A Smart Attendance System using Convolutional Neural Network

Manoj Kumar V[1*], Hariharan M[2], Aravinthan B[3], Ragavendra Vignesh S[4]

[1]*Assistant Professor, Department of Information Technology, Saranathan college of Engineering, Tamil Nadu, India.*
[2,3,4] *UG Student, Department of Information Technology, Saranathan college of Engineering, Tamil Nadu, India.*

[*]Corresponding author

Abstract

In surveillance scenarios, identifying a person captured on image or video is one of the key tasks. This implies matching faces on every still photos and video sequences. A Automatic face recognition for still photos with high quality will do satisfactory performance, aside from video-based face recognition it's exhausting to achieve similar levels of performance. Compared to still pictures face recognition, there are several disadvantages of video sequences. First, pictures captured by CCTV cameras are typically of poor quality. The background level is higher, and pictures could also be blurred because of movement or the topic being out of focus. Second, image resolution is generally lower for video sequences. If the topic is incredibly far away from the camera, the actual face image resolution can be as low as 64 by 64 pixels. Last, face image variations, like illumination, expression, pose, occlusion, and motion, are additional serious in video sequences. The approach can address the unbalanced distributions between still images and videos in a robust way by generating multiple abridges to connect the still images and video frames. So, in this project, we can implement still to video matching approach to detect the face from videos using Grassmann manifold learning approach and also recognize the faces using neural network algorithm to know unknown matches. Finally provide SMS alert at the time unknown matching in real time attendance environments.

Keywords: Grassmann, CNN, Image processing.

1. Introduction

In imaging science, image process is process of pictures mistreatment mathematical operations by mistreatment any sort of signal process that the input is a picture, a series of pictures, or a video, like a photograph or video frame; the output of image process is also either a picture or a group of characteristics or parameters associated with the image. Most image- processing techniques involve treating the image as a two-dimensional signal and applying commonplace signal-processing techniques to that. Images are processed as three-dimensional signals with the third- dimension being time or the coordinate axis. Image process typically refers to digital image process, however optical and analog image processes are also attainable. This article is regarding general techniques that apply to any or all of them. The acquisition of pictures (producing the input image within the 1st place) is brought up as imaging.

Closely associated with image process area unit special effects and laptop vision. In special effects, pictures area unit manually made up of physical models of objects, environments, and lighting, instead of being acquired (via imaging devices such as cameras) from natural scenes, as in most animated movies. Computer vision, on the other hand, is usually thought of high-level image process out of that a machine/computer/software intends to decipher the physical contents of a picture or a sequence of pictures (e.g., videos or 3D full-body resonance scans). In fashionable sciences and technologies, pictures conjointly gain abundant broader scopes thanks to the ever growing importance of scientific mental image (of usually large- scale advanced scientific/experimental data). Examples embody microarray information in genetic analysis, or time period multi- asset portfolio commerce in finance. Image analysis is that the extraction of vital data from pictures; in the main from digital pictures by means of digital image method techniques. Image analysis tasks may be as easy as reading bar coded tags or as refined as distinguishing an individual from their face. Computers square measure indispensable for the analysis of huge amounts of information, for tasks that need complicated computation, or for the extraction of quantitative data. On the opposite hand, the human cortical area is a superb image analysis equipment, particularly for extracting higher-level data, and for several applications — as well as medication, security, and remote sensing — human analysts still can't be replaced by computers. For this reason, several necessary image analysis tools like edge detectors and neural networks area unit galvanized by human seeing models.

Image piece of writing encompasses the processes of neutering pictures, whether they are digital photographs, traditional photochemical photographs, or illustrations. Traditional analog image piece of writing is understood as pic retouching, victimization tools like Associate in Nursing airbrush to change images, or writing illustrations with any ancient art medium. Graphic code programs, which might be broadly speaking sorted into vector graphics editors, formation graphics editors, and 3D modelers, area unit the first tools with that a user could manipulate, enhance, and transform images. Many image piece of writing programs are accustomed render or produce laptop art from scratch. Raster pictures square measure keep during a laptop within the sort of a grid of image parts, or pixels. These pixels contain the image's color and brightness data. Image editors will amendment the pixels to boost the image in some ways. The pixels may be modified as a bunch, or one by one, by the delicate algorithms among the image editors. This article principally refers to icon graphics editors, that square measure typically accustomed alters images and different formation graphics. However, vector graphics software system, like Adobe artist,

CorelDraw, Xara Designer professional, Pixel Style pic Editor, Inkscape or Vector, square measure accustomed produce and modify vector pictures, that square measure keep as descriptions of lines, Bezier curves, and text rather than pixels. It is easier to format ionize a vector image than to vectorize a formation image; the way to move vectorizing a raster image is that the focus of abundant analysis in the field of computer vision. Vector pictures may be changed a lot of simply, as a result of they contain descriptions of the shapes for straightforward arranging. They are additionally ascendible, being rasterizable at any resolution. Many graphics applications square measure capable of merging one or a lot of individual pictures into one file. The orientation and placement of every image may be controlled. When choosing a formation image that's not rectangular, it needs separating the perimeters from the background, additionally called silhouetting. This is the digital analog of surgery the image from a physical image. Clipping ways could also be accustomed add silhouetted pictures to vector graphics or page layout files that retain vector information. Alpha compositing, permits for soft semitransparent edges once choosing pictures. There square measure variety of the way to silhouette a picture with soft edges, as well as choosing the image or its background by sampling similar colors, choosing the perimeters by formation tracing, or changing a clipping path to a formation choice. Once the image is chosen, it's going to be traced and affixed into another section of identical file, or into a separate file. The selection may additionally be saved in what's called Associate in nursing alpha channel. A popular way to create a composite image is to use 66 transparent layers. The background image is employed because the bottom layer, and therefore the image with elements to be else square measure placed during a layer higher than that. Using a picture layer mask, all but the parts to be merged are hidden from the layer, giving the impression that these parts have been added to the background layer. Performing a merge during this manner preserves all of the peel information on each layer to a lot of simply modify future changes within the new unified image.

2. Steps of Image Processing

2.1. Image Acquisition

This is the primary step or method of the elemental steps of digital image process. Image acquisition might be as easy as being given a picture that's already in digital type. Generally, the image acquisition stage involves preprocessing, like scaling etc.

2.2. Image Enhancement

Image sweetening is among the best and most appealing areas of digital image process. Basically, the thought behind sweetening techniques is to bring out detail that's obscured, or simply to highlight certain features of interest in an image. Such as changing brightness & contrast etc.

2.3. Color Image Processing

Color image process is a neighborhood that has been gaining its importance as a result of the many increase within the use of digital pictures over the web. This may embody color modeling and process during a digital domain etc.

2.4. Wavelets and Multi-resolution Processing

Wavelets are the muse for representing pictures in varied degrees of resolution. Images subdivision is turn into smaller regions for knowledge compression and for paramedic illustration.

2.5. Compression

Compression deals with techniques for reducing the storage needed to avoid wasting a picture or the information measure to transmit it. Particularly within the uses of net it's substantially necessary to compress knowledge.

2.6. Morphological Processing

Morphological process deals with tools for extracting image elements that a helpful within the illustration and outline of form.

2.7. Segmentation

Segmentation procedures partition a picture into its constituent elements or objects. In general, autonomous segmentation is one among the foremost tough tasks in digital image process. A rugged segmentation procedure brings the method a protracted manner toward in resolution of imaging issues that need objects to be known on an individual basis. Segmentation is a classifier which helps to fragment each character from a word present in a given image or page. The objective of the segmentation is to extract every character from the text gift within the image. After playacting Segmentation, the characters of the string will be separated and it will be used for further processing. Different character segmentation techniques has been proposed until like, Dissection Techniques, Recognition Based Hidden Markov Models and Non-Markov Approaches, Holistic Strategies. By dissection is supposed the decomposition of the image into a sequence of sub pictures victimization general options. The structure consists of a set of states plus transition probabilities between states. A method stemming from ideas utilized in machine vision for recognition of occluded objects. A holistic method acknowledges a whole word as a unit.

2.8. Representation and Description

Representation and outline nearly always follow the output of a segmentation stage that typically is raw we knowledge, constituting either the boundary of a region or all the points in the region itself. Choosing an illustration is barely a part of the answer for reworking information into a type appropriate for ensuing laptop process. Description deals with extracting attributes that lead to some quantitative info of interest or a basic for differentiating one category of objects from another.

2.9. Object recognition

Recognition is that the method that assigns a label, such as, "vehicle" to an object based on its descriptors.

2.10. Knowledge Base

Data could also be as easy as particularization regions of a picture wherever the data of interest is thought to be situated, thus limiting the search that has to be conducted in seeking that information. The cognition can also be quite tough, like associate degree interconnected list of all major doable defects in an exceedingly materials review downside or a picture info containing high-resolution satellite pictures of a locality in reference to change-detection applications.

2.11. Feature Extraction

In pattern recognition and in image method, feature extraction could be a special type of dimensional reduction. Transforming the computer file into the set of options is termed feature extraction. If the options extracted ar rigorously chosen it's expected that the options set can extract the relevant info from the computer file so as to perform the specified task victimization this reduced illustration rather than the total size input.

3. Existing System

The term multi-view face recognition, in a very strict sense, solely refers to things wherever multiple cameras acquire the topic (or scene) at the same time Associate in Nursing a rule [1] collaboratively utilizes the no heritable images/videos. But the term has oft been accustomed acknowledge faces across cause variations. This ambiguity doesn't cause any downside for recognition with (still) pictures; a gaggle of images at the same time gaga multiple cameras and people gaga one camera however at completely different read angles area unit equivalent as so much as cause variations area unit involved. However, within the case of video information, the 2 cases diverge. While a multi-camera system guarantees the acquisition of multi-view information at any moment, the prospect of getting the equivalent information by employing a single camera is unpredictable. Given a try of face pictures to verify, they give the impression of being up within the assortment to "align" the face part's look in one image to an equivalent cause and illumination of the other image. This methodology also will need the poses and illumination conditions to be calculable for each face pictures. This "generic reference set" plan has conjointly been accustomed develop the holistic matching rule, wherever the ranking of look-up results forms the premise of matching live.

4. Proposed System

In the video scenes, human faces can have unlimited orientations and positions, so its detection is of a variety of challenges to researchers. The planned feature is developed victimization the spherical harmonic illustration of the face, texture mapped onto a sphere. The texture map for the full face is built by back-projecting the image intensity values from every of the views onto the surface of the spherical model. A promising approach to handle cause variations and its inherent challenges is that the use of multi-view knowledge. In video primarily based face recognition[5], nice success has been created by representing videos as linear subspaces that typically belong a special form of non- Euclidean house referred to as Grassmann manifold. From the angle of manifold learning, our technique is often considered performing arts a geometry-aware spatiality reduction from the initial Grassmann manifold to a lower-dimensional, a lot of discriminative Grassmann manifold

wherever additional favorable classification will be achieved. And also provide neural network algorithm to classify faces with improved accuracy in attendance system. Finally provide voice, SMS and Email based alert system with real time implementation.
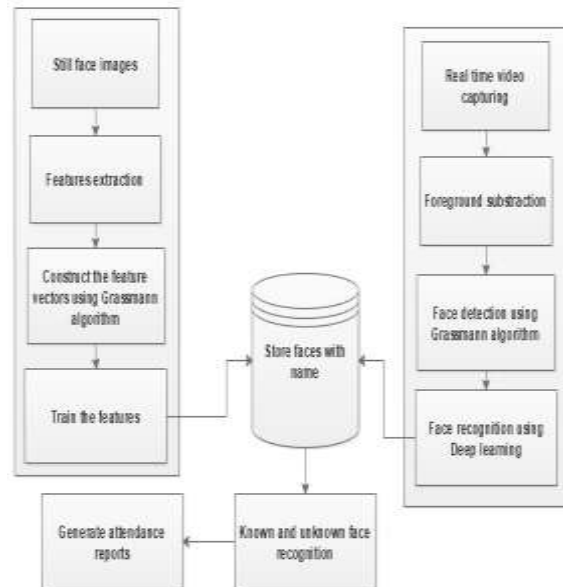
5. System Architecture

Fig.1. System Architecture
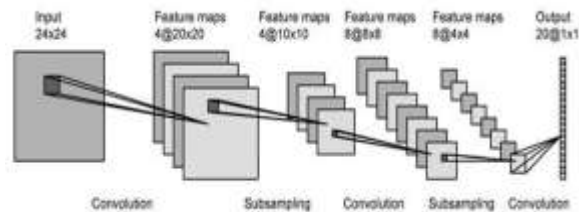
6. Proposed Methodology

Fig.2. CNN Architecture

7. MODULES

7.1. Face image acquisition

Description: In this module, admin can train multiple faces. Face may be captured through web cameras or uploaded as still pictures. In this image, user faces without occlusion, straight pose and normal light conditions

7.2. Features extraction

This module, facial features are extracted. And constructed as feature vectors. Facial features include nose part, eye parts and lip part. These values are stored is in the form of matrix. Grassmann algorithm is used in this process

7.3. Register the faces

Face registration is the process of transforming different sets of data into one coordinate system. Facial features are stored with labels. Labeling the faces using their names.

7.4. Face classification

This module is known as login phase or testing phase. Input is in the form of real time video capturing. Video images are splited into still images. Face detection is done in the process. Matching the features using deep learning algorithm.

7.5. Alert system

Matching the testing face with database still faces. If the feature vectors are matched means, face image is labeled as known faces. If the feature vectors are not matched means, considered as unknown faces. Create alert for unknown labeling to admin. And generate the reports for all students.

8. TECHNIQUES USED

Object Detection victimization Haar feature- based cascade classifiers [2] may be a smart object detection technique planned by Paul Viola and Archangel Jones in their paper, "Rapid Object Detection uses a Boosted Cascade of Simple Features" in 2001. It is machine learning primarily based approach wherever a cascade performs is trained from heaps of positive and negative pictures. It is then accustomed discover objects in alternative pictures.

Here we will work with face detection. Initially, the algorithmic rule desires heap of positive pictures (images of faces) and negative pictures (images while not faces) to coach the classifier. Then we need to extract features from it. For this, haar options shown in below image area unit used. They are just like our convolutional kernel. Each feature may be a single worth obtained by subtracting total of pixels underneath white parallelogram from total of pixels underneath black parallelogram.
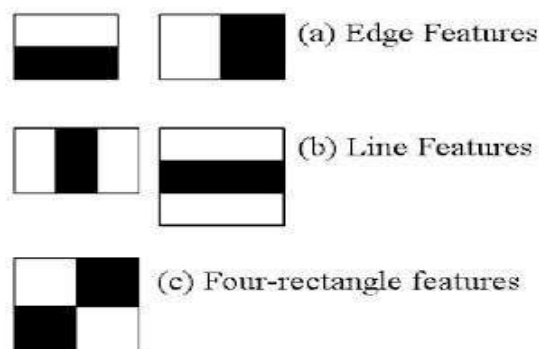

Fig.3. Feature Extraction

We apply every and each feature on all the coaching pictures. For each feature, it finds the simplest threshold which can classify the faces to positive and negative. But obviously, there will be errors or misclassifications. We choose the options with minimum error rate, which implies they're the options that best classifies the face and non-face pictures.

Fig.4. Edges pointering

First we need to load the required XML classifiers. Then load our input image (or video) in grayscale mode

Fig.5. Results

9. Psuedo Code

Step 1: Read the face boundaries from video frames
Step 2: Boundary values are constructed as feature vectors
Step 3: Predict the tangent vector
Step 4: Calculate distance values from query frames with still images
Step 5: If distance value is equal to zero or less than the minimum means, match found
Step 6: Otherwise send alert about unknown person

10. Implementation

10.1. Face Image Acquisition

A digital video surveillance system is a surveillance system capable of capturing images and videos that can be compressed, stored or sent over communication networks. Digital video surveillance systems can be used for nearly any environment. The main difference between a digital video surveillance system and an analog video surveillance system is that a digital video surveillance system is capable of capturing and storing the video signal in a digital format. Since the data is captured in a digital medium, this eliminates the need for any conversion. Most digital video surveillance solutions are capable of being managed from anywhere and provide interoperability. The cameras are networked and footage is digitally archived, which is considered economical for most businesses. A face recognition system is a computer application capable of identifying or verifying a person from a digital image or a video frame from a video source. One of the ways to do this is by comparing selected facial features from the image and a face database. Recognition algorithms can be divided into two main approaches, geometric, which look at distinguishing features, or photometric, which is a statistical approach that distills an image into values and compares the values with templates to eliminate variances. Face recognition from image or video is a popular topic in biometrics research.
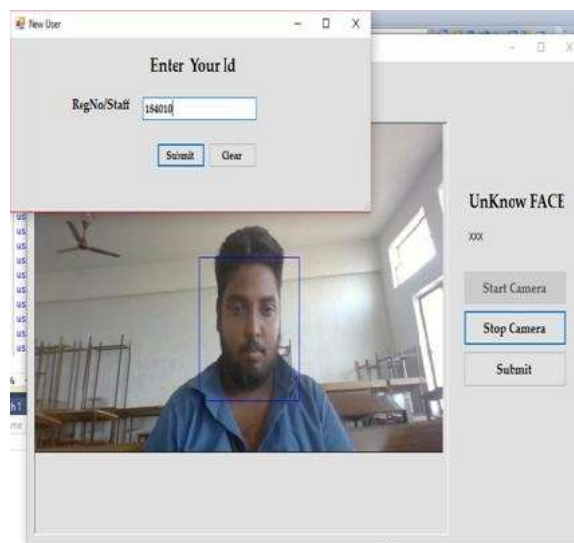
Fig.6.Image acquistion

10.2. Features Extraction

Applying human visual property in the recognition of faces, people can identify face from very far distance, even the details are vague. That means the symmetry characteristic is enough to be recognized. Human face is made up of eyes, nose, mouth and chin etc. There are differences in shape, size and structure of those organs, so the faces are differ in thousands ways, and we can describe them with the shape and structure of the organs so as to recognize them. One common method is to extract the shape of the eyes, nose, mouth and chin, and then distinguish the faces by distance and scale of those organs. The other method is to use deformable model to describe the shape of the organs on face subtly

Fig.7.Extraction of features from face

10.3. Register the Face

Face registration is the process of transforming different sets of data into one coordinate system. Facial features are stored with labels. Image registration or image alignment algorithms can be classified into intensity-based and feature-based. Face recognition systems identify people by their face images. Face recognition systems establish the presence of an authorized person rather than just checking whether a valid identification (ID) or key is being used or whether the user knows the secret personal identification numbers (Pins) or passwords. The following are example. To eliminate duplicates in a nationwide voter registration system because there are cases where the same person was assigned more than one identification number.

Fig.8.Detail window for registering user information

10.4. Face Classification

Face recognition have gained a great deal of popularity because of the wide range of applications such as in entertainment, smart cards, information security, law enforcement, and surveillance. It is a relevant subject in pattern recognition, computer vision, and image processing. Human's face is a prominent feature in machine learning and computer vision system. A face conveys various information including gender, age, ethnicity etc. Face information is applicable in many sectors like biometric authentication and intelligent human-computer interface. Since our main concern of this paper is gender classification from human faces so a proper localization of human face area is necessary. For both face detection and gender classification purpose selection of color space for detecting skin region is a main concern. Extracting two sets of data for both male and female and separate them accurately is a challenging job. So we need to select a better classifier to improve the classification performance. Face identification is a one-to-many matching process that compares a query face image against all the template images in a face database to determine the identity of the query face. The identification of the test image is done by locating the image in the database that has the highest similarity with the test image.

Fig.9.face classification

10.5. Alert System

In many of the access control applications, such as office access or computer logon, the size of the group of people that need to be recognized is relatively small. The face pictures are also caught under natural conditions, such as frontal faces and indoor illumination. The face recognition system of this application can achieve high accuracy without much co-operation from user. Face recognition technology is used to monitor continuously who is in front of a computer terminal. It allows the user to leave the terminal without closing files and logging out. When the user leaves for a predetermined time, a screen saver covers up the work and disables the mouse & keyboard. When the user comes back and is recognized, the screen saver clears and the previous session appears as it was left. Any other user who tries to logon without authorization is denied

Fig.10.Alert generation for unregistered face

11. Conclusion

In the modern era of computing, everything gets automated with the help of computers. In such a situation why we have to follow the same odd procedure in college also. In order to realize this huge vision. First one is attendance; in college every period attendance was taken. A lot of time gets wasted because of this. In order to reduce the time wastage. We planned to automate the process of attendance through face recognition technology A faculty have to take a photograph of students in order to mark attendance. Through face recognition attendance for every student will get mapped automatically. Automation is a word which became an inevitable thing. Ranging from small tasks to large herculean task, everything gets automated.

In such a situation, now the time came to automate the process something in the field of education, particularly in colleges. For attendance we are going to use face recognition technology through which every person face gets mapped with the registered faces in the database. As a result attendance will be marked for every student.

REFERENCES

[1]. Y. Yan, F. Nie, W. Li, C. Gao, Y. Yang, and D. Xu, "Image classification by cross- media active learning with privileged information," IEEE Transactions on Multimedia, vol. 18, no. 12, pp. 2494– 2502, 2016.

[2]. X. Chang, F. Nie, S.Wang, Y. Yang, X. Zhou, and C. Zhang, "Compound rank-k projections for bilinear analysis," IEEE Transactions on Neural Networks and Learning Systems, vol. 27, no. 7, pp. 1502– 1513, 2016.

[3]. S. Liao, Y. Hu, X. Zhu, and S. Z. Li, "Person re-identification by local maximal occurrence representation and metric learning," in Proc. CVPR, 2015, pp. 2197– 2206.

[4]. Y. Yang, Z. Ma, A. G. Hauptmann, and N. Sebe, "Feature selection for multimedia analysis by sharing information among multiple tasks," IEEE Transactions on Multimedia, vol. 15, no. 3, pp. 661–669, 2013.

[5]. Y. Yang, F. Nie, D. Xu, J. Luo, Y. Zhuang, and Y. Pan, "A multimedia retrieval framework based on semi- supervised ranking and relevance feedback," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 34, no. 4, pp. 723–742, 2012. Netherlands, 2002.