Rishabh Software

Guide: Mr. Kshitij Tripathi

I42 Vraj Kotwala
(622079)

Internship Summary

Guides at Rishabh Software: Dr. Falguni Ranadive, Mr. Gagan Dubey

How Rishabh Software is different?

i. Rishabh Software provides software, recruitment and engineering services and solutions to mid-sized enterprises globally.
ii. Since its inception in 1999, Rishabh Software has proudly delivered over 1000+ successful projects to its customers across 23 countries including USA, UK, Europe, Canada, Australia.
iii. It is both product-based and service-based company.

Domain

Internship on Machine Learning as a domain

Project Definition

Emotion Recognition with Convolutional Neural Network

i. Emotion recognition is the process of identifying human emotion.
ii. The software can detect 7 different emotions (**Angry, Sad, Happy, Disgusted, Fearful, Neutral and Surprised**).
iii. It uses your system's webcam to capture your face.
iv. The captured emotions are stored into the database with timestamp to perform further statistical analysis.

Tools & Technology used

i. Anaconda to maintain dependencies in Python 3
ii. Spyder 3.0 IDE
iii. Jupyter IDE

Packages used with python 3:

i. MySQL-Connector
ii. NumPy
iii. OS
iv. OpenCV 2
v. Pandas
vi. Tensorflow 2
vii. Time

Introduction

Master Algorithm for **Emotion Recognition System**

i. First, I use **Haar cascade** to detect faces in each frame of the video. I used it to **collect data**.
ii. The region of image containing the face is resized to **48 x 48** and is passed as input to the 1st layer of ConvNet. This is to **train the model**.
iii. Again, I use **Haar cascade** to capture faces in each frame from web cam. The network outputs a list of **Softmax Scores** for the seven classes.
iv. The emotion with maximum score is displayed on the screen.
v. The emotions will be stored into the database.

Ingredients of Training Algorithm

Data

Collection of Dataset.

Model

Create a model. Find some coefficients. A model, loosely speaking, is a simplification of some thing or process.

Objective Function

This function, taking data and model parameters as arguments, can be evaluated to return a number. Classification.

Optimization Algorithm

A function that minimizes an error or one that maximizes reward over punishment. By using Adam Algorithm.

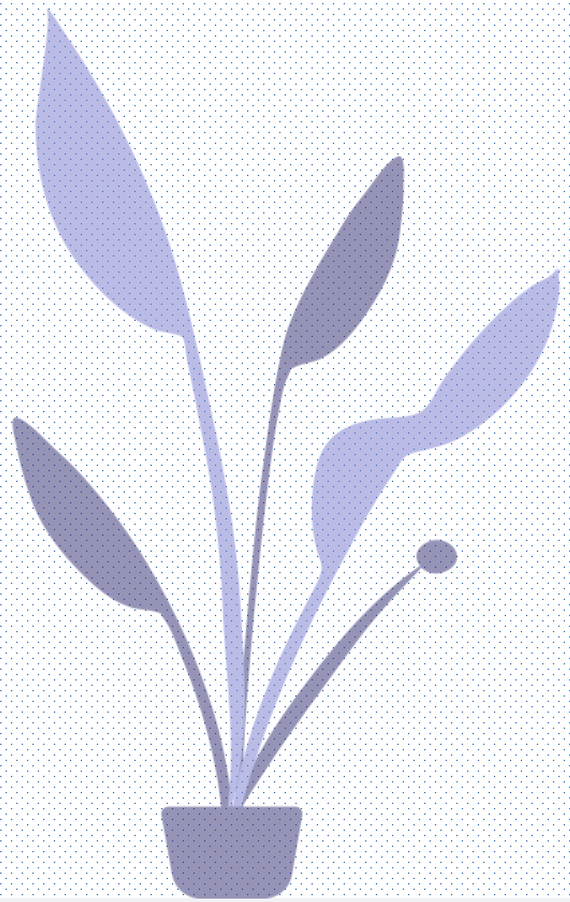
Training an algorithm involves 4 ingredients

i. First we must prepare a certain amount of **data** to train with. Usually this is historical data which is readily available, like I have used the FER 2013 dataset which was readily available.

ii. Second we need a **model**. The simplest model we can train is a linear model. That would mean to find some coefficients multiply each variable with them and sum everything to get the output.

iii. The third ingredient is the **objective function**. So far we took data fed it to the model and obtained an output. We want this output to be as close to reality as possible. That's where the objective function comes in. It estimates how correct the models outputs are. On average the entire machine learning framework boils down to optimizing this function. For example, if my function is measuring the prediction error of the model I would want to minimize this error or in other words minimize the objective function.

iv. Our final ingredient is the **optimization algorithm**. It consists of the mechanics through which we vary the parameters of the model to optimize the objective function.

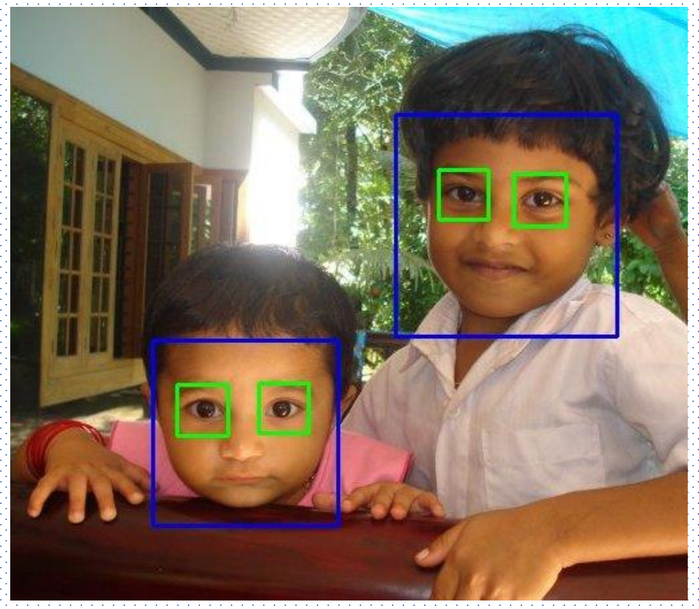


Dataset Collection

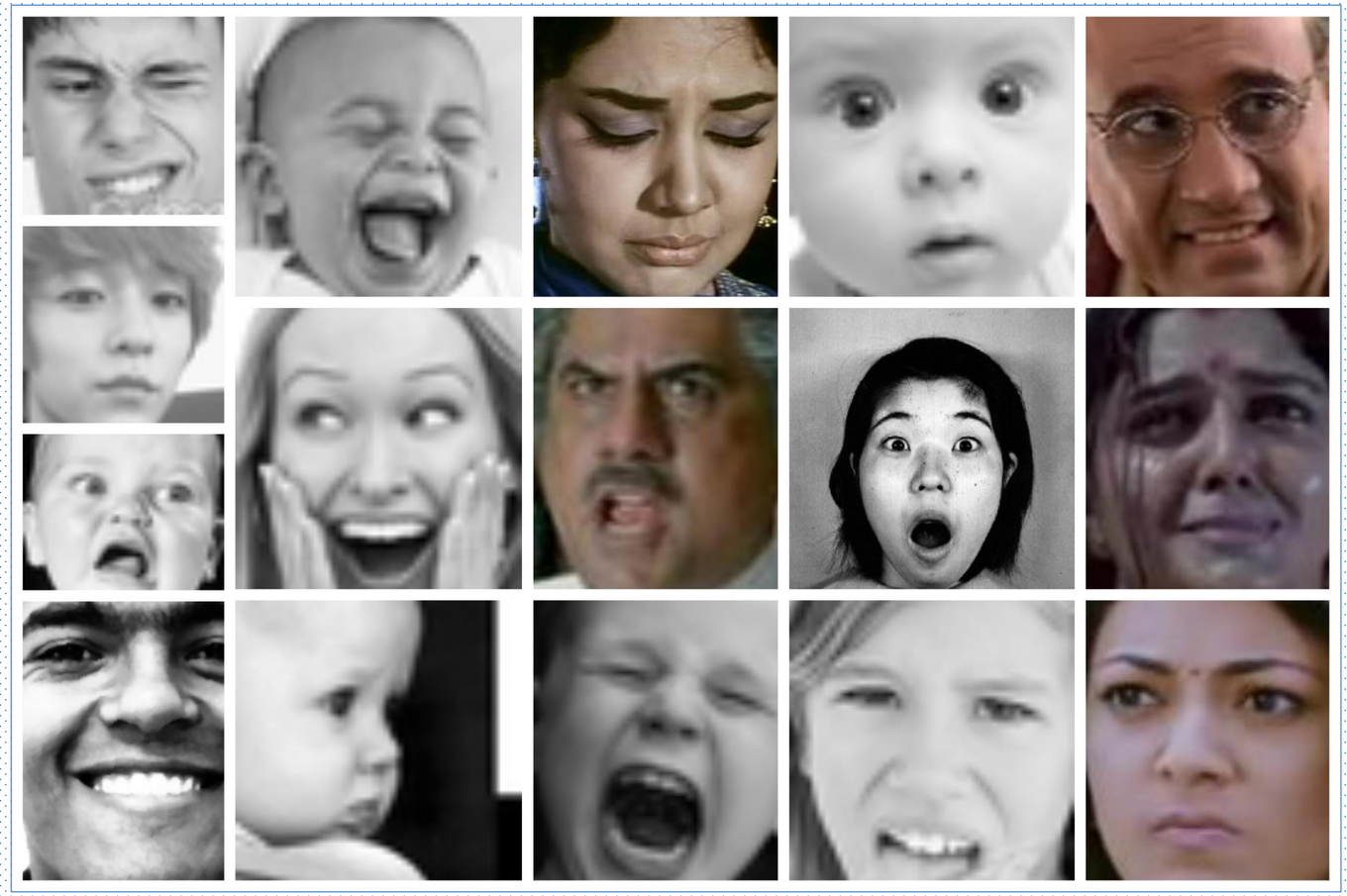
i. Neural networks, and deep networks in particular, are known for their need for large amounts of training data.
ii. I used the videos from YouTube to collect the data.
iii. For gathering the dataset, I had made a python script, which used **haar-cascade.xml** file to identify the faces from the YouTube videos.

What is Haar-Cascade?

- **Haar-Cascade** is a machine learning object detection algorithm used to identify objects in an image or video.



Collected Dataset merged with FER-2013 Dataset



Samples from the collected dataset combined with FER-2013

Ingredients of Training Algorithm

Data

Collection of Dataset.

Model

Create a model. Find some coefficients. A model, loosely speaking, is a simplification of some thing or process.

Objective Function

This function, taking data and model parameters as arguments, can be evaluated to return a number. Classification.

Optimization Algorithm

A function that minimizes an error or one that maximizes reward over punishment. By using Adam Algorithm.

Training of model – Using Deep Convolutional Neural Network:

What is a neural network?

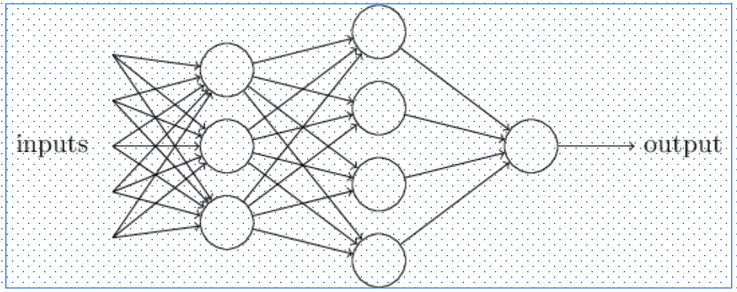
i. Creating a machine learning algorithm ultimately means building a model that outputs correct information.

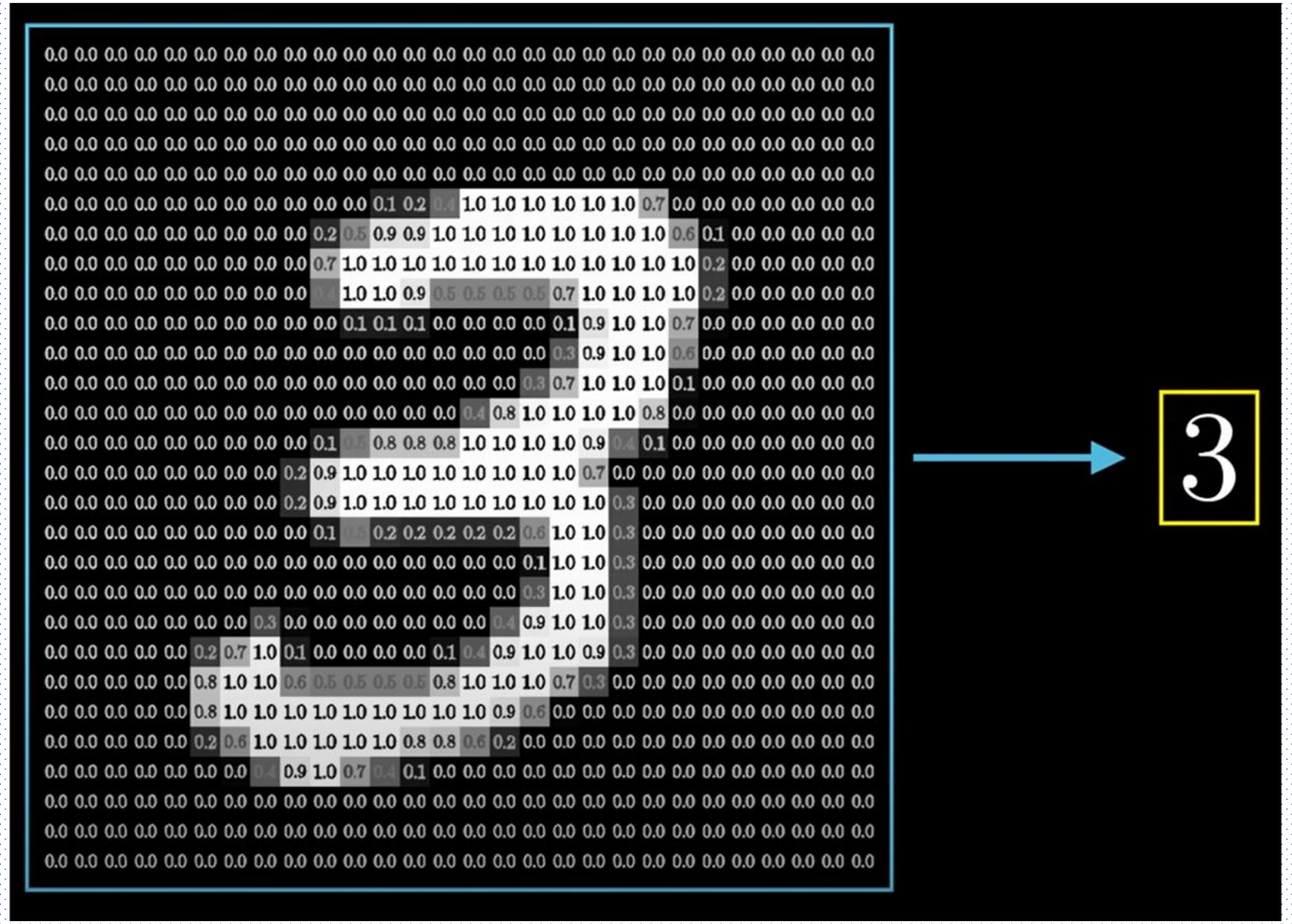
ii. Given that we've provided input data for now think of this model as a black box we feed input and it delivers an output.

iii. We must train the model for that.

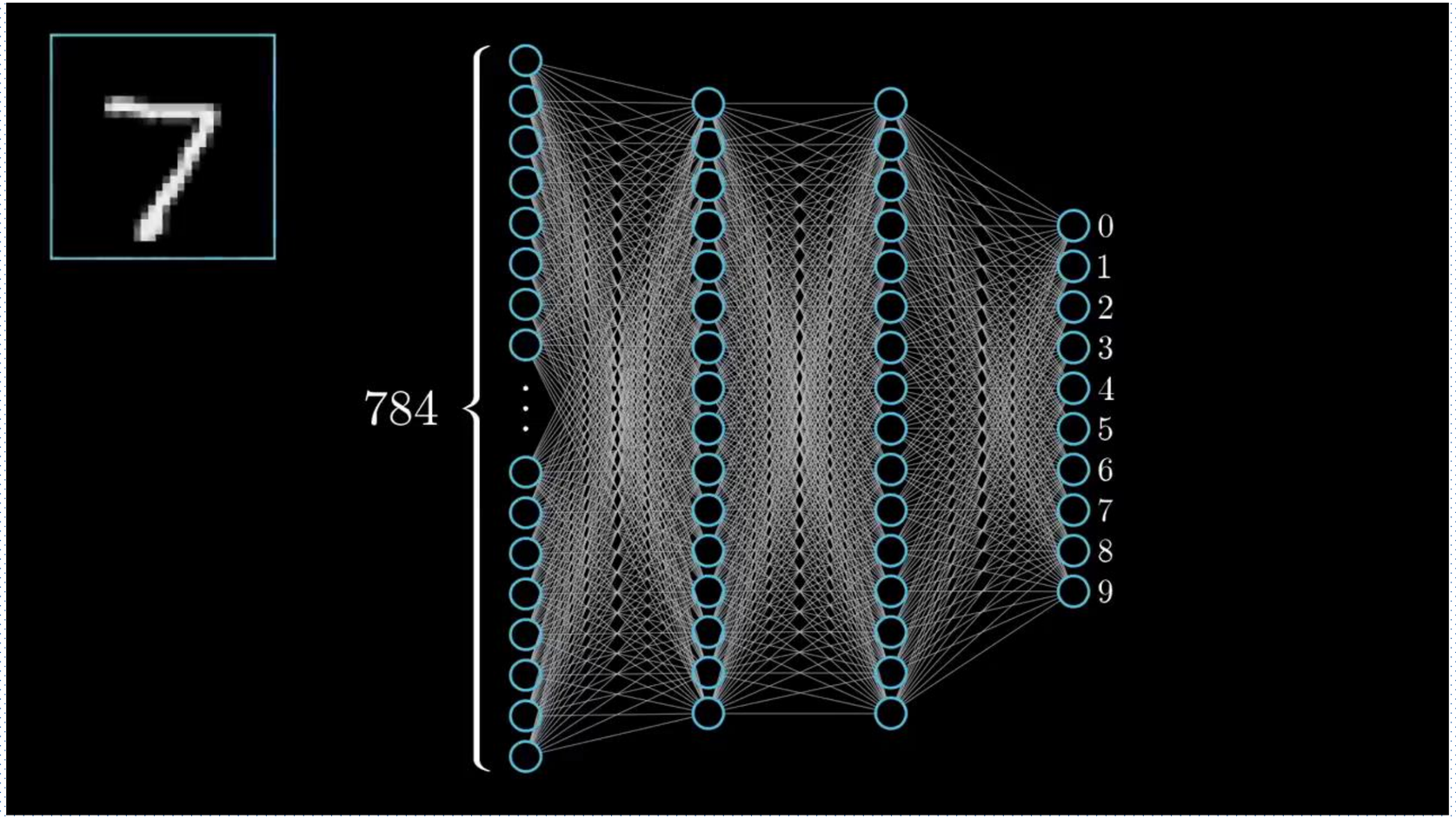
iv. Training is a central concept in machine learning as this is the process through which the model learns how to make sense of the input data.

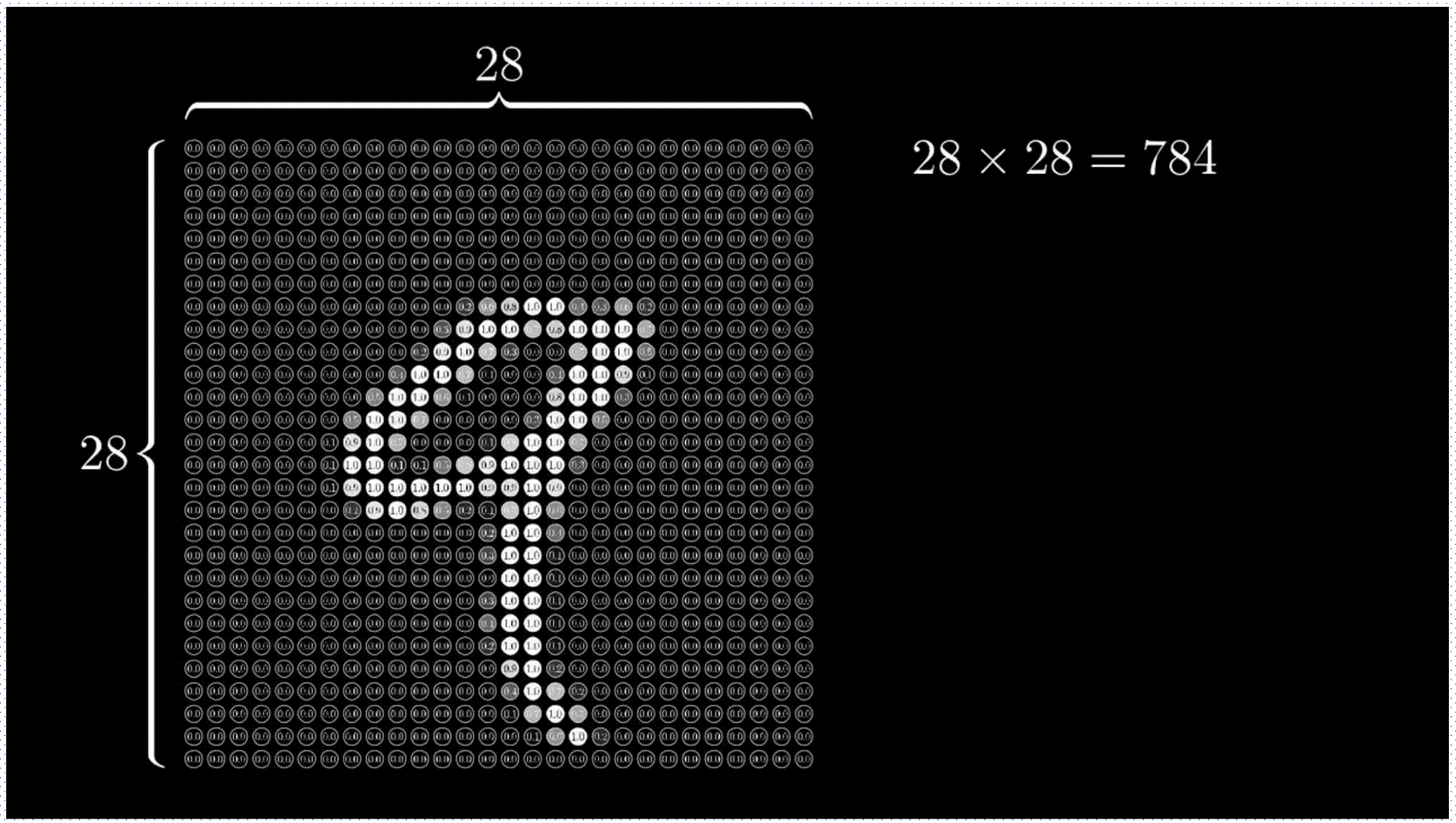
v. Once we have trained our model we can simply feed it with data and obtain an output.




3

Neural Network to learn & recognize digits




28
28
28 × 28 = 784

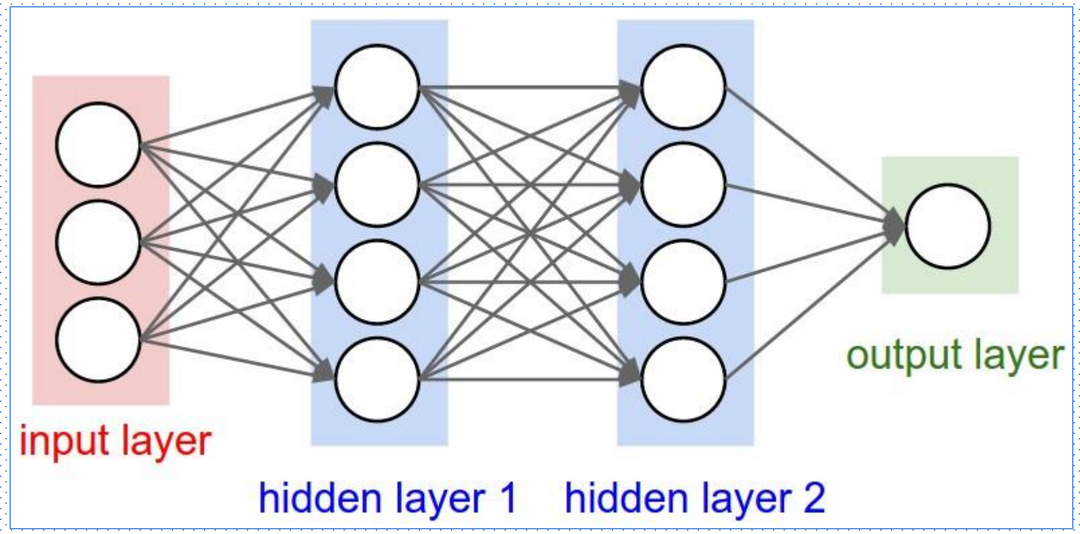
Layer

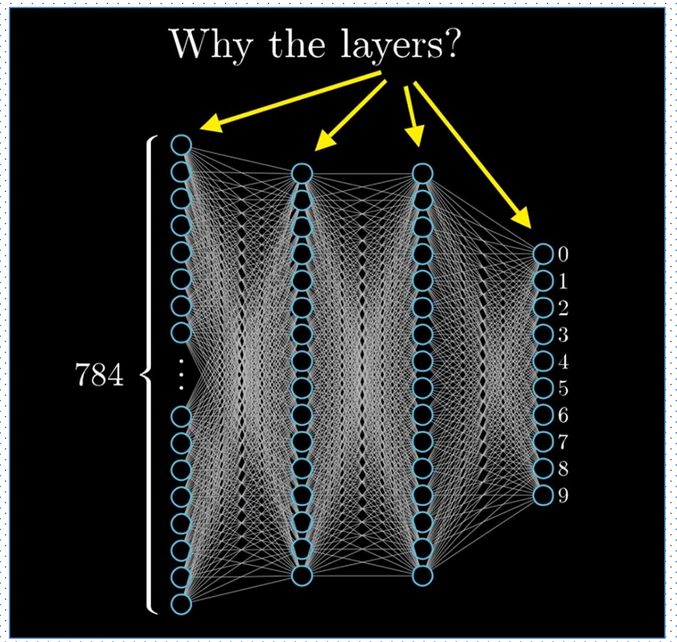
i. A layer is the highest-level building block in deep learning.

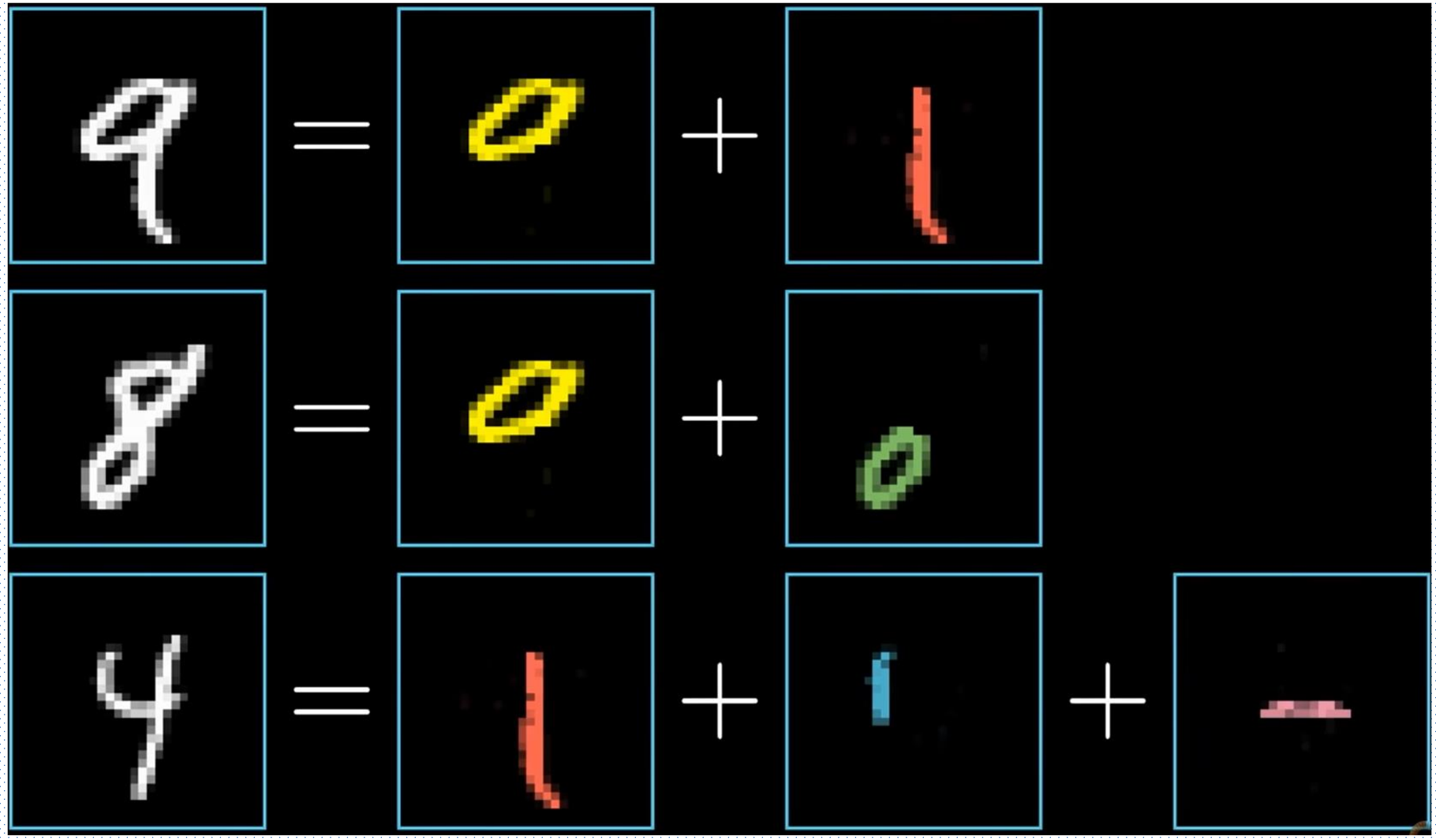
ii. A layer is a container that usually receives weighted input, transforms it with a set of mostly non-linear functions and then passes these values as output to the next layer.

iii. Single layer neural networks are very limited for simple tasks, deeper NN can perform far better than a single layer.

iv. Hidden layers, simply put, are layers of mathematical functions each designed to produce an output specific to an intended result.







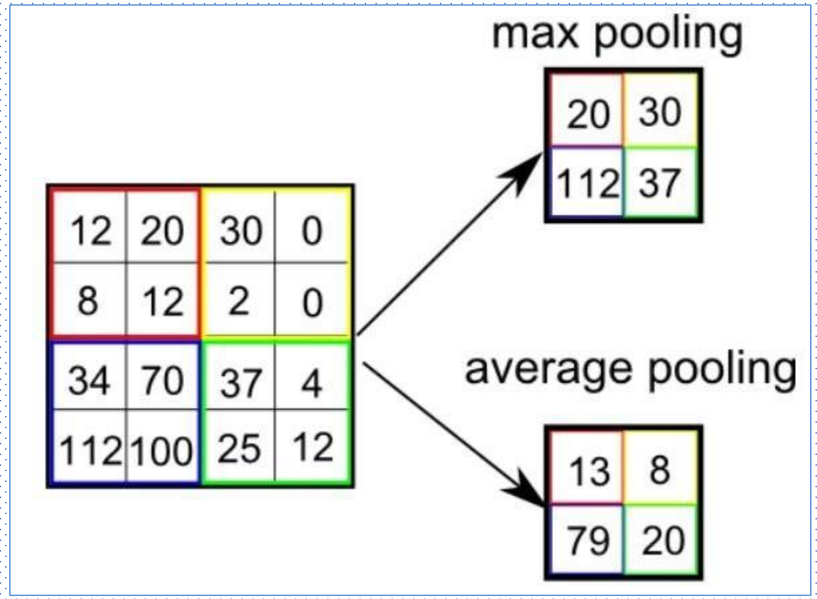
Pooling Layer

i. A pooling layer is another building block of a CNN.

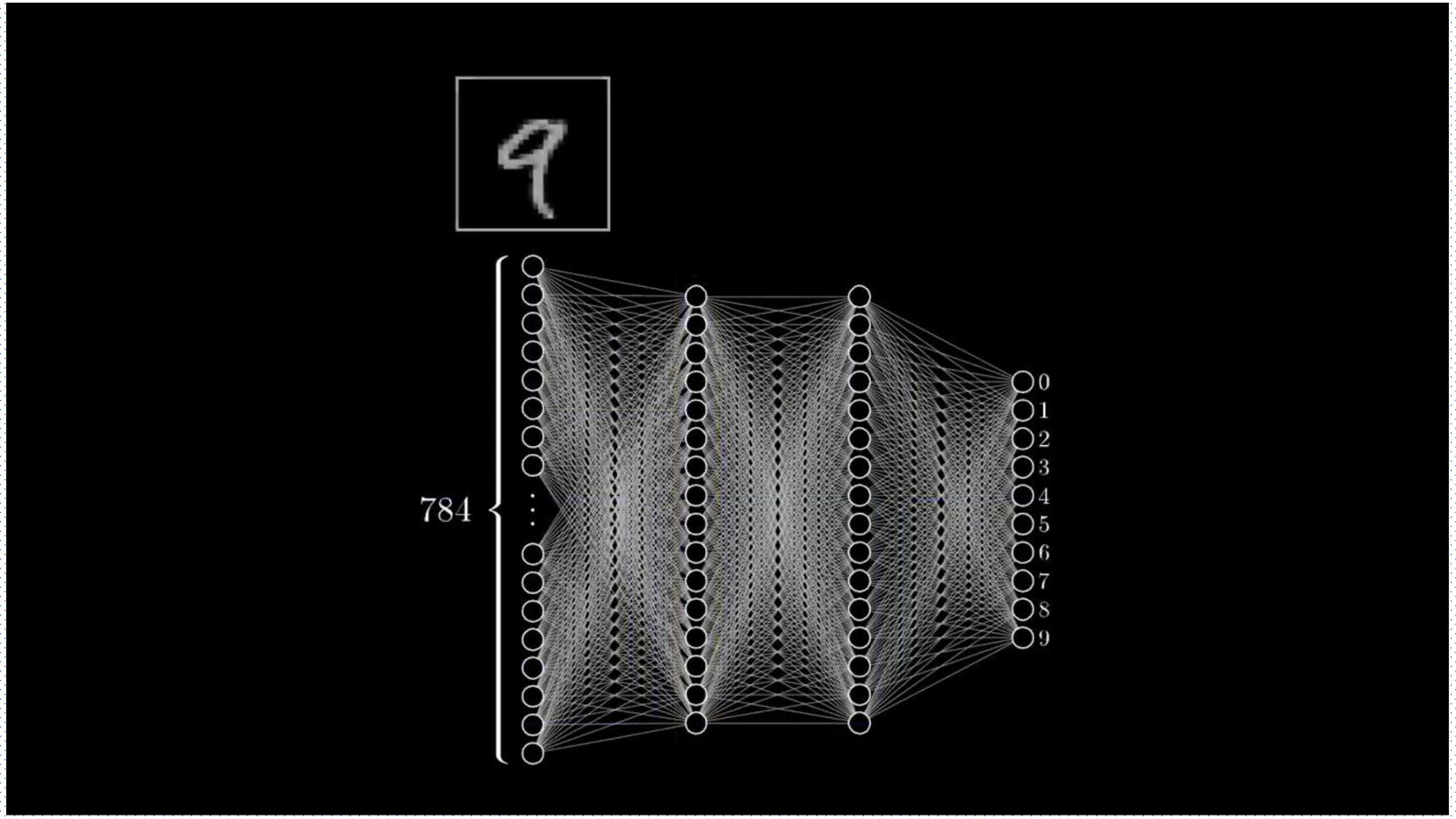
ii. Its function is to progressively reduce the spatial size of the representation to reduce the amount of parameters and computation in the network.

iii. Pooling layer operates on each feature map independently.

iv. The most common approach used in pooling is max pooling.



Graphical representation of hidden layers (arbitrary)

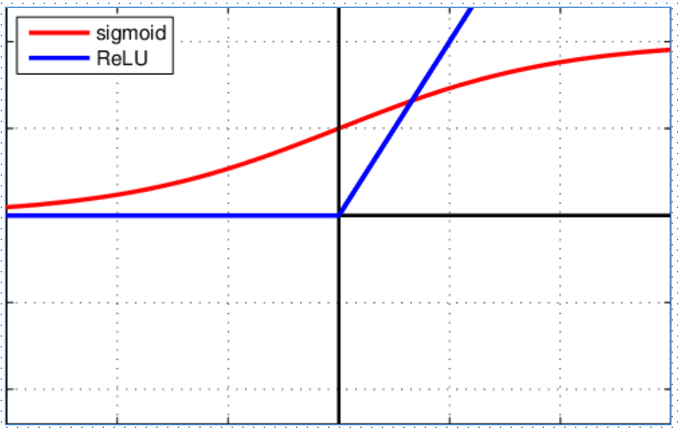


Rectifier

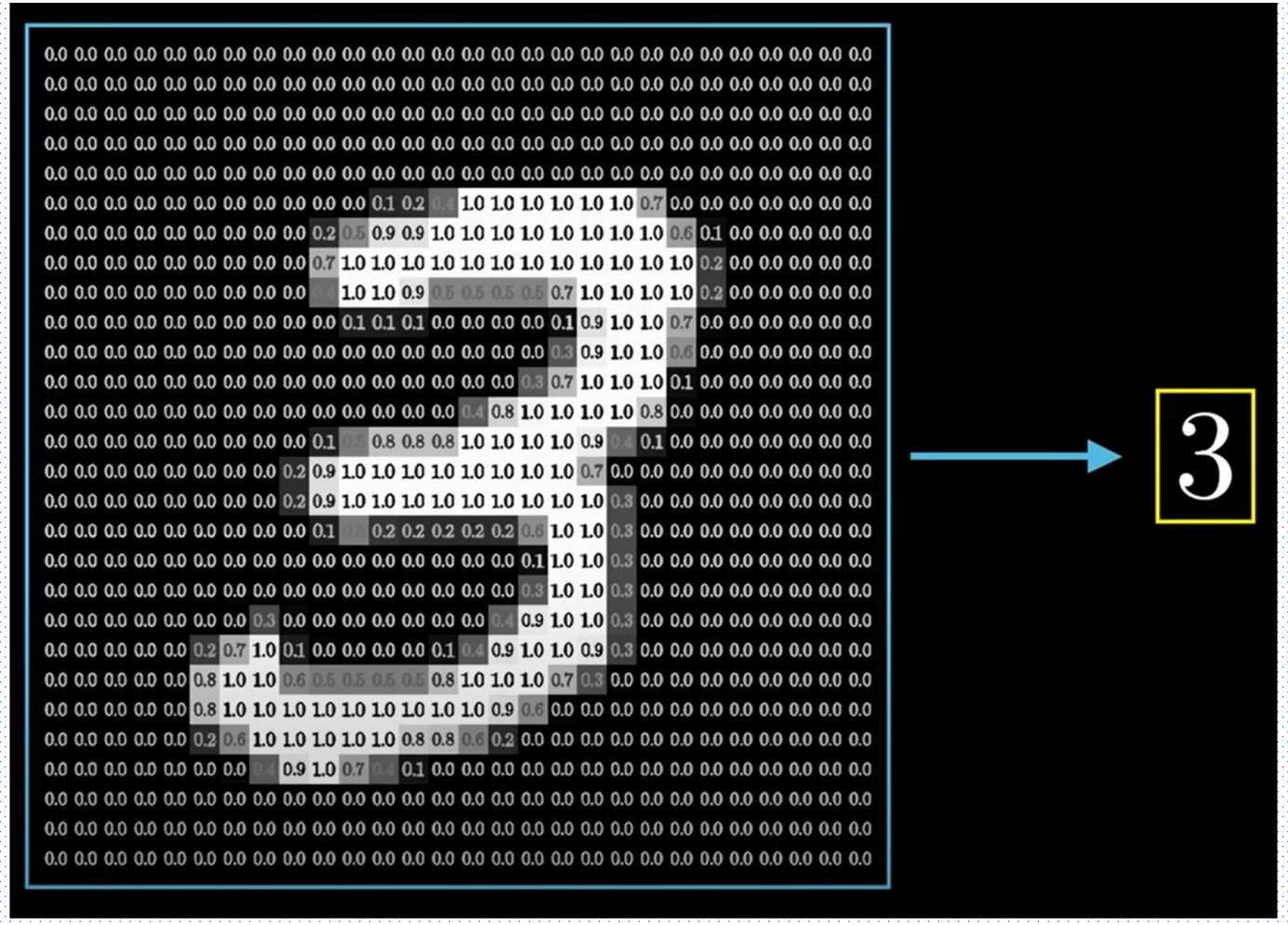
i. Mainly implemented in *hidden layers* of the Neural network.

ii. We need some function that pushes the real number line in between 0 and 1. Common function that does this is called sigmoid function.

iii. A unit employing the rectifier is also called a rectified linear unit (**ReLU**). Rectified linear units find applications in computer vision and speech recognition using deep neural nets.

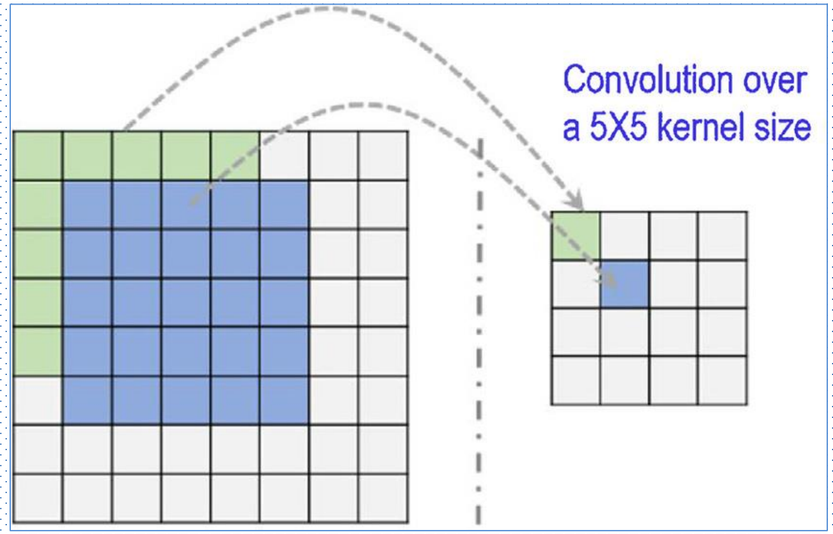


Activation Function



Kernel Size

i. Kernel Size: In image processing a Kernel is simply a 2-dimensional matrix of numbers. While this matrix can range in dimensions.

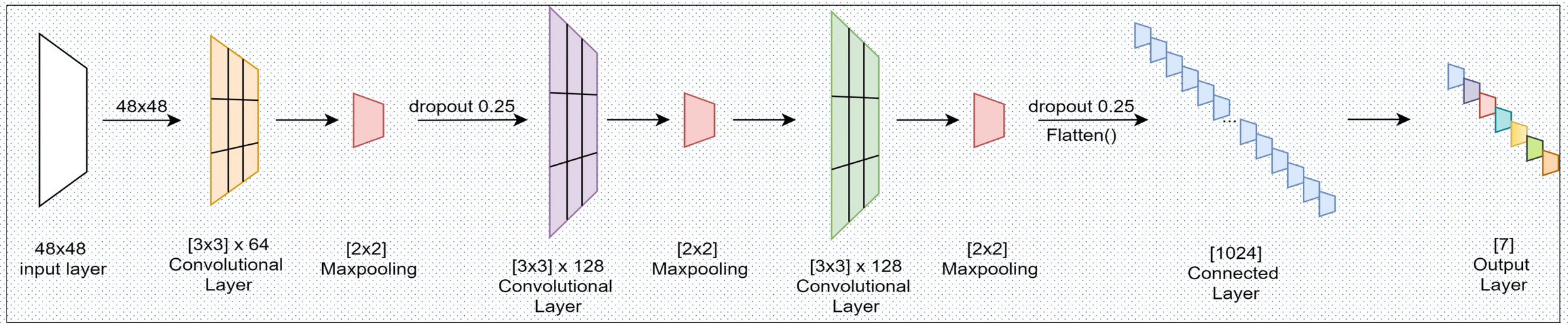


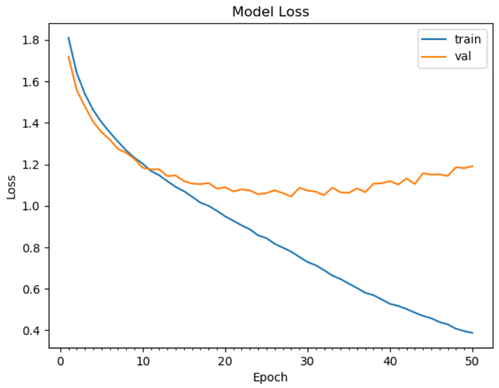
What are Softmax Scores?

Softmax functions are multi-class sigmoid, meaning they are used in determining probability of multiple classes at once. Since the outputs of a Softmax function can be interpreted as a probability
(i.e. they must sum to 1), a Softmax layer is typically the final layer used in neural network functions.

Implementation

Neural Network Architecture





The software detects emotions on all faces in the webcam feed. With a simple 4-layer CNN, the test accuracy reached 73.2% in 50 epochs.

Ingredients of Training Algorithm

Data

Collection of Dataset.

Model

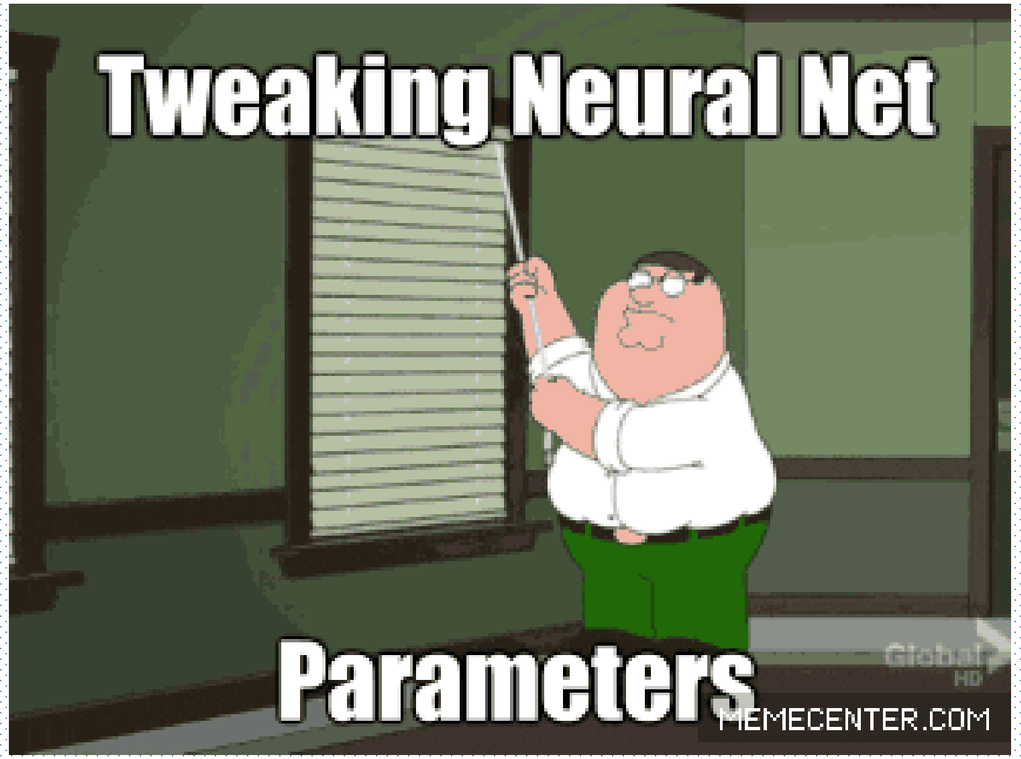
Create a model. Find some coefficients. A model, loosely speaking, is a simplification of some thing or process.

Objective Function

This function, taking data and model parameters as arguments, can be evaluated to return a number. Classification.

Optimization Algorithm

A function that minimizes an error or one that maximizes reward over punishment. By using Adam Algorithm.


Tweaking Neural Net
Parameters
MEMECENTER.COM

Ingredients of Training Algorithm

Data

Collection of Dataset.

Model

Create a model. Find some coefficients. A model, loosely speaking, is a simplification of some thing or process.

Objective Function

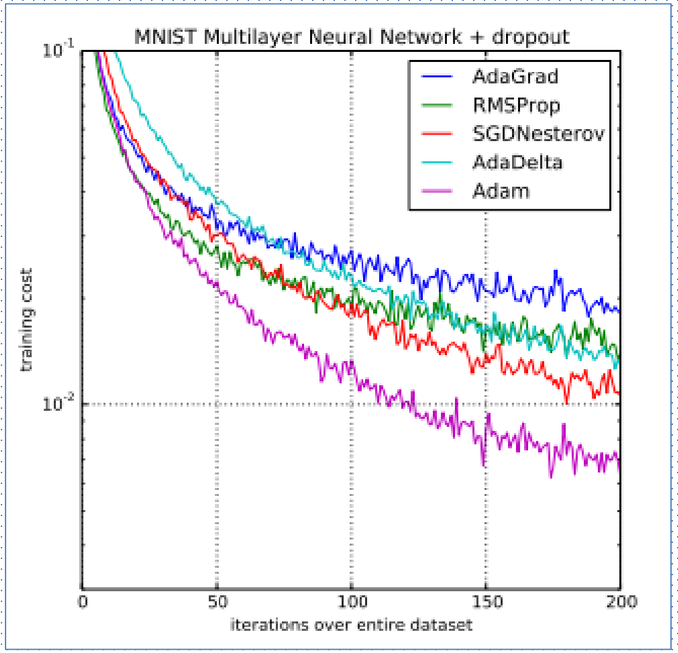
This function, taking data and model parameters as arguments, can be evaluated to return a number. Classification.

Optimization Algorithm

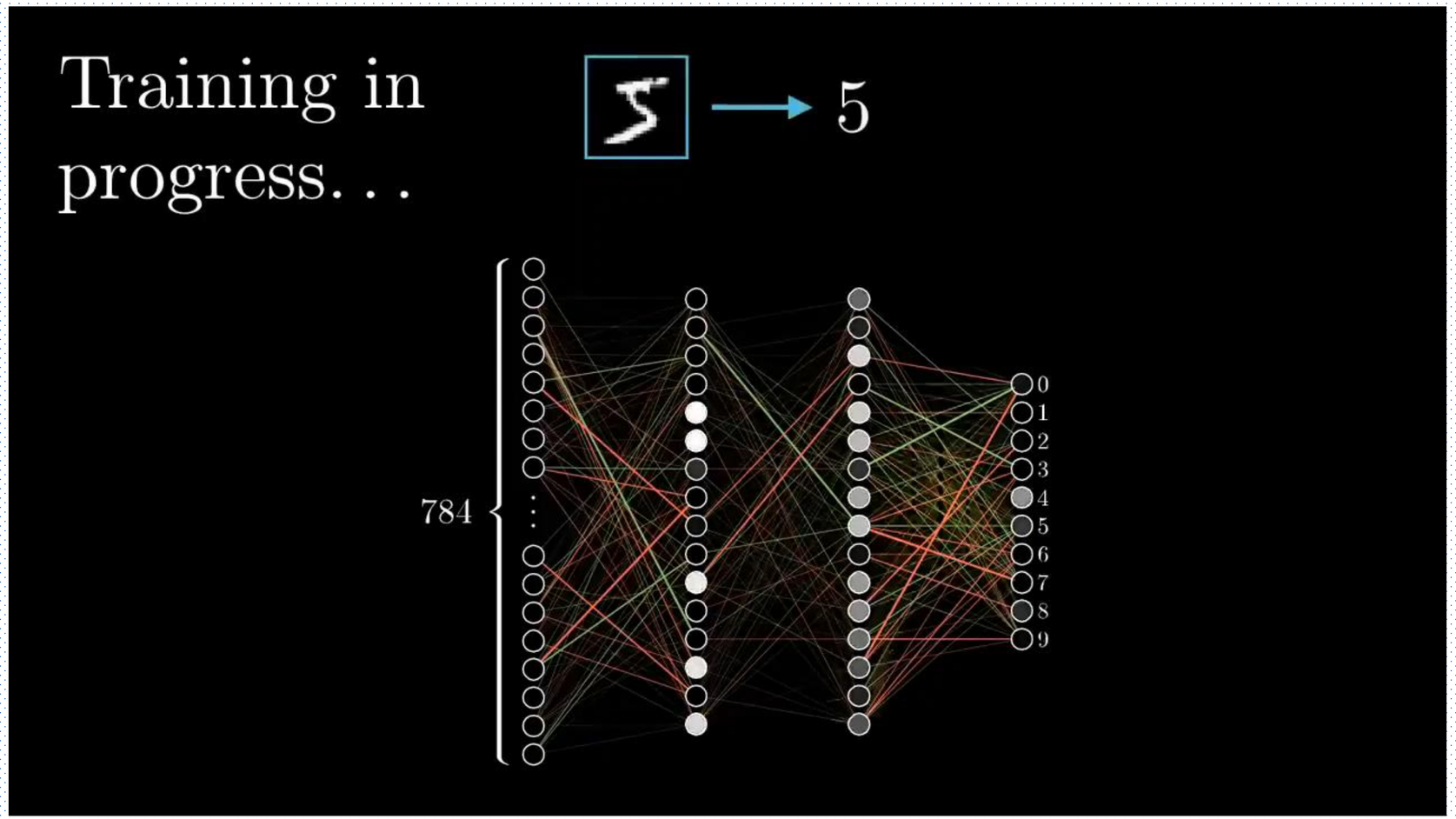
A function that minimizes an error or one that maximizes reward over punishment. By using Adam Algorithm.

Adam – Optimizer function

i. **Adam** is an **optimization** algorithm that can be used instead of the classical **stochastic gradient descent** procedure to update network weights iterative based in training data.



Training



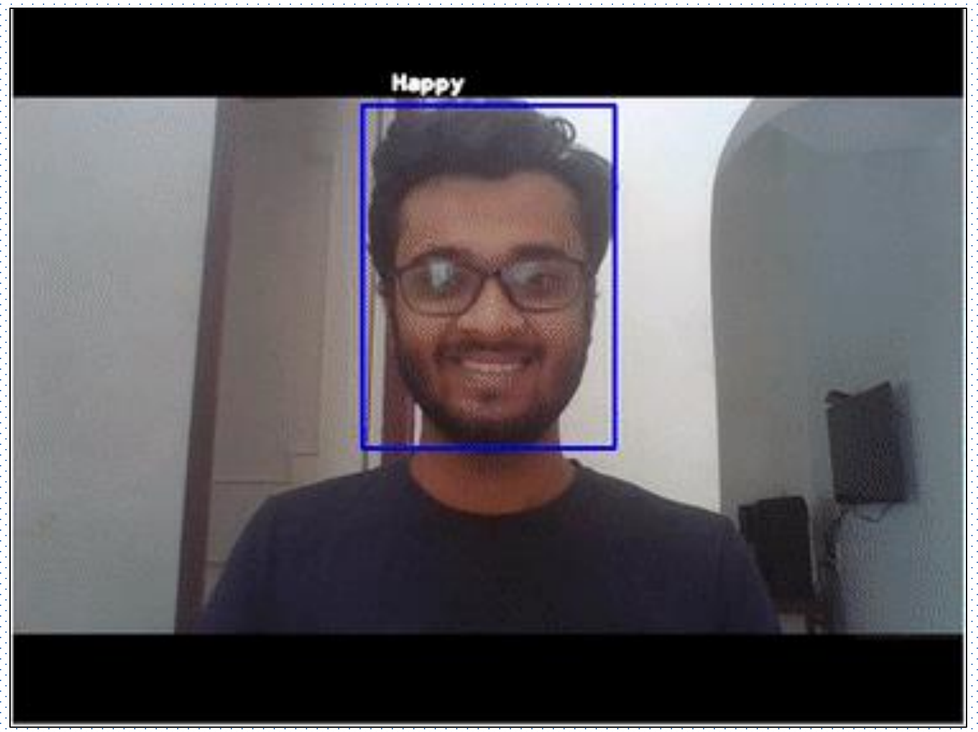
Output or Testing the model

i. With use of the OpenCV face recognition, the appearing face(s) from real-time video is tracked, extracted, and scaled to usable 48x48 input.

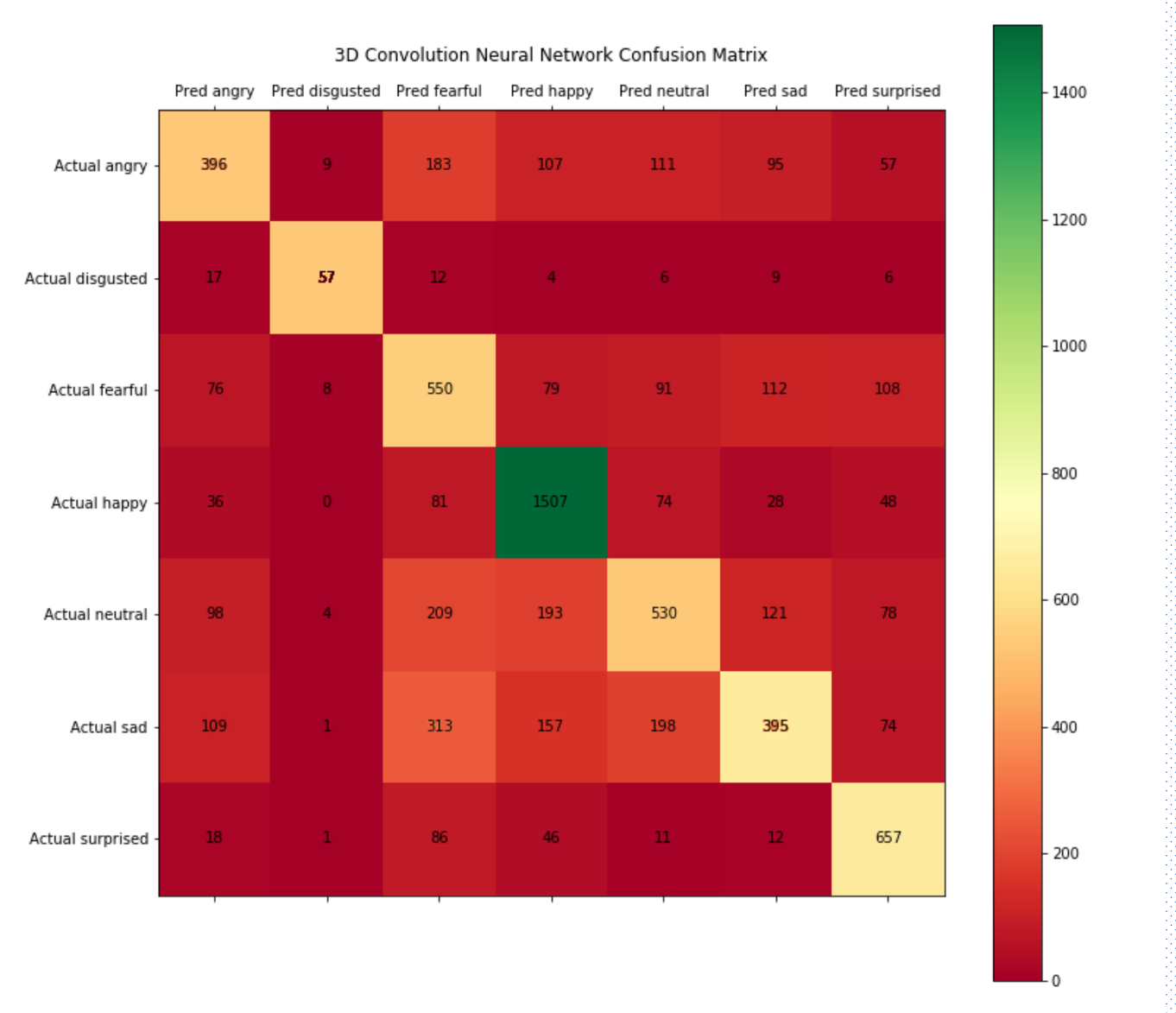
ii. This data is then fed to the input of the neural network model, which in its turn returns the values of the output layer.

iii. The output with the highest value is assumed to be the current emotion of the user, and is depicted by a green square box on the face(s).

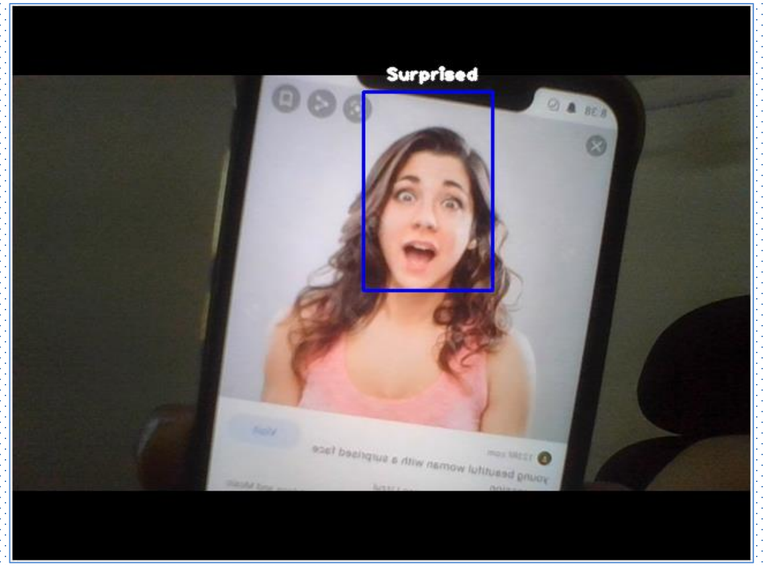
iv. Though, it encounters problems when shadows are present on the face of the subject. All emotions are easily recognized when acted by the user.

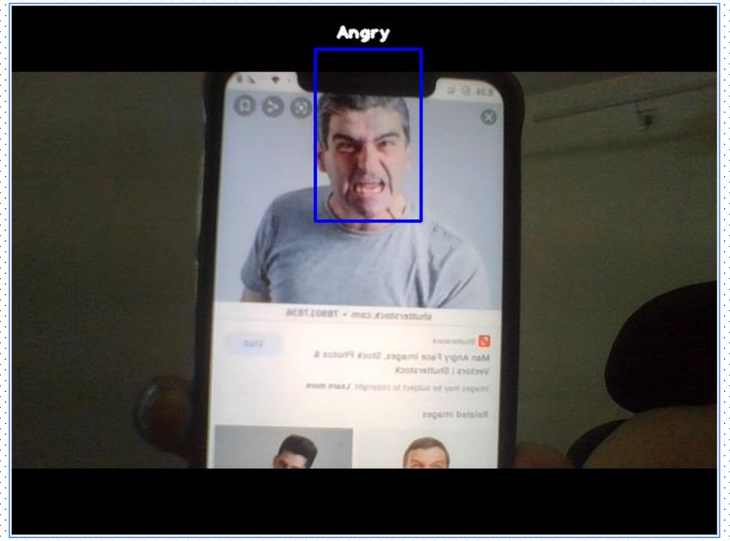


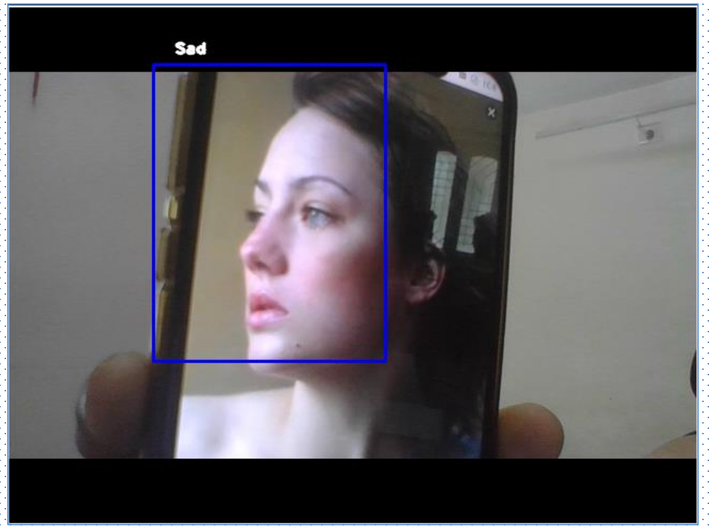
Confusion Matrix for Validation

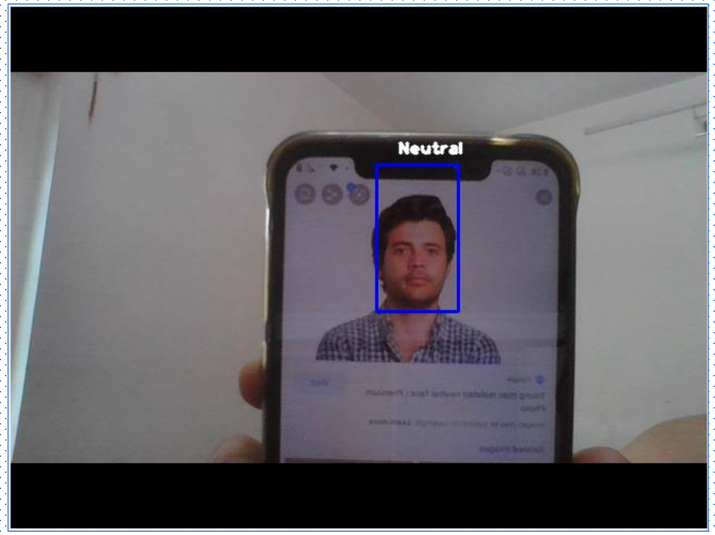


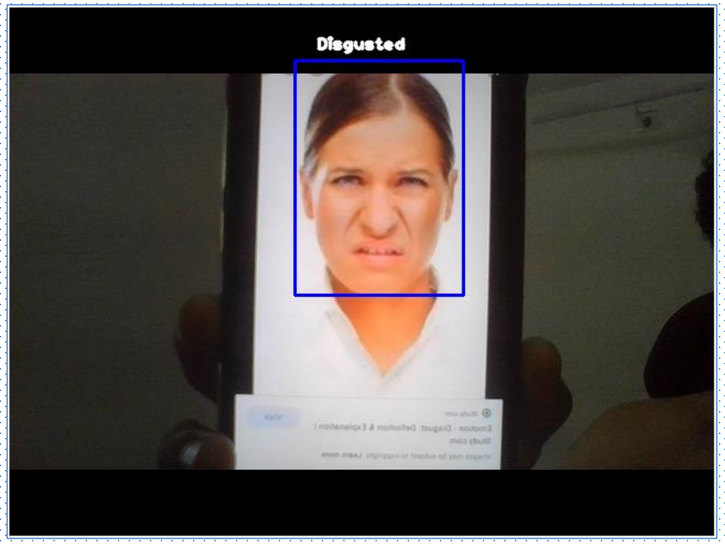
The system recognizing different emotions

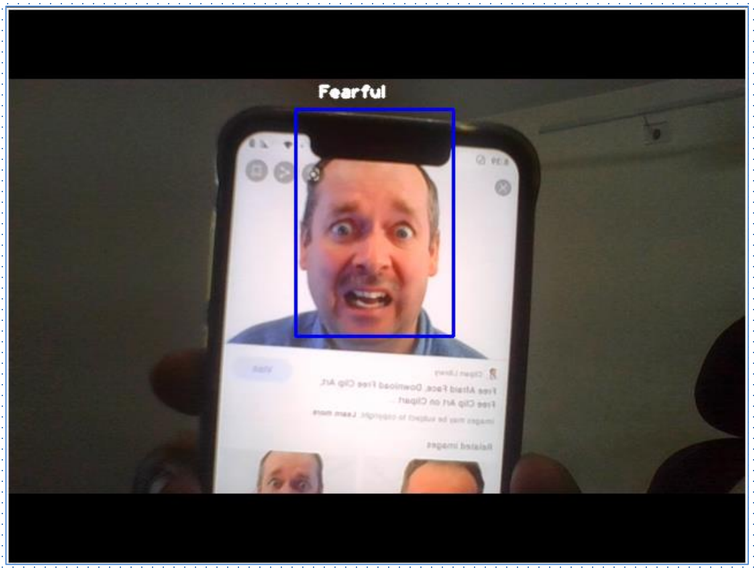












Limitations

i. This does not imply it is reading your innermost feelings -- it only reads what you express outwardly.
ii. The accuracy is very low, as there is a lot of complexity in human faces and a human face can show multiple feelings at a time.
iii. Deploying the project is costly for Mass Surveillance

Data Mining

i. Data mining is a process of discovering patterns in large data sets involving methods at the intersection of machine learning, statistics, and database systems.

ii. While the detected emotions were collected and stored into the database, I performed analysis on the MySQL database

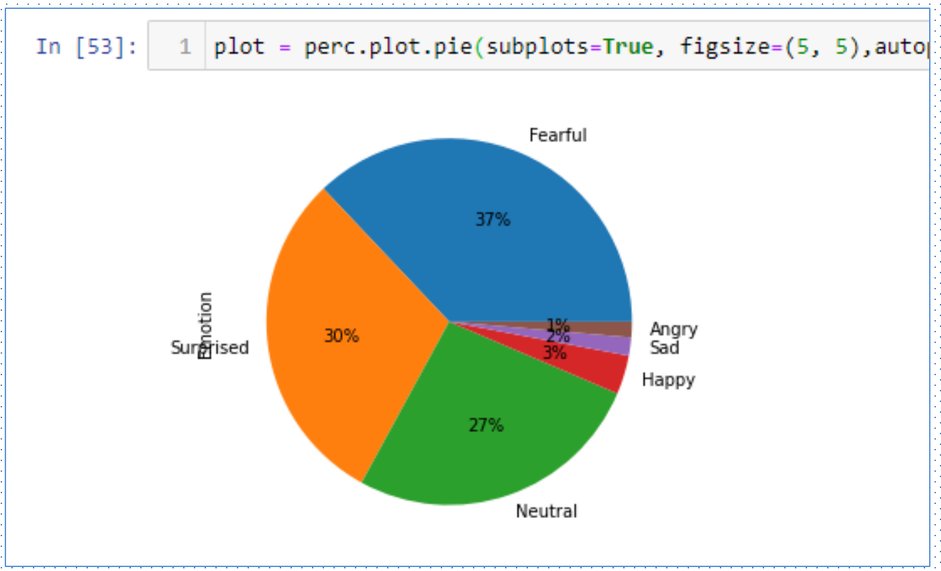
Results from Data Mining

i. The insights showed what I actually felt.

ii. Due to Covid19 situation I could not perform the tests in the real world, so I performed it on myself.

iii. Whenever I coded, my webcam recognized my emotions and stored it into the database.

iv. Found these fascinating insights after interpreting the database:
37% of the time I'm nervous - “whether my code will work or not!” And, 30% of the time I'm surprised – “why this code worked?!”



Timeline

(Work division with respect to time)

January

Data Collection

February

i. Data Collection
ii. Trainings on Regression, Clustering, Logistic Regression

March

i. Theory and study of research papers
ii. Creating Neural Network for MNIST data.
iii. Training of the Model

April & May

i. Testing Generating Report using Pandas and SQL Tweaking and Optimizing the Model
ii. Testing

Thank You

I42 Vraj Kotwala
(622079)