982.

[16] Vapnik, Vladimir N. (1992). "A training algorithm for optimal margin classifiers.” *Proceedings of the fifth annual workshop on Computational learning theory – COLT '92*. p. 144.

[17] Sepp Hochreiter; Jürgen Schmidhuber (1997). "Long short-term memory". *Neural Computation*. **9** (8): 1735–1780. doi:10.1162/neco.1997.9.8.1735. PMID 9377276.

[18] Felix A. Gers Jürgen Schmidhuber, and Fred Cummin(2000), **“**Learning to Forget: Continual Prediction with LSTM, “Massachusetts Institute of Technology, Neural Computation, 2000 12:10, 2451-2471.

[19] Wang, J. X., Kurth-Nelson, Z., Tirumala, D., Soyer, H., Leibo, J. Z., Munos, R., Blundell, C., Kumaran, D., and Botvinick, M. Learning to reinforcement learn. 2016.

[20] Vecoven N, Ernst D, Wehenkel A, Drion G (2020), Introducing neuromodulation in deep neural networks to learn adaptive behaviors. PLoS ONE 15(1): e0227922.

[21] Hebb, Donald (1949). The Organization of Behavior. New York: Wiley. ISBN 978-1-135-63190-1.