# Capstone Project

## Face Emotion Recognition

By : Mohd Atif Ansari

# Face emotion recognition

## Introduction:

Facial emotion recognition is the process of detecting human emotions from facial expressions. The human brain recognizes emotions automatically, and software has now been developed that can recognize emotions as well. This technology is becoming more accurate all the time, and will eventually be able to read emotions as well as our brains do. AI can detect emotions by learning what each facial expression means and applying that knowledge to the new information presented to it. Emotional artificial intelligence, or emotion AI, is a technology that is capable of reading, imitating, interpreting, and responding to human facial expressions and emotions.


Fearful


Angry


Sad


Happy


Disgusted


Surprised

# WHAT KIND OF EMOTIONS CAN BE DETECTED AND RECOGNIZED?

| Emotion | Facial Expression |
| --- | --- |
| Anger | Lowered and burrowed eyebrows<br>Intense gaze<br>Raised chin |
| Joy | Raised corners of mouth into a smile |
| Surprise | Dropped jaw<br>Raised brows<br>Wide eyes |
| Fear | Open mouth<br>Wide eyes<br>Furrowed brows |
| Sadness | Furrowed brows<br>Lip corner depressor |

# Data Preparation:

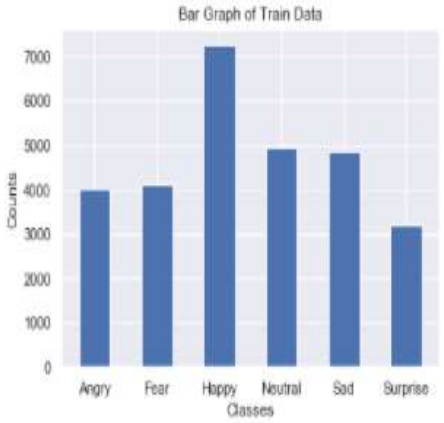
• The original FER2013 dataset in Kaggle is available as a single csv file. We can convert it to any format and use it for training/testing and provided this as the dataset in the previous section.

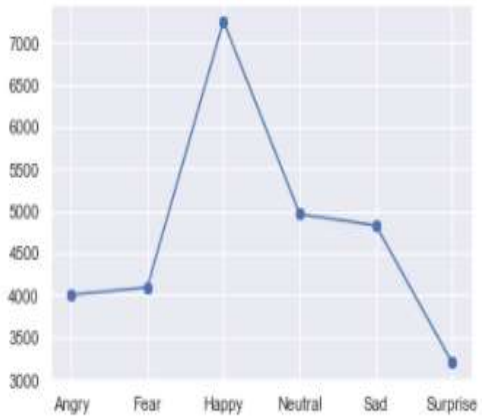
• In case you are looking to experiment with new datasets, you may have to deal with data in the csv format.

# Exploratory Data Analysis:

## Train:

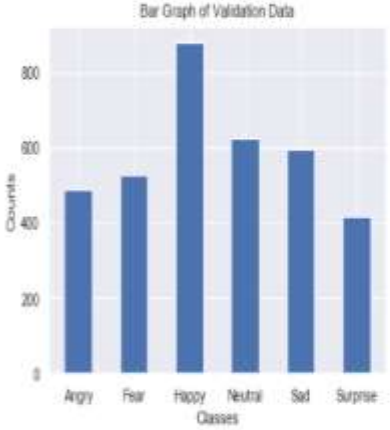
The Bar Graph and the scatter plot of the train data are as shown below:

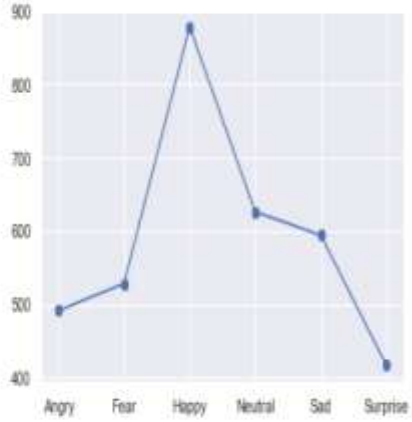



## Validation:

The Bar Graph and the scatter plot of the train data are as shown below:




- There are total 35887 files in this data set, out of which 28709 files are for training, the remaining files are for testing.

1. **KNOWLEDGE BASE**
2. This base contains images that are used for comparison and recognizing emotion variations. The images are stored in the database. Every time an input is given to the system, it finds a relevant image from its knowledge base by comparing the stored pictures and the input to come up with an output.
3. **PREPROCESSING AND RESIZE**
4. This step enhances the input and removes different types of noises. After that, the input image will be resized, typically with the use of the eye selection method.
5. **DIFFERENCE MEASUREMENTS**
6. During this step, the system will find any differences between the input image and the stored images and will finally lead to the emotion recognition step.
7. **EMOTION RECOGNITION**
8. This is the final step of the process. The comparison is made, and the final output is given depending on the differences found.

# Algorithm:

## Tools we need:

I am using the Visual studio code (VSCODE) for this project.

First, the **haar cascade** method is used to detect faces in each frame of the webcam feed.

The region of image containing the face is resized to **48x48** and is passed as input to the CNN.

The network outputs a list of **SoftMax scores** for the seven classes of emotions.

The emotion with maximum score is displayed on the screen.

# IMPLEMENTATION USING:

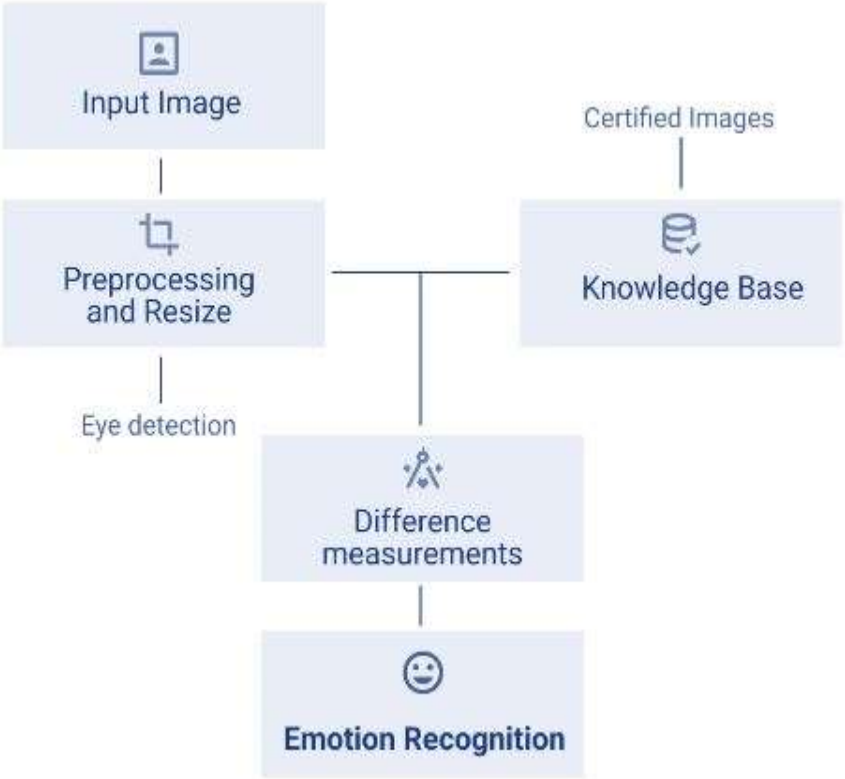
PYTHON , OPENCV , TENSORFLOW

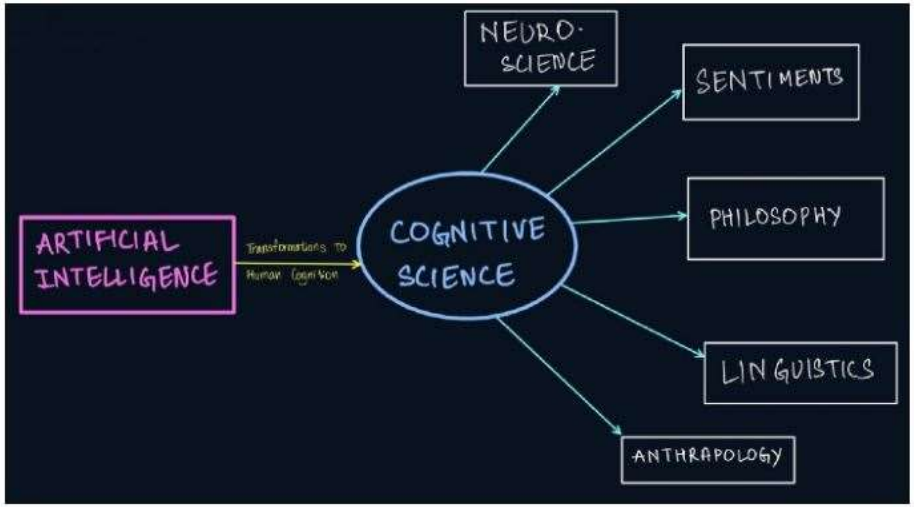
Import the library :

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt

from keras.layers import Flatten, Dense
from keras.models import Model
from keras.preprocessing.image import ImageDataGenerator , img_to_array, load_img
from keras.applications.mobilenet import MobileNet, preprocess_input
from keras.losses import categorical_crossentropy
import cv2
from tensorflow.keras.models import model_from_json
```

# How Face emotion work :



# Cognitive science:

AI

In terms of computing systems, cognitive science is the study of scientific processes that occur in the human brain. It is responsible for examining the functions of cognition, namely, perception of thoughts, languages, memory of the brain, reasoning, and processing received information. At a broader level, it is the study of intelligence and behavior.

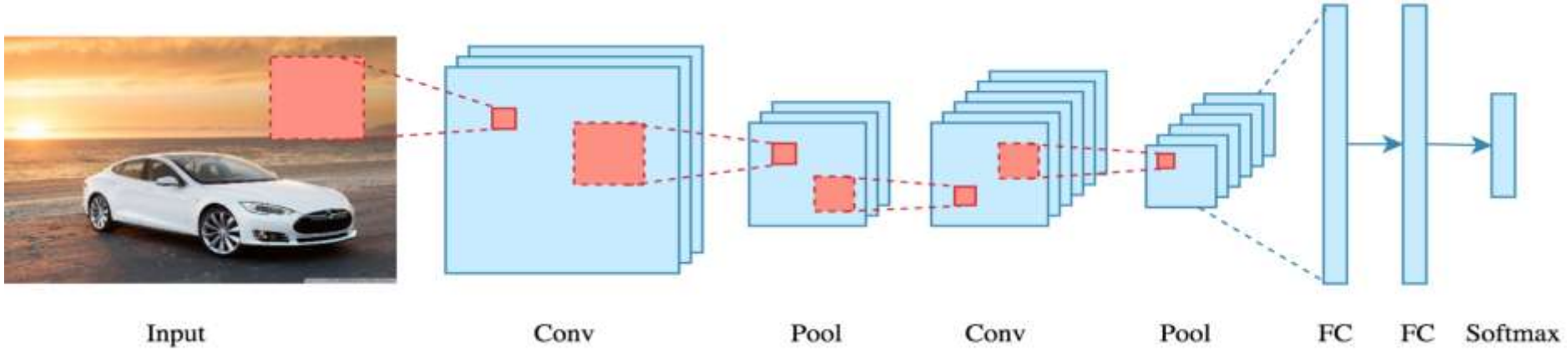
The goal of Cognitive Science is to study the human brain and understand its principles of intelligence. This is done with the hope that by building computer systems from the knowledge of human intelligence, machines will be able to mimic learning and develop intelligent behavior patterns like humans.

Cognitive Science operates in three distinct levels of analysis:

1.The computational theory: At this level, the goals of the analysis are specified and fed to the computer system. This could be an imitation of speech or an understanding of emotions.
2.Representation and algorithms: In general Machine Learning terms, this is the training stage. Here the ideal input and output scenarios are presented to the machine and algorithms are put in place that will ultimately be responsible for transforming input to output.
3.The hardware implementation: This is the final cognition phase. It is the enactment of the algorithm in the real world and analysis of its working trajectory against a human brain.

# What exactly is a CNN?

In deep learning, a **convolutional neural network** (**CNN/ConvNet**) is a class of deep neural networks, most commonly applied to analyze visual imagery. Now when we think of a neural network we think about matrix multiplications but that is not the case with ConvNet. It uses a special technique called Convolution. Now in mathematics **convolution** is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other.



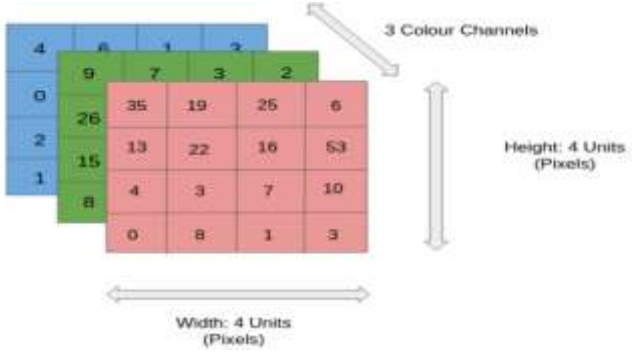

Bottom line is that the role of the ConvNet is to reduce the images into a form that is easier to process, without losing features that are critical for getting a good prediction.

# So there are various steps to achieve this CNN algorithm, they are as follows :

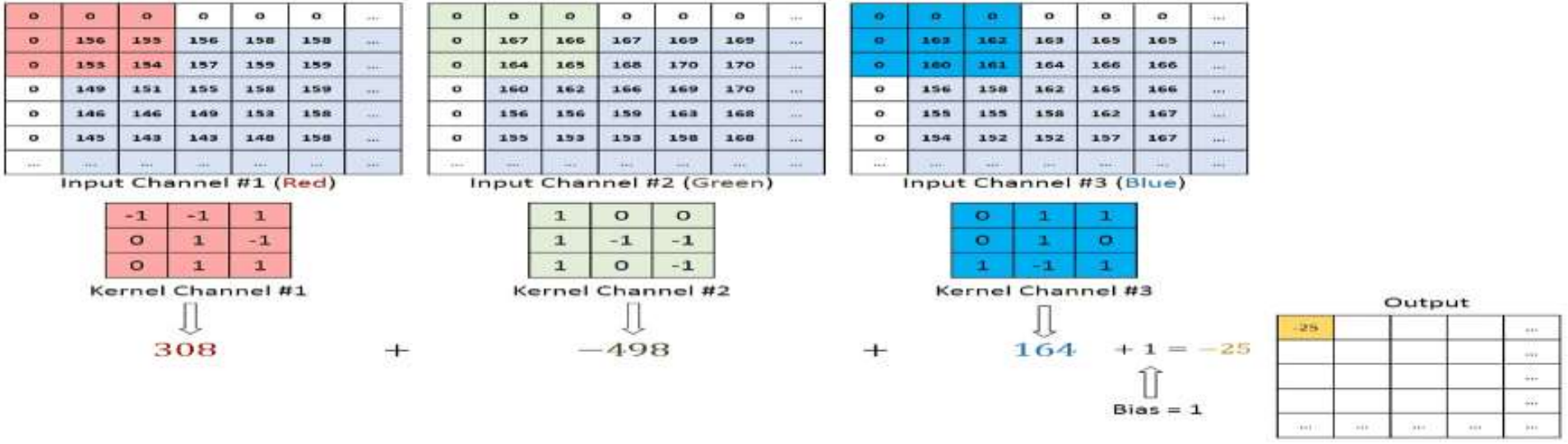
1. Convolution layer
2. Activation function
3. Pooling
4. Flattening
5. Full connection

# How does it work?

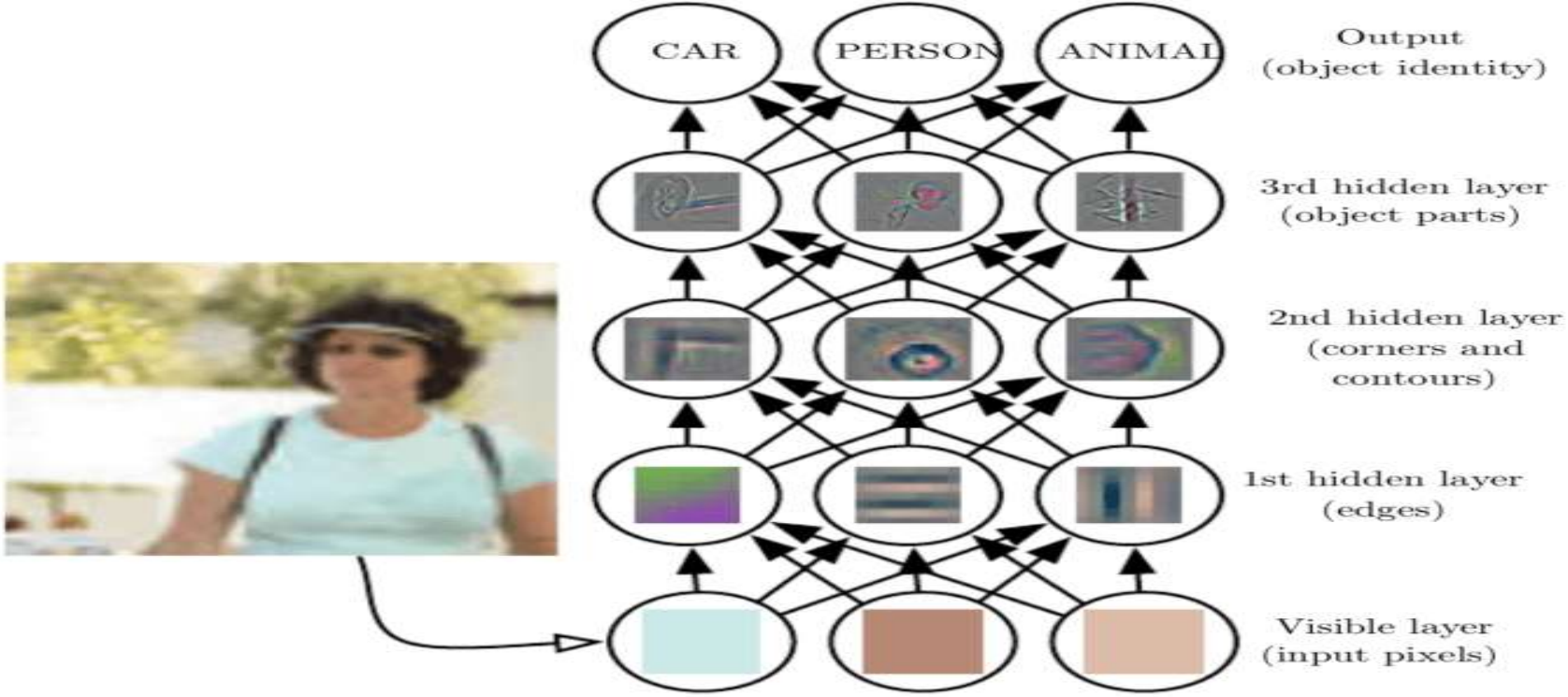
An RGB image is nothing but a matrix of pixel values having three planes whereas a grayscale image is the same but it has a single plane. Take a look at this image to understand more.




- In the case of RGB color, channel take a look at this animation to understand its working :



Based on the activation map of the final convolution layer, the classification layer outputs a set of confidence scores (values between 0 and 1) that specify how likely the image is to belong to a "class." For instance, if you have a ConvNet that detects cats, dogs, and horses, the output of the final layer is the possibility that the input image contains any of those animals.

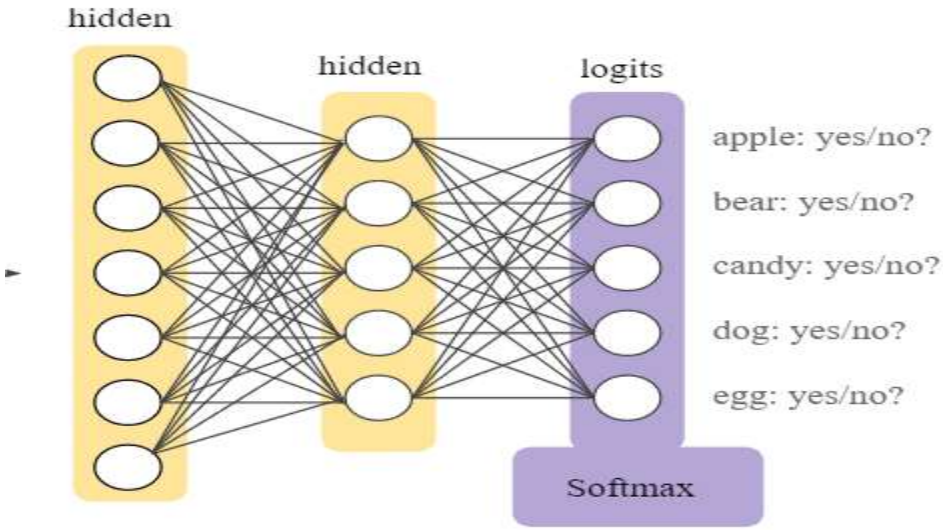


# Fully Connected Output layer:

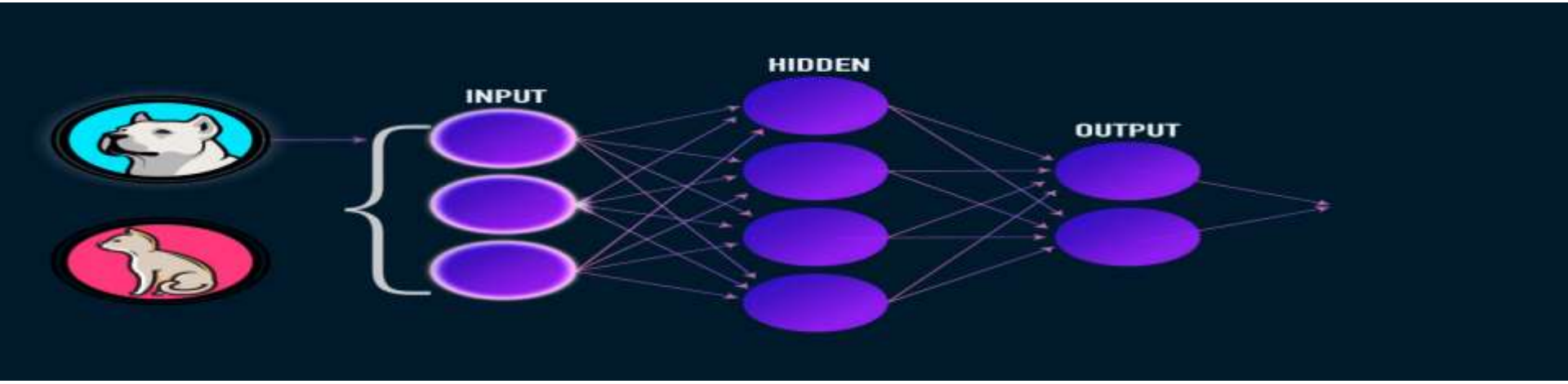
The final layer of the CNN model contains the results of the labels determined for the classification and assigns a class to the dataset (input)

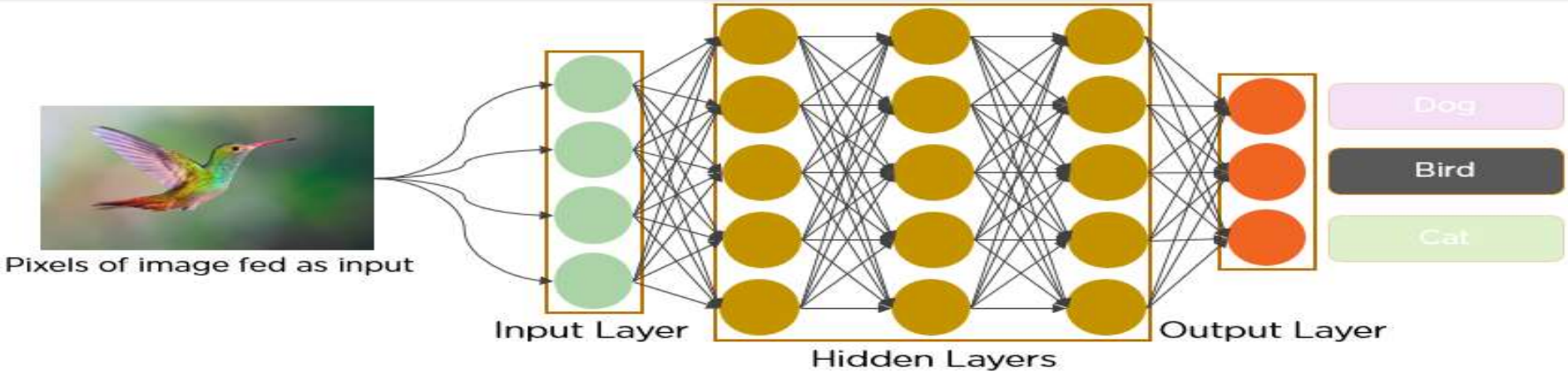
# SoftMax:

The reason why SoftMax is useful is that it converts the output of the last layer in your neural network into what is essentially a probability distribution. It is mainly used to normalize neural networks output to fit between zero and one. It is used to represent the certainty “probability” in the network output.

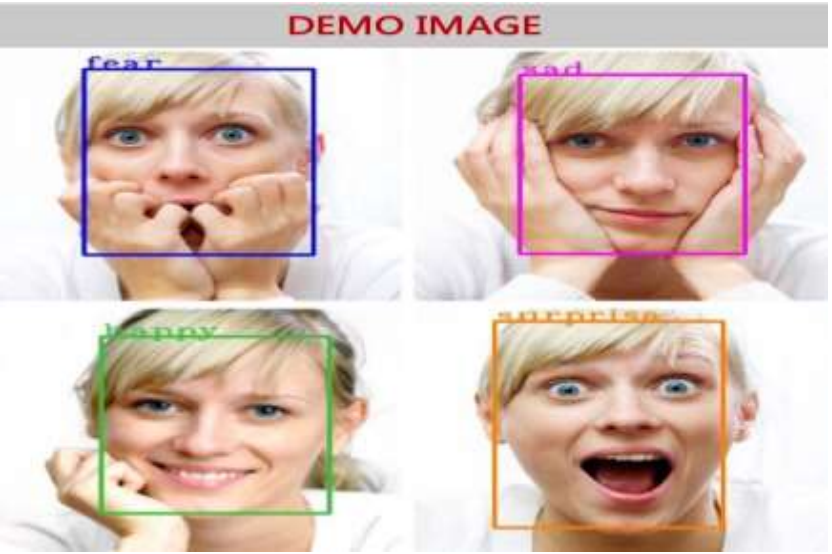


# The below animation will explain how CNN works :



# TYPES OF FACIAL EXPRESSION:





**Results:** Of **8** people detected, **1 (13%)** is happy.

# DIFFERENT FIELDS FOR EMOTION RECOGNITION

- Choosing what model or training data to use is not the end of it, as there are different inputs that can be used or given to the system to analyze. Below are the main fields for emotion recognition:

**Video** : It is one of the major datasets in Affective Computing that are available, and lots of research is still ongoing to understand how to use them in emotion recognition. An example of modern software for this field is cameras on mobile phones.

**Image** : Just like video, it is one of the available major sets in emotion recognition. Some research stated that classified emotions in images could be applied in automatic tagging of pictures with emotional categories and sorting video sequences into various genres.

**Speech** : Usually, speeches are transcribed into texts to be able to analyze them, but this method would not be applicable in emotion recognition. Various research is still ongoing to explore the usage of speech instead of transcribed texts for emotion recognition applications.

**Conversation** : Emotion recognition in conversation focuses on acquiring emotions from discussions between two or more persons. Datasets under this are commonly from free samples from social platforms. However, there are still a lot of challenges in this field, including the presence of sarcasm in a dialogue, the emotion shift on the interlocutor, and conversational-context modeling.

# WHAT IS EMOTION RECOGNITION USED FOR TODAY?

Nowadays, emotion recognition is used for various purposes that some people do not even notice on a daily basis. Here are some of the areas that would show that emotion recognition is beneficial:

**SECURITY MEASURES :** Emotion recognition is already used by schools and other institutions since it can help prevent violence and improves the overall security of a place.

**HR ASSISTANCE :** There are companies that use AI with emotion recognition API capabilities as HR assistants. The system is helpful in determining whether the candidate is honest and truly interested in the position by evaluating intonations, facial expressions, keywords, and creating a report for the human recruiters for final assessment.

**CUSTOMER SERVICE :** There are systems launched nowadays that are installed in customer service centers. Using cameras equipped with artificial intelligence, the customer's emotions can be compared before and after going inside the center to determine how satisfied they are with the service they've received. And if there is a low score, the system can advise the employees to improve the service quality.

- **DIFFERENTLY ABLED CHILDREN :** There is a project using a system in Google Glass smart glasses that aims to help autistic children interpret the feelings of people around them. When a child interacts with other people, clues about the other person's emotions are provided using graphics and sound.
- **AUDIENCE ENGAGEMENT :** Companies are also using emotion recognition to determine their business outcomes in terms of the audience's emotional responses. Apple also released a new feature in their iPhones where an emoji is designed to mimic a person's facial expressions, called Animoji.
- **VIDEO GAME TESTING :** Video games are tested to gain feedback from the user to determine if the companies have succeeded in their goals. Using emotion recognition during these testing phases, the emotions a user is experiencing in real-time can be understood, and their feedback can be incorporated in making the final product.

- **HEALTHCARE :** The healthcare industry sure is taking advantage of facial emotion recognition nowadays. They use it to know if a patient needs medicine or for physicians to know whom to prioritize in seeing first.

# WHAT ARE THE CHALLENGES OF EMOTION DETECTION & RECOGNITION?

Just like any developing technology, emotion recognition is not perfect and has its imperfections and challenges. One of the challenges is that datasets are labeled by the people, and different persons can read and interpret emotions in different ways. Also, some visible visual cues like furrowed eyebrows can mean other emotions aside from anger, and other cues may be subtle hints of anger, although they are not obvious.

Another issue faced by this technology is when detecting emotions from people of different colors. There are models that detect more anger in black people. This means that training sets need to be more diverse, and experts are already doing what they can to fix this.

# ARE EMOTION DETECTORS EFFECTIVE?

- The emotion recognition technologies are not near to perfect. Although they can truly detect emotions, there are still issues and challenges they face and produce. For example, a system can consider subtle emotions and expressions more alarming than those which actually are. Also, since it automatically links facial expressions to certain emotions, it cannot distinguish which ones are genuine and which are not and could be deceived easily.

- AI also cannot easily understand the differences in cultures when expressing emotions and therefore making it hard to produce correct conclusions.

- Overall, biases in emotion detection like these, if not addressed, can lead to severely wrong assumptions and cause major misunderstandings. Risks of misinterpretation like these should be considered, and further improvements shall be made.

# SUMMARY: IS THERE A FUTURE OF ARTIFICIAL INTELLIGENCE (AI) EMOTION RECOGNITION

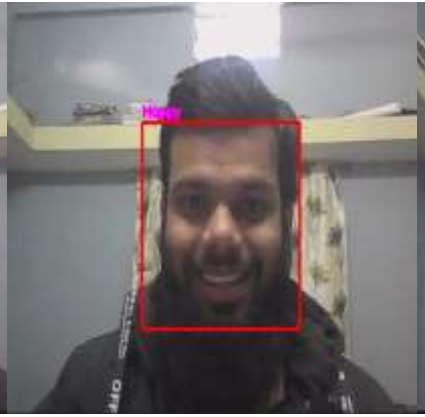
- **A study published in 2015 showed that recognition results improve when confounding factors are removed from the input images and the lighting is adequate.**

- AI Emotion Recognition technology is continuously being studied and improved to solve and provide solutions to arising risks and issues and to avoid any cultural and ethical problems in the society.

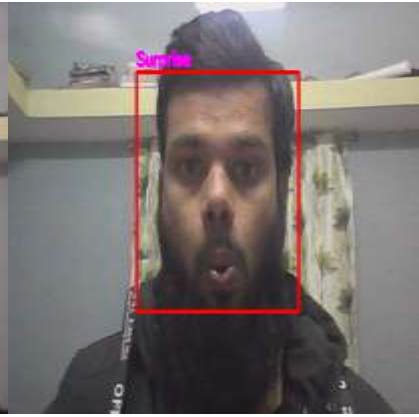
# Result :

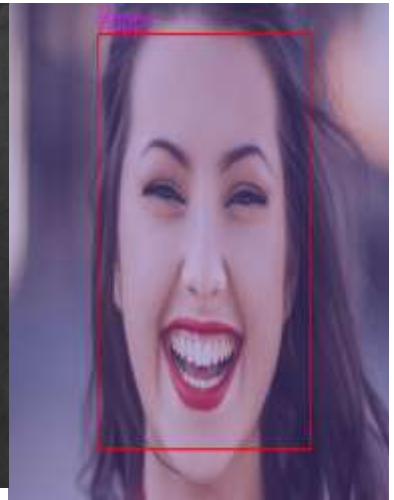

Neutral

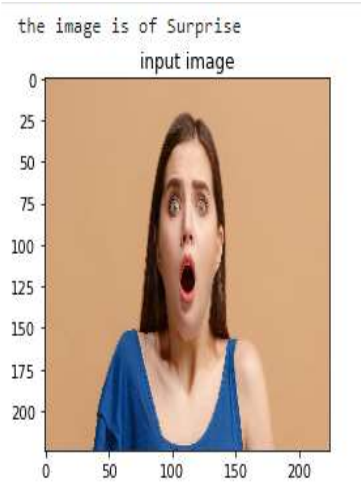

Angry


Happy

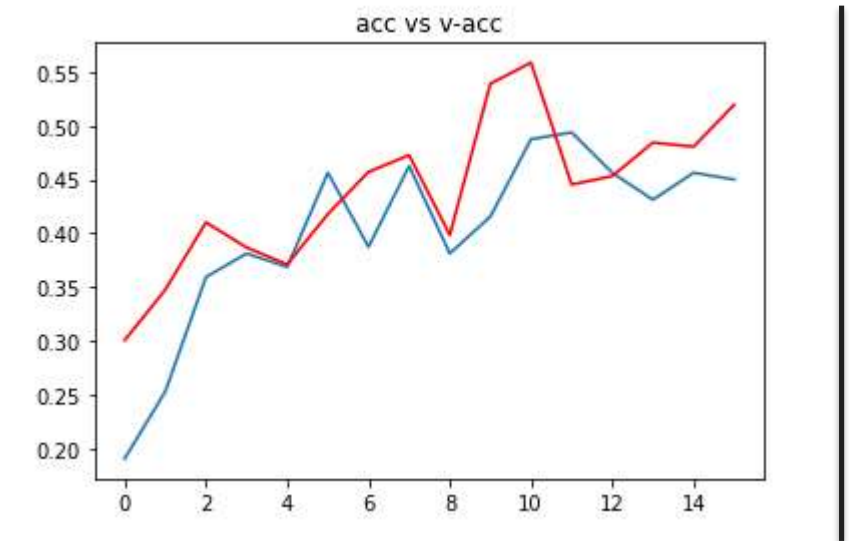

Sad

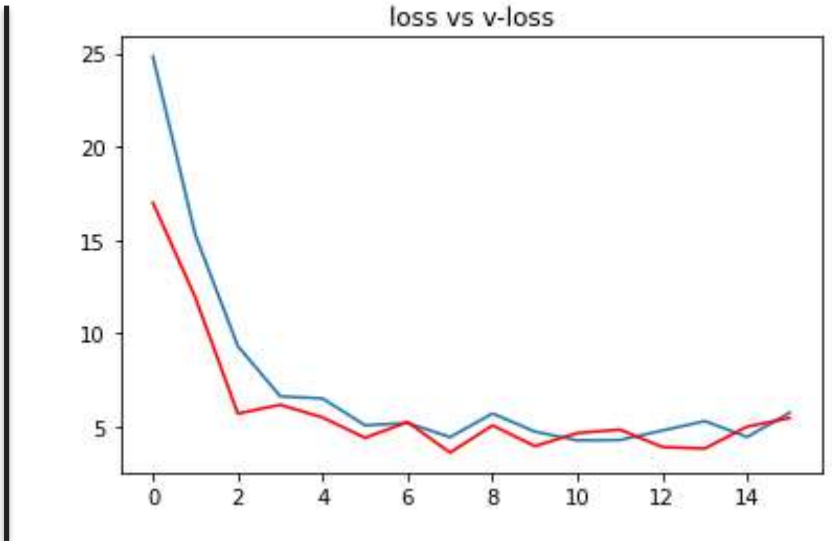

Surprise




the image is of Surprise
input image
0
25
50
75
100
125
150
175
200
0
50
100
150
200

```
10/10 [==============================] - ETA: 0s - loss: 4.7687 - accuracy: 0.4563
Epoch 00013: val_accuracy did not improve from 0.55859
10/10 [==============================] - 45s 5s/step - loss: 4.7687 - accuracy: 0.4563 - val_loss: 3.8812 - val_accuracy: 0.4531
Epoch 14/30
10/10 [==============================] - ETA: 0s - loss: 5.2748 - accuracy: 0.4313
Epoch 00014: val_accuracy did not improve from 0.55859
10/10 [==============================] - 42s 4s/step - loss: 5.2748 - accuracy: 0.4313 - val_loss: 3.8038 - val_accuracy: 0.4844
Epoch 15/30
10/10 [==============================] - ETA: 0s - loss: 4.4247 - accuracy: 0.4563
Epoch 00015: val_accuracy did not improve from 0.55859
10/10 [==============================] - 43s 4s/step - loss: 4.4247 - accuracy: 0.4563 - val_loss: 4.9813 - val_accuracy: 0.4805
Epoch 16/30
10/10 [==============================] - ETA: 0s - loss: 5.7268 - accuracy: 0.4500
Epoch 00016: val_accuracy did not improve from 0.55859
10/10 [==============================] - 38s 4s/step - loss: 5.7268 - accuracy: 0.4500 - val_loss: 5.4386 - val_accuracy: 0.5195
Epoch 00016: early stopping
```



# Conclusion:

Sentiment Analysis and Face Detection, individually have numerous use-cases in today's world. We see object detection algorithms in public parking lots, traffic monitoring systems, etc. that take images of people driving vehicles to keep records. Sentiment Analysis is furthermore used in therapy where physical meetings of the therapist and their patient are not possible. The study of human cognition has also evolved medicines**.** On the technological front, virtual assistants, profile evaluation assistants, and automation bots are built to mimic the actions of humans and replace them with the hope of increasing accuracy and decreasing errors. It is therefore a very important part of the Artificial Intelligence inspired world we live in today. A more engrossing and complicated approach to computer vision is by using cloud-based algorithms like Azure Cognitive Services or Deep Learning mechanisms, which we have not covered in this story, but could come in handy for complex scenarios.

Through this story, we have learned the below:

•Cognitive Science is the study of human thought processes and aims at delivering human reactions and emotions to machines through algorithms.

•Computer Vision is the branch of Artificial Intelligence that focuses on implementing Cognitive Science in the real world by working with human data in the form of images.

•Image Processing is a part of all Computer Vision algorithms that helps algorithms understand images, process them, work with them as numeric vectors and perform necessary operations.

We used the power of Artificial Intelligence to work on Cognitive Science and dealt with human faces, this space is generally referred to as ***Computer Vision.*** We were able to extract emotions out of photos and videos of human faces.

# Reference:


- 1) https://www.almabetter.com/
- 2) https://www.wikipedia.org
- 3) https://www.kaggle.com/
- 4) https://github.com/