BY: AKSHAY KAMATH

60004210116

B Divison

Task 3 :

CV:Facial Recognition and Emoji Recognition

DATA SET: **FER 2013**

The data consists of 48x48 pixel grayscale images of faces. The faces have been automatically registered so that the face is more or less centred and occupies about the same amount of space in each image.

Summary

1. ***Image Pre-Processing Steps***
2. ***Techniques used for classification***
3. ***Conclusion***

Image Pre-Processing:

Image pre-processing is a method to perform operations on an image to extract information from it or enhance it.

Since images exist in different formats, i.e., natural, fake, grayscale, etc., we need to take into consideration and standardize them before feeding them into a neural network.

The aim of this process is to improve the image data(features) by suppressing unwanted distortions and enhancement of some important image features so that our Computer Vision models can benefit from this improved data to work on

GRAYSCALE CONVERSION:

Grayscale is simply converting images from colored to black and white. It is normally used to reduce computation complexity in machine learning algorithms.

Since most pictures don't need color to be recognized, it is wise to use grayscale, which reduces the number of pixels in an image, thus, reducing the computations required.

We can execute the code below to convert the original image to grayscale:

```
gray_image = skimage.color.rgb2gray(image)
plt.imshow(gray_image, cmap = 'gray')
```

NORMALIZATION:

Also referred to as *data re-scaling*, it is the process of projecting image data pixels (intensity) to a predefined range (usually `(0,1)` or `(-1, 1)`).Normalization is usually applied to convert an image's pixel values to a typical or more familiar sense.Its benefits include:

- Fairness across all images
- Provides a standard learning rate - Since high pixel images require a low learning rate and low pixel images high learning rate, re-scaling helps provide a standard learning rate for all images.

DATA AUGMENTATION:

Data augmentation is the process of making minor alterations to existing data to increase its diversity without collecting new data.

It is a technique used for enlarging a dataset. Standard data augmentation techniques include ***horizontal & vertical flipping, rotation, cropping, shearing***

Performing data augmentation helps in preventing a neural network from learning irrelevant features. **This results in better model performance.**

There are two types of Augmentation:

- ***Offline Augmentation*** :Used for small datasets,applied in data preprocessing steps.
- ***Online Augmentation*** :Used for large dataset.It is normally applied in real time.

- We use **Keras's** ImageDataGenerator class to augment our data. This is because it provides a quick and easy way to augment your images.
- In addition, it supports augmentation techniques such as **flips, rotations, brightness change**

CLASSIFICATION TECHNIQUES: This step categorizes detected objects into predefined classes by using a suitable classification technique that compares the image patterns with the target patterns.

1. Support Vector Machines

It is a supervised machine learning algorithm used for both regression and classification problems.
When used for classification purposes, it separates the classes using a linear boundary.

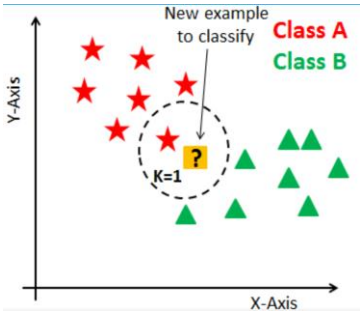
2. Decision Trees

It is also a supervised machine learning algorithm, which at its core is the tree data structure only, using a couple of if/else statements on the features selected. Decision trees are based on a hierarchical rule-based method and permits the acceptance and rejection of class labels at each intermediary stage/level.This method consists of 3 parts:

- Partitioning the nodes
- Finding the terminal nodes
- Allocation of the class label to terminal node

3. K Nearest Neighbor

The k-nearest neighbor is by far the most simple machine learning algorithm.
This algorithm simply relies on the distance between feature vectors and classifies unknown data points by finding the most common class among the k-closest examples.

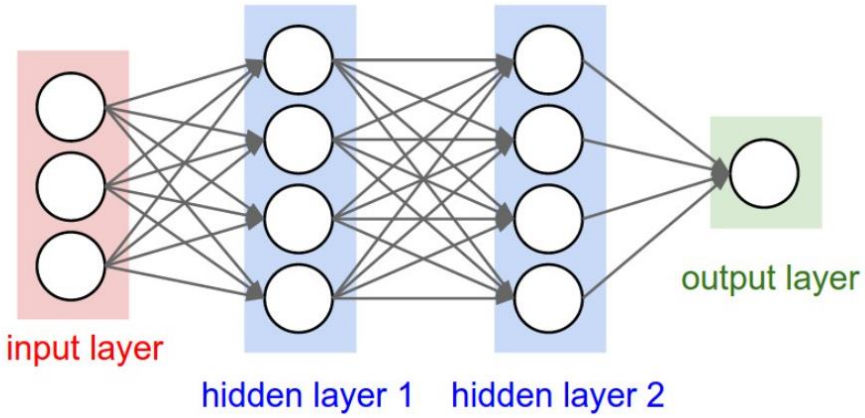


Here we can see there are two categories of images and that each of the data points within each respective category are grouped relatively close together in an n-dimensional space.

In order to apply the k-nearest Neighbor classification, we need to define a distance metric or similarity function.

4. Artificial Neural Networks

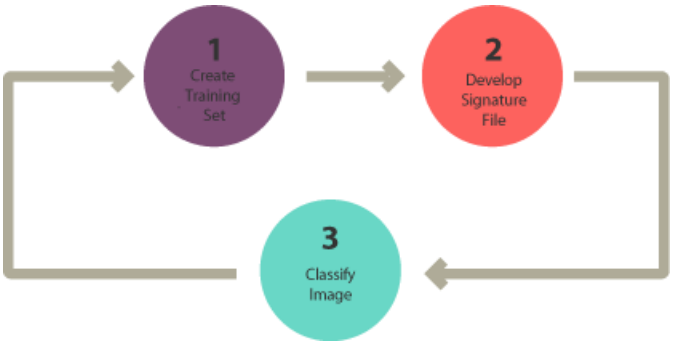
Artificial Neural Networks are statistical learning algorithms and are used for a variety of tasks, from relatively simple classification tasks to computer vision and speech recognition.
ANNs are implemented as a system of interconnected processing elements, called nodes, which are functionally analogous to biological neurons.The connections between different nodes have numerical values, called weights, and by altering these values in a systematic way, the network is eventually able to approximate the desired function.



In supervised classification, you **select representative samples** for each land cover class. The software then uses these **"training sites"** and applies them to the entire image.

The three basic steps for supervised classification are:

- Select training areas
- Generate signature file
- Classify



Supervised classification uses classification algorithms and regression techniques to develop predictive models. The algorithms include

- Linear regression
- Logistic regression
- Neural Networks
- Decision Tree
- Support Vector Machine
- Random Forest
- Naive Bayes
- and K-Nearest Neighbor.

CONCLUSION:

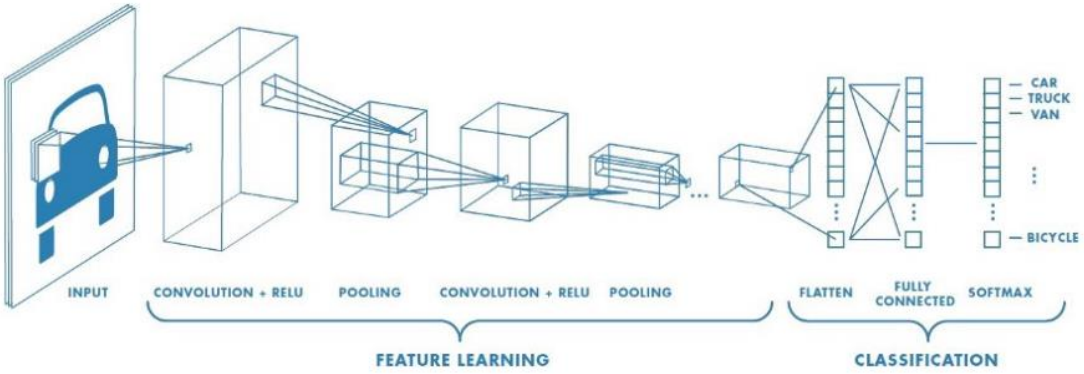
Why I am chosing CNN over other Classification Techniques:

The Convolutional Neural Network (CNN or ConvNet) is a subtype of the Neural Networks that is mainly used for applications in image and speech recognition. Its built-in convolutional layer reduces the high dimensionality of images without losing its information

CNN works by extracting features from the images. Any CNN consists of the following:

1. The input layer which is a grayscale image
2. The Output layer which is a binary or multi-class labels

3. Hidden layers consisting of convolution layers, ReLU (rectified linear unit) layers, the pooling layers, and a fully connected Neural Network



- CNN's are **really effective for image classification as the concept of dimensionality reduction suits the huge number of parameters in an image**

- The best thing is there is no need for feature extraction. The system learns to do feature extraction and the core concept of CNN is, it uses convolution of image and filters to generate invariant features which are passed on to the next layer.

- CNNs are very effective in reducing the number of parameters without losing on the quality of models.

- CNNs are trained to identify the edges of objects in any image

CLASSIFIER	ACCURACY	PRECISION	RECALL	ROC
SVM	85.68%	0.86	0.87	0.86
Decision Trees	84.61%	0.85	0.84	0.82
KNN	86.32%	0.86	0.86	0.88
ANN(for 100 epochs)	83.10%	0.88	0.87	0.88
CNN(for 300 epochs)	91.11%	0.93	0.89	0.97

The above data also shows that '**CNN**' gives the most accuracy amongst other algorithms,Hence I would use this algorithm over other.