**CHAPTER 1**

**INTRODUCTION**

* 1. **Introduction**

Coronavirus is a global pandemic caused by the severe acute respiratory syndrome (SARS-CoV-2) virus that has shaken the entire world. This contagious illness was eradicated in Wuhan, China, in December 2019. The World Health Organization (WHO) declared the pandemic a public health emergency on a global scale. The entire world jumped at the chance to figure this out. According to World Health Organization (WHO) reports, there have been 615,310,890 confirmed cases of COVID-19, with 6,524,568 deaths. As of September 27, 2022, 12,677,499,928 vaccine doses had been administered [1]. The COVID-19 pandemic drove almost all nations to go into lockdown. The clinical symptoms of severe COVID-19 contamination are bronchopneumonia, which causes fever, hack, dyspnea, and respiratory failure. That’s reason many people were compelled to stay at home and maintain social distance.

The rapid spread of COVID-19 puts healthcare systems under a lot of stress, but if a reliable screening tool for COVID-19 infections is developed, this spread might be greatly curtailed. The primary screening method for detecting COVID-19 cases is reverse transcriptase-polymerase chain reaction (RT-PCR) which can detect virus from respiratory specimens. Chest X-ray screening is an alternate way for detecting COVID-19. Chest X-ray (CXR) imaging is performed and examined by radiologists to check for visual signs associated with SARS-CoV-2 virus infection. It is a less time-consuming method than RT-PCR testing with a goal to approach more infected patients. Computer-aided diagnostic systems that can help radiologists analyze radiography images more efficiently and accurately to discover COVID-19 cases are highly desired.





**Fig. 1.1: Normal X-ray [2] Fig. 1.2: Covid-19 X-ray [2]**

These are X-ray images of a normal and Covid-19 infected person. The hazy gray areas on a chest x-ray indicate that some fluid is partially filling the air spaces inside the lungs.

**1.2 Overview**

The COVID-19 epidemic continues to have a terrible impact on the global population's health and well-being. A Successful screening of infected individuals is a critical step in the fight against COVID-19, with radiological imaging employing chest radiography being one of the primary screening modalities. The most recent data mining and machine learning techniques, such as the Convolutional Neural Network (CNN), are used in collaboration with X-ray images of the lungs for the accurate and speedy identification of the illness, assisting in alleviating the problem of testing kit scarcity. The purpose of this survey is to categorize and review the literature on computer analysis of chest images in a concise manner.

The aim of this research is to apply radiographic imaging to identify areas infected by the COVID-19 virus, along with applying convolutional neural network architecture. In the coming chapters, the proposed methodology and implementations will be discussed in details. The challenges of the Coronavirus pandemic will be discussed in this chapter. The motivations and objectives for this research will be discussed later.

**1.3 Challenges**

The fragmentation of health services, the concentration of medical technology and human resources in a few urban hospitals, the underfunding of primary healthcare and epidemiological surveillance, and the lack of coordination between different levels of care all pose challenges to national response action coordination in the majority of countries. It is a monumental challenge to maintain essential health services for women, kids, and teenagers while minimizing the epidemic's effects [3].

The usual incubation time is five days, however, it can range from two to fourteen days. The majority of adults and children infected with SARS-CoV-2 have flu syndrome (90%) and mild symptoms. Some persons, especially the elderly and those with cardiovascular or lung illness, diabetes, or hypertension, may develop severe diseases such as respiratory failure, multiple organ failure, and death.

The latest machine-learning techniques are excellent at efficiently diagnosing diseases. In order to classify the chest X-ray images as Normal and COVID-19 positive several challenges arise that must be handled carefully, such as:

* Shortage of data
* Unavailability of CNN model
* Unavailability of source code
* Uneven data distribution
* Datasets not publicly available
* Individual communities limit their resources.

**1.4 Motivations**

This virus has had a massive global impact. People all across the world have been wondering when this contagious sickness will take their life. This virus has killed about 6,524,568 individuals [1]. That is why it continues to motivate researchers to work harder to discover the best solution for eradicating this pandemic. Researchers are still dedicating their lives to finding a more effective strategy to combat this infection. Through several research projects, it has done brilliantly in the case of humanity.

A convolutional neural network is a form of artificial neural network that is used to extract and classify features. These computer-aided diagnostic methods have helped radiologists in classifying COVID-19-positive and COVID-19-negative patients. Convolutional neural networks are artificial neural networks that extract and categorize information. These computer-aided diagnostic tools have supported radiologists in identifying patients as COVID-19 positive or COVID-19 negative. Because chest X-ray imaging is considered basic equipment in most healthcare systems, it is available and accessible in many clinical sites. CXR imaging is more widely available than other types of screening.

**1.5 Objectives**

The main objective of this research is to develop a convolutional neural network for classifying chest X-ray images as COVID-19 positive or negative. There are other objectives as well-

* Implement CNN for image classification
* Learning how convolutional neural network works
* To increase accuracy than the existing one
* Analyze performance
* To apply a different model on the dataset
* Result comparison with other CNN models

**1.6 Thesis Organization**

The rest of our thesis work is organized as follows:

**Chapter 2 - Literature Review**

This is the

**Chapter 3 - Background/Methodology**

**Chapter 4 – Implementation**

**Chapter 5 – Result & Analysis**

**Chapter 6 - Conclusion**

**1.7 Conclusion**

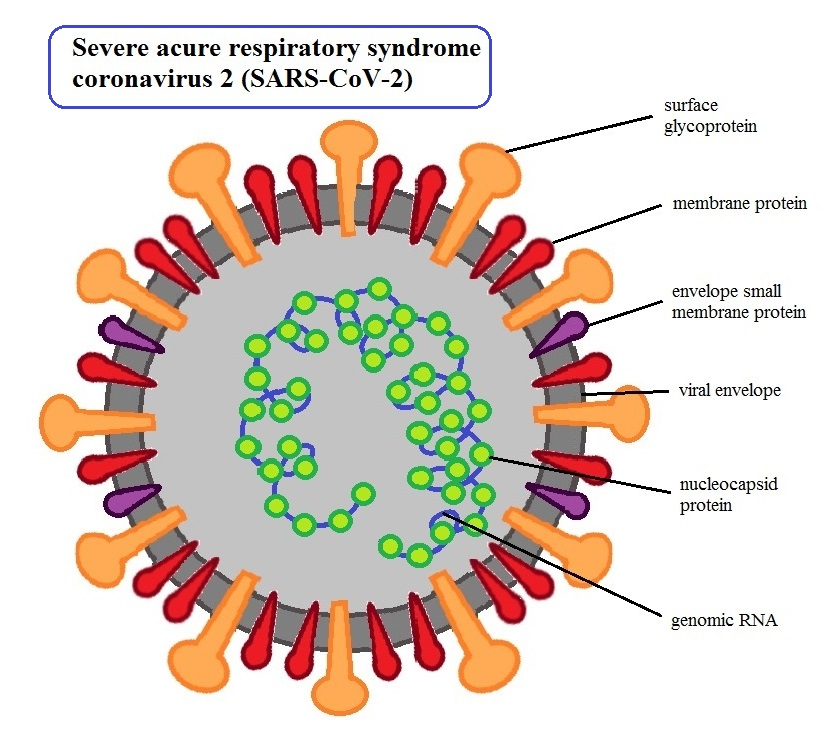
**CHAPTER 2**

**BACKGROUND STUDY & LITERATURE REVIEW**

This chapter discusses some previous works on COVID-19 detection using various machine learning algorithms. This chapter begins with a brief description on SARS-Cov-2 virus.

* 1. **SARS-CoV-2 virus**

SARS-CoV-2 is a virus of the coronavirus family. The SARSr-CoV species is a member of the genus ***[Betacoronavirus](https://en.wikipedia.org/wiki/Betacoronavirus" \o "Betacoronavirus)*** and of the subgenus ***Sarbecovirus*** (**SAR**S **Be**ta**co**rona**virus**)[6].The virus is assumed to transmit from person to person by droplets emitted when an infected person coughs, sneezes, or speaks. It can also be transferred by contacting a surface with the virus on it and then touching one's lips, nose, or eyes, however this is less likely [5]. Recent studies have discovered that SARS-CoV-2 may replicate effectively in human intestine organoids and epithelial cells. As a result, SARS-CoV-2 might spread through the digestive system [7].



**Fig 2.1: SARS-CoV-2**[8]

**2.2 Literature Review**

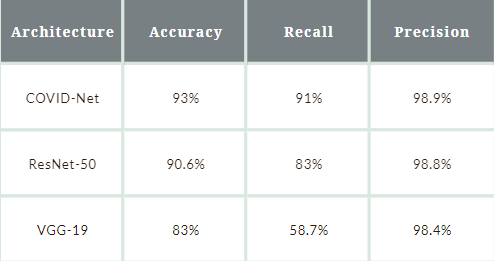
A paper titled "COVID-Net: A Tailored Deep Convolutional Neural Network Design for Detection of COVID-19 Cases from Chest X-Ray Images" authored by Linda Wang, Zhong Qiu Lin, and Alexander Wong. This paper was published in 2020 on November 11. This paper proposes an efficient deep convolutional neural network, to classify COVID-19 cases from chest X-rays. In this process, the identification and extraction of hazy gray areas are the most crucial stages. They performed the segmentation by applying a small dose of ionizing radiation to produce pictures of the inside of the chest and extracted the infected hazy gray areas. They compared their experimental result with other deep neural network architectures and demonstrated the efficiency of their model.

The steps they followed to perform the task are:

* Data pre-processing
* Data augmentation
* PEPX(projection expansion projection extension module)
* Train the model
* Result evaluation

**Contribution:**

**Table 2.1: Comparison with different models [9]**



**Limitation:**

* Training method could be more improved
* More data could be used

Next, a paper titled "DeepCOVID-XR: An Artificial Intelligence Algorithm to Detect COVID-19 on Chest Radiographs Trained and Tested on a Large U.S. Clinical Data Set" was authored by Ramsey M. Wehbe, Jiayue Sheng, Shinjan Dutta, Siyuan Chai, Amil Dravid, Semih Barutcu, Yunan Wu, Donald R. Cantrell, Nicholas Xiao, Bradley D. Allen, Gregory A. MacNealy, Hatice Savas, Rishi Agrawal, Nishant Parekh, Aggelos K. Katsaggelos [10]. This paper was implemented by a deep learning approach named DeepCOVID-XR. The authors compared their CNN model with thoracic radiologist experts' interpretations. DeepCOVID-XR is a CNN model that is an ensemble of six different CNN architectures. This model ensemble consists of DenseNet-121, ResNet-50, InceptionV3, Inception-ResNetV2, Xception, and EfficientNet-B2 [10].

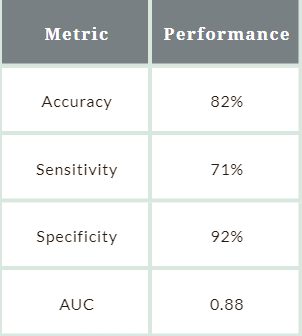
The CNNs in this ensemble were fine-tuned using transfer learning on our clinical training set of chest radiographs from the COVID-19 period after being pretrained on a sizable publically accessible data set of over 100000 chest radiographs from the National Institutes of Health. Hyperparameter optimization using the validation data set[10-11].

The steps they followed to perform the classification task are:

* Data collection
* Image pre-processing
* Applying CNN models
* Model ensembling
* Weighted average prediction

**Contribution:**

**Table 2.2: Performance analysis of CNN model**

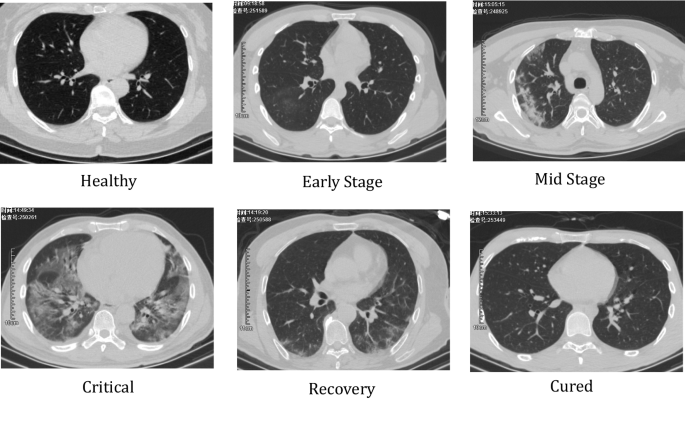


This research has acquired 82% of testing accuracy while having 71% of sensitivity and 92% of specificity.

**Limitation:**

* Limited sensitivity
* Implementation is not clarified
* Sampling error

The next research title is "COVID-VIT: Classification of Covid-19 from CT chest images based on vision transformer models" was authored by Xiaohong Gao, Yu Qian, Alice Gao[12]. In this research, they have implemented a vision transformer model based on attention models and DenseNet that is built upon conventional convolutional neural network (CNN). This was applied over a small dataset.



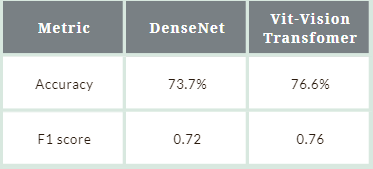
**Fig 2.2: Chest CT-images [13]**

In this study, flattened patches were projected linearly onto CT scans. Following their transmission through the transformer encoder, the multilayer perceptron head, and slice voting, the COVID and Non-COVID pictures were separated. These are the steps they followed to perform the classification are:

* Data collection
* Image pre-processing
* Augmentation
* Training
* Prediction result

**Contribution:**

**Table 2.3: Comparison with different models [12]**



Vision transformer acquires 76.6% of accuracy comparing with DenseNet.

**Limitation:**

* Less number of data
* Implementation is not discussed properly
* Too slow for clinical operation
* Training methodology could be more improved
  1. **Accuracy Comparison of Papers**

Table 2.4: Comparison for classification of COVID-19



**2.4 Conclusion**

In this chapter, some previous works performed in this field are discussed along with their

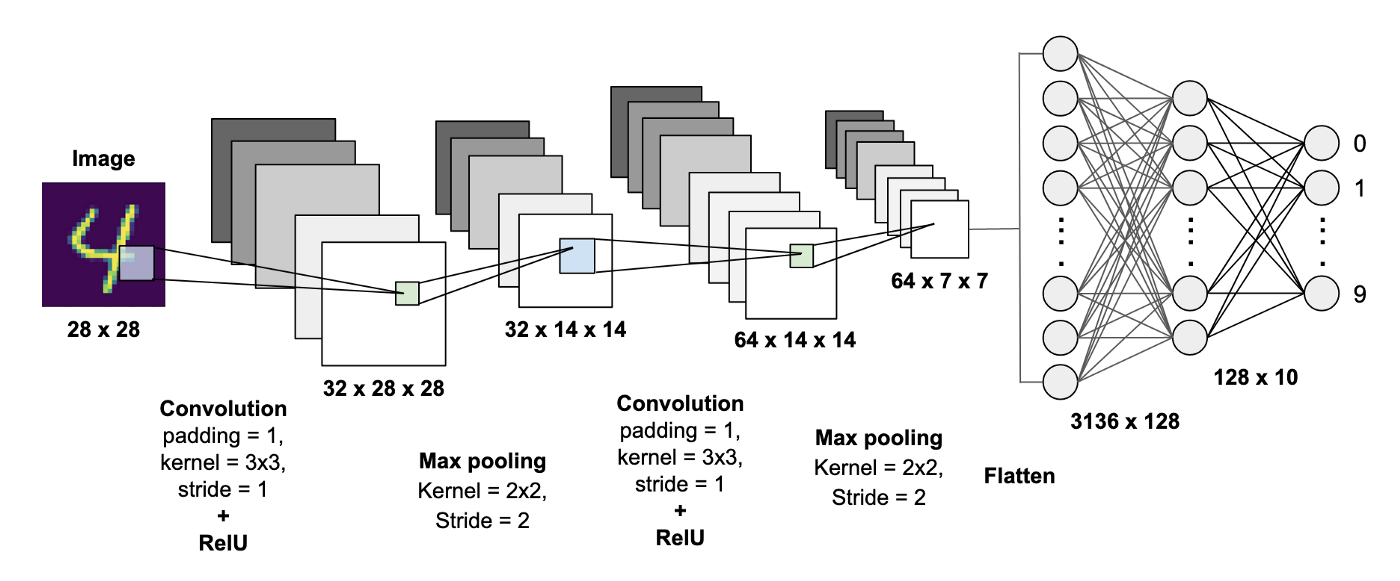
limitations and comparison among them. All of these contributions have had a significant and innovative effect on the area of computer-aided approaches.

**CHAPTER 3**

**THEORETICAL BACKGROUND**

This chapter contains an in-depth explanation of Convolutional Neural Networks. It covers an overview of the overall architecture of the CNN model, as well as its several layers with some specific parameters and how it works.

**3.1 Introduction**

****

**Fig. 4.1: A CNN model to classify handwritten digits [14]**

Convolutional Neural Network is a type of deep neural network that is used for computer vision or studying visual imagery. It is a specific type of deep neural network that is best suited for feature

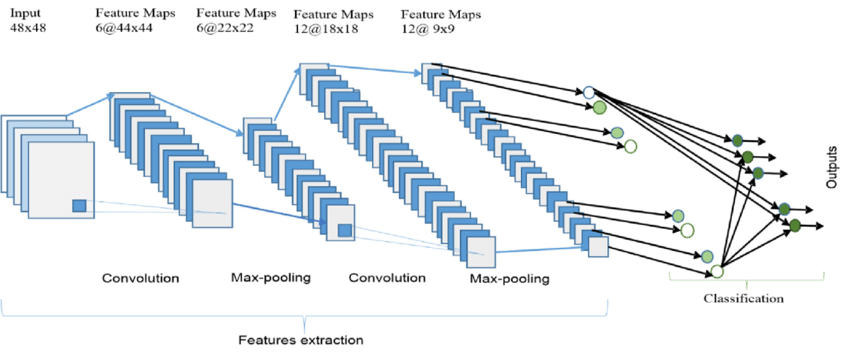
extraction. Convolutional neural network is one of the most widespread types of neural networks used to conduct image identification and classification tasks. CNNs are frequently employed in domains such as object identification, facial recognition, and so on.

**3.2 Convolutional Neural Network**

Convolutional neural networks and regular neural networks are quite similar to one another. Both have nodes that correspond to neurons in the brain that have biases and weights that can be learned. Each node processes the necessary actions after receiving the necessary information. While nodes in the same layer are not linked, every node is connected to every node in the layer below it. The output layer, which is the next completely linked layer, describes the final class scores in classification scenarios.

It is a feed-forward artificial neural network. The models are called “feed-forward” as the learning

flows in the forward direction through the model. Feed-forward neural networks are also known as multi-layer perceptrons (MLPs).Convolutional neural networks provide a more scalable method to image classification and object identification problems by utilizing linear algebra concepts, notably matrix multiplication, to find patterns within an image. However, they can be computationally intensive, necessitating the use of graphics processing units (GPUs) to train models.



**Fig. 3.2: Basic CNN architecture [16]**

**3.3 Basic CNN structure**

Regular neural networks and convolutional neural networks have many characteristics. They both have nodes that resemble neurons in the brain with biases and weights that can be learned. Each node receives input from other nodes and uses that information to carry out the required actions. While all nodes in the layer below are connected to one another, nodes in the same layer are not connected to one another. The fully-connected layer is known as the output layer.

For image classification, CNN models take an image as input, then perform the necessary

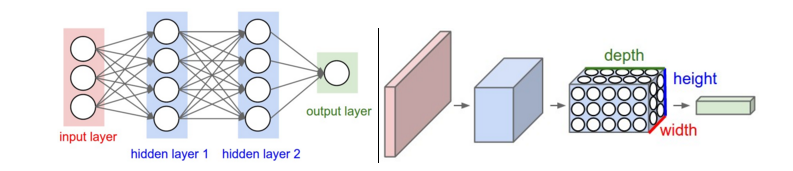
operations to extract features from it, and then use those features to classify the image as one of

the given categories. The models see the image differently from us. All they assume is the pixel

array which depends on the resolution of the images. Depending on the resolution of the image,

the model will concede h x w x d number of pixels where h represents height, w represents the

width and d means the dimension of the images [17].



**Fig. 3.3: Hidden CNN process [17]**

**3.4 Layers in CNN**

A simple Convolutional Neural Network model has several types of layers. Each of them performs a definite task and passes the output to the following layer. The various types of CNN layers are:

* Convolution
* Activation
* Pooling
* Normalization
* Fully Connected

**3.4.1 Convolution Layer**

The convolution layer is first for performing feature extraction from an image given as input. It conserves the correlation among the pixels by acquiring image features employing little squares of data supplied as input. Convolution is a mathematical procedure that uses two inputs, like a filter and an image matrix. The output of the operation proceeds into the feature map. Different kinds of operations like edge detection, sharpening, or blurring on an image can be performed by convolution layer using filters.



**Fig. 3.4: Convolution layer [18]**

Convolutional Layer employs a group of filters that develop new skills throughout training. These filters allow different kinds of characteristics to be retrieved from the input images. The filters are often represented as matrices of size (MxMx3), with a smaller dimension but a depth that is comparable to the input.

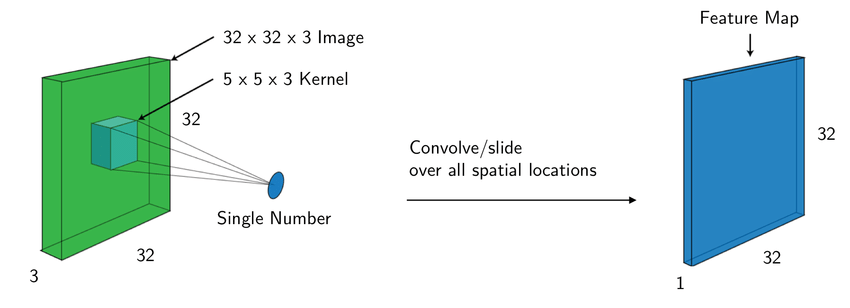


Fig. 3.5: Feature map generation using kernel [19]

The filter then convolves (slide) along with the height and width of the input plus, a dot product

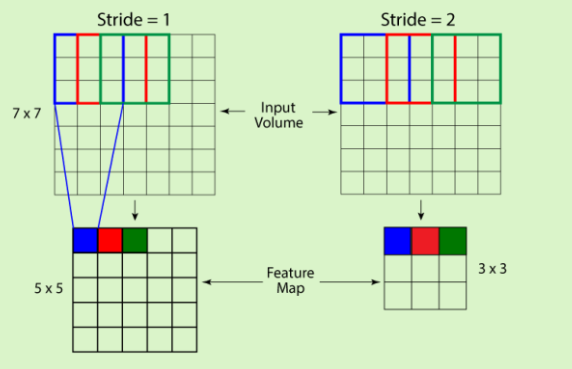
gets calculated to yield an activation map. Various filters that detect different features get

convolved over the input image. Also, a collection of activation maps is generated, which is

supplied to the following layer of the Convolutional Neural Network model.

**3.4.2 Strides**

Stride represents the number according to which the filter shifts over the input image matrix. When the value of the stride is 1, then the filter shifts by 1 pixel at a time. When it is 2, then the filter shifts by 2 pixels at once, and accordingly.



**Fig. 3.6: Stride in CNN [26]**

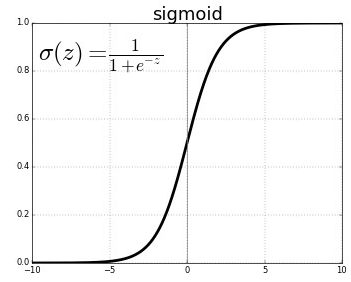
For avoiding overlapping among the cells, the stride size can be increased. Because of an

increased stride size, the filter will slide over with a larger interval.

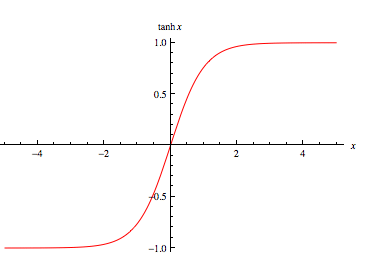
**3.4.3 Activation Function**

Activation functions are often implemented at the model's end or in the space between the networks. An activation function aids in determining whether a certain neuron will produce output or not. Only the ReLU activation function is covered here. ReLU is used more frequently than any other activation method worldwide.

**Sigmoid**

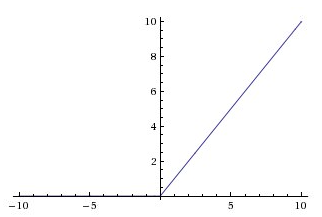


**Fig. 3.7: Sigmoid function [21]**

**tanh**

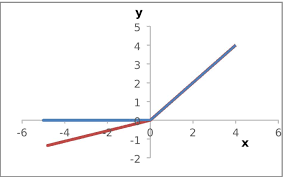


**Fig. 3.8: tanh function [21]**

**ReLU**



**Fig. 3.9: ReLu function [21]**

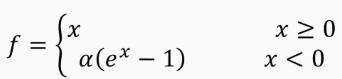
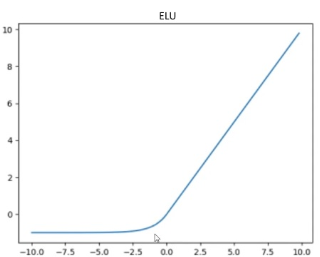
**Leaky ReLU**



**Fig. 3.10: Leaky ReLu function [22]**

**Maxout**

*h(x) = max (w1.x + b1, w2.x + b2, …, wn.x + bn****)* [21]**

**ELU**

**Fig. 3.11: ELU function [23]**

Here, we only discuss the ReLU activation function. Among all the other activation functions,

ReLU is used more vastly worldwide. One of the most prominent advantages ReLU holds is, it

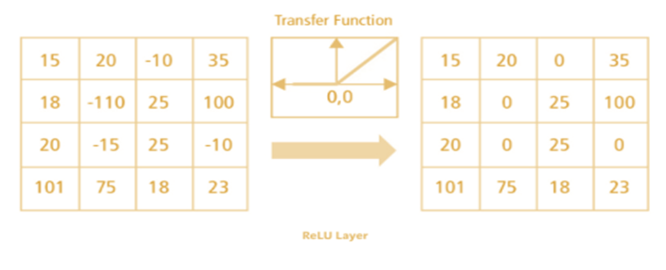
doesn't allow all neurons to be activated at a time. Equation of ReLU activation function is:

**ƒ(x) = max (0, x)**

The equation of the ReLU activation function shows that it transforms the negative values into zero. Because of which the neurons do not give output. This characteristic makes the computation very efficient due to the presence of a low number of neurons at a time. In practice, the ReLU

activation function focalizes six times quicker compared to the sigmoid and the tanh activation

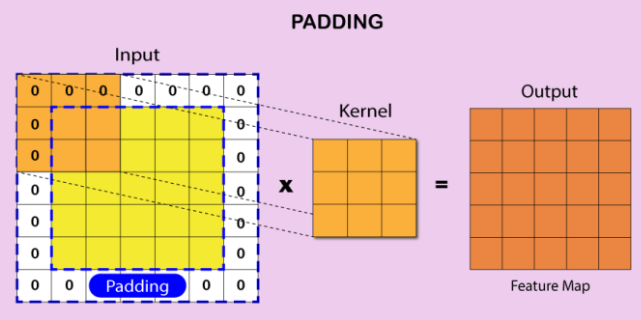
functions [24].



**Fig. 3.12: ReLU activation operation[25]**

**3.4.4 Padding**

Padding is the process of making the input picture fit in order to precisely position the filter on it. Padding is used to make the size of the output equal to the dimension of the input by adding zeros to the input frame of the matrix. Padding provides extra room for the kernel to cover the picture and is accurate for image analysis. Because to padding, information on the edges of images is retained in the same way as it is preserved in the center of the image.



**Fig. 3.13: Padding process [26]**

There are two ways:

* Pad the image with zeros (also known as zero-padding) to make it fit.
* Drop the portion of the input image file if the filter doesn't fit. This technique is known as valid padding. It holds only the valid portion of the input image.

**3.4.5 Pooling Layer**

In a CNN model, the pooling layers are used in between the Convolution layers. The Pooling layers

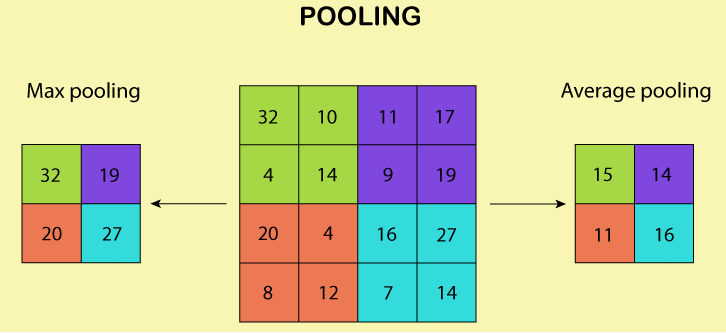
are responsible for reducing the number of parameters and computational complexity of the model. The pooling layer functions on each feature map individually. This decreases the resolution of the feature map by decreasing the height and breadth of the feature maps, but keeps the features of the map necessary for categorization. This is known as down-sampling [26].

There are different types of pooling, like:

* Average Pooling
* Max Pooling
* Sum Pooling

**Max-pooling:** It chooses the most elements from the feature map. The resultant max-pooled layer contains significant elements of the feature map. It is the most commonly used strategy since it produces the best outcomes.

**Average pooling:** It includes calculating an average for each region of the feature map.



**Fig. 3.14: Pooling layer [26]**

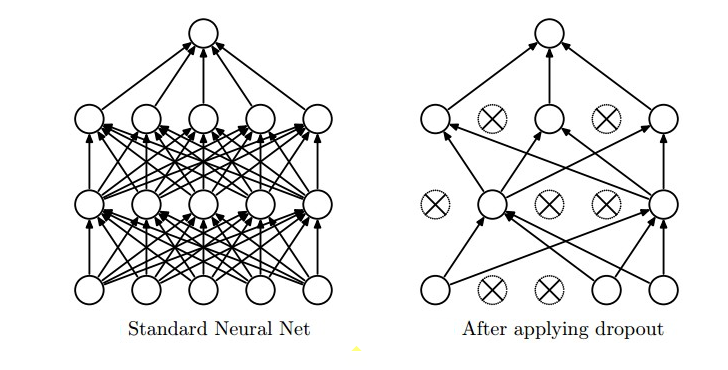
**Sum pooling:** Sum pooling takes the summation of all the values present within the filter window.It is the the average value of a pattern's presence in a certain region.

**3.4.6 Normalization layer**

Various normalization strategies have been proposed to be employed in Convolutional Neural Network architectures to implement inhibition systems experienced by the live brain. However, these layers have gone out of favor due to their poor contribution, if any.

**3.4.7 Dropout**

Dropout is a regularization method used in neural networks. In each stage of the approach, randomly selected nodes in the network are disregarded throughout the training phase. Dropout may be used on any or all of the network's hidden levels in addition to the visible or input layer. On the output layer, it is not utilized [27].

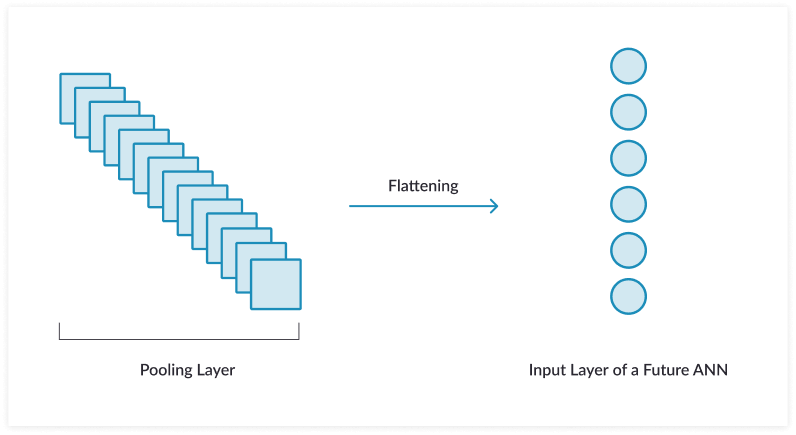
****

**Fig. 3.15: Dropout layer [28]**

In machine learning, overfitting is a well-known issue. The dropout regularization method is used to prevent overfitting. The dropout concept is pretty easy to put into practice. With a probability of p for each iteration, some neurons are momentarily deactivated or "dropped" during training. This suggests that all inputs and outputs into and out of those neurons will be disabled during that iteration. Each iteration involves randomization of the disabled neurons with probability p, making it possible for a disabled neuron in one iteration to become active in a subsequent iteration. The dropout rate, or hyperparameter p, usually has a value of approximately 0.5, resulting in 50% of the neurons being lost.

**3.4.8 Flattening**

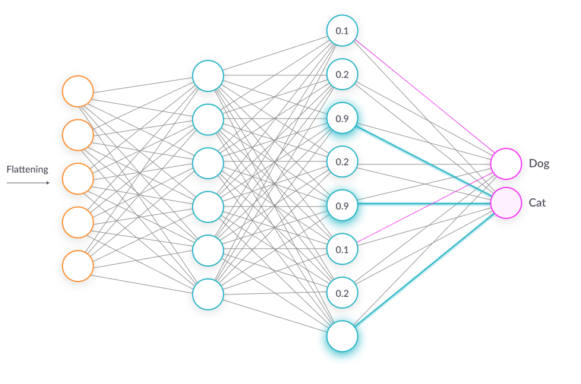
Flattening is used to convert the output matrix into a 1-D vector, which is then fed into the fully connected layer**.**

****

**Fig. 3.16: Flattening [29]**

**3.4.9 Fully connected layer**

The fully connected layer is made up of neurons that are completely connected to every output from the previous layer. Matrix multiplication and a bias offset can be used to compute the activation of the fully connected layer. This layer uses the Softmax or Sigmoid activation function to do categorization.



**Fig. 3.17: Fully connected layer [29]**

**3.5 Famous CNN Architectures**

There are numerous architectures in the domain of Convolutional Neural Networks that have been implemented to perform various classification tasks. Some of them are:

* AlexNet
* VGG-16
* VGG-19
* Inception and GoogLeNet
* ResNet
* DenseNet
* Squeeze Net

**3.6 Advantages of CNN**

* When compared to earlier methods, CNN's main advantage is its ability to automatically identify significant elements without human intervention. Using several images of cats and dogs as an example, it may teach itself the essential characteristics of each class.

* Because of its capacity to manage massive amounts of unstructured data, convolutional neural networks have transformed the business**.**
* In fields like speech recognition, natural language processing, and picture classification where there is a lot of huge, unstructured data, CNN has great success. It is computationally effective and more powerful than machine learning techniques**.**

**3.7 Conclusion**

An example of a neural network that accepts pictures as inputs is the convolutional neural network (CNN). Several layers are layered on top of one another to make up a convolutional neural network model. CNN helps to satisfy the two key requirements for object detection. By combining the collective representational power of multi-layer neural networks with multi-stage properties, they are a greater recognition percentage and a shorter training time. The Convolutional Neural Network is introduced in general terms in this chapter.

**CHAPTER 4**

**DATASET DETAILS**