**CHAPTER 1**

**INTRODUCTION**

* 1. **Overview**

Image classification and Object Recognition is probably the most well-known problem in computer vision. As a human, our cognitive ability to make decisions depends upon our visual aids, providing such abilities to the computer would allow them to experience the same power. Fig 1.1 shows the overview of an image analysis task. Image classification consists of classifying an image into one of many different categories. For e.g. an image of an animal can be classified into its appropriate category. In recent years a lot of research has been done in the field of image classification and also the accuracy to classify the images has been improved with the implementation of new deep algorithms. Object detection is one more computer technology related to [computer vision](https://en.wikipedia.org/wiki/Computer_vision) and [image processing](https://en.wikipedia.org/wiki/Image_processing) that is useful in detecting occurrences of semantic objects of a confident class, for e.g. humans, buildings, or cars in digital images and videos. Such accuracy in image classification and object recognition is obtained with the help of Deep Learning models. Deep Learning is an artificial intelligence function that copies the workings of the human mind in processing data and creating patterns which is useful for decision making. It is a subset of machine learning in [Artificial Intelligence (AI)](https://www.investopedia.com/terms/a/artificial-intelligence-ai.asp) that has networks which are capable of learning unsupervised from data that is unstructured or unlabelled. Common techniques comprise of deep learning based methods such as “convolutional neural networks”, and “feature-based methods using edges”, “gradients”, “histogram of oriented gradients (HOG)”, “Haar wavelets”, and “linear binary patterns”.

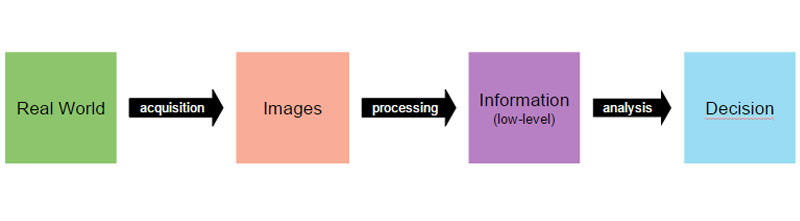


Fig 1.1 Image analysis steps

* 1. **Problem Definition**

Today, people show their emotions, remain it anger, love, disgust or sympathy, through images they share on the social media sites. These images can have a positive connotation or a negative connotation and it is very important to know the sentiments of people towards the type of images that are shared on the social media websites. As a part of this project, I aim to predict the sentiments of images which will be classified under two main categories, i.e. positive and negative as shown in the fig 1.2(a) and 1.2(b) respectively. The positive category further includes the images that depict love, peace, and joy. The negative category further includes the images that depict war, hate, poverty, disgust, and fear.

Fig 1.2 (a) negative connotation Fig 1.2(b) positive connotation

* 1. **Motivation**

A lot of research has been done in the field of sentiment analysis of text. As people use social media platform more and more, they upload vast amount of images on to the social platform. So it becomes necessary to know the sentiments, people try to convey using their images. Recently, social media users are progressively using images to express their views and share their involvements. Sentiment analysis results from such large chunks of data can help understand the user sentiments towards different topics and events. Hence the prediction sentiments from visual content is similar to that of textual sentiment analysis. I used Convolutional Neural Networks (CNN) for the sentiment analysis of the images.

* 1. **Scope and Objective of Research**

Online social networking platforms are giving more and more suitable services to their users. In today’s world, social networking platforms one of the most famous sources for people to gain facts or other kind of knowledge on all traits of their lives. Also, every online social network user is a provider of information. Social website users tend to share their involvements express their opinions on almost all events and subjects. Among all of this huge chunks of data generated, we are solely interested in the opinion or sentiment of people regarding particular topics and events. In the past, a lot of research and work have been dedicated for the prediction of box office revenue generation for movies using online sentiments of people. These predictions have been successful in proving that online user’s opinions or sentiments are related with our real-world activities. However, right now all of these works are done in the field of textual analysis and not in the field of sentiment analysis of images. Indeed, online social network providers are competing with each other by providing easier access to their increasingly powerful and diverse services. Figure 1.4 shows example images related to the 2012 United States presidential election. Clearly, images in the top and bottom rows convey opposite sentiments towards the two candidates.

****

Figure 1.3 Examples of Flickr images related to the 2012 United

States presidential election. [1]

**CHAPTER 2**

**LITERATURE REVIEW**

**2.1 Study of existing systems**

Sentiment analysis is a very challenging task. A lot of previous work tries to address the different kinds of issues, prevalent in the sentiment analysis task, using both handcrafted features and neural networks. Here, I briefly describe some of the most influential work that addresses this issue. “Visual Sentiment Prediction with Deep Convolutional Neural Networks” proposed by Xu et al. [2] use CNN’s pre-trained models on Object recognition data to perform sentiment analysis on images which are collected from social websites like Twitter and Tumblr. “Robust image sentiment analysis using progressively trained and domain transferred deep networks” by You et al. [3] uses VGG-ImageNet and its architectural variants to study sentiment analysis on Twitter and Flickr datasets. “Recognizing image style” by Karayev et al. [4] experiment with handcrafted features like L\*a\*b colour space features, GIST and saliency features on Flickr style data, Wikipaintings and AVA Style data. [6] “Emotion based classification of natural images” by Dellagiacoma et al. [7] uses MPEG7 colour and edge descriptors to perform emotion classification and compares the results with Bag of Emotions method.

**2.2 Review of Literature and Findings**



|  |  |
| --- | --- |
| TITLE | Emotion Detection and Sentiment Analysis of Images. [2] |
| AUTHOR | Vasavi Gajarla, Aditi Gupta [2] |
| YEAR | 2016 |
| FINDINGS | The authors of the above research paper used VGG-Imagenet, fine-tuning models like RESNET, Places205-VGG16 for the sentiment analysis and emotion detection. They used the dataset from Flickr by using the Flickr API. They used around 10,000 images out of which 75% were used in training set and 25% were used in the test set. |

Table 2.1 Literature review 1



|  |  |
| --- | --- |
| TITLE | Robust Image Sentiment Analysis Using Progressively Trained and Domain Transferred Deep Networks [1] |
| AUTHOR | Quanzeng You and Jiebo Luo, Hailin Jin and Jianchao Yang [1] |
| YEAR | 2015 |
| FINDINGS | The authors of the above research paper used Visual Sentiment Analysis with regular CNN and Visual Sentiment Analysis with Progressive CNN. They compared the CNN architectures using half a million Flickr images. |

Table 2.2 Literature review 2



|  |  |
| --- | --- |
| TITLE | Image Sentiment Analysis[3] |
| AUTHOR | Jayesh Mandhyani, Latika Khatri, Varsha Ludhrani, Raveena Nagdev, Prof. Sunita Sahu [3] |
| YEAR | 2017 |
| FINDINGS | The authors of the above research paper proposed their own system by combining the features of SetniBank, RCNN, and SetniStrenght. In this method they worked with the mid-level features of the images. They have used the R-CNN for object detection in sentibank. |

Table 2.3 Literature review 3



|  |  |
| --- | --- |
| TITLE | Sentribute: Image Sentiment Analysis from a Mid-level Perspective.[4] |
| AUTHOR | Jianbo Yuan, Sean Mcdonough , Quanzeng You, Jiebo Luo [4] |
| YEAR | 2015 |
| FINDINGS | In this paper they have demonstrated Sentribute, a novel image sentiment prediction algorithm based on mid-level attributes. Asymmetric bagging approach is employed to deal with unbalanced dataset. To enhance prediction performance, they introduce eigenface-based emotion detection algorithm, which is simple but powerful especially in cases of detecting extreme facial expressions, to dealing with images containing faces and obtain a distinct gain in accuracy over result based on mid-level attributes only. [4] |

Table 2.4 Literature review 4



|  |  |
| --- | --- |
| TITLE | Visual Sentiment Prediction with Deep Convolutional Neural Networks. [10] |
| AUTHOR | Can Xu, Suleyman Cetintas, Kuang-Chih Lee, Li-Jia Li [10] |
| YEAR | 2014 |
| FINDINGS | In this work, the authors offered a new visual sentiment prediction framework that performs image understanding with Convolutional Neural Networks (CNN). Specifically, it performs transfer learning from a CNN with millions of parameters, which is pre-trained on large-scale data for object recognition. Experiments conducted on two real-world datasets from Twitter and Tumblr demonstrate the effectiveness of the proposed visual sentiment analysis framework. [10] |

Table 2.5 Literature review 5



|  |  |
| --- | --- |
| TITLE | Social Media Sentiments Using Latent Correlation Analysis. [11] |
| AUTHOR | Miss. Sonali B. Gaikwad, Prof. Vikas N. Dhakane. [11] |
| YEAR | 2017 |
| FINDINGS | In this paper a novel approach has been proposed which exploits latent correlation analysis of multi-views, and also works with sentiment analysis depending on visual and textual contents in social media. A hybrid approach of aggregating sentiments have been proposed from textual and visual contents. [11] |

Table 2.6 Literature review 6

**CHAPTER 3**

**ANALYSIS OF EXISTING WORK**

**3.1 Applications of existing system**

Sentiment analysis, with either image or text, is applied by the social networking websites who have a large number of users uploading vast amount of images or videos. It is mainly used in social media monitoring and VOC to track customer reviews, survey responses, competitors, etc. It is also used for (1) businesses and organisations which require consumer opinions to do with products they market and services they produce, (2) individuals who make decisions to purchase products or services based upon word of mouth or on-line reviews, or to find public opinion, e.g. concerning politics or local issues, (3) on-line advertising where in social media, an organisation may place an advertisement in response to a favourable review of a product, or a rival product could be advertised upon receipt of a bad review, and (4) opinion retrieval for general searches of opinions. [12]

**CHAPTER 4**

**PROJECT WORK**

**4.1 Introduction to the proposed system**

**4.1.1 Deep learning**

Computer Vision is a hot buzz word in this era of technology. Computer Vision is a field of Artificial Intelligence and Computer Science that aims at giving computers a visual understanding of the world. It is one of the main components of machine understanding. Applications of computer vision are vast. Applications range from tasks such as industrial [machine vision](https://en.wikipedia.org/wiki/Machine_vision) systems which, say, inspect bottles speeding by on a production line, to research into artificial intelligence and computers or robots that can understand the realm around them. Deep learning is a technique through which we can attain high performing results in the field of Computer Vision. Deep learning is a machine learning technique that teaches computers to do what comes naturally to humans: learn by example. Deep Learning has contributes tremendously in ground breaking research and discovery such as driverless cars, which are able to identify different kinds of signs on the street and also differentiate between people and other non-living objects. It also has its application in electronic devices like phones, tablets, TVs, and hands-free speakers, where it is responsible for the voice control. It’s achieving results that were not possible before. In deep learning, a computer model is designed such a way that it directly learns to classify objects from text, sound or images. Deep learning models are capable of achieving state-of-the-art accuracy, and can even possibly do better than human performance. For this purpose, deep learning model have to be provided with large set of labelled data and they are trained with the help of neural networks which contain many layers. Deep learning requires large amounts of labelled data. For example, driverless car development requires millions of images and thousands of hours of video. Deep learning requires substantial computing power. High-performance GPUs have a parallel architecture that is efficient for deep learning.

**4.1.2 How do Deep Learning works**

Mostly, deep learning models make a use of neural networks and hence they are called deep neural networks. (Fig 4.1).

The term “deep” depicts the number of hidden layers used in the neural network. The more number of hidden layers the more deep is the neural network. Previously, the neural networks only contained 2-3 hidden layers, while deep networks can have as many as 150.

Deep learning models are trained by using large sets of labelled data and neural network architectures that learn features directly from the data without the need for manual feature extraction.

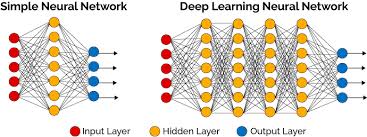


Fig 4.1 Simple Neural Network vs. Deep Learning Neural Network

**4.2 Algorithm of proposed system**

Deep Learning has achieved tremendous popularity recently due to its ability to learn from experiences and also its speciality to iterate over more hidden layers. There are several algorithms which contribute to this speciality of Deep Learning. Some of them are as follows:

1. Artificial Neural Network (ANN)
2. Convolution Neural Network (CNN)
3. Recurrent Neural Network (RNN)
4. Radial basis function Neural Network (RFNN)
5. Modular Neural Network (MNN)

**4.2.1 Convolution Neural Network (CNN)**

Convolution Neural Networks have a range of applications in image and video recognition, recommender systems and natural language processing. CNNs, similar to neural networks, are consisted of neurons with learnable weights and biases. Each and every neuron in the CNN works as an input/output combination. It receives several inputs, sums the input and then pass it through an activation function, and then answers as an output. The whole CNN network has a loss function and all the steps which are applicable to the neural network pretty well applied to the CNN’s too. Convolutional Neural Networks (CNN) have been very fruitful in recognition of document. The main difference between a normal neural network and a CNN is that, the CNN consists of several convolution layer. The working of convolution layers is explained in the later part of this chapter. Along with the convolution layers, the CNN also may consists several pooling layers and at last there is a fully connected layers which is connected to the output layer. The working of pooling and fully connected layers have also been explained later in this chapter. CNN being a supervised learning algorithm, back-propagation is used to learn parameters of different layers. As the CNN may contain huge number of layers for the classification purpose, it has only been applied to relatively small images in the literature. With the increasing computational power of GPU, it is now possible to train a deep convolutional neural network on a large scale image dataset. Indeed, in the past several years, CNN has been successfully applied to “scene parsing”, “feature learning”, “visual recognition and image classification”.

An overview of the CNN can be depicted in the image (fig 4.2).

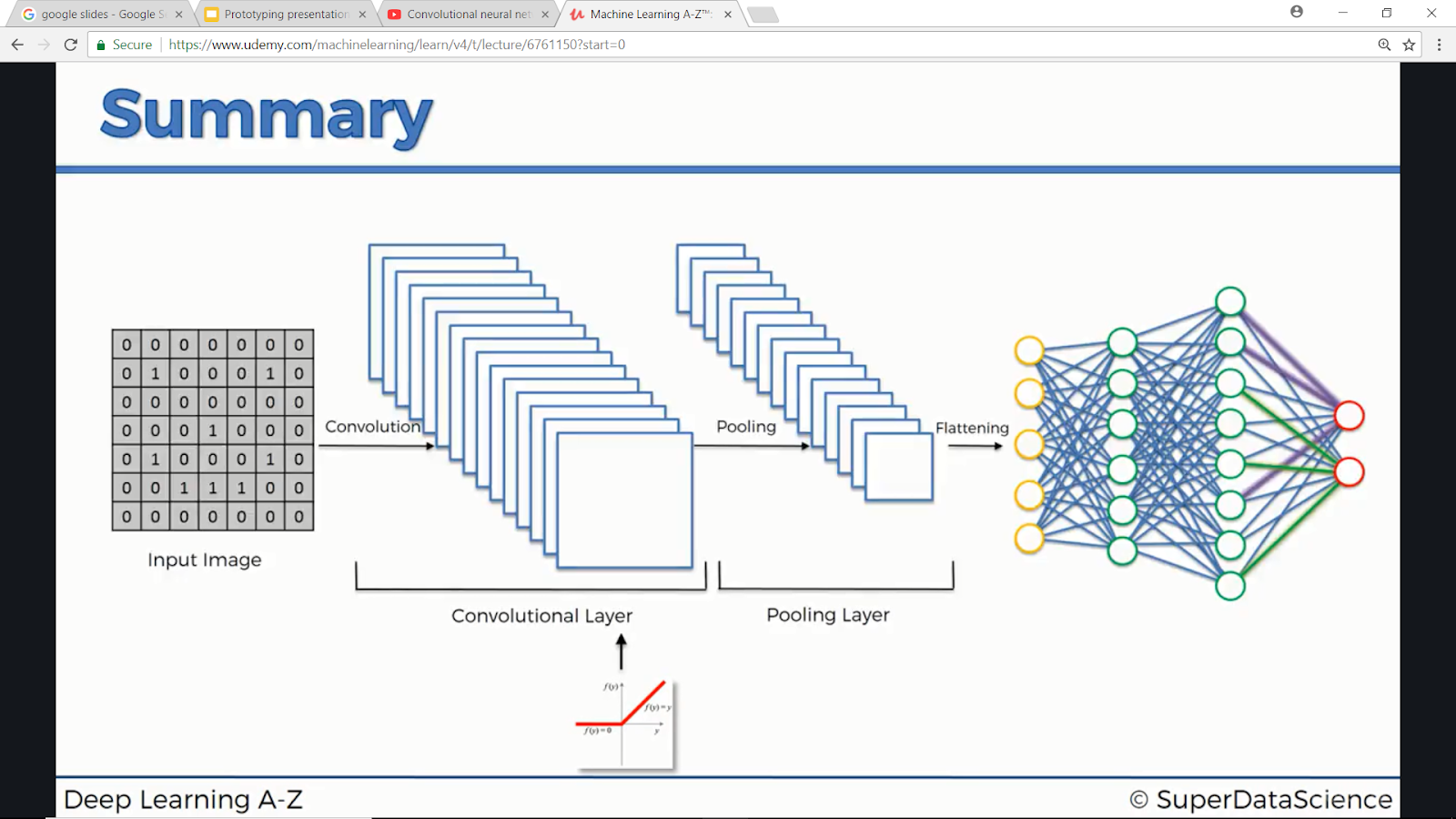


Fig 4.2. An overview of CNN

**4.3 Working of the proposed system**

There are many different ways in which a CNN can be implemented according to the requirements of the neural network. Here, I mention the basic steps which are involved in every CNN.

**Step – 1 the Convolution:**

ConvNet derive their name from the [“convolution” operator](http://en.wikipedia.org/wiki/Convolution). The main aim of the convolution is to extract features from the input image. Convolution preserves the spatial relationship between pixels by learning image features using small squares of input data. Every image can be considered as a matrix of pixel values. In CNN terminology, the 3×3 matrix is called a ‘filter‘ or ‘kernel’ or ‘feature detector’ and the matrix formed by sliding the filter over the image and computing the dot product is called the ‘Convolved Feature’ or ‘Activation Map’ or the ‘Feature Map‘. This filter is slide over the actual image by 1 pixel, called the ‘stride’ and for every position element wise multiplication, between the two matrices, is computed and the multiplication outputs is added to get the final integer which forms a single element of the output matrix. [9]

**Figure 4.3** shows an example of convolution step.

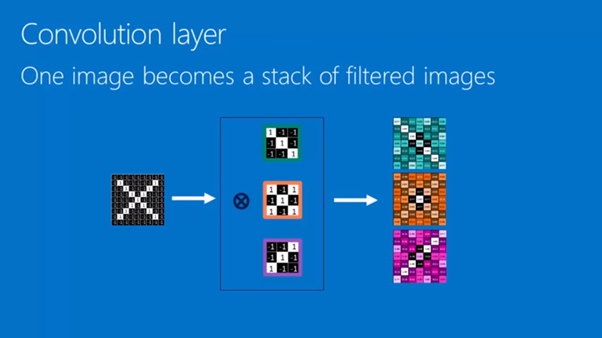


Fig 4.3 Convolution step

**Step -2 Max/Min Pooling:**

Spatial Pooling, also called subsampling or downsampling reduces the dimensionality of each feature map but retains the most important information. Spatial Pooling can be of different types: Max, Average, Sum etc. In case of Max Pooling, we define a spatial neighbourhood, for example, a 2×2 window, and take the largest element from the rectified feature map within that window. Instead of taking the largest element we could also take the average or sum of all elements in that window. In practice, Max Pooling has been shown to work better. [9]

**Figure 4.4** shows an example of Max Pooling operation on a Rectified Feature map (obtained after convolution layer) by using a 2×2 window.

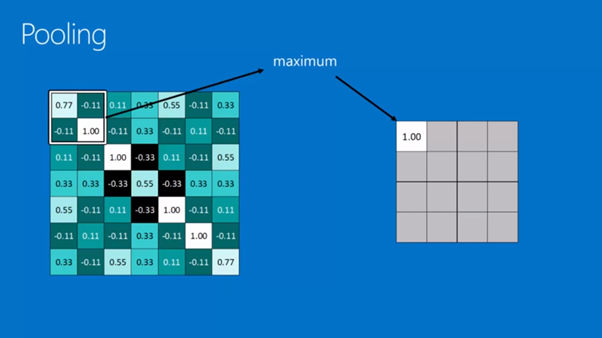


Fig 4.4 Max Pooling step

**Step -2 Fully Connected Layers:**

The Fully Connected layer is a traditional Multi Layer Perceptron that uses a softmax activation function in the output layer. The term “Fully Connected” implies that every neuron in the previous layer is connected to every neuron on the next layer. The output from the convolutional and pooling layers represent high-level features of the input image. The purpose of the Fully Connected layer is to use these features for classifying the input image into various classes based on the training dataset. Apart from classification, adding a fully-connected layer is also a cheap way of learning non-linear combinations of these features. Most of the features from convolutional and pooling layers may be good for the classification task, but combinations of those features might be even better .The sum of output probabilities from the Fully Connected Layer is 1. This is ensured by using the [Softmax](http://cs231n.github.io/linear-classify/#softmax) as the activation function in the output layer of the Fully Connected Layer. The Softmax function takes a vector of arbitrary real-valued scores and squashes it to a vector of values between zero and one that sum to one. [9]

Fig 4.5 shows an example of a fully connected layer.

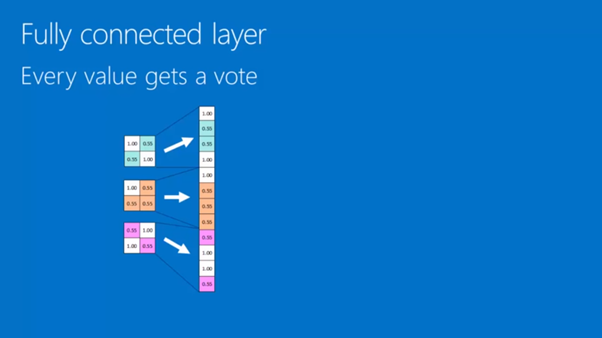


Fig 4.5 Fully Connected Layer step

**CHAPTER 5**

**IMPLEMENTATION DETAILS**

**5.1 Deciding the emotion category**

I am inspired by the 6 emotion categories defined by the famous psychologist Ekman i.e. Happiness, Sadness, Fear, Disgust, Anger and Surprise. Out of the following 6 categories, I divided into two categories. The first category, named positive, contains all the emotions like happiness, joy, and peace. The second category, named negative, contains all the emotions like fear, horror, and terror.

**5.2 Data collection**

Nowadays, there are a lot of social networking platforms where people upload a huge amount of images. I decided to take the images from three different websites and then apply the algorithms to see which has the most accuracy. Flickr is one such social website where photographers from all around the world upload the images. I decided to use the image dataset from Flickr, Pixbay and Google Images.

**5.2.1 Flickr Dataset**

**Step 1**: Querying the image metadata.

I used the Flickr’s API service to query for images using each of the emotional categories - Happiness, Joy, and Peace, Fear, Horror, and Terror - as search query parameters to collect the image metadata, like server and farm ID numbers on which the image is stored, after sorting the result set by interestingness to make sure that images fitting to the emotional categories are retrieved on priority.

**Step 2:** Downloading the images from Flickr server.

Once the image metadata is retrieved, it was possible to use that metadata to directly access the images from Flickr servers instead of going through the Flickr API.

For this project, from Flickr Dataset, I collected a total of 3370 images. Out of the total, each category comprised of 1685 images. Each category was then divided into training and testing set. The training set contains 1348 (80%) and the test set contains 337 (20%) of the total images in each category.

**5.2.2 Pixbay and Google Images Dataset:**

Pixbay and Google Images is a website that contains images form a wide variety of categories. I used web scraping to download the images of the corresponding categories. Python has inbuilt libraries, like Beautiful Soup, HTML Parser, etc. which helps in scraping the images from the website.

From Pixbay and Google Images Dataset, I collected a total of 2814 images. Out of the total, each category comprised of 1407 images. Each category was then divided into training and testing set. The training set contains 1125 (~ 80%) and the test set contains 282 (~20%) of the total images in each category.

* 1. **Libraries used**

Python is a very strong language with great library support for the neural networks. It is an open source language and also its neural network libraries are available for free to run on host computers. The most popular library used to run the neural network on the computer is Keras, which is built upon python.

* + 1. **Keras**

Keras is a minimalist Python library for deep learning that can run on top of Theano or TensorFlow. It was developed to make implementing deep learning models as fast and easy as possible for research and development. It runs on Python 2.7 or 3.5 and can seamlessly execute on GPUs and CPUs given the underlying frameworks. It is released under the permissive MIT license. [5]

* + 1. **NumPy**

NumPy is a library for the [Python programming language](https://en.wikipedia.org/wiki/Python_(programming_language)), adding support for large, multi-dimensional [arrays](https://en.wikipedia.org/wiki/Array_data_structure) and [matrices](https://en.wikipedia.org/wiki/Matrix_(math)), along with a large collection of [high-level](https://en.wikipedia.org/wiki/High-level_programming_language) [mathematical](https://en.wikipedia.org/wiki/Mathematics) [functions](https://en.wikipedia.org/wiki/Function_(mathematics)) to operate on arrays. [6]

**5.3.3 Scikit-Learn**

Scikit-learn is a free software machine learning library for the Python programming language. It features various classification, regression and clustering algorithms including support vector machines, random forests, gradient boosting, k-means and DBSCAN, and is designed to interoperate with the Python numerical and scientific libraries NumPy and SciPy. [7]

* + 1. **Matplotlib**

Matplotlib is a plotting library for the Python programming language and its numerical mathematics extension NumPy. It provides an object-oriented API for embedding plots into applications using general-purpose GUI toolkits like Tkinter, wxPython, Qt, or GTK+. [8]

* 1. **CNN Pre-trained models**

**5.4.1 Multiple Convolution layer**

The idea was to simply run the neural network with a multiple convolutional layers. This layer contained four convolution layers along with four other max pool layers and one dense and one fully connected layers. I applied the multiple convolution layer to all of my dataset and then checked the performance. All the images used were of the size 100x100 and the mode of channel was RGB.

Fig 5.1 shows the neural network for the Multiple Convolution Layer and also the number of epochs run for the same.

****

Fig 5.1 Epochs of Multilayer Convnet

* + 1. **VGG ImageNet 16**

VGG-Imagenet model is pre trained model which was developed during the **ImageNet Large Scale Visual Recognition Challenge 2012 (ILSVRC2012)** is classifies on 1000 categories. I experimented with fine-tuning the VGG-ImageNet model.I fine tuned the model to perform semantic analysis with images. Every input image passed into the network has a dimension of (100x100x3). VGG Imagenet 16 consists of 16 layers as shown by the summary of the model in fig 5.2.

The following are the steps involved in this experiment:

**Step 1:**

Pass the data (both train and test) as input to VGG-ImageNet model.

**Step 2:**

Store the activations from second-to-last fully connected layer of the network as feature vectors.

The training accuracy obtained from VGG16 ImageNet was about 65%.

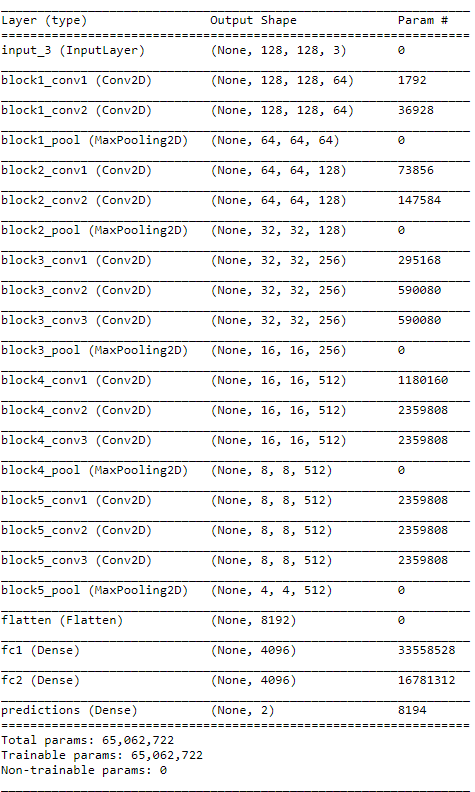
****

Fig 5.2 VGG Imagenet 16 model summary

Fig 5.3 shows the neural network for the VGG Imagenet 16 and also the number of epochs run for the same

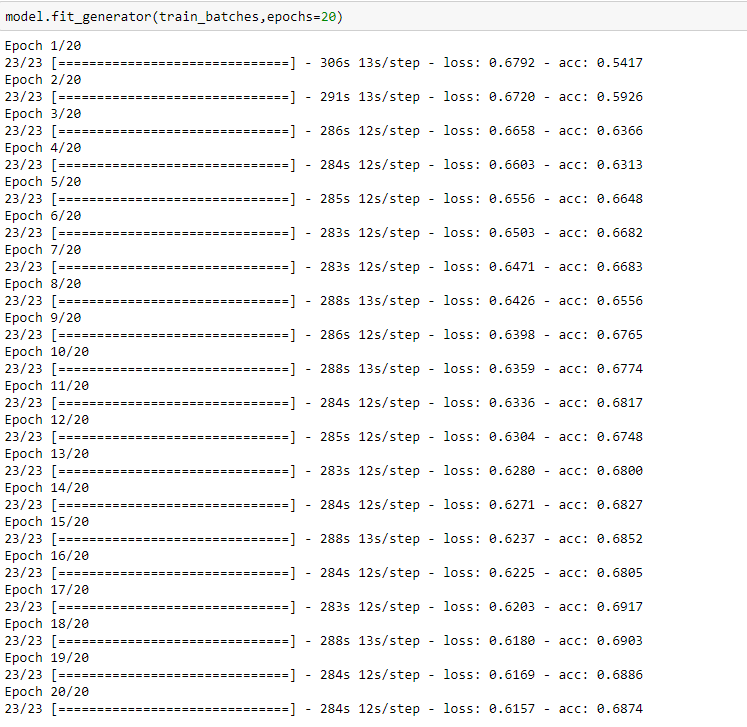
****

Fig 5.3 Epochs of VGG Imagenet 16 model

**CHAPTER 6**

**EXPERIMENTALS RESULTS**

**6.1 Flickr Dataset**

**6.1.1 Multilayer Convolution Neural Network**

On applying the multilayer neural network to the flickr dataset, I observed a training accuracy of X% and testing accuracy of Y%.

Fig 6.1 shows the confusion matrix for the same. Where out of total 562 test cases, 274 are predicted to be correct by the neural network.

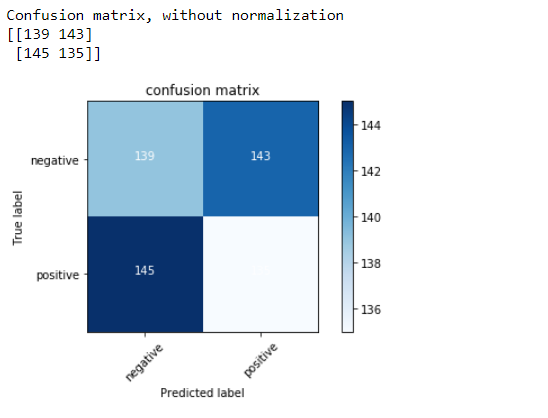


Fig 6.1

**6.1.2 VGG Imagenet 16 Neural Network**

On applying the multilayer neural network to the flickr dataset, I observed a training accuracy of X% and testing accuracy of Y%.

Fig 6.2 shows the confusion matrix for the same. Where out of total 562 test cases, 274 are predicted to be correct by the neural network.

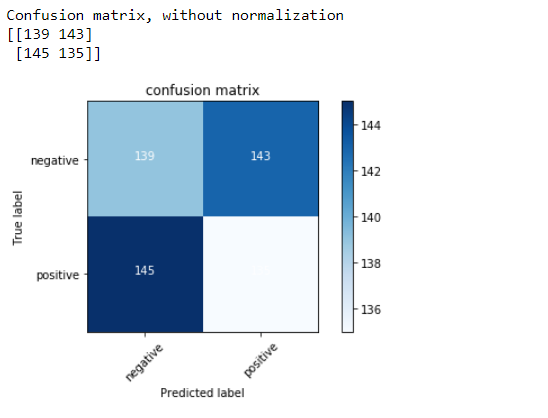


Fig 6.2

**6.2 Pixbay Dataset**

**6.2.1 Multilayer Convolution Neural Network**

On applying the multilayer neural network to the flickr dataset, I observed a training accuracy of X% and testing accuracy of Y%.

Fig 6.3 shows the confusion matrix for the same. Where out of total 562 test cases, 274 are predicted to be correct by the neural network.

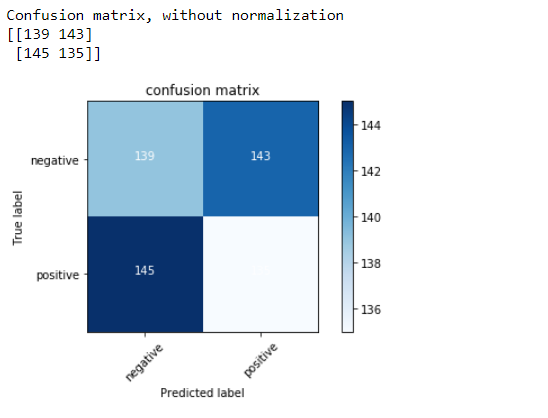


Fig 6.3

**6.2.2 VGG Imagenet 16 Neural Network**

On applying the multilayer neural network to the flickr dataset, I observed a training accuracy of X% and testing accuracy of Y%.

Fig 6.4 shows the confusion matrix for the same. Where out of total 562 test cases, 274 are predicted to be correct by the neural network.

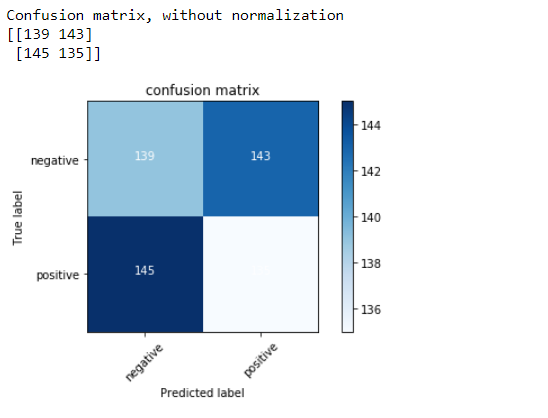


Fig 6.4

**6.3 Google Image Dataset**

**6.3.1 Multilayer Convolution Neural Network**

On applying the multilayer neural network to the flickr dataset, I observed a training accuracy of X% and testing accuracy of Y%.

Fig 6.5 shows the confusion matrix for the same. Where out of total 562 test cases, 274 are predicted to be correct by the neural network.

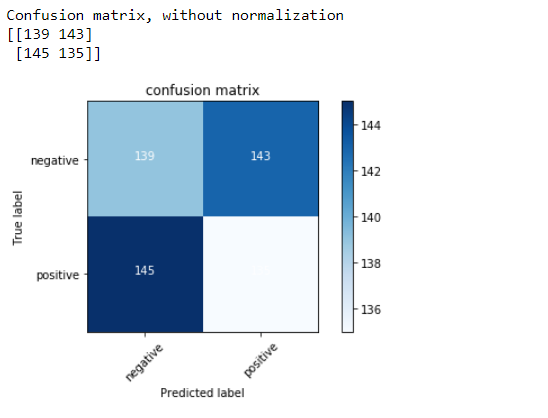


Fig 6.5

**6.3.2 VGG Imagenet 16 Neural Network**

On applying the multilayer neural network to the flickr dataset, I observed a training accuracy of X% and testing accuracy of Y%.

Fig 6.6 shows the confusion matrix for the same. Where out of total 562 test cases, 274 are predicted to be correct by the neural network.

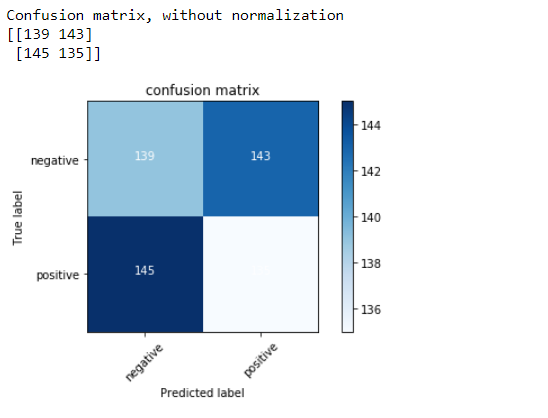


Fig 6.6

**6.4 An overview of the algorithms performance.**

**Flickr Dataset**

|  |  |  |  |
| --- | --- | --- | --- |
| Sr. No | Model Used | Training accuracy  (%) | Testing accuracy  (%) |
| 1 | Multilayer CONVNET (4 layers) | 67 | 53 |
| 2 | Multilayer CONVNET (5 layers) | 71 | 48 |
| 3 | VGG ImageNet 16 | 49 | - |

Table 6.1 Flickr dataset statistics

**Pixbay Dataset**

|  |  |  |  |
| --- | --- | --- | --- |
| Sr. No | Model Used | Training accuracy  (%) | Testing accuracy  (%) |
| 1 | Multilayer CONVNET (4 layers) | 82 | 48 |
| 2 | Multilayer CONVNET (5 layers) | 71 | 48 |
| 3 | VGG ImageNet 16 | 68 | 48 |

Table 6.2 Pixbay dataset statistics

**Google Images Dataset**

|  |  |  |  |
| --- | --- | --- | --- |
| Sr. No | Model Used | Training accuracy  (%) | Testing accuracy  (%) |
| 1 | Multilayer CONVNET (4 layers) | 87 | 48 |
| 2 | Multilayer CONVNET (5 layers) | 71 | 48 |
| 3 | VGG ImageNet 16 | 75 | 51 |

Table 6.3 Google Images dataset statistics

**CHAPTER 7**

**CONCLUSION**

**CHAPTER 8**

**FUTURE ENHANCEMENT**

As people upload a huge amount of data onto social networking websites, this method can be useful to understand and comprehend the sentiments or nature of humans towards a specific action or event. This method currently is made to work only on the images. Although, as the discovery of new and fast algorithms along with strong and fast working GPU’s takes place, it can be made to work on videos also. People, along with images tend to upload videos, or gifs, to show their behaviour or sentiment towards something. This method can also be used as a scanner to avoid people from sharing illegal videos or images, for e.g. drugs, sexual violence, nsfw etc. to the social networking websites which can have an adverse effect on the youth of the society.

**CHAPTER 9**

**REFERENCES**

[1]. You, Quanzeng, et al. "Robust Image Sentiment Analysis Using Progressively Trained and Domain Transferred Deep Networks." *AAAI*. 2015.

[2]. Karthika, I., and S. Priyadharshini. "Survey on Location based sentiment analysis of Twitter data." (2017).

[3]. Mandhyani, Jayesh, et al. "Image Sentiment Analysis." International Journal of Engineering Science 4566 (2017).

[4]. Jianbo Yuan, Sean Mcdonough, Quanzeng You, Jiebo Luo. “Sentribute: Image Sentiment Analysis from a Mid-level Perspective”

[5]. Keras https://machinelearningmastery.com/introduction-python-deep-learning-library-keras/

[6]. Numpy https://en.wikipedia.org/wiki/NumPy

[7]. Scikit Learn https://en.wikipedia.org/wiki/Scikit-learn

[8]. Matplotlib https://en.wikipedia.org/wiki/Matplotlib

[9]. Working of Convnets https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/

[10]. Xu, Can, et al. "Visual sentiment prediction with deep convolutional neural networks." arXiv preprint arXiv:1411.5731(2014).

[11]. Gaikwad, Miss Sonali B., and Vikas N. Dhakane. "Social Media Sentiments Using Latent Correlation Analysis." (2017).

[12] Sentiment Applications https://www.quora.com/What-are-the-applications-of-sentiment-analysis-Why-is-it-in-so-much-discussion-and-demand