Criminal Identification based on Sentiments using Facial Features

Akansha chaudhary[1], Abhishek Kumar Ojha[2], Akarmah Khan[3], Mr. Pushpendra Singh[4]

[1,2,3]B.Tech students, Department of Information Technology, Inderprastha Engineering College
[4]Associate Professor, Department of Information Technology, Inderprastha Engineering College

---

Abstract: The "Criminal Identification based on Sentiments using Facial Features" is a software that can collate all sources of information to predict and identify in advance people planning terrorist acts. As an artificial super intelligence, its objective is to predict and prevent imminent criminal activity and does so by analysing immense amount of surveillance data.

This software can analyse data from the surveillance cameras, electronic communication and audio input and predict acts of crime.

In this, we will focus on the following: 1. Recognize faces of the peoples.

2. Recognize the speech of the peoples.

3. Recognize and analyse their emotions.

1. Introduction

The idea of the project has been taken from the TV Series named PERSON OF INTEREST which is an American science fiction crime drama television series in which a computer system were developed for the U.S. government named "THE MACHINE" that is capable of collating all sources of information to predict and identify—in advance—people planning terrorist acts. As an artificial super intelligence, its objective is to predict and prevent imminent terrorist attacks and does so by analysing immense amounts of surveillance data.
The Machine sorts through all available information and categorizes persons of interest into relevant (national security risk) and irrelevant (ordinary risk) cases. It categorizes each POI based on their actions to determine whether they are the victim or the perpetrator. It uses various machine learning techniques to determine the identity, location, and intentions of monitored persons by infiltrating domestic organizations such as the National Security Agency and foreign agencies including Interpol ("No Good Deed") to analyse their databases and data from various sources such as video footage, phone calls (landline, VOIP, mobile), GPS, electronic transactions, e-mail, social media, etc.
The Machine was created with the ability to simulate the outcomes of different scenarios to aid it in making choices and to better fulfil its purpose. Finch taught it to how to play chess and the importance of making good decisions.

It can evaluate the outcomes of different strategies by way of simulating them. In a mere fraction of a second, the Machine can create and process thousands of highly detailed, highly-accurate simulations. These simulations are displayed on a decision tree and produce varying outcomes.

2. Face recognition

A Face recognition is a method of identifying or verifying the identity of an individual using their face. This system can be used to identity people in photos, videos or in real-time.

3. Algorithm Implemented-

VIOLA-JONES ALGORITHM FOR FACE RECOGNITION- The Viola-Jones algorithm first detects the face on the grayscale image and then finds the location on the colour image. With smaller steps, several boxes detect **face**-like features (Haar-like features) and the data of all those boxes put together, helps the algorithm determine where the face is.

The algorithm has four stages-

- **Haar Feature Selection-** Haar-like features are named after Alfred Haar, a Hungarian mathematician in the 19th century who developed the concept of Haar wavelets (kind of like the ancestor of Haar-like features). The features below show a box with a light side and a dark side, which is how the machine determines what the feature is.
- **Creating an integral image-** The integral image plays its part in allowing us to perform these intensive calculations quickly so we can understand whether a feature of several features fit the criteria.
- **Adaboost training-** The algorithm learns from the images we supply it and can determine the false positives and true negatives in the data, allowing it to be more accurate.
- **Cascading classifiers-** Cascading is another sort of "hack" to boost the speed and accuracy of our model. So, we start by taking a sub window and within this sub window, we take our most important or best feature and see if it is present in the image within the sub window. If it is not in the sub window, then we do not even look at the sub window, we just discard it

4. Face Verification and Identification

We investigated both the face identification and substantiation tasks by means of a single positive teaching example with the one-shot resemblance measure on the YouTube Video Face Dataset. Moreover, the One-Shot resemblance was presented in the context of face recognition for situations where used to produce state-of-the-art results with shallow learning methods, additionally, they also investigated its use with pre-trained deep mock-ups. They illustrated the significance of combining framework into the recognition engine by using temporal data from past frames to forecast the uniqueness of the face in the present video frame. They further exhibited/revealed the difference between performances on face verification. The figure below shows a flowchart of the proposed architecture for face verification.

5. Similarity-based measure in image database for crime investigation:

Face recognition systems are not magical boxes that work mysteriously. Essentially, they comprise of a capturing interface that is the camera and a database system to store the uprooted quality of facial image. It's interesting to know that most face recognition systems have not been able to yield 100% accuracy and yet no adequate research has been carried out on the actual body recognition. Body Part Recognition system (BPR) will help to substantiate the result from a traditional face recognition system. In fact, the face recognition system will be a component of BPR. This will help to wipe out inconsistent Body Image Characteristics (BIC) that are not like bodies under inspection. In a face recognition system, face images taken are used to substantiate queries by image content in a Face Image Database (FID), all images must be analysed before they are stored. Eventually, those attributes can be extracted and stored in the database independently. These features are then used to look out the FID and to decide which image meet the required query selection criteria. The user may specify several objects with complex shapes and interrelationships and may ask for all images containing self-similar. The term self-similar means that at every time the retrieved images from FID are semblances of what the user's query. The database searching must distribute the self-similar ones so that all images ranging up to a specified correlation will be retrieved, concentrate on analysis of query for numerous correlation factors in FID searching. A density of original face image content, comparison cost that is a similarity-based measurement and a variance of sample means that sample size is essential with the timescale comparison

6. Speech Recognition

Speech recognition is the ability of a machine which identify alphabetic words or spoken

language and then it will convert those words in the machine-readable format. And the speech recognition software is the technology which transform spoken words into the text it has limited Waterbury words and it is only identify or recognize these words if they spoke very louder voice and very neat and clearly

Algorithm Implemented-

Mel-Frequency Cepstral Coefficients- Mel Frequency Cepstral Coefficients (MFCCs) The Mel frequency cepstral coefficients (MFCCs) of a signal are a small set of features (usually about 10-20) which concisely describe the overall shape of a spectral envelope.

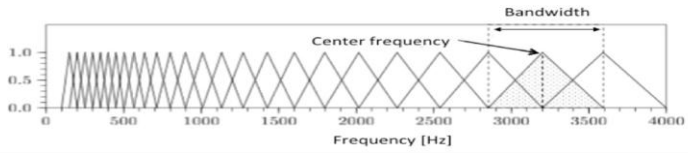
Steps of Algorithm-

- **Fourier Transform-**

$$c_{\tau,k}^{(1)} = \left| \frac{1}{N} \sum_{j=0}^{N-1} f_j \, exp\left[-i \, 2\pi \, \frac{jk}{N} \right] \right| \qquad k = 0, 1, ..., (N/2) - 1$$

Where N is the number of sampling points within a speech frame and the time frame τ

- **Mel-Frequency spectrum-** The spectrum is filtered with Nd different band-pass filters and the power of each frequency band is computed. Figure shows a typical filter-bank with 25 triangular band-pass filters.



- **Logarithm-**The third processing step computes the logarithm of the signal, to mimic the human perception of loudness because experiments showed that humans perceive loudness on a logarithmic scale.

$$c_{\tau,j}^{(3)} = log(c_{\tau,j}^{(2)}) \qquad j = 0, 1, ..., N_d$$

- **Cepstral Coefficients-**The fourth processing step tries to eliminate the speaker dependent characteristics by computing the cepstral coefficients. From the Source-Filter model is known, that the signal is the convolution of the speaker dependent source signal and the filter signal. To suppress the source, signal the cestrum is computed.

7. Emotion Recognition

Emotion recognition is the process of identifying human facial expressions, most typically from facial status and scales.

Algorithm Implemented- VIOLA-JONES ALGORITHM FOR EMOTION RECOGNITION-

The Viola-Jones algorithm is a broadly used mechanism for object detection. The principle property of this calculation is that preparation is moderate, yet location is quick. This algorithm implements Haar basis feature filters, so it does not use duplications.

Steps of Algorithm-

- **Haar Feature Selection-**The four features matched by this algorithm are then sought in the image of a face. Shown below-

- **Creating An Integral Image** An image representation called the integral image evaluates rectangular in consistent time, which gives them a significant speed advantage over increasingly modern elective highlights.
- **Adaboost Training-** The speed with which highlights might be assessed does not enough make up for their number, notwithstanding., however. This calculation develops a "solid" classifier as a straight mix of weighted basic "weak "classifiers. Each powerless classifier is an edge work dependent on the component.
- **Cascading Classifier-** In falling, each stage comprises of a solid classifier. So, all the features are assembled into a few phases

where each stage has certain number of features.

The activity of each stage is to decide if a given sub-window is unquestionably not a face or might be a face.

```
F(0) = 1.0; D(0) = 1.0; i = 0
while F(i) > Ftarget
  increase i
  n(i) = 0; F(i)= F(i-1)

  while F(i) > f × F(i-1)
    increase n(i)
    Evaluate current cascaded classifier on validation set to determine F(i) and D(i)
    decrease threshold for the ith classifier (i.e. how many weak classifiers need to accept for strong classifier to accept)
      until the current cascaded classifier has a detection rate of at least d × D(i-1) (this also affects F(i))
  N = Ø
  if F(i) > Ftarget then
    evaluate the current cascaded detector on the set of non-face images
    and put any false detections into the set N.
```

8. Conclusion:

In this report, I have discussed the various types of technologies like voice recognition, trajectory motion within the confinement of facial recoganizability and how technology has evolved over the years. To this end, I strongly believe facial recognition technology still requires adequate research in that none of the current technologies can produce 100% accuracy and some of the technologies can be fooled.

9. References

1.Person of Interest (TV Series 2011–2016) - IMDb (http://www.imdb.com/title/tt1839578/)
2.Essentials of Machine Learning Algorithms (with Python and R Codes) (https://www.analyticsvidhya.com/blog/2015/08/common-machine-learning-algorithms/)
3.Machine Learning Project in Python Step-By-Step (http://machinelearningmastery.com/machine-learning-in-python-step-by-step/)
4.Pattern Recognition - The Journal of the Pattern Recognition Society (https://www.journals.elsevier.com/pattern-recognition)
5.Face Recognition (http://www.face-rec.org/)