Review of Different Deep Learning Approaches for Image Classification

[1] Aashish Bohra, [2] Neha Sankhla, [3] Shrwan Ram
[1,2] Computer Science Department, MBM Engineering College, Jodhpur, India
[3] Associate Professor Computer Science Department MBM Engineering College
Jodhpur, India

Abstract : As a major breakthrough in artificial intelligence, deep learning has achieved very impressive success in solving grand challenges in many fields including speech recognition, natural language processing, computer vision, image and video processing, and multimedia. Deep Learning is the subpart of machine learning; it uses neural networks which stimulate by the human brain's structure and working. It is a new approach for data analysis and prediction and has become very popular recently. Deep Learning has achieved much higher success than machine learning in many applications, one of the reasons for that is, machine learning not able to process large amount of data effectively but also not able to extract the features from the data automatically. Since Deep Learning has been the core topic in machine learning and convolutional neural network have become state of the art methods for image classification over the last couple of years. Convolutional neural network has won numerous competitions in recent years. It has outstanding results in image recognition. In this paper, different deep learning approaches will be reviewed which have been used in the field of image classification and localization.

IndexTerms **- Artificial Intelligence, machine learning, deep learning, image classification, neural network, convolutional neural network, artificial neural network**

I. INTRODUCTION

Recent advances in deep learning made possible tasks such as image and speech recognition. Deep learning is a subset of machine learning algorithms that is very good at recognizing patterns, but generally requires a large amount of data. Deep learning is excellent in the recognition of objects in images, since it is implemented using 3 or more layers of artificial neural networks where each layer is responsible for extracting one or more characteristics of the image.
Neural network: a computational model that works in a similar way to neurons in the human brain. Each neuron takes an input, performs some operations and then passes the output to the next neuron. Representation of neural network (src), we will teach the computer to recognize the images and classify them into one of these categories:

To do this, we must first teach the computer how a cat, a dog, a bird, etc. looks before we can recognize a new object. The more cats see the computer, the better it will be to recognize cats again. This is called supervised learning. We can carry out this task by labeling the images, the computer will begin to recognize the patterns present in the images of cats that are absent from others and will begin to develop their own cognition. We are going to use Python and TensorFlow to write the program. TensorFlow is an open source deep learning framework created by Google that gives developers granular control over each neuron (known as a "node" in TensorFlow) so that it can adjust weights and achieve optimal performance. TensorFlow has many integrated libraries (some of which we will use for image classification) and has an incredible community, so you can find open source implementations for virtually any deep learning topic.

II. MACHINE LEARNING

Machine Learning Algorithm is a step by step process to get information from the given set of data, without relying on a fix program. This information is useful to predict output for a given input.
Machine Learning Algorithms find patterns inside the set of data, i.e. given to it. These Patterns are useful to make better decisions and predictions. Today ML algorithms are useful in many areas to make better decisions like, in medical diagnosis, stock trading, and energy load forecasting, etc. These algorithms are used by media sites to recommend movies and songs to users and also use by retailers to recommend products to customers, based on their purchasing behavior [1].

A. Machine Learning Algorithms

Mainly two types of Machine Learning Algorithms are defined based on their learning capability. Which are shown below in Fig 1.

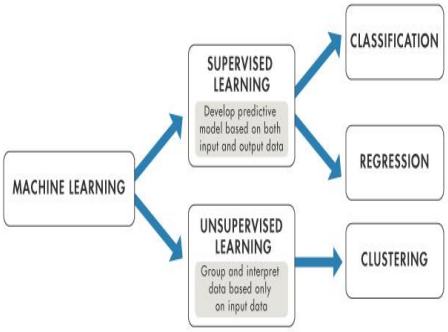

Fig 1: Machine Learning Architecture

Supervised Machine Learning

In this type of learning sample labeled inputs and outputs training datasets are given to the system initially as shown in Fig 2.

/* Historic Data */ Labeled (input, output) ------> (input,?) /*Present Data*/

A supervised machine learning algorithm analyses and studies training data sets and produces a function that use to map new inputs, which are given to the system and predict output for the given input, based on analysis of training datasets. Supervised machine learning architecture defines that how supervised machine learning algorithm works in practical scenario.

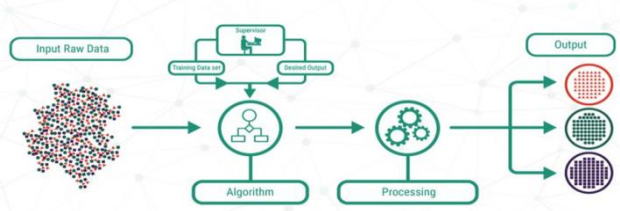

Fig 2: Supervised Machine Learning Architecture

Unsupervised Machine Learning

In this type of learning, sample unlabeled inputs (unlabeled training datasets) are given to the system initially, without any output as shown in Fig 3.

/*Historic Data*/ input (Unlabeled), output (not present)) ----------> (input,?) /*Present Data*/

Than an unsupervised machine learning algorithm analyses training data sets and finds pattern behind the data, based on size, shape, color etc., and cluster the data as similar pattern data should be in same cluster and finally produces a function that use to map new inputs, when inputs given to the system and predict outputs for the given inputs, based on to which cluster it belongs.

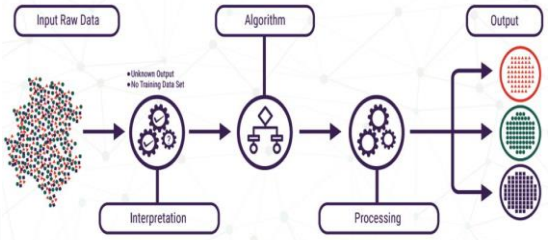

Fig 3: Unsupervised Machine Learning Architecture

III. DEEP LEARNING ALGORITHMS

Deep learning is the subpart of machine learning; it uses neural networks which stimulate by the human brain's structure and working.

Deep learning, a method initialized by the human brain's functionality. In deep learning, networks of artificial neurons analyze large dataset to automatically find out underlying patterns, without human interference [2].

A computer trains to classify pictures, text and sound in deep learning. The computer is trained with large image datasets and then it changes the pixel value of the picture to an internal representation, where the classifier can detect patterns on the input image.

Deep learning algorithms are used to create intelligent machines which use techniques inspired by the human brain's ability to learn. It is the subfield of machine learning; it deals directly with data and can handle complex datasets.

Deep learning is a new approach for data analysis and prediction; it has become very popular recently. Deep learning has achieved much higher success than machine learning in many applications, one of the reasons for that is, machine learning not able to process large amount of data effectively as well as it also not able to extract the features from the data automatically.

Deep learning is being recognized as an essential tool for artificial intelligence research with various applications in several areas such as recognizing speech, recognizing an object and classifying an image.

In particular, deep learning is preferred in the classification of images because it can provide efficient results. For image classification, in deep learning the computer is trained with large image datasets and then it changes the pixel value of the picture to an internal representation where the classifiers can detect patterns on the input image.

Idea of deep learning is to implement learning algorithms that capture human brain. Neural Network is use to implement deep learning, neural network is nothing but the set of artificial neurons or perceptron. These neurons are similar to biological neurons.

A. Neural Network

Neural network receives an input and passes it through a number of hidden layers as shown in Fig 4. A hidden layer has number of neurons, where a neuron is fully interlinked to each neuron in the previous layer. In neural network, each layer functions independently.

The last layer in neural network is called ‘output layer’, which represents the class to which input belongs. Neural network has different hidden layers, where a single layer categorizes some kind of information which passes to the next layer. Neural network combines the features from different hidden layers, and predicts the output using training dataset.

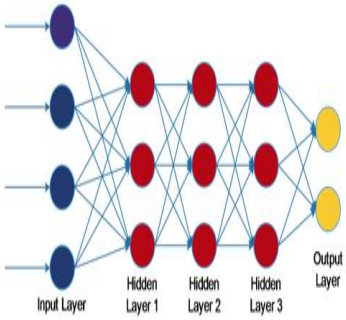

Fig 4: Neural Network Architecture

IV. ARTIFICIAL NEURAL NETWORK

The study of the human brain is thousands of years old. With the advent of modern electronics, it was only natural to try to harness this thinking process. The first step toward artificial neural networks came in 1943 when Warren McCulloch, a neurophysiologist, and a young mathematician, Walter Pitts, wrote a paper on how neurons might work. They modeled a simple neural network with electrical circuits. [3]
An artificial neuron is a computational model inspired by natural neurons. Natural neurons receive signals through the synapses located in the dendrites or membrane of the neuron. When the signals received are strong enough (exceeds a certain threshold), the neuron is activated and emits a signal through the axon. This signal could be sent to another synapse and could activate other neurons. The complexity of real neurons is highly abstract when artificial neurons are modeled. These consist basically of inputs (such as synapses), which are multiplied by the weights (intensity of the respective signals) and then calculated by a mathematical function that determines the activation of the neuron. Another function (which can be the identity) calculates the output of the artificial neuron (sometimes depending on a certain threshold). [4]

The ANNs combine artificial neurons to process the information process of information storage as patterns, use those patterns and then solve problems encompassing a new field in information technology. This field, as mentioned above, does not use traditional programming, but involves the creation of massive parallel networks and the training of these networks to solve specific problems. This field also uses words very different from traditional computer science, words like behave, react, self-organize, learn, generalize and forget.
When we talk about a neural network, more popularly we should say "Artificial Neural Network (ANN), ANN are computers whose architecture is based on the brain, usually consisting of hundreds of simple processing units that are connected together in a network complex communication Each unit or node is a simplified model of a real neuron that sends a new signal or triggers if it receives a sufficiently strong input signal from the other nodes to which it is connected as shown in Fig 5.

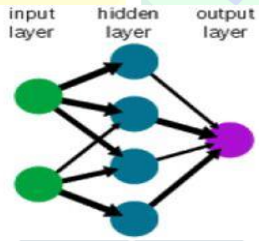

Fig 5: Layers of Neural Network [3]

Traditionally neural network was used to refer as network or circuit of biological neurons, but modern usage of the term often refers to ANN. ANN is mathematical model or computational model, an information processing paradigm i.e. inspired by the way biological nervous system, such as brain information system.

ANN is made up of interconnecting artificial neurons which are programmed like to mimic the properties of biological neurons. [5] These neurons working in unison to solve specific problems. ANN is configured for solving artificial intelligence problems without creating a model of real biological system.

ANN is used for speech recognition, image analysis, adaptive control etc. These applications are done through a learning process, like learning in biological system, which involves the adjustment between neurons through synaptic connection. Same happen in the ANN.

A. Process of ANN

The other parts of the art of the use of neural networks revolve around the innumerable ways in which these individual neurons can be grouped. This grouping occurs in the human mind in such a way that information can be processed dynamically, interactively and self-organized. Biologically, neural networks are constructed in a three-dimensional world from microscopic components. These neurons seem capable of interconnections almost without restrictions. That is not true for any proposed, or existing, network made by man. Integrated circuits, which use current technology, are two-dimensional devices with a limited number of layers for interconnection. This physical reality restricts the types and scope of artificial neural networks that can be implemented in silicon.
Currently, neural networks are the simple grouping of primitive artificial neurons. This grouping is produced by creating layers that are then connected to each other. How these layers connect is the other part of the "art" of engineering networks to solve real-world problems.

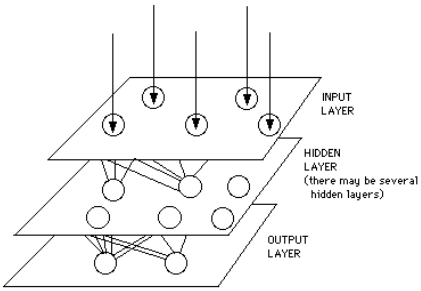

Fig 6: A Simple Neural Network Diagram

Basically, all artificial neural networks have a similar structure or topology as shown in Fig 6. In that structure, some of the neurons connect with the real world to receive their inputs. Other neurons provide the real world with the results of the network. This output could be the particular character that the network thinks it has scanned or the particular image it thinks it is seeing. All the rest of the neurons are hidden from view. But a neural network is more than a bunch of neurons. Some early researchers simply tried to connect the neurons in a random way, without much success. Now, it is known that even the brains of snails are structured devices. One of the easiest ways to design a structure is to create layers of elements. It is the grouping of these neurons into layers, the connections between these layers and the addition and transfer functions that make up a functioning neural network. The general terms used to describe these characteristics are common to all networks.

Although there are useful networks that contain only one layer, or even one element, most applications require networks that contain at least the three normal types of layers: input, hidden, and output. The entry layer. neurons receives the data either from input files or directly from electronic sensors in real-time applications. The output layer sends information directly to the outside world, to a secondary computer process or to other devices, such as a mechanical control system. Between these two layers there can be many hidden layers. These internal layers contain many of the neurons in several interconnected structures. The inputs and outputs of each of these hidden neurons simply go to other neurons. In most networks, each neuron in a hidden layer receives signals from all the neurons in a layer above it, typically an input layer. After a neuron performs its function, it passes its exit to all the neurons in the layer below it, providing an advance path to the exit.

These lines of communication from one neuron to another are important aspects of neural networks. They are the glue to the system. They are the connections that provide a variable force to an input. There are two types of these connections. One causes the addition mechanism of the next neuron to be added, while the other causes it to be subtracted. In more human terms, one excites, while the other inhibits. Some networks want a neuron to inhibit the other neurons in the same layer. This is called lateral inhibition. The most common use of this is in the output layer. For example, in text recognition, if the probability that a character is a "P" is 0.85 and the probability that the character is an "F" is 0.65, the network wants to choose the highest probability and inhibit to everyone else. You can do it with lateral inhibition. This concept is also called competition. The classification takes time difference to be classified because the data training set numbers depend on the algorithm of the neural network model and the number of classes [6].

V. CONVOLUTIONAL NEURAL NETWORKS

Convolutional Neuronal Networks are very similar to ordinary neural networks. They are formed by neurons that have weights and learning biases. Each neuron receives some inputs, performs a point product and, optionally, follows it with non-linearity. The entire network still expresses a unique distinguishable scoring function: from unprocessed image pixels on one end to class scores on the other. The common problem in the classification of the image with deep learning is a lower performance due to an excessive adjustment. To increase performance, avoid over-adjustment, use large data sets. CNN has fewer connections and hyper parameters that make the CNN model easy to train and perform slightly worse than other models [7]. CNN get information directly from the image, thus it eliminate manual feature extraction [8].

CNN, a different type of artificial neural network with feed forwarding, which initialized using visual cortex. In CNN, in a layer a neuron is interconnected to a small area before it, not as neural network where all the neurons are fully interconnected, so CNN handle fewer amounts of weights and also less number of neurons. A simple convolutional neural network concept is shown in Fig 7.

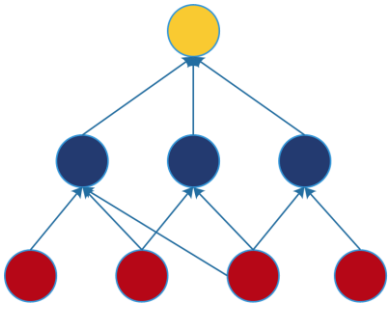

Fig 7: Convolutional Neural Network Concept

VI. WORKING ARCHITECTURE OF CNN

CNN has mainly four layers,

- Convolutional Layer
- Pooling Layer
- Flattening Layer
- Fully Connected Layer

For every convolutional neural network above layers are the basic parts. Let see the idea behind each of the above layer.

Convolutional Layer

The Convolutional layer is the core building block of a Convolutional Network that does most of the computational heavy lifting. This layer has a set of filters, where each filter has a small receptive area. The primary aim of convolutional layer is to extract features from the input image. [9]

The brain view: If you are a fan of the brain / neuron analogies, each input in the 3D output volume can also be interpreted as an output of a neuron that looks only at a small region at the input and shares parameters with all the neurons on the left space (since all these numbers result from applying the same filter). Now we analyze the details of the connectivity of neurons, their disposition in space and their scheme of exchange of parameters.

Local Connectivity: When it comes to high-dimensional entries, such as images, as we saw earlier, it is not practical to connect the neurons to all the neurons in the previous volume. Instead, we will connect each neuron only to a local region of the input volume. The spatial extent of this connectivity is a hyper parameter called the receptive field of the neuron (equivalently, this is the size of the filter). The extent of the connectivity along the depth axis is always equal to the depth of the input volume. It is important to emphasize again this asymmetry in the way we treat the spatial dimensions (width and height) and the depth dimension: The connections are local in space (along the width and height), but always complete throughout all the depth of the input volume.
Convolutional layer take an image as a input, and slide each filter over input image by 1 pixel [also called stride] and for every position, compute element wise multiplication [between the 2 matrices] and sum up the outputs of multiplication to find the final value that forms a single element of a feature map.
For example, consider a 7*7 image whose pixel values are -1 and 1, also consider a 3*3 filter as shown below in Fig 8,

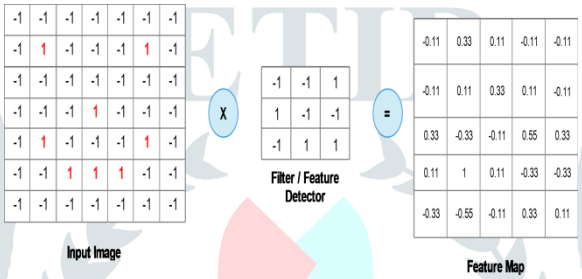

Fig 8: Convolutional Layer

In above, slide 3*3 filter over 7*7 input image by 1 pixel [stride] and perform some computation to get a 5*5 matrix known as feature map or activation map. Here a filter acts as a feature detector which extract feature from the input image.
Above feature map size is 5*5 pixels which generated with {[(W-F+2P) / S] + 1}, Where W is the size of image, F is the size of filter, P is the zero padding and S represents stride value.
For above convolutional layer value of W = 7, F = 3, P = 0 and S = 1, so feature map size will be 5, which is 5 * 5 matrix.

- Relu

Relu is an additional operation used after convolutional layer operation. It replaces negative elements by zero (as shown in Fig 9) in feature map or activation map. The main aim of Relu is to define non-linear values in CNN.

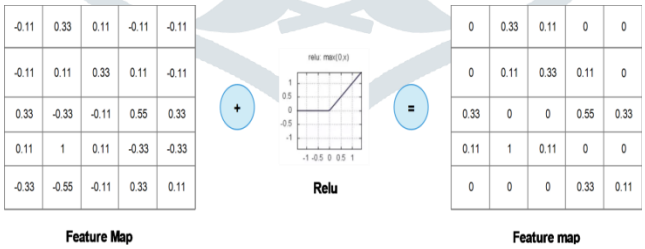

Fig 9: Rectified Linear Unit Process

Pooling Layer

It is an essential building block of CNN. It used to minimize the size of each rectified feature map without removing important information. There are different types of pooling present as: Max, Average and Sum pooling.
Among all pooling, Max pooling shown to work better. In max pooling, a window [for example, 2*2 window] slides over the activation map and get the largest value present in that 2*2 window.
A max pooling example shown in Fig 10 below, where polling operation performed on a feature map with 2*2 sliding window,

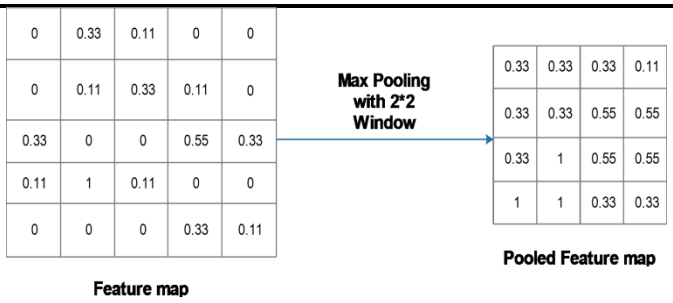

Fig 10: Pooling Layer Process

In above Fig 10 a 2*2 window slides over the feature map and obtain the largest element in each field, it minimizes the size of a feature map.
For any CNN, Convolutional, Relu and Pooling are the essential building parts. For more feature extraction we can use more than one convolutional layer, which implement similarly as above that apply convolutional, relu and then pooling layer.
A pooling layer output is given as an input to the next layer that is flattening layer.

Flattening Layer

In flattening a resultant two dimensional matrix translates into a single long continuous linear vector as shown in Fig 11. Flattening arrange the pooled map into column for artificial neural network presentation. The flattening process is used to present fully connected layer, which means to combine all the founded features of the previous convolutional layer.

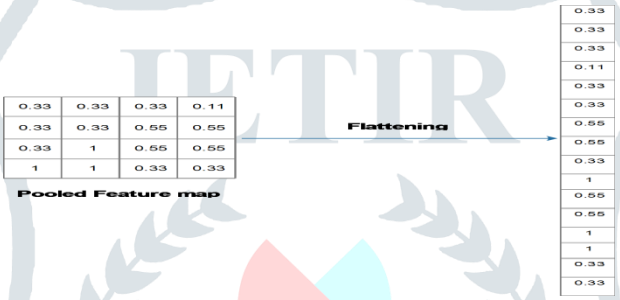

Fig 11: Flattening Process

Fully Connected Layer

The fully connected layer is old multilayer method which uses a classifier in the output layer for classification. 'Fully Connected' means that a neuron in a layer is interconnected to each neuron present in the last layer.
This layer gets the inputs from previous layer [flattening] and generates output as N dimensional vector, where N is the number of classes in output layer. Convolutional and pooling layer obtains features of high quality from a training image. The main aim of fully connected layer is to get these high quality features for classifying image into different classes depends on the training dataset. We can use an image classifier support vector machine in place of fully connected layer. But fully connected layer is good to create user to user trainable model. The fully connected layer is explained in the Fig 12, which is shown bellowed. Each input is being subsequently connected to hidden layers and correspondingly output is being generated.

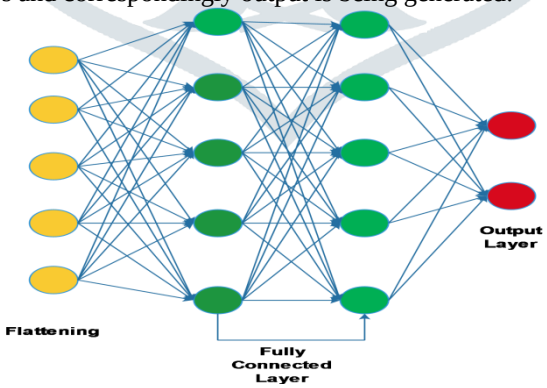

Fig 12: Fully Connected Layer

Features obtained through previous CNN layers might be good for image classification but combination of that features provides better results with fully connected layer.

VII. CONCLUSION

Depending on the nature of the application and the strength of the internal data patterns, you can usually expect a network to be trained quite well. This applies to problems where relationships can be quite dynamic or non-linear. Artificial neural networks provide an analytical alternative to conventional techniques that are often limited by strict assumptions of normality, linearity, variable independence, etc. Because an artificial neural network can capture many types of relationships, it allows the user to model phenomena quickly and relatively easily, otherwise it may have been very difficult or impossible to explain in any other way. Nowadays, discussions of neural networks are happening everywhere. His promise seems very bright, since nature itself is proof

that this kind of thing works. However, its future, in fact, the key to all technology, lies in the development of hardware. Currently, the development of more neural networks is simply demonstrating that the main one works.

You need an image processing technology to help in the classification process, which are the defect and the non-defect. In this paper, we use one of deep learning architecture that is the convolutional neuronal network.

Convolutional neural network proved to be very efficient regarding classifying images. In the preparation of the convolutional neural network architecture model, initializing the parameter configuration accelerates the network training process. In this work, a deep-learning study based on Convolutional Neural Network and Artificial Neural Network has been presented. Various advantages and disadvantages of Artificial Neural Network and Convolutional Neural Network have been discussed. In this the working of Artificial Neural Network has been explained that how the idea of basic neural network have evolved over the years. In the experiments conducted it has been seen that Artificial Neural Network are effectively being used in Natural Language recognition.

The other deep learning approach discussed is the Convolutional Neural Network. They are the most efficient neural network among others. Various layers of the convolutional neural network have been discussed. Among the layers, convolution layer is the core building block of the whole convolutional neural network.

A number of different images which contains two types of animals, namely cat and dog are used for classification. Convolutional Neural Network is being effectively used in Image Classification and Pattern Recognition. It is shown that the convolutional neural network with higher layer performs classification process with much higher accuracy.

Convolutional Neural Network also used in various other applications like cancer tissue detection, vehicle detection in traffic management, human image detection, robotics etc. Convolutional Neural Network approach is the best technique for image recognition and detection of correct image as per image dataset.

REFERENCES

1. Narendra Kumar, Nidhi Kharkwal, Rashi Kohli and Shakeeluddin Choudhary (ICICCS 2016),” Ethical Aspects and Future of Artificial Intelligence”,.
2. Yongkweon Jeon, Jaeyoon Yoo, Jongsun Lee, and Sungroh Yoon “NC-link A New Linkage Method for Efficient Hierarchical Clustering of Large-Scale Data”, IEEE (2017)
3. Ms. Sonali. B. Maind, Ms. Priyanka Wankar “Research Paper on Basic of Artificial Neural Network”, International Journal on Recent and Innovation Trends in Computing and Communication Volume: 2 Issue: 1, ISSN: 2321-8169 pp. 96-100
4. Kh Tohidul Islam, Ghulam Mujtaba, Dr. Ram Gopal Raj, Henry Friday Nweke (ICE2T 2017) ,“Handwritten Digits Recognition with Artificial Neural Network”
5. Tao-ran Cheng, Peng-cheng Wen and Yang Li “Research Status of Artificial Neural Network and Its Application Assumption in Aviation” 2016 12th International Conference on Computational Intelligence and Security
6. Nur Anis Mahmon and Norsuzila Ya’acob “A Review on Classification of Satellite Image Using Artificial Neural Network (ANN)” IEEE 5th Control and System Graduate Research Colloquium 2014
7. Hasbi Ash Shiddieqy, Farkhad Ihsan Hariadi, Trio Adiono “Implementation of Deep-Learning based Image Classification on Single Board Computer”, International Symposium on Electronics and Smart Devices (ISESD), ISBN 978-1-5386-2779-2, pp. 133-137, 2017
8. Laila Ma’rifatul Azizah, Sitti Fadillah Umayah, Slamet Riyadi, Cahya Damarjati, Nafi Ananda Utama “Deep Learning Implementation using Convolutional Neural Network in Mangosteen Surface Defect Detection”, ICCSCE, ISBN 978-1-5386-3898-9, pp. 242-246, 2017
9. Sameer Khan and Suet-Peng Yongy “A Deep Learning Architecture for Classifying Medical Images of Anatomy Object” Proceedings of APSIPA Annual Summit and Conference 2017