WWW.IJRPR.COM

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Comparison of Hog & SVM With FACENET & MTCNN

L. Nikhila[1], K. Ravi Teja[1], A. Shiva[1], S. Mohini[1], D. Yugandhar[2]

[1]Student, Department of Electronics and Communication Engineering, Aditya Institute of Technology and Management ,Tekkali, Srikakulam, Andhra Pradesh, India, 532201

[2]Professor, Department of Electronics and Communication Engineering, Aditya Institute of Technology and Management ,Tekkali, Srikakulam, Andhra Pradesh, India, 532201

Abstract:

This paper presents a comparative study of two popular methods for face detection and recognition: Histogram of Oriented Gradients (HOG) with Support Vector Machine (SVM), and Face Net with MTCNN (Multi-Task Cascaded Convolutional Neural Network). The objective is to evaluate the performance of these techniques in terms of accuracy, computational complexity, and robustness to variations in lighting, pose, and occlusion. The experiments were conducted on a dataset images, and the results indicate that FaceNet with MTCNN outperforms HOG with SVM in terms of accuracy and robustness, but at the cost of increased computational complexity. This study contributes to the ongoing efforts to develop more efficient and effective face recognition systems for various applications, including security, surveillance, and biometric authentication.

1. INTRODUCTION

Face detection and recognition are critical tasks in many applications, including surveillance, security, and human-computer interaction. Various approaches have been proposed to tackle these challenges, ranging from traditional methods based on handcrafted features and machine learning algorithms, such as Histogram of Oriented Gradients (HOG) with Support Vector Machine (SVM), to more recent deep learning techniques, such as Face Net with Multi-Task Cascaded Convolutional Neural Network (MTCNN).

In this paper, we present a comparative study of HOG with SVM and Face Net with MTCNN for face detection and recognition. The objective is to evaluate the performance of these techniques in terms of accuracy, computational complexity, and robustness to variations in lighting, pose, and occlusion. HOG with SVM is a popular method that extracts features based on the gradient orientation and magnitude of the image pixels and uses SVM for classification. On the other hand, Face Net with MTCNN is a state-of-the-art deep learning approach that learns a high-dimensional embedding space for faces and uses MTCNN for face detection and alignment.

Therefore, in this study, we aim to provide a comprehensive evaluation of these two methods and their trade-offs, which can guide the selection of the most appropriate method for a given face detection and recognition task. The experiments were conducted on a dataset of facial images from various sources.

2. METHODOLOGY:

The system is designed for the attendance of the different organizations to reduce the fault of an existing manual system. Data of students are added manually by an administrator. Our system is divided into two parts. The main focus of this project is to give a smart way of an attendance taking system in organizations for the attendance of students. and the second is the backend of the project which consists of logic and based on python machine learning.

The system is designed to handle pose and illumination problems in recognition of faces. This paper focuses on finding an exact match of the face from 40 subjects in AT & T facial database in addition to user defined subject, based on HOG features and matching. It shows the overall system architecture which covers the entire workflow that includes setting up of the database, feature extraction using HOG, facial landmark detection, building up of classifier model and feature matching.

A. Feature extraction and classification

Both training and testing phase requires undergoing the same type of feature extraction technique which in this case is done by using HOG features.

Algorithm: Feature-based facial recognition

Input: Images of persons S ={1,2,...41} each with S_i={1,2,...,10} where i ={1,2,....41} forming a total of 410 images

Output: Matched subject in the database based on training features F and labels generated by means of HOG feature and Harris feature matching of different facial components

Training set: T = i_1,i_2....., i_9 from 41 subjects

forming a total of 369 images

Testing set: Te = i_{10} each from 41 subjects forming a total of 41 images.

These images used to test the system are not included in the training set.

1. Create database by accessing files from original AT & T database in which original images in .pgm format are converted to .jpg format to match with inclusion of real-time images
2. Include 10 new user images by detecting faces using Viola-Jones face detector
3. Partition the database into training and testing set
4. Extract facial features for all images in the training dataset T
5. Train the dataset T using HOG feature descriptor
6. Choose a test image from the test dataset Te
7. Extract feature vector F using HOG feature descriptor
8. Predict the class of test input using SVM classifier based on matching data available in training features and labels
9. Perform feature matching to display the corresponding matching between a test image and trained images.

The model built in the training phase is used to predict the class of the subject by using SVM classifier. SVM can be used most frequently as a binary classifier. However, it can be used to classify multi-class data. The labels and features are used to train a model which can be used as an input to the SVM classifier for finding an exact match of the given test data against relevant subjects available in the database. SVM is trained to correct even possible misclassification when an unknown test input is given for recognition.

III. RESULTS AND DISCUSSION

The larger dimension of feature vector accommodates magnitude of all possible edge information in different orientations. the classifier result retrieving all matched subjects from the database against the given test input. Classifier output showing matched results from the database The conventional method of evaluating a classification model belonging to a supervised learning technique is to formulate confusion matrix over test data. The outcome of the classification results in the form of providing them in confusion matrix

This shows an overall accuracy of 90.2439% computed.

FACE NET ALGORITHM:

METHODOLOGY:

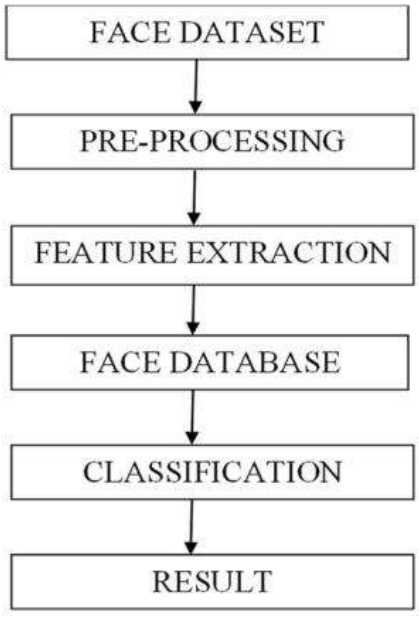
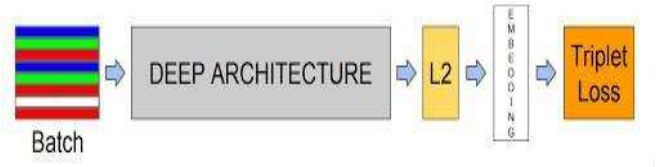


Fig 1: methodology of system

The above figure shows the methodology of our system. In this methodology, they are six (6) steps that involve in resulting a output. The first step of



The methodology is face dataset. In the face dataset we have taken eighty-six (86) images as a input. The second step is pre-processing. In this, it pre-processes the dataset by using three methods. They are: -

- Detection of an image: - In pre-processing, firstly it determines the position of the image.
- Cropping of an image: - In cropping of an image, it checks weather the image is in boundary box or not.
- Resizing of an image: - It resize the raw image into the particular pixels.

The third step of methodology is feature extraction. In this feature extraction, Face Net model is used. It extract feature based on the facial landmarks. The fourth step of methodology is face database. In the face database, it collects and store some face characteristics of an input image pixel after pre-processing. If these pixel matches with the pixels in the database.

The fifth step of the methodology of classification. In the classification SVM is used, because the SVM is one of the most popular supervised algorithms. It classifies the pixel by the measurement of the facial landmarks.

By taking the facial landmarks as a nodes. We take these nodes to calculate. X and Y are the two points represents the facial landmarks. The final step of methodology is result. In this result, based on the vector detection it calculates the how many times the image vector is matched that many times it displays the name of a person in a image.

This suggested strategy addresses the automation of a face detection and identification attendance system. This method consists of three modules:

1) Face Detection
2) Face Training
3) Face Recognition

FaceNet is deep neural network used for extraction features from an image of person's face. It was published in 2015 by Google researchers Schroffetal. This is also called as One-shot learning approach. FaceNet is a Method that ptimizes its embedding using deep convolutional neural network rather than intermediate bottleneck layers, as was the case with pervious deep learning approaches.

FaceNet uses batch as input and deep architecture that uses deep CNN, L2 normalization, and face embedding to produce the result. FaceNet was also pursed by triplet loss during training process. The anchor, the positive and the negative are three main components of triplet loss training methods.

MTCNN:

Multi-task cascaded convolutional neural network (MTCNN) is a framework developed as a solution for both face detection and face alignment. The process consists of three stages of convolutional neural networks that are able to recognize faces and landmark location such as an eyes, nose and mouth.

RESULT:

```
Running on device: cpu
Face detected with probability: 0.999994
Face detected with probability: 1.000000
Face detected with probability: 0.999989
Face detected with probability: 0.999997
Face detected with probability: 0.999995
Face detected with probability: 0.998926
Face detected with probability: 0.999990
Face detected with probability: 0.999995
Face detected with probability: 1.000000
Face detected with probability: 0.710982
Face detected with probability: 0.999974
Face detected with probability: 0.999858
Face detected with probability: 0.999653
Face detected with probability: 0.999999
Face detected with probability: 0.999965
Face detected with probability: 0.999980
Face detected with probability: 0.999998
Face detected with probability: 1.000000
Face detected with probability: 0.998885
Face detected with probability: 0.999997
Face detected with probability: 0.999998
Face detected with probability: 0.999265
Face detected with probability: 0.999816
Face detected with probability: 1.000000
Face detected with probability: 0.999960
Face detected with probability: 0.999785
Face detected with probability: 0.999999
Face detected with probability: 1.000000
Face detected with probability: 0.999983
Face detected with probability: 1.000000
Face detected with probability: 0.999994
Face detected with probability: 1.000000
Face detected with probability: 0.999972
Face detected with probability: 1.000000
Face detected with probability: 1.000000
Face detected with probability: 1.000000
Face detected with probability: 0.999993

Dataset accuracy:  95.45454545454545 % correct!!
```

CONCLUSION:

The choice between using the HOG-SVM or the FaceNet-MTCNN approach for face recognition depends on several factors such as the application, the dataset, and the available computational resources. The HOG-SVM approach is a traditional computer vision method that relies on handcrafted features extracted from the face image, followed by a classifier such as SVM. This approach can work well for face recognition tasks with small datasets and limited computational resources, but it may not perform as well as deep learning methods for larger datasets and more complex applications. On the other hand, the FaceNet-MTCNN approach uses deep learning methods to learn features directly from the face images and can achieve state-of-the-art performance on face recognition tasks. However, this approach requires a large amount of training data and computational resources.

Overall, if one have limited data and computational resources, the HOG-SVM approach may be a good option for face recognition. But if you have access to a large amount of data and computational resources, the FaceNet-MTCNN approach is likely to produce better results.++++++++++++++++

REFERENCES:

1. Create database by accessing files from original AT & T database in which original images in .pgm format are converted to .jpg format to match with inclusion of real-time images

2. Include 10 new user images by detecting faces using Viola-Jones face detector

3. Partition the database into training and testing set

4. Extract facial features for all images in the training dataset T

5. Train the dataset T using HOG feature descriptor

6. Choose a test image from the test dataset Te

7. Extract feature vector F using HOG featuredescriptor

8. Predict the class of test input using SVM classifier based on matching data available in training features and labels

9. Perform feature matching to display the corresponding matching between a test image and trained images.