FACIAL RECOGNITION USING CONVOLUTIONAL NEURAL NETWORKS AND K NEAREST NEIGHBORS CLASSIFICATION

Dr. G.SRINIVASA RAO, M.Tech, Ph.D[1], B. NAGENDRA BABU[2], V. JAGADEESH[3], D. RAJESWARI [4]

[1]Professor, Dept. of Computer Science & Engineering, LINGAYAS INSTITUTE OF MANAGEMENT AND TECHNOLOGY, A.P., India.

[2,3,4,] Student, B. Tech (CSE), LINGAYAS INSTITUTE OF MANAGEMENT AND TECHNOLOGY, A.P., India.

ABSTRACT: With the social media boom in today's world, People constantly uploading photos of themselves along with their friends and family on various social media platforms such as Facebook, Instagram, Twitter, Google+, etc. if they want to see all the photos in a categorized form such as photos with a particular person. There should be an identification algorithm present to identify the faces and categorize the photos by cropping them. So in order to solve this problem we use convolutional neural networks (CNN) for Facial detection and Recognition, after recognition we further compare the cropped faces with the faces which we taken before and find the name of particular person this can be done by using classification algorithms in Machine Learning. In my face recognition project, a computer system will be able to find and recognize human faces fast and precisely in images with Multiple Faces. Convolutional Neural Networks (CNN) are further advance then the other because, the hidden network layers were expanded based on the input coordinates and by analysing in a very deep it recognizes the face accurately and at the end it compresses the layers and give output in form of a stack and the no of output values are equal to the no of inputs.

Keywords: Multi-Tasking Convolutional Neural Networks (MTCNN), VGGFace, K-Nearest Neighbors.

1. INTRODUCTION

Face Recognition is a recognition technique used to detect faces of individuals whose images saved in the data set and identifying a detected object as a known or unknown face. Often the problem of face recognition is confused with the problem of face detection, Face Recognition on the other hand is to decide if the "face" is someone known or unknown.

1.1 FACIAL RECOGNITION

There are many ways for face recognition. Here we use OpenCV for face recognition. In face recognition, the image first prepared for pre-processing and then trained the face recognizer to recognize the faces. After teaching the recognizer, we test the recognizer to see the results.
The OpenCV face recognizer are of three types:

Eigen Faces Face Recognizer

Eigen Faces face recognizer views at all the training images of all the characters as a complex and try to deduce the components. These components are necessary and helpful (the parts that grab the most variance/change) and discard the rest of the images, This way it not only extracts the essential elements from the training data but also saves memory by rejecting the less critical segments.

Fisher Faces Recognizer

Fisher faces algorithm, instead of obtaining useful features that represent all the faces of all the persons, it removes valuable features that discriminate one person from the others. This features of one person do not dominate over the others, and you have the features that distinguish one person from the others.

Local Binary Patterns Histograms

We know that Eigen faces and Fisher faces are both affected by light and in real life; we cannot guarantee perfect light conditions. LBPH face recognizer is an improvement to overcome this drawback. The idea is not to find the local features of an image. LBPH algorithm tries to find the local structure of an image, and it does that by comparing each pixel with its neighbouring pixels.

1.2 FACE RECOGNITION METHODS

1.2.1 Geometric Based / Template Based

Face recognition algorithms classified as geometry based or template based algorithms. The template-based methods can be constructed using statistical tools like SVM [Support Vector Machines], PCA [Principal Component Analysis], LDA [Linear Discriminant Analysis], Kernel methods or Trace Transforms. The geometric feature based methods analyze local facial features and their geometric relationship. It is also known as a feature-based method.

Geometric features in the face can be represented in Fiducial Points, Feature Choosing, Feature Set Optimization

Fiducial Points

Fiducial points are points that are used as points of reference or measure. Determining of fiducial points can be a fundamental step to recognize a face. A few important fiducial points are the eyes, lip edges, nose, chin etc. Using the fiducial points, we can either obtain an outline of the entire face or develop a relationship between the Fiducial points themselves to act as a medium to recognize a face. Lot of research has been ongoing in this regard. In this paper the Fiducial points and their existing relationships are studied using a Support Vector Machine with a Radial Basis Function kernel. New images when tested showed a high accuracy of correct results in terms of the actual positions of the fiducial points in the image. Further classification of the fiducial points is done using an Adaboost classification to improve the accuracy.

Feature Choosing

Geometric features may be presented by segments, perimeters and areas of some figures formed by the detected points. We studied different subsets of the features looking for the most robust features but due to the great quantity of them we cannot report about strong results of the searching yet. Hence, to present our progress in comparison with known recognition results. It includes 15 segments between the points and the mean values of 15 symmetrical segment pairs. The explored feature set is not the best.

Feature Set Optimization

Once a feature set has been obtained it can be optimized by the presenting technique. How to choose the optimal feature subset? The point was to find the feature space with the maximal distances between the clusters and minimal ones between the patterns of one cluster. In our case all database images of the same person were considered as one cluster. To evaluate the effectiveness of every feature subset the F value was calculated:

$$F = \sqrt{\frac{\sum_{i=1}^{k} \left(M_{D_i} - D_i\right)^2}{\sum_{i=1}^{k} \left(M_{M_i} - M_i\right)^2}}$$

where M i and D i are mean and variance of the feature values for k images of the i-th person, MD i and MM i are mean of Di and MI, respectively. The lowest F value corresponds to the better feature set. To validate this technique, the distances between the clusters were computed every time of the feature space changing.

1.2.2 Piecemeal / Wholistic

The relation between the elements or the connection of a function with the whole face not undergone into the amount, many researchers followed this approach, trying to deduce the most relevant characteristics. Some methods attempted to use the eyes, a combination of features and so on. Some Hidden Markov Model methods also fall into this category, and feature processing is very famous in face recognition. Piecemeal approach is a type of approach in facial feature detection which deals with minimalism; the idea is to use very few facial features detected, instead of waiting to get all features. It assumes it is a face as long as those few features have been detected.

1.2.3Appearance-Based / Model-Based

The appearance-based method shows a face regarding several images. An image considered as a high dimensional vector. This technique is usually used to derive a feature space from the image division. The sample image compared to the training set. On the other hand, the model-based approach tries to model a face. The new sample implemented to the model and the parameters of the model used to recognize the image.

The appearance-based method can classify as linear or nonlinear. Ex- PCA, LDA, IDA used in direct approach whereas Kernel PCA used in nonlinear approach. On the other hand, in the model-based method can be classified as 2D or 3D Ex- Elastic Bunch Graph Matching used.

1.2.4Template / Statistical / Neural Networks Based

Template Matching

In template matching the patterns are represented by samples, models, pixels, textures, etc. The recognition function is usually a correlation or distance measure. Template Matching method uses pre-defined or parameterized face templates to locate or detect the faces by the correlation between the templates and input images. Ex: - A human face can be divided into eyes, face contour, nose, and mouth. Also, a face model can be built by edges just by using edge detection method.

Statistical Approach

In the Statistical approach, the patterns expressed as features. The recognition function in a discriminant function. Each image represented regarding d features. Therefore, the goal is to choose and apply the right statistical tool for extraction and analysis.

In statistical approach, each image is represented in terms of the features. It is viewed as a point (vector) in a dimensional space. Therefore, the goal is to choose and apply the right statistical tool for extraction and analysis of the underlying manifold. These tools must define the embedded face space in the image space and extract the basic functions from the face space. This would permit patterns belonging to different classes to occupy disjoint and compacted regions in the feature space. Consequently, a line, curve, plane or hyper plane that separates faces belonging to different classes could be defined

There are many statistical tools, which used for face recognition. These analytical tools used in a two or more groups or classification methods. These tools are as follows:

Principal Component Analysis [PCA]: One of the most used and cited statistical method is the Principal Component Analysis. A mathematical procedure performs a dimensionality reduction by extracting the principal component of multi-dimensional data.

The basic concept behind Eigen face method is information reduction, when an evaluation of a small image is done; there is a great amount of information present. This method generates base-faces and then represents any image being analysed by the system as a linear combination of the base faces. Once the base faces have been chosen, the problem has been reduced to a standard classification problem. Euclidian distance measure is used here for classification. This Facial recognition method can be broken down into the following components: Generate the Eigen faces; Project training data into face-space to be used with a classification method; Evaluate a projected test element by projecting it into face space and comparing to training data.

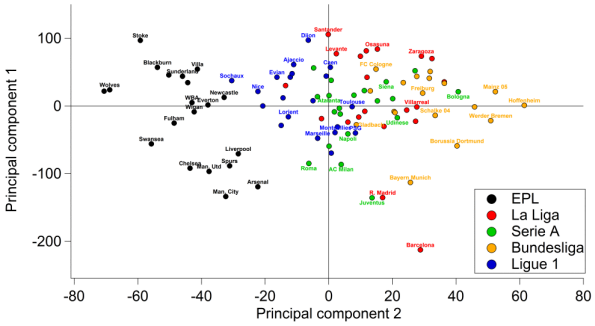

Fig 1.2.4: Principal Component Analysis

Discrete Cosine Transform [DCT]: It signifies a series of data points regarding a sum of cosine functions different oscillating frequencies. The Discrete Cosine Transform is based on Fourier discrete transform and therefore, by compacting the variations it can be used to transform images and allowing an efficient dimensionality reduction.

1.3 NEURAL NETWORKS

Neural Network has continued to use pattern recognition and classification. Kohonen was the first to show that a neuron network could be used to recognize aligned and normalized faces. There are methods, which perform feature extraction using neural networks. There are many methods, which combined with tools like PCA or LCA and make a hybrid classifier for face recognition. These are like Feed Forward Neural Network with additional bias, Self-Organizing Maps with PCA, and Convolutional Neural Networks with multi-layer perception, etc. These can increase the efficiency of the models.

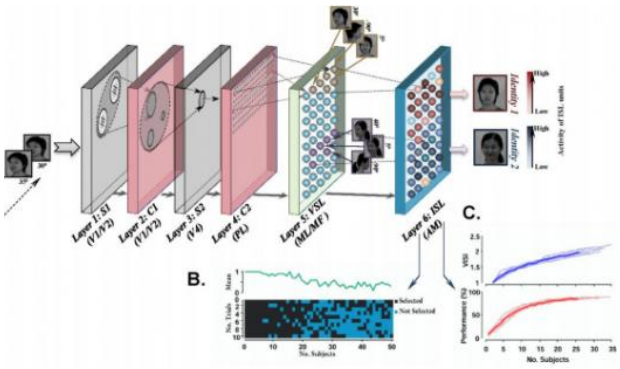

Fig 1.3: Neural networks

1.3.1 Neural Networks with Gabor Filters

The algorithm achieves face recognition by implementing a multilayer perceptron with a back-propagation algorithm. Firstly, there is a preprocessing step. Each image normalized in phases of contrast and illumination. Then each image is processed through a Gabor filter. The Gabor filter has five orientation parameters and three spatial frequencies, so there are 15 Gabor wavelengths.

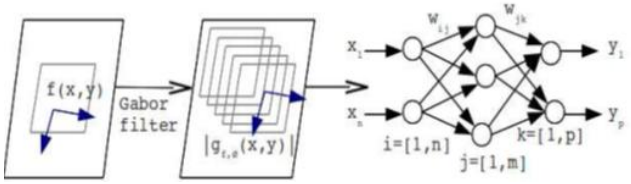

Fig 1.3.1 Neural Network with Gabor Filter

1.3.2 Neural Networks and Hidden Markov Models

Hidden Markov Models are a statistical tool used in face recognition. They have used in conjunction with neural networks. It generated in a neural network that trains pseudo 2D HMM. The input of this 2D HMM process is the output of the ANN, and It provides the algorithm with the proper dimensionality reduction.

1.3.3 Fuzzy Neural Networks

The fuzzy neural networks for face recognition introduce in 2009. In this a face recognition system using a multilayer perceptron. The concept behind this approach is to capture decision surfaces in nonlinear manifolds a task that a simple MLP can hardly complete. The feature vectors are obtained using Gabor wavelength transforms.

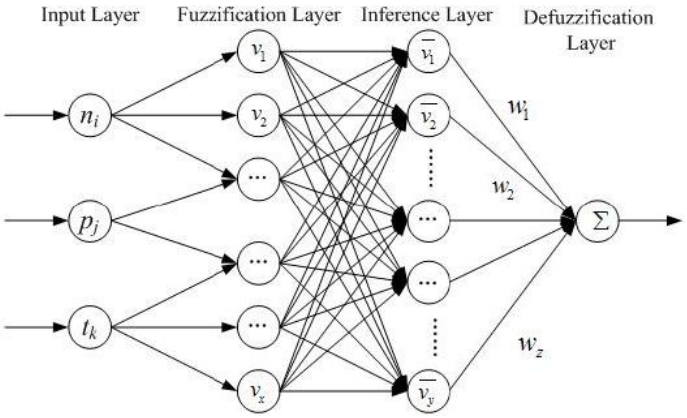

Fig 1.3.3: Fuzzy Neural Networks

1.4 DEEP LEARNING

Deep learning, as the name suggests, is a sub sect of machine learning. Deep Learning mostly involves using deep artificial neural networks (algorithms/computational models loosely inspired by the human brain) to tackle machine learning problems. Going back to the example I gave earlier, state-of-the-art image classification solutions today use deep learning. Note: sometimes, using decision trees and/or other machine learning algorithms may also be referred to as deep learning, but for the most part deep learning involves the use of neural networks.

So, what is a neural network? Here's an analogy: imagine a neural network as a series of doors one after another and think of yourself as the 'input' to the neural network. Every time you open a door, you become a different person. By the time you

open the last door, you have become a very different person. When you exit through the last door, you become the 'output' of the neural network. Each door, in this case, represents a layer.
A neural network, therefore, is a collection of layers that transform the input in some way to produce an output.

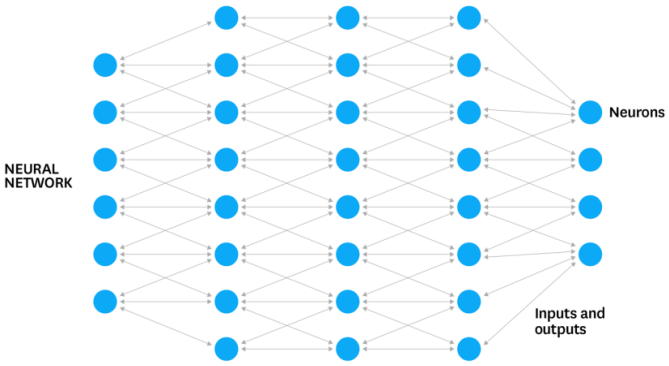

Fig 1.4: CNN Layers

Each layer in the neural network consists of 'weights' and 'biases' these are just numbers that augment the input. The overall idea of a neural network is that it takes in some input (usually a collection of numbers that represent something, e.g. Red-Green-Blue values of pixels in an image), applies some mathematical transformations to the input using the weights and biases in its layers and eventually spits out an output. If you've taken some linear algebra class before, you can look at the input, output and weights as matrices. The input matrix gets transformed by a series of matrices and that becomes your output.

A deep neural network is just a neural network with many layers (as you stack layers on top of another, the neural network keeps getting 'deeper'). How many is many? Well, there's a VGG16 neural network architecture (used for image classification) that consists of 16 layers and then there's the ResNet architecture (also used for image classification) that consists of 152 layers. So, the range is pretty wide. The basic idea of deep learning is using neural networks with multiple layers.

2. LITERATURE SURVEY

Face recognition is the process of recognizing the face of a relevant person by a vision system. It has been a crucial human-computer interaction tool due to its usage in security systems, access-control, video surveillance, commercial areas and even it is used in social networks like Facebook as well. After rapid development of artificial intelligence, face recognition has once again attracted attention due to its nonintrusive nature and since it is main method of person identification for human when it is compared with other types of biometric techniques.

KaipengZhang and Zhan PengZhan proposed a deep cascaded multi-task framework which exploits the inherent correlation between detection and alignment to boost up their performance. In particular, our framework leverages a cascaded architecture with three stages of carefully designed deep convolutional networks to predict face and landmark location in a coarse-to-fine manner. In addition, we propose a new online hard sample mining strategy that further improves the performance in practice. Our method achieves superior accuracy over the state-of-the-art techniques on the challenging FDDB and WIDER FACE benchmarks for face detection, and AFLW benchmark for face alignment, while keeps real time performance. method can achieve very fast speed in joint face detection and alignment. It takes 16fps on a 2.60GHz CPU and 99fps on GPU (NVidia Titan Black). Implementation is currently based on un-optimized MATLAB code.

Musab Coskun and Aysegul Ucar proposed a modified Convolutional Neural Network (CNN) architecture by adding two normalization operations to two of the layers. The normalization operation which is batch normalization provided accelerating the network. CNN architecture was employed to extract distinctive face features and Softmax classifier was used to classify faces in the fully connected layer of CNN. In the experiment part, Georgia Tech Database showed that the proposed approach has improved the face recognition performance with better recognition results. In the designed CNN with Beta23 version of MatConvNet software tool. After pre-processing stage, size of each image was changed as 16x16x1, 16x16x3, 32x32x1, 32x32x3, 64x64x1, and 64x64x3. 66% of images were assigned as training set, 34% as test set. We empirically tried different scenarios by making changes in image size, learning rate, batch size, and etc. CNN was trained for 35 epochs. Performance of the proposed CNN was evaluated according to top-1 and top-5 errors. Top-1 error rate checks if the top class is the same as the target label and top-5 error rate checks if the target label is one of your top five predictions. The prominent features of the proposed algorithm is that it employs the batch normalization for the outputs of the first and final convolutional layers in training stage and that makes the network reach higher accuracy rates. In fully connected layer step, Softmax Classifier was used to classify the faces. The performance of the proposed algorithm was tested on Georgia Tech Face Database. The results showed satisfying recognition rates according to studies in the literature.

HuiWang and DA. Bell with three others proposed a novel kNN type method for classification that is aimed at overcoming these shortcomings. Our method constructs a kNN model for the data, which replaces the data to serve as the basis of classification. The value of k is automatically determined, is varied for different data, and is optimal in terms of classification accuracy. The construction of the model reduces the dependency on k and makes classification faster. Experiments were carried out on some public datasets collected from the UCI machine learning repository in order to test our method. The experimental results show that the kNN based model compares well with C5.0 and kNN in terms of classification accuracy, but is more efficient than the standard kNN. Experimental results carried out on six public datasets show that the kNN Model is a quite competitive method for classification. Its average classification accuracy on six public datasets is comparable with C5.0 and kNN. Also the kNN Model significantly reduces the number of the data tuples in the final model for classification with a 90.41% reduction rate on average.

3. PROPOSED METHODOLOGY

INTRODUCTION

Face recognition is a major challenge encountered in multidimensional visual model analysis and is a hot area of research. The art of recognizing the human face is quite difficult as it exhibits varying characteristics like expressions, age, change in hairstyle etc. Face recognition plays a crucial role in applications such as security system, credit card verification, identifying criminals in airport, railway stations etc. Although many methods have been proposed to detect and recognize human face developing a computational model for a large data base is still a challenging task. That is why face recognition is considered as high level computer vision task in which techniques can be developed to achieve accurate results. Few popular methods

known for face recognition are neural network group based tree, neural nets, artificial neural networks and principal component analysis

3.1 PROPOSED METHODOLOGY

The proposed method implement an efficient Face Detection and Recognition technique which is independent of variations in features like color, hairstyle, different facial expressions etc using Viola Jones algorithm, PCA and ANN. The proposed methodology uses the BioID Face Database as the standard image data base. The dataset consists of 1521 gray level images with resolution of 384*286 pixel and frontal view of a face of 23 different persons. The test set features a large variety of illumination, background and face size representing real world conditions

Steps

- PRE PROCESSING
- FACE DETECTION USING VIOLA JONES ALGORITHM
- FACE RECOGNITION USING ARTIFICIAL NEURAL NETWORKS

3.1.1 Viola Jones Algorithm

Viola-Jones detector was chosen as a detection algorithm because of its high detection rate, and its ability to run in real time. Detector is most effective on frontal images of faces and it can cope with 45° face rotation both around the vertical and horizontal axis. The three main concepts which allow it to run in real time are the integral image, Ada Boost and the cascade structure. The Integral Image is an algorithm for cost-effective generation of the sum of pixel intensities in a specified rectangle in an image. It is used for rapid computation of Haar-like features. Calculation of the sum of a rectangular area inside the original image is extremely efficient, requiring only four additions for any arbitrary rectangle size. AdaBoost is used for construction of strong classifiers as linear combination of weak classifiers.

3.1.2 Artificial neural Networks

All neurons are connected by a path to carry electrical signals referred to as synapses. They communicate through these paths and approximately there are 100 billion neurons in a brain. Each cell has inputs and outputs.

In a similar way the computer created artificial network has inputs for inserting the data, outputs for providing the network output and hidden layer for processing the data and training of the network

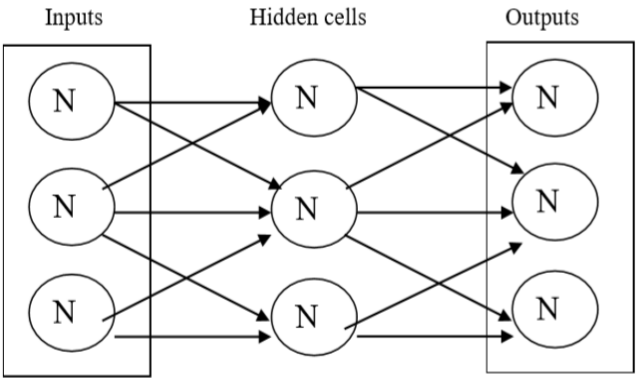

Fig 3.1.2: Artificial Neural Networks

3.2 Disadvantages

- Long training times for deep networks, which are the most accurate architecture for most problems. This is especially true if you're training on a CPU instead of a specialized GPU instance.
- Need lots of data, especially for architectures with many layers. This is a problem for most ML algorithms, of course, but is especially relevant for ANNs because of the vast number of weights and connections in ANNs.
- Architectures have to be fine-tuned to achieve the best performance. There aremany design decisions that have to be made, from the number of layers to the number of nodes in each layer to the

activation functions, and an architecture that works well to some one problem very often does not generalize well.

4. RESULTS AND DISCUSSION TESTING THE MODEL

First webcam was loaded and each frame was taken and detection was performed then the detected face was cropped and loaded into the VGGFace model and the face encodings from the model was sent to the K Nearest Neighbors Classification. Classification algorithm calculates the Euclidean Distance between the trained data and testing data and checks whether the probability is greater than 0. If the probability greater than 0 then the target label name was displayed at the top of the bounding box.

OUTPUT

Fig 1: Label1

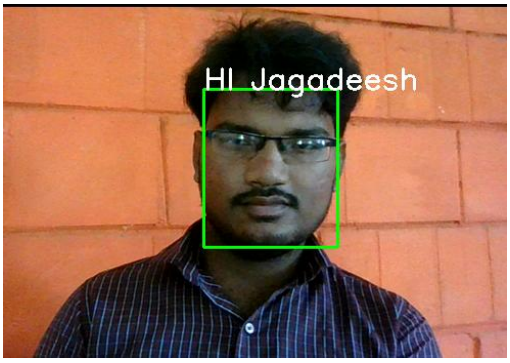

Fig 2: Label2

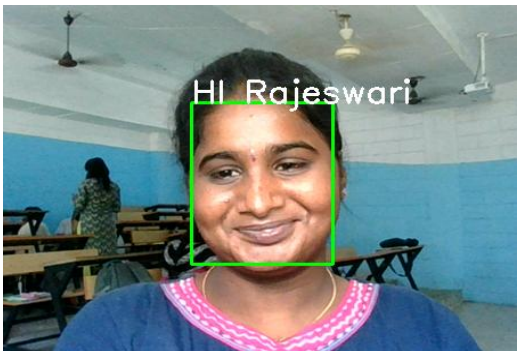

Fig 3: Label3

CONCLUSION

Face recognition technology has come a long way in the last twenty years. Today, machines are able to automatically verify identity information for secure transactions, for surveillance and security tasks, and for access control to buildings etc. These applications usually work in controlled environments and recognition algorithms can take advantage of the environmental constraints to obtain high recognition accuracy. However, next generation face recognition systems are going to have widespread application in smart environments, where computers and machines are more like helpful assistants.

To achieve this goal computer must be able to reliably identify nearby people in a manner that fits naturally within the pattern of normal human interactions. They must not require special interactions and must conform to human intuitions about when recognition is likely. This implies that future smart environments should use the same modalities as humans, and have approximately the same limitations. However, substantial research remains to be done in making person recognition technology work reliably, in widely varying conditions using information from single or multiple modalities. Face recognition can add significant value to our lives, and we seek to advance these technologies in ways that respect the

worldwide principals of freedom, justice, rights to privacy, transparency and continuous improvement.

SCOPE FOR THE FUTURE WORK

Facial recognition is a very effective tool that can help law enforcers recognize criminals and software companies are leveraging the technology to help users access their technology. This technology can be further developed to be used in other avenues such as ATMs, accessing confidential files, or other sensitive materials. This can make other security measures such as passwords and keys obsolete. It can also be used in Human Robot Interaction for using Robot as out daily life member with certain restrictions.

Another way that innovators are looking to implement facial recognition is within subways and other transportation outlets. They are looking to leverage this technology to use faces as credit cards to pay for your transportation fee. Instead of having to go to a booth to buy a ticket for a fare, the face recognition would take your face, run it through a system, and charge the account that you've previously created. This could potentially streamline the process and optimize the flow of traffic drastically.

REFERENCES

- Kaipeng Zhang, Zhanpeng Zhang and Zhifeng "Joint Face Detection and Alignment using Multi-task Cascaded Convolutional Networks" IEEE International Joint Conference2016
- P. Viola and M. J. Jones, "Robust real-time face detection. International journal of computer vision," vol. 57, no. 2, pp. 137-154, 2004
- M.Mathias, R. Benenson, M. Pedersoli, and L. VanGool, "Face detection without bells and whistles," in European Conference on Computer Vision, 2014, pp. 720-735.
- M. Köstinger, P. Wohlhart, P. M. Roth, and H. Bischof, "Annotated facial landmarks in the wild: A large-scale, real-world database for facial landmark localization," in IEEE Conference on Computer Vision and Pattern Recognition Workshops, 2011, pp. 2144-2151.
- J. Zhang, S. Shan, M. Kan, and X. Chen, "Coarse-to-fine auto-encoder networks (CFAN) for real-time face alignment," in European Conference on Computer Vision, 2014, pp. 1-16.
- X. Xiong, and F. Torre, "Supervised descent method and its applications to face alignment," in IEEE Conference on Computer Vision and Pattern Recognition, 2013, pp. 532-539.
- S. Yang, P. Luo, C. C. Loy, and X. Tang, "WIDER FACE: A Face Detection Benchmark". arXiv preprint arXiv:1511.06523.
- Musab Coşkun, Özal Yıldırım and Yakup Demir, "Face Recognition Based on Convolutional Neural Network"IEEE International Joint Conference2017
- S. G. Bhele and V. H. Mankar, "A Review Paper on Face Recognition Techniques," Int. J. Adv. Res. Comput. Eng. Technol., vol. 1, no. 8, pp. 2278–1323, 2012.

- V. Bruce and A. Young, "Understanding face recognition," Br. J. Psychol., vol. 77, no. 3, pp. 305–327, 1986.
- S. N. Borade, R. R. Deshmukh, and S. Ramu, "Face recognition using fusion of PCA and LDA: Borda count approach," in 24th Mediterranean Conference on Control and Automation, MED 2016, 2016, pp. 1164– 1167.
- C. Zhou, L. Wang, Q. Zhang, and X. Wei, "Face recognition based on PCA image reconstruction and LDA," Opt. - Int. J. Light Electron Opt., vol. 124, no. 22, pp. 5599–5603, 2013.
- Gongde Guo, Hui Wang, David Bell, Yaxin Bi, and Kieran Greer, "KNN Model-Based Approach in Classification" by the European Commission project ICONS, project no. IST-2001-32429.
- D. Hand, H. Mannila, P. Smyth.: Principles of Data Mining. The MIT Press. (2001)
- F.Sebastiani.: Machine Learning in Automated Text Categorization. In ACM Computing Surveys, Vol. 34, No. 1, March (2002) pp.1-47.
- G. Gates.: The Reduced Nearest Neighbour Rule. IEEE Transactions on Information Theory, 18, 431-433, (1972)
- C.M.Bishop.: Neural Networks for Pattern Recognition. Oxford University Press, UK (1995)

