International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 1 Jan. 2023, pp: 829-841 www.ijaem.net ISSN: 2395-5252

IoT and AI for Nature Conservation: Nyungwe forestry management and real time monitoring system.

Leopord Uwamahoro[1], Isaac Sikubwabo[2],Dr.Anselme Ndikumana, PhD[3],Pierre Clement Cyemezo[4], Jean deLa Paix Muragijimimana[5].

[1,5]Assistant Lecturer at Rwanda Polytechnic Integrated Regional College Kitabi/RP-IPRC KITABI,
[2]Snr Registration Assistant at UNHCR Innovation
,[3]Synchromedia Laboratory, École de TechnologieSupérieure, Université du Québec, Montreal, QC H3C 1K3, Canada
,[4]Assistant Lecturer at Kepler University .

Date of Submission: 17-01-2023 Date of Acceptance: 27-01-2023

ABSTRACT
Our research is a junction of two areas. Artificial intelligence with deep learning applied to image and audio recognition, and Internet things with cloud based.It investigates the methods of proving the support on forest management and security insurance in natural forests which are most the time used by rebellions for insecurity preparation in most of African countries and most affected by human activities. It will focus on real time data collection from big forests, their transmission and their classification and predictive analysis for finding the insight and proving alerts to the concerned people.The idea is real time monitoring of natural forests though IoT collected data and their classification using deep learning technology. We first of all identified the architecture to use while building our intelligent system. The question resides on the difference in classification accuracy, classification correctness, and run time for a specific number of inputs between AlexNet, VGG, ResNet. This research presents a comparative study among the stated three most popular image classification algorithms using a convolution neural network. It presents the results and interpretations of comparison when a classifier is applying the three different algorithms. In the results, the models for AlexNet, VGG, ResNet algorithms were provided with accuracy greater than 90% for both. The computing time was ranging between 0.05 seconds 0.875 seconds. The results showed that the AlexNet is the best algorithms with a good accuracy and very smallest time cost and it will be used for next step of Nyungwe forest real time monitoring system.

Keywords: Artificial intelligence, Convolution neural network, image recognition.

I. INTRODUCTION

Natural forest is very crucial and invaluable ecological, economical and climate dependency resources. The management and conservation of these bequeathed resources should be regular and based on accurate, reliable data and fast insight retrieval for quickly safeguarding the resource its self and alerting about the unauthorized use of it. Natural forests are mostly affected by human behavior and mostly used for rebellions organization for security destabilizationof neighboring country. Many examples can be given in almost of countries of African countries and we can mention example of Nyungwe forest which was used by FLN rebellion in 2018 and killed people in Nyabimata sector, Nyaruguru, District, south province in Rwanda .A real time monitoring of ongoing activities in natural forests would allow forests conservation and unauthorized groups of people identification and localization. Recently My research interest was directed to deep learning technology, by using recurrent neural (Long Short Term Memory) my research was limited find the hidden pattern time series data by developing predictive models to fight against the climate change.The further step was evaluating image classification algorithms and developing the classification models for image recognition using convolution networks and simple monitoring of motion detection basing on IoT to All these researches are very trending and this research will combine these technologies to provide an extent of controlling natural forest and their conservation as well.The real time monitoring will provide the way and on time monitoring activities that affects natural resources like, deforestation, firing as well as improving the security in natural forests by monitoring all the groups lactated in the forest. It will save the life of people and reduces the insecurity especially in the African countries that are handicapped by rebellions with different camps in forests.

Recently many services including banking, self-driving cars, and user identification are adopting the application of image detection and classification for automaticity and security purpose. The application can be extended and reach to community services to grant the security of community and contribution to the conservation of the nature.The misuse of natural forests has a great impact on live people, biodiversity and on the economy in general.How many strategies like deploying solders in natural forest (Nyungwe forest case), and others people to ensure the security of the forests, the real time monitoring for the forests is needed.The purpose of this research is provide a mean using Artificial intelligent deep learning model based of processing Real time Images, audios from Internet of things to facilitate the real time monitoring of natural forest. The image recognition should be done with high accuracy and within the affair instant of time. However, there are various classification algorithms and models such as AlexNet, VGG, ResNet. It is necessary to choose the right algorithms and for image detection and classification, which is faster and more accurate. As the Idea indeed we need to first of all carry out a carrying out a comparison of the three most popular image architectures: AlexNet, VGG, ResNet(Tricks,2020) and get the best performer for image detection, classification and fast algorithms for the development of the intelligent systems.The main objective of this research is to assess the existing image classification and detection algorithms for the purpose of determining the best performing algorithm for a best intelligent system.We access the existing algorithms for image classification and detection, We implement the top existing algorithms for image classification and detection and evaluate the performance of top existing algorithms for image classification and detection that will be later used of building an intelligent system for real time monitoring.

1.1 The evolution of deep learning

Artificial intelligence is trending and has various applications in problem-solving. The applications of artificial intelligence include speech recognition and image recognition. The artificial intelligence can identify the hidden patterns and helps the human being to obtain new insights from data. In the beginning, the computers were designed to perform simple and complex calculations, and their architecture permitted them for the storage of not only data but also the instructions being used by the processor to manipulate those data. This had grown up to where the computer can process data according to a structural model of the scenario in the real-world, expressible in mathematical expressions. At that time, the computer did not learn but was merely following and executing instructions. The next step was to put in place a set of instructions that would allow the computer to learn from experience, to extract its own rules from a large amount of data by discovering the hidden patterns and use those rules for classification or making the prediction. This was the starting point of machine learning which finally led to the field that is collectively defined as artificial intelligence (AI). Another major development came with the implementation of algorithms that were loosely modeled basing on neurons of the brain architecture, with numerous interconnecting nodes sharing weighted inputs among them, organized in computational layers. This is called deep meaning. Artificial intelligence already revolutionized many aspects of modern life like self-driving cars and biomedical research and authentication application.

1.2 Theoretical Perspectives

The theories of deep learning and Artificial Intelligence are discussed in this section. Due to the limited time, only the theories related to deep learning, neural network, and to convolution networks and convolution neural network were discussed. This is part of the methodology used in this research.

Machine learning is the art of the study of algorithms that can be learnt through example and experience (Brownlee, 2019). Machine learning can be classified into two wide learning tasks: Supervised and Unsupervised (Brownlee, A Tour of Machine Learning Algorithms, 2020).

- **Supervised vs unsupervised learning**

Supervised learning is when you have the input data and output data.The remaining task is the mapping function between input and output. It is called supervised as the teacher is there to adjust

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 1 Jan. 2023, pp: 829-841 www.ijaem.net ISSN: 2395-5252

until the desired output is acquired. When you have only the input without output, it can be called unsupervised. The purpose of unsupervised learning is to learn more about the data structure (Srivastava, 2017).

- **Machine learning**

Machine learning is one sub-set of Artificial Intelligence, which seeks the automation of the discovery of regularities in data by using the computer algorithms and generalize those into new which are similar data. In general, machine learning seeks to make less pre-assumption than the traditional statistical method and requires greater data. Machine learning techniques can be generally split into supervised learning, unsupervised learning (MD, 2019).

- **Traditional Machine Learning**

The term traditional means something that has been used for many years ago. There are four machine learning techniques that are considered traditional including clustering, classification, regression, and market basket analysis (P.E.Visscher, 2018).

1.2.1 Artificial Intelligence

Artificial intelligence, also called machine intelligence, is the intelligence that can be gotten by the machine in contrast to the intelligence of human beings or other animals. This can have a variety of applications in problem-solving like speech recognition, image recognition, and robotization (Mohamed, 2018).

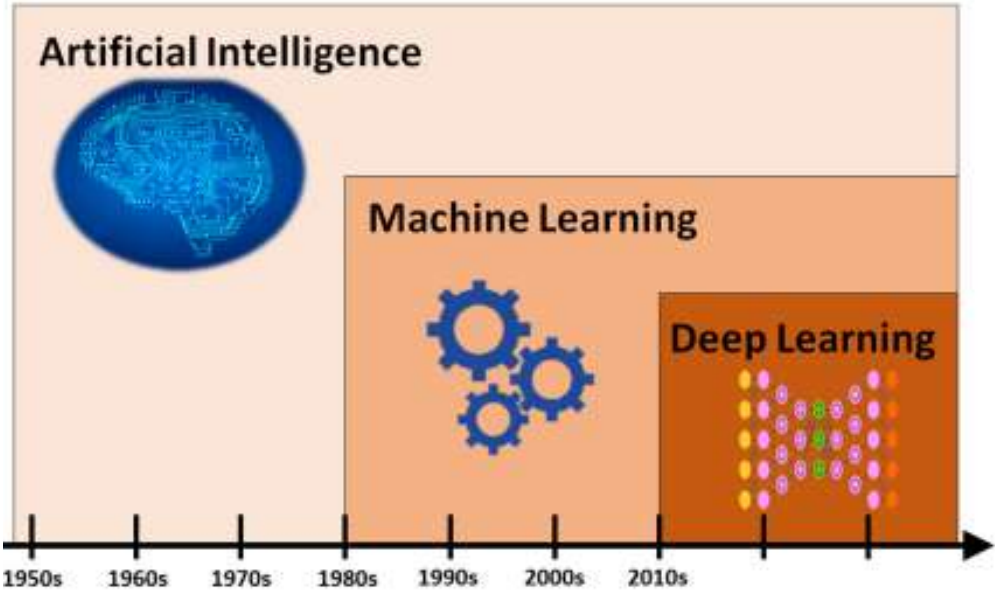

Fig 1The relationship between machine artificial intelligence, machine learning, and deep learning (MD, 2019)

As illustrated in the figure above, the artificial intelligence evolved from ninety fifties. Its technology got different names from time to time. Between ninety eighties to two hundred and teen, it was called machine learning. Now it is using deep learning technology.

1.2.1.1 Computer vision

Computer vision systems comprise such equipment as survey satellites, robotic navigation systems, smart scanners, and remote sensing systems. The principal aim of computer vision is to reconstruct and interpret natural scenes based on the content of images captured by digital cameras [190]. A natural scene is that part of the visual field that is captured either by human visual perception or by optical sensor arrays. (Peters, 2017)

1.2.1.2 Deep learning

Deep learning is a subset of machine learning and artificial intelligence, which renders the computation of multi-layer neural networks feasible (MD, 2019).

Deep learning is a subset of a subset of artificial intelligence, which comprises most logic and rule-based systems designed to solve problems. Within AI, you have machine learning, which uses a sequence of algorithms to go through data to

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 1 Jan. 2023, pp: 829-841 www.ijaem.net ISSN: 2395-5252

make and ameliorate the decision-making process. In addition, within machine learning, you finish deep learning, which can add up data using multiple layers of abstraction. (Algorithmia, 2016).

Deep learning is a class of machine learning that can perform better on data. It has various networks including recurrent neural network, feed-forward, and Convolution neural network. The application of deep learning has reached image processing by applying convolution neural networks (Amitha Mathew,Amudha Arul, 2021).

1.2.1.2.1 Artificial Neural Network

According to (Robert,2018) ANN is a computing system composed of several simple, tightly interconnected processing units, which process data by their dynamic state response to external inputs. The neural network is designed to the belief that it will imitate the human brain. The Artificial Neural Network is composed of the nodes which are the same as the neurons of a human being.
The following figure shows the shows a simple artificial neural network.

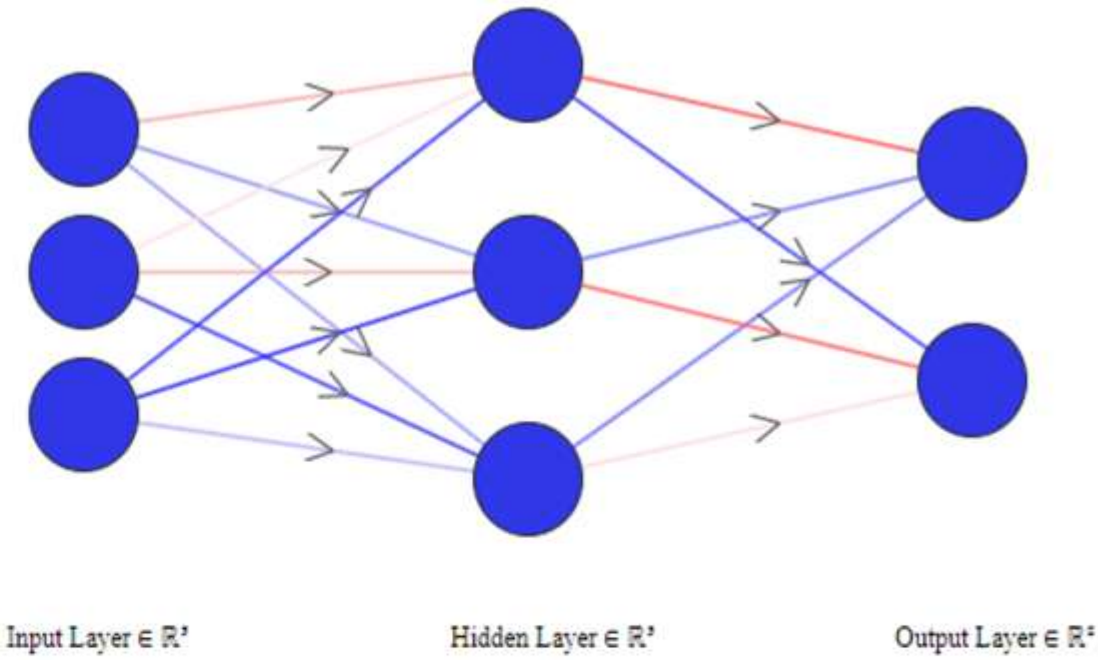

Fig 2 Simple Artificial Neural Network (source: Author)

The node is connected to the other nodes. The node operates on data and the output is passed to the other nodes. The output of the node is called activation or node value. We can distinguish two types of neural networks namely feed-forward and recurrent neural networks. For the feed-forward neural network, the data is unidirectional and the output is independent of the previous output. In contrast to recurrent neural networks, the current out dependent on the previous output by applying backpropagation (Brownlee, 2020).

1.2.2 Perceptron

A perceptron is a neural network unit (an artificial neuron) that does some computations to detect features or hidden patterns in the input data (Davis, 2020). Perceptions are subdivided into Single layer and Multilayer. Single-layer Perceptron can learn only single linearly separable patterns. Multilayer Perceptions or feedforward neural networks are composed of two or more layers and it has greater processing power.
The Perception algorithm learns the weights for the input signals in order to extract a linear decision. This helps to distinguish between the two linearly separable classes of one and negative one (Davis, 2020).

1.2.3 Perceptron function

Perception is a function that matches its input “x,” which is multiplied with the learned weight coefficient and the output value(x)”is generated.

$$P(x)=\begin{Bmatrix} 1 \text{ if} w.a+b>0 \\ 0 \text{ otherwise} \end{Bmatrix}$$

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 1 Jan. 2023, pp: 829-841 www.ijaem.net ISSN: 2395-5252

“w” = weights" = bias and "a" = vector of input values. The output can be 1 changed to 1and -1 according to the activation function used. The activation function of perceptron can be sign function, step function, and sigmoid function (Davis, 2020).

All neural network uses the principle of the perceptron. When you have many perceptron connected, the neural network is called a multilayer perceptron also known as a feed-forward neural network. This occurs when the current output does not depend on the previous outputs. When the current output depends on the previous output, the network is called a recurrent neural network. For solving the vanishing gradient problem, other recurrent neural networks like convolution neural network (CNN) and Long-short Term Memory Recurrent Neural network are used for different purposes. In this research, the focus was only on convolution neural network which is part of the methodology used in the research.

1.2.1.2 Neural Networks

The neural network is a structure used for complex computational tasks. It is composed of artificial neurons imitating the neurons of the brain of a human being. Those neurons are also called nodes and they are staked one to next each other to form the network. In general, the network of the nodes is staked into three layers namely: the input layer, hidden layers, and output layers. The input layer accepts the input(X_n)multiplies them with weight(W_n) and add the bias, the output is passed to the transfer matrix then in activation function and the output is produced as shown in the following figure (Biswal, 2020).

Motivation

The purpose of my research is wide, by using the intelligence system we can be able to monitor all activities and natural forest, monitor all natural resources and biodiversity in big forests. Accessing the image recognition,detection and classification architecture is one step of big idea on developing a real time intelligence system that will process the IoT based real-time collected data.The projects will reduce misuse of natural resources especially Nyungwe forest in Rwanda.

II. MATERIALS AND METHODS

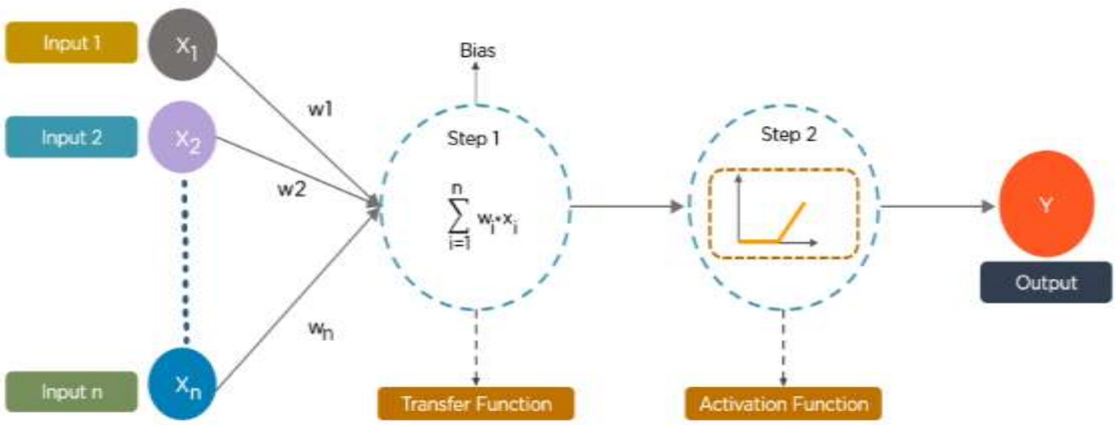

Fig 3.Illustration of how neural network works (Biswal, 2020)

The transfer function translates the input signal into the output signal. It applies one of the following functions: used, Unit step (threshold), sigmoid, piecewise linear, and Gaussian. Unlike the transfer function which makes a transformation, the activation function checks if input meets a certain threshold and generates one to activate a certain cell or zero to set off the cell (Klaus, 2019).Activation function use one of the following functions:Sigmoid function,function,tanh, nReLu function, sofmax function.Their formulas are shown respectively as follow.

$$s(a) = \frac{1}{1 + e^{a}}$$

$$t(a) = \frac{1 - e^{-2a}}{1 + e^{-2a}}$$

$$r(a) = \ln(1 + e^{a})$$

The decision taken by the above equations are defined as following:

$$s(a) = \begin{cases} a(a \geq 0) \\ b.(e^{x} - 1)(otherwise) \end{cases}$$

$$r(a) = \begin{cases} a(a \geq 0) \\ b.(e^{x} - 1)(otherwise) \end{cases}$$

$$t(a) = \begin{cases} \frac{1 - e^{-2a}}{1 + e^{-2a}}(a < 0) \\ \max(0, a)(otherwise) \end{cases}$$

Where s,r,t represent sigmoid function,Relu function and tanh function respectively is bias and a is the vector of input value.

When working with a conventional CNN for image classification, the output layer has N neurons or nodes, where N is the number of image classes you want to classify. You want each output neuron to represent the probability that you have observed each image class. The sigmoidfunction is fine for representing a probability. Its domain of definition is every one real numbers, but its variety is 0 to 1 (Brownlee, 2020).

For network layers that are not output layers, you could also employ the sigmoid. In hypothesis, any non-linear transfer function will work in the internal layers of a neural network. Nevertheless, there are sensible reasons not to employ the sigmoid. Some of those reasons are (Chris, 2020):

1. Sigmoid requires a reasonable amount of computation.
2. The slope of the sigmoid function is very thin when the input is far from zero, which slows gradient descent learning downward.
3. Contemporary neural networks have a lot of layers, and if you have more than a few layers in a neural network by means of sigmoid functions between them, it isfairly possible to finish up with a zero-learning rate.

The ReLU function solves many of sigmoid's troubles. It is uncomplicated and quick to compute. at whatever time the input is positive, ReLU has a gradient of -1, which provides a physically powerful gradient to descend. ReLU is not restricted to the range 0-1, so if you used your output layer, it would not be certain to be able to symbolize a probability (Davis, 2020).

2.2.1.2 Convolution Neural Networks (CNNs)

CNN's, also called ConvNets, comprise multiple layers and are basically used for image processing and object detection. Yann LeCun developed the first CNN in 1988 when it was known as LeNet. It was used for detecting characters like ZIP codes and digits (Mohamed, 2018). CNN is a multi-layered neural network with a unique architecture designed to extract increasingly complex features of the data at each layer to determine the output. It is widely applied in image recognition.

The architecture of Convolution Neural Networks

A CNN is composed of stacking of multiple building blocks: convolution layers, pooling layers (e.g., max-pooling), and fully connected (FC) layers. A model's performance under particular kernels and weights is calculated with a loss function through forward propagation on a training dataset, and learnable parameters, i.e., kernels and weights, are updated in line to the loss value through back propagation with gradient descent optimization algorithm. ReLU, rectified linear unit (Mohamed, 2018).

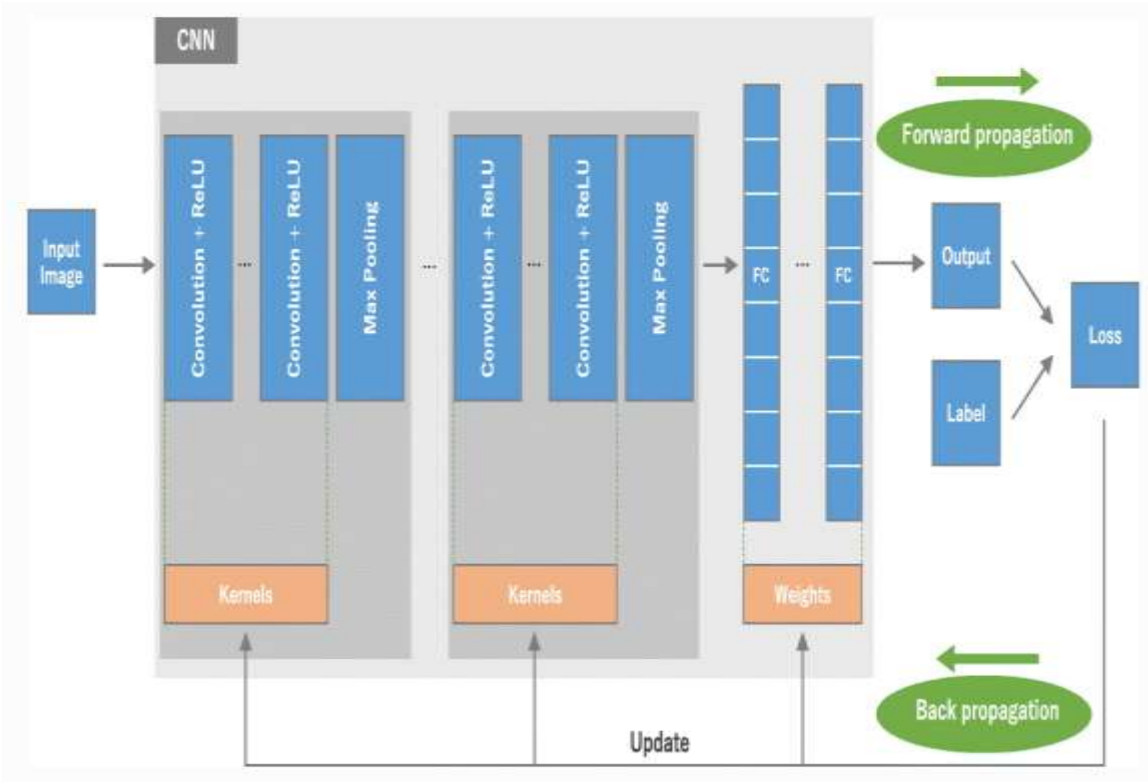

Fig 4 Layers in convolution neural network architecture (Rikya,2018)

Layers in convolution neural network

The convolution layer in a convolution neural network is of the convolutions and ReLufunction. MaxPooling and fully connected.

Convolution layer

The convolution layers in the convolution network are a central layer that makes it to be called so. This layer performs an operation called convolution. In its context, the convolution operation is the multiplication of the inputs with the weights. The multiplication is done by a two-dimensional array of the input matrix and a two-dimension matrix of the weight called filter or kernel (Brownlee, How Do Convolutional Layers Work in Deep Learning Neural Networks?, 2020).

Max pooling layer

The max pooling layer is added to reduce the dimensionality of the image by reducing the number of pixels in the width and height of the images of the output from the convolution layer. Creating ConvNets often goes hand in hand with pooling layers. More specifically, we most of the time see additional layers like max pooling, average pooling, and global pooling (Chris, 2020).

Full connection layer

The Role of a fully connected layer is to consider the results of the convolution/pooling process and use them in classifying the image into a label (Team, 2018)

In digital images, pixel values are all stored in a two-dimensional (2D) grid, and a miniature grid of parameters known as the kernel, an optimizer attribute extractor, is applied at each and every image position, which renders CNNs very much efficient for image processing, since a feature may take place anywhere in the image. As one layer feeds its output into the next layer, extracted features may hierarchically and gradually become more complex. The method of optimizing parameters such as kernels is called training, which is performed for the purpose to minimize the difference between outputs and ground truth labels through an optimization practice called back propagation and gradient descent (Peters, 2017).

III. EXPERIMENTATION AND RESULTS

This chapter consists of the processes and results of all experiments related to the comparison of VGG,alextnet and resnet.The pet images were used to differentiate the algorithms using different models.

Experimental environment

In this section, we started by discussing the experimental part of the research. First, we are

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 1 Jan. 2023, pp: 829-841 www.ijaem.net ISSN: 2395-5252

discussing the environmental characteristics. Then We describe the selected tools, their parameters, and the selected datasets. Finally, we will discuss processing and optimization methods in the implementation of the methods within the implementation environment.

The main tasks for the experimental part were to implement the top existing image classification algorithm by using a convolution neural network, to test how well a model of these algorithms trained on a certain dataset can generally perform accurately and with less associated time cost. Major problems noticed during the experimentation of deep learning methods like convolution neural network is computing resources usage like memory and CPU utilization by the training process. As this experience was not expected, the memory runs out during the process was annoying and rendered the delay in implementation. During the training, we used to switch on and off my personal computer to ease to free the memory and CPU from running processes.

To all the classification model, all experiments were conducted on a machine with Intel(R) Celeron(R)

i7-CPU N2840 @2.16GHz. The installed memory is 4.00GB (3.86BGUSABLE) IN WINDOW 10PROFESSIONAL 64 BITS. The model was using by taking memory utilization at 72% to 98% for the training. The models were implemented using anaconda navigator Python3.7.9,PyCharm 4.1 The other alternative tool to python is MATLAB.The python was selected due to its free package and extensible libraries. The torch vision from PyTorch was used for the transformation and manipulation of the model architecture. The CUDA which is parallel computing were set. For model optimization, we used the torch for calibrating the losses between 0 and 1. To load the image, we chose to combine both PIL and Pillow in my working environment to facilitate both simple and complex image processing functionality. The models accepted the input of the images in form of NumPy array. To make my image processing possible, we adopted cud toolkit in the environment. Due to the limitation in memory usage and slowness, we chose to use both Keras and Tensor flow backend. The back propagation was set to Rectilinear Rectifier Unit. During a

training rectilinear Rectifier Unit cause weights update in a way that a neuron will never be active,unassisted on the input’s values, and so that neuron will be unutilized. If this occurs, then the gradient flowing through the unit will forever be zero from that point on and then jumped to check the following neuron. That reduces the vanishing gradient as the neuron will be jumped and used when the loss is reefed to the network.

3.1The experimental setup

To reach the performance and accuracy, we set different environment setups to help in constructing and running my model. These setups include different hyper parameters .The vgg model is made of three layers with different blocks. The input layer is composed of convolution layers with the number of pixels in width which equals to the number of pixels in height. This convolution layer is in two dimensions with the form of (224 x224) pixels. The hidden layers are composed of four blocks. Each block of hidden has a two-dimension convolution layer staked with a max-pooling layer. The max-pooling layer will produce an output less than the output of the convolution layer. Like the convolution is, the max-pooling layer is a two-dimension shape with half of the pixels of convolution layers in both width and height. The formalized was set to 0.456x0.0456 and loss of std[0.24x0.25]. To create the output of the in one single of one-dimensional array I the flatten layer. And the out layer is a dense layer that produces a consolidated output. In our training, we have fixed the optimizer to Adam and the learning rate 0.5. The AlexNet model is made of three layers with different blocks. The input layer is composed of convolution layers with the number of pixels in width which equals to the number of pixels in height. This convolution layer is in two dimensions with the form of (227 x227) pixels. The hidden layers are composed of three blocks. Each block of hidden is having a two-dimension convolution layer staked with a max-pooling layer. The max-pooling layer will produce an output which is less than the output of the convolution layer. Like the convolution is the max-pooling layer is the two-dimension shape with half of the pixels of convolution layers in both width and height. The normalize was set to 0.456x0.0456 and loss of std[0.27x0.27]. To create the out of the network in one and single of a one-dimensional array, the flatten layer was used. And the output layer is a dense layer that produces a consolidated output image. In my training, the optimizer was fixed to Adam and learning rate 0.5. This configuration is almost the same with vggnet though the main difference resides on a number of blocks in the hidden layer. Because of the small number of layers, the learning rate is set to 0.1

The Resnet model is composed of five blocks. The input layer is SoftMax layer

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 1 Jan. 2023, pp: 829-841 www.ijaem.net ISSN: 2395-5252

concatenated to full connection of 1024 full connections between the neurons. The all max-pool layers can be counted from one to 7 with a form of [1X512] pixel, in contrast with the two firs model, the number of pixels in height and number of pixels to the width is different. The output layer is a convolution layer without a dense layer. The normalize is set to 456x0.0456 and loss is determined by std [0.12x0.12]. The learning rate was fixed to 0.5.

3.2 Creating Image Labels

In this section, a dictionary was created that has key and value. This dictionary was for the purpose of helping when checking the accuracy of the model. After the training, each algorithm was tested with different images including dog images, not dog images .As the results, each of the inputs to the model, the type of image entered were displayed and classified to prove if the model has got a kind of artificial intelligence. One thousand labels were created. Dog names include both lower and uppercase letter.

3.3Indexing

To make calculation possible, the following indices helped in determining the results of classification

index 0 = Pet Image Label, index 1 = model Label (ex: English foxhound), index 2 = 0/1 where 1 = labels match , 0 = labels don't match index 3 = 0/1 where 1= Image Label is a dog, 0 = Pet Image Label isn't a dog .index 4 = 0/1 where 1= model Label is a dog, 0 = Model Label isn't a dog

When the image is entered, the model will check its list and characterize the inputted image. It will match zero when the entered image does not match any image and one when the image matches with one of the images. The index two indicates that none dog image is entered.

Results statistics

This section discusses how the statistics of the results were calculated while implementing the models. To calculate the results, we just created a dictionary with the statistic's name as key and value which is simply the value of the statistic name. The results were computed by the counter that displays the percentage of statistics. The counted was the number of dog images, number and not dog image, number of dog images correctly recognized by the trained model, number of not dot images correctly recognized by the model, number of dog images not recognized by the model, and number of not dog images classified as the dog. We considered the number of correctly recognized dog images Total Number of dog images .We could compute the percentage of correctly recognized “dog" images as the rate of correctly recognized dog images and Total Number of dog images times hundred. Percentage of correctly not recognized “dog" images as the rate of not Correctly recognized dog images and Total Number of dog images times hundred. Percentage of correctly recognized “not dog" images as the rate of Correctly recognized not dog images and Total Number of dog images times hundred. The percentages (and a total number of images) can be automatically generated from the counter; so, these values should be calculated after the counter has been calculated by iterating through the results statistics dictionary.

Timing

We have integrated the timing program into the implementation of models. The timing program allows comparing the time costs associated with using various algorithms to solve a problem using given computing resources. Our results of time format were basically looking like hh:mm:ss where:hh is two digit hours indicator,mm is two-digit minutes indicator,ss is two digit seconds indicator and 3600 seconds in an hour 60 seconds in a minute.

3.4Results

For each algorithm, the model performance was provided. The quality indicators like a true positive, false negative, true negative, and false negative were provided. The cross-validation methods were used to test the accuracy of the model. The classifier was initialized to randomly calculate the percentage of dog images that are correctly classified in a dog breed, dog images that are not correctly classified as a dog, The none dog image that is correctly identified, and none dog images identified. The elapsed time taken by the model was identified. To be more precised group of 40 images were uploaded and the average standard performance measure for each model of each algorithm was plotted. Quality indicators like True positive, True Negative, False positive, and false negative were simultaneously plotted.

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 1 Jan. 2023, pp: 829-841 www.ijaem.net ISSN: 2395-5252

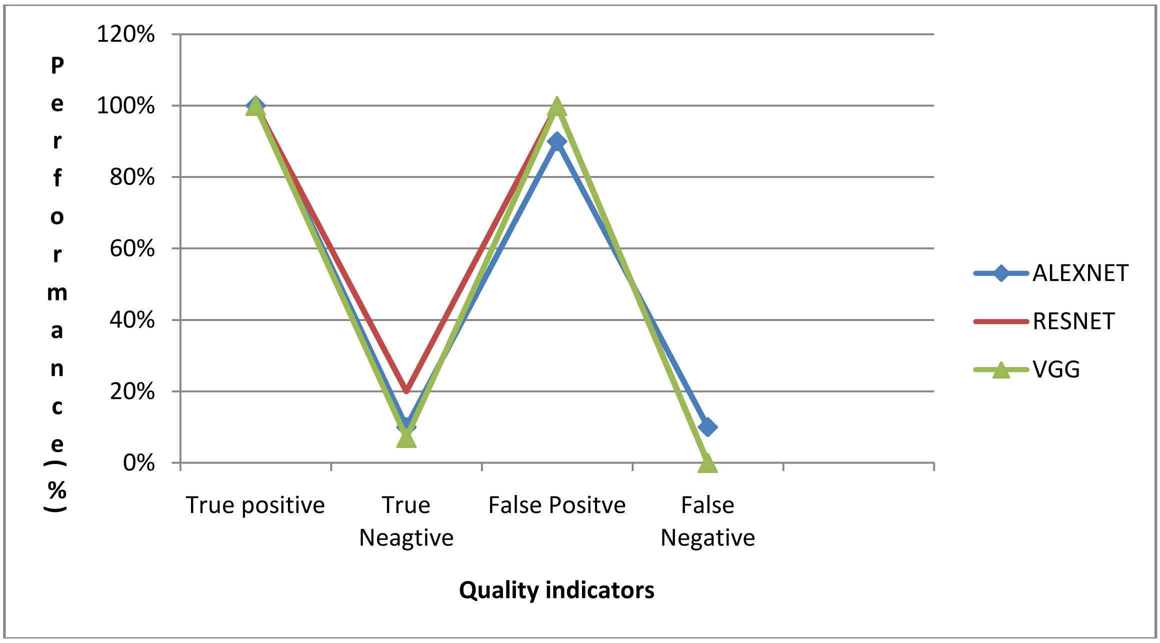

Fig 5.Quantitative indicators of classification models(Source: Research experimentations)

In the figure above, the presentation shows that VGG net models' low values in all negative values, which means the number of not images that were normally known and the models did not classify them as expected. Positive values mean that images were correctly classified by the model. For the positive, the VGG shows the small values. In contrast, the ResNet has high values for the negative and small value for the positive. In summary, the VGG Net has got good performance when we consider these four basic performance indicators.

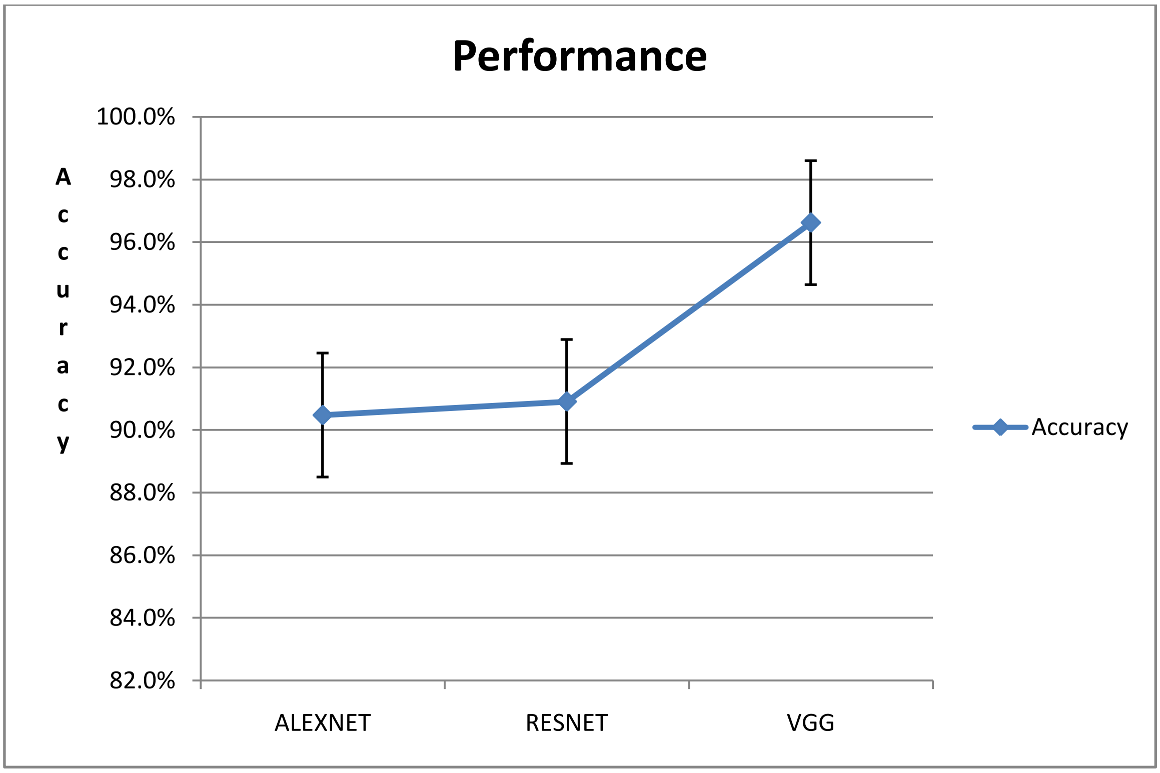

Fig 6.Performance of models in terms of accuracy (Source: Research experimentations)

International Journal of Advances in Engineering and Management (IJAEM)
Volume 5, Issue 1 Jan. 2023, pp: 829-841 www.ijaem.net ISSN: 2395-5252

According to the above figure, VGGNET has also a good performance in terms of accuracy. When the grows big, it means that the performance of the models is good. The Alex Net has a low accuracy which means low performance in classification. The best of all used models, the accuracy is above 90% which really shows improvement in image classification and detection or in computer vision in general.

From the above indicators, VGG Net has an improved performance. The remaining task resides on computational time used in completing a task when implementing VGG Net Algorithms.

The following graph shows the computational time when all the models are implemented. The input has been varied from one to forty images. The time used varies from 0.05 seconds to 35 seconds. As the number of images increases the time increases. The results shown are when forty various images are loaded once.

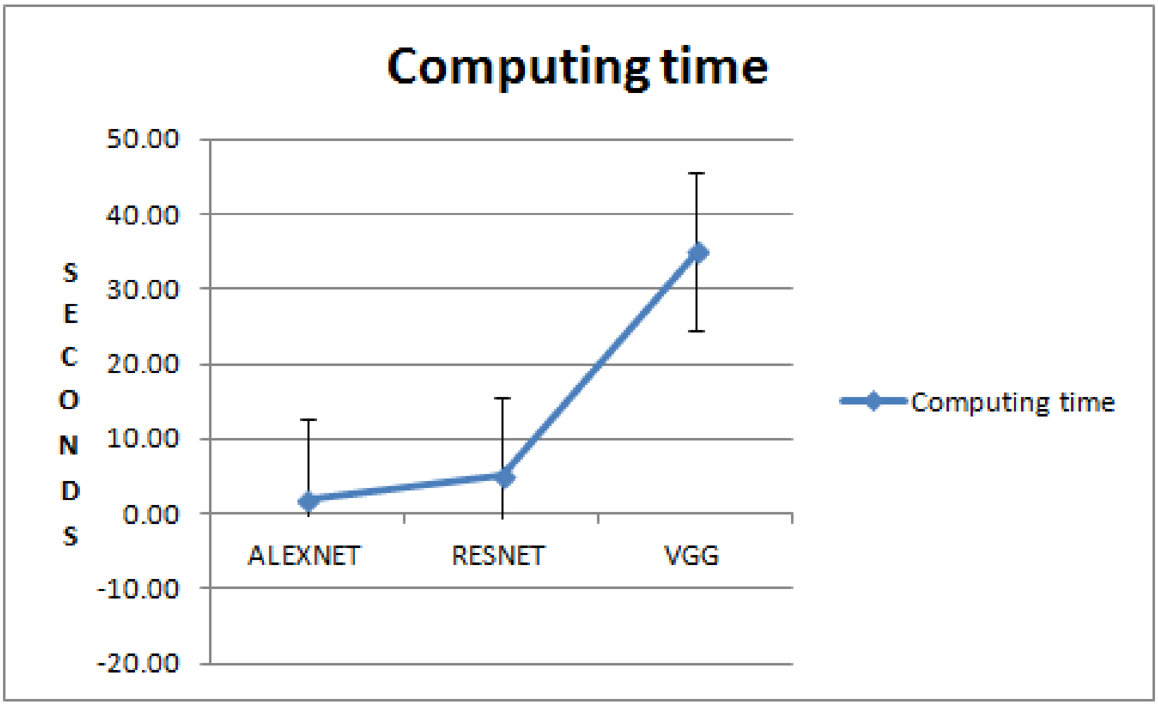

Fig 7.**Response time of CNN Models(Source: Research experimentations)**

In the figure above, the computing time for both models of each algorithm is plotted. Despite the performance of VGG Net is not good compared to the Alex Net and Resnet. When the images were loaded, VGG Net is taking around thirty-five seconds to know the forty images. It is also a bit long in responding disregarding it is the most accurate.

IV. CONCLUSION

To experiment with a convolution networks method in practice, a working Python implementation was created. It was found out that the most challenging part of implementing an image classification model training is the network itself due to the usage of memory, CPU, and power consumption. The convolution neural network is composed of input, output, and hidden layers disregarding which algorithm is being implemented. The implementation went through a different trial but it was recommended to use CUDA in implementing parallelism of back and forward propagation as the solution for the memory usage and rescue of associated time cost.

This project showed that image classification is possible in accuracy and faster manner. The Tensor flow backend was used to achieve the best results. The performance of the top existing image classification algorithms was assessed. The different strategies used worked generally and the outcome was as expected. Our research was about image classification algorithms based on convolution neural network networks. The models were built, the best algorithm reached an accuracy of 96% and the last was 90%. The response time of all the models was computed and the best was about2 seconds. According to the observation as the accuracy increase, the model takes too long to respond. These are due to the increase in the number of computation resources use like increased layers in CNN models, weight, and parameters.

Depending on the application and purpose of the usage of the algorithm, we can recommend the best algorithm in image recognition and computer vision. For the security and authentication where we need more accuracy, we recommend to use VGG Net. But the other systems like attendance or registration where we do not need much security, we can use AlexNet as it very fast and its accuracy is fair. Generally, we can recommend AlexNet as the best performer algorithm to use for Nyungwe forest real time monitoring system.

REFERENCES

[1]. Alom,Z.(2018). The History Began from AlexNet: A Comprehensive Survey on Deep Learning Approaches, research gate

[2]. Vilardi, A.L. (2016). very deep convolutional neural networks for face identification, Barcelona UniversitatPolitècnica de Catalunya, Spain.

[3]. Das, S. (2017).CNN Architectures: LeNet, AlexNet, VGG, GoogLeNet, ResNet [web log post],retrieved from medium.com)

[4]. Khan,A.(2019). A Survey of the Recent Architectures of Deep Convolutional Neural Networks, Pattern Recognition Lab, DCIS, PIEAS, Nilore, Islamabad 45650, Pakistan

[5]. Zhang, A.,Lipton, Z. C. , Li, M.,&Smola, A.J. (2020). Dive into Deep Learning,Github

[6]. tricks, CV .(2020).ResNet, AlexNet, VGGNet, Inception: Understanding various artecture of convolutional neural networks[web log post] retrieved from cv-tricks.com

[7]. Mwai, C. (2019).Rwandan firm debuts on global Artificial Intelligence scene[web log post], retrieved from newtimes.co.rw

[8]. Chan, T. (2019).African App Launchpad initiative 2019 for young African startups[web log post], retrieved from adalidda.asia.

[9]. Bizimungu, J.(2019).Sophia, the famous humanoid robot has a message for us[web log post], retrieved from newtimes.co.rw

[10]. Kasoni, R. (2019).Image detection, recognition and image classification with machine learning[web log post], retrieved from medium.com

[11]. Data science(2018).Difference Between Model and Algorithm[web log post], retrieved from financetrain.com

[12]. Chadalawada, S.K. (2020).Real Time Object Detection and Recognition Using Deep Learning Methods, Faculty of Computing, Blekinge Institute of Technology, 371 79 Karlskrona, Sweden

[13]. Castillo, J.F. (2019).Supervised person annotation in video sequences, UniversitatPolitècnica de Catalunya,Spain

[14]. Tejasri, N. (2018).Material Recognition Using Deep Learning Techniques, Asian Institute of Technology

[15]. James, L. (2018).The 4 Convolutional Neural Network Models That Can Classify Your Fashion Images, towardsdatascience

[16]. Shu, M. (2019).Deep learning for image classification on very small datasets using transfer learning, Iowa State University Ames, Iowa,Greece

[17]. Nampally, T. (2018).Material Recognition Using Deep Learning Techniques, Jawaharlal Nehru Technological University Hyderabad Telangana, India

[18]. Xin,.M & Wang,.Y(2019). Research on image classification model based on deep convolution neural network, EURASIP Journal on Image and Video Processing

[19]. Manjula,K.Vijayarekha, K.,&Vimaladevi, P. (2017).review on classification algorithms in image processing , International Journal of Innovative Trends in Engineering & Research

[20]. Christopher, T. (2019).An introduction to Convolutional Neural Networks, [web log post],retrieved from towardsdatascience.com

[21]. OSDS(2020), Using the CNN Architecture in Image Processing, retrieved from medium.com

[22]. Stanford Vision Lab(2016). About ImageNet,Stanford University, Princeton University

[23]. Xinhua. (2019, May 15). Humanoid robot Sophia addresses Africa technology summit in Rwanda. Xinhua.

[24]. Jessica, L. (2019, February 2). Stanford Dogs dataset. Retrieved October 2, 2020, from Kaggle website: https://www.kaggle.com/jessicali9530/stanford-dogs-dataset

[25]. Algorithmia. (2016). eep Learning AI: Why Deep Learning Matters and What's Next for Artificial Intelligence. AICPA SOC: AICPA SOC.
[26]. Amitha Mathew,Amudha Arul. (2021). eep Learning Techniques: An Overview. Research Gate, 9-54.
[27]. Artificial Intelligence Definition, Ethics and Standards. (n.d.).
[28]. Artificial Intelligence Definition, Ethics and Standards. (n.d.).
[29]. Biswal, A. (2020, October 14). simplilearn. Retrieved December 17, 2020, from simplilearn website: https://www.simplilearn.com/tutorials/deep-learning-tutorial/deep-learning-algorithm
[30]. Brownlee, J. (2019, Noember 11). Different Types of Learning in Machine Learning. Retrieved December 21, 2020, from machinelearningmastery: :http//machinelearningmastery.com
[31]. Brownlee, J. (2020, August 14). A Tour of Machine Learning Algorithms. Retrieved 2020 December, 2020, from machinelearningmastery: https://machinelearningmastery.com/
[32]. Brownlee, J. (2020, April 17). How Do Convolutional Layers Work in Deep Learning Neural Networks? Retrieved December 18, 2020, from machinelearningmastery: https://machinelearningmastery.com
[33]. Chris. (2020, January 30). What are Max Pooling, Average Pooling, Global Max Pooling and Global Average Pooling. Retrieved December 18, 2020, from machinecurve: www.machinecurve.com
[34]. Davis, A. (2020, Juanuary 10). Deep Learning with Keras and TensorFlow. Retrieved December 28, 2020, from simple learn: https://www.simplilearn.com/
[35]. Jessica, L. (2019, February 2). Stanford Dogs dataset. Retrieved October 2, 2020, from Kaggle website: https://www.kaggle.com/jessicali9530/stanford-dogs-dataset
[36]. Klaus Debes,Alexander Koenig, Horst-Michael Gross. (2019). Transfer Functions in Artificial Neural Networks . In T. F. Networks, Klaus Debes (pp. 2-13). ilimenau: Ilmenau,Germany .
[37]. MD, K. S. (2019). Application of mobile health, telemedicine. Research Gate, 52-63.
[38]. Mohamed, Z. (2018). Artificial Intelligence Definition, Ethics and Standards. Electronics laws, standars and practice, 18-71.
[39]. P.E.Visscher. (2018). Forecasting Time Series with. Utrecht: Scientific Computing Group Mathematical InstituteUtrecht University.
[40]. Peters, J. F. (2017). Basics leading to the machine. Reseach Gate, 1-85.
[41]. Rikiya Yamashita, Mizuho Nishio, Richard Kinh Gian Do & Kaori Togashi . (2018). Convolutional neural networks: an overview and application in radiology. springer open, 611–629.
[42]. SRIVASTAVA, P. (2017). Essentials of Deep Learning : Introduction to Long Short Term Memory. Analytics vidya, 1-10.
[43]. Team, S. (2018, August 18). Convolutional Neural Networks (CNN). Retrieved December 18, 2020, from www.superdatascience: www.superdatascience.com