Surveillance System with Impediment of Face Discovery Using Python

Er. Sital Prasad Mandal (Ph.D. Scholar)\textsuperscript{1}, Sunil Sharma\textsuperscript{2}
Email: info.sitalmandal@gmail.com\textsuperscript{1}
Email: sunilsharmaxt@gmail.com\textsuperscript{2}
Teaching Assistant\textsuperscript{1,2}
Department Of Computer Science & Application
Mechi Multiple Campus, Bhadrapur, Jhapa, Nepal

Abstract

In recent years, It is an active area of research and development with there ongoing and active work on computer vision. This paper presents an empirical literature review of the existing research on human facial recognition as up-to-date. This research start providing the reader with a brief on the uses and working of face recognition. The most recent approaches to facial recognition are then reviewed in the literature. The performance and drawbacks of these face recognition algorithms are also tested through use of the Labelled Wild Face (LWF) face dataset images as well as real time images. With the increasing use of real-time facial recognition systems, it is important to examine the accuracy of these systems given the reliability and 94.23% accuracy are achieved for facial recognition by the proposed model.

Keywords: Neural Networks, Graph Matching, Facial Recognition, Combined Classifiers.

Introduction

Face detection and face awareness are amongst the most necessary lookup subjects in computer vision, as many purposes use faces as objects of biometric technology. In the recognition of faces is a significant research issue that spans a wide range of fields and disciplines. This is due to the Facial Recognition (FR) is a mortal that's necessary for effective dispatches and relations among individuals. In expansion to having operations similar as bankcard recognizable proof, get to control, shots looking, security observing, and observation frameworks, FR also has these operations. In the first formal approach to faces classify was proposed. The author suggested collecting biographies in the form of angles, determining their norm, and also grading other biographies according to how different they're from the norm. As a result of this, a vector of measures are compared to different dataset. FR are now being demonstrated in real world settings due to advances in technology. FR's rapid-fire development can be attributed to a number of factors a system for assessing the performance of FR algorithms, the vacuity of large trained data of facial images, and the active development of algorithms. The FR problem can be described as follows in the literature Compare the faces in a database to static( still) or videotape images of a scene to identify or corroborate one or further people in the scene. The proposed device is predicted to be used for concurrently discovering and figuring out faces in actual time. This project's primary goal is to develop a real-time FR system that can recognize or confirm people based on video frames. First, we must determine whether the face is actually present in the frame in order to identify it. Select as a Region of Interest( ROI) and process it for facial recognition if it's there.

Statement of the Problem

Over the decade, vision experimenters and man-made product manufacturers have improved the overall performance of automatic FR algorithms across a wide range of sensitive FR tasks. With mortals now performing facial recognition tasks in the most real-world security situations,
it's unclear whether the use of algorithms improves security or poses a greater threat. The real challenge of face and knowledge discovery technologies is to implicitly handle all eventualities where the actors don't cooperate and the opt-in range is unlimited. There are several foundations that change facial expressions. These sources of variation in facial appearance can be classified into two groups: Intrinsic Variables and Outward Factors. *Intrinsic Variables:* The facial expressions are anticipated to directly reflect the existent's passions and are not told by the observer. These factors can be classified into two types intrapersonal and relational. Intrapersonal factors are responsible for altering the facial features of the same person, analogous as age, appearance, and physical attributes like beard, specs, and makeup. In distinction, relational factors regard for the variations in facial appearance among different individualities, including their personality traits and exposure. *Outward Factors:* The face is altered through the interplay between light, the face itself, and the observer. These variables encompass light, posture, scale, resolution, noise, and so on.

**Justification/Rationale of the Study**

The following are typical issues and challenges that facial recognition systems may encounter while identifying and recognizing faces:

* Naturally find the face: The initial step in a facial system is to detect or locate a face within a photo or video. However, it is not always easy to determine the position of the head within a sequence of images. For instance, in a crowded public area, such as a video surveillance system, identifying a face can be challenging due to constant movement. Furthermore, the complexity of the background can also make facial detection more difficult.

![Figure 1: Face Detection](image)

* Façade: Variations in facial orientation within an image are also a significant challenge for facial recognition systems. The position of a face changes with the bystander's viewing angle and the gyration of the head. These changes in posture pose a significant challenge for directly relating the input image. While a facial recognition system can handle cases with minor gyration angles, it becomes increasingly delicate when the gyration angle is advanced. If the image in the trained data only shows the front-facing perspective of the face, this may differ from the input image, resulting in flawed identification or no recognition at all.
Figure 2: Variations in Facade

**Ageing:** Human face is not a static, rigid entity. It evolves with time, and as an individual ages, their appearance changes, which can affect facial recognition systems, as illustrated in the figure.

Figure 3: Ageing variations

**Impediment:** Impediment refers to a blockage or obstruction. In facial recognition systems, when the complete face or a series of images are not available as input due to the presence of blocking objects such as glasses, facial hair, mustaches, etc., it is referred to as an impediment, which is a major challenge in facial recognition, as shown in the figure. In applications of real world, common for individuals to be talking on phone or wearing hats, glasses, or have their faces covered with objects for various reasons. Such situations can seriously affect the recognition system's identification process.

Figure 4: Partial occlusion in images

3.5 **Low Resolution:** The low-resolution issue arises in facial recognition systems, resolution of the image to be honored is lower than 16x16. This issue is current in numerous surveillance operations, similar as standalone camera systems in supermarkets and banks, CCTV cameras in public thorough fares, etc., where images captured by surveillance cameras generally have a small face area and cannot give sufficient resolution for FR. When an existent's face is far down from the camera, the face area will be lower than 16x16, performing in a low-resolution face
image that contains limited information, and utmost of the details are lost. This can significantly reduce the recognition rate.

Figure 5: Typical frame from a surveillance video (CAVIAR database)

These issues can be expected framework shortcomings utilized in face acknowledgment, like camera bending, foundation clamor, wasteful capacity, inappropriate procedures and so forth. More than that there can be network issues due to ecological circumstances.

Objective of the Study

The aim of this study is to examine the historical evolution of FR, which is one of the most challenging issues in the field of image recognition. A human face is not only a three-dimensional object, but it is also a non-rigid body. Additionally, facial images are often captured in natural environments, where the background can be complex, and lighting conditions can be intense. Face detection is a crucial issue in FR. Before recognizing a face, it is necessary to locate the face accurately and determine its position. In practical systems, segmentation is usually the most challenging task to solve. The FR problem can be formulated into two categories in the problem statement.

A. Detection of Real Time Face: Detection of Real Face refers to the process of recognizing a face in multiple frames captured by a video device. Compared to detecting faces in static images, real-time face identification is less complicated, but requires more advanced hardware. This is due to the fact that humans are always in motion, blinking, twitching, gesturing, etc.

B. Face Detection in Photos/Video: On the other hand, most detection systems, it extract a portion of the entire face, excluding non-essential elements such as background and hair. This is often done on static images by dragging a cursor across the image, after which the system verifies if there is a face inside the window (Brunelli and Poggio, 1993). Unfortunately, the search area for a face in a static image is quite wide.

Review of the Literature

S. M. M, A. Geroge et al. [1] The paper compares deep learning-based face popularity algorithms like R-CNN, Fast R-CNN, Faster R-CNN, and YOLO.V3, with YOLO.V3, a one-shot recognition method, being suggested for real-time face reputation. The study uses a unique dataset of a famous person's face, showing that YOLO.V3 is faster than R-CNN and its variations, making it suitable for real-time applications.

H. Aung et al. [2] The paper presents a methodology that uses the YOLO algorithm and the VGG16 convolutional neural network to identify face blends in live video. The VGG16 grid is used for extraction, and the YOLO.V2 environment is adjusted for visibility. The method accurately recognizes faces in various lighting, skin, and shadow situations, provides high-quality pictures, real-time video streaming, and allows for immediate use and planning. It enhances conversation skills for face recognition.
Aanchal Singh et al. [3] The important problem of facial recognition in Computer Aided Design (CAD). With the benefit of more practice, it offers a brand-new, effective way to recognize faces that is based on VGG16 rules. For the test, you have to look at pictures of faces in JPEG, GIF, and PNG forms, which means you need to know how to tell the difference between facial features. The results showed that the first optimized scheme got 83.11% of the tests right, while the second optimized scheme got 92.8% of the tests right.

B.R. Maale et al. [4] proposed a methodology of a three-step forward recognition method based on Haar Cascade Classifiers (HCC). It uses 1000 pictures to get 94% accuracy in recognizing faces, with a focus on dance styles like Kathak and Kuchipudi. The study shows how useful necklace cascade classification is for real-time face recognition and gives information about different uses. A big part of how well this system works is the HCC, which use the cascade design and the AdaBoost Algorithm. The study is mostly about Kathak and Kuchipudi dance and gives data and suggestions on whether the program can work.

M.A.Hoque et al. [5] proposed a methodology to improve security monitoring, give us a separate facial recognition system that works with real-time streaming video. The system utilizes an ATMega328p microcontroller-based Arduino Uno with a Pan-Tilt device, and an open-source computer vision system with OpenCV for image processing. The HCC method is used to recognize faces. Different situations showed that the method could work well. It can find faces in live video and send out alerts, so it can be used to make restricted places safer, like prisons and border posts.

C. Jianfei and Z. Changming [6] The paper proposes a methodology to improve the ResNet model by adding more shortcut links, switching up convolution methods, and increasing the dataset size. Real-world results from tests on the CIFAR-10 and CIFAR-100 datasets show that these changes enhance the model's accuracy. However, comparing the approach to existing methods or standards may be challenging.

Yaswanthram et al. [17] The system checks the person's identity by comparing their face's patterns, features, and shape to a database that has already been saved. This is called face recognition. Face recognition technology like this is used in biometric systems in companies to make sure people are who they say they are. Lessening the number of dimensions has a bigger effect on how well machine learning systems can recognize faces. The methods tested included Random Forests, SVM, Linear Regression, Logistic Regression, and KNN. It was found that Logistic performs better, with a 97% success rate without Principle Component Analysis (PCA). With PCA, Logistic Regression gets a 93% success rate.

Raju et al. [18] Nowadays Accidents on the roads have become more common because drivers are careless and don't pay attention. By taking out facial traits, you can make a smart driver alerting system that can help keep car accidents from happening. It was made that driver monitoring device with an camera and alarm. The user's face is what the camera sees. Changes in the driver's face expression are tracked and used to figure out if the driver is antsy. The warning goes off to let the driver know something is wrong.

Aishwarya et al. [19] A very important way to make computers smarter is to teach them to recognize human emotions. A method for detecting emotions from facial expressions is created. To get the face features, Histogram Of Oriented Gradients (HOG) and HCC algorithms are used. Extracted feature is used to sort feelings into groups, such as angry, disgusted, scared, happy, neutral, and sad. With a success rate of 65.5%, the HOG method is more reliable than other methods that are already out there.
Kalturi et al. [20] Face recognition technology can be used for many things. One of the best things about these kinds of tools is that they can identify and verify users. A smart door unlocking model built on AI that uses face recognition was made for blind. This text message goes to the registered phone number and it sets alarm for safety. Images of approved users are used to teach the model how to open the door. Unauthorized users will not be able to get in.

Gesture Tek [7] is a renowned organization that specializes in pattern recognition. Their depth shadowing software allows relations with simple hand movements rather of a remote control, keyboard, or touch screen. still, this technology is presently only used on particular computers. also, a cons of this system is that the 3D seeing is a single depth seeing camera, which is precious due to the ray used for depth seeing.

The Prime Sensor technology developed by Prime Sense [8] enables gesture recognition through depth information obtained by an IR sensor. The technology is integrated into the Prime Sensor Reference Design, which is grounded on the SoC. This SoC includes largely resemblant sense that takes an infrared pattern as input and produces a VGA size image of the scene. The target followership of the proposed system, called "Controlling Mobile bias via Gesture Recognition," is mobile phone druggies. With its stoner-friendly design that doesn't bear any physical contact, mobile phones equipped with this system would be an ideal.

[9] NSW is a company that creates computer-related products and services, including tackle, software, and IT results. They've created Gesture Software (3) which can identify the color and shape of the hand and distinguish hand's stir from different objects background. NSW uses software as input system for controlling the computer's cursor as the hand moves.

[10] The German company Evoluce has introduced, a Multi-touch HD TV Screen that incorporates gesture technology. This technology allows druggies to operate multi-touch operations by performing hand gestures above the screen without physically touching it. These gestures include rotating, scrolling, stretching, or rotating, which are snappily detected and restated into conduct on the screen, similar as pinch & drone, screen transition, or operation redundancy.

This section provides a summary of the fundamental techniques used for human FR, which typically apply to front-facing faces. Additionally, it outlines the benefits and drawbacks of each approach. The integrated software is designed to identify a person's facial image from a dataset. Software have variations in intensity, expression and pose, within a certain range. A condition of image dataset was used to test the software. Image features around image shape and points of image features of the eye and eyebrows were used to create a signature of the image. The existing literature indicates that there is a research gap in comparing faces in a trained data or video images of a scene for the purpose of identifying or verifying individuals. This approach allows for a claim to be verified without the need to refer to the entire dataset of images, known as model images. An input image, known as a probe image, is compared to a model image of the person whose identity is claimed. To achieve this, a neural network is trained using supervised learning with sample of "non-face" and "face" images to distinguish between the two in a picture, as seen in a face detection system.

**System Implementation**

*Understand what is OpenCV*

The implementation of this project utilizes OpenCV-Python, which is a Python-based open-source object-oriented language. Python is known for its effectiveness in areas such as Application Programming Interface (API), Low-Level Programming, Object Linker Embedding,
Simulation, Platform Independence And Network Configuration. One of the benefits of Python is that when a program is developed, it does not need to be compiled before each run.

OpenCV are available in different OS such as Windows, Android, and iOS. It also supports various programming languages including Java, Python, and more. OpenCV-Python have many features of the OpenCV API and it provide an excellent libraries.

Implementation using Python

The OpenCV-Python library is utilized for image and video processing and can be used for various applications, such as facial detection, photo editing, and robotic vision. The FR library is used to aid in the development of code. The library uses machine learning algorithms to identify faces within an image by breaking down the task into smaller, easier-to-solve classifiers. The process of how machines recognize faces differs from human facial recognition, making it an interesting topic to investigate. This essay focuses on exploring FR with minimal facial data and will use a combination of OpenCV-Python and Object-Oriented Programming Language (OOP) to properly classify and identify a human face.

In order to detect faces, classifiers are used in which an algorithm decides whether an is positive image mean a face or negative mean not a face. Fortunately, OpenCV offers two pre-trained detection classifiers that can be utilized. To train a classifier, thousands of images with and without faces are required.

Types of Classifiers

- **a. HAAR Features:** 'HAAR features' extraction is a method used in computer vision to identify objects in images or videos. It involves analyzing small rectangular regions in the image and calculating the difference between the sums of pixel values in each region. This technique is often used in face detection algorithms as Figure 6.

\[ \text{Figure 6: A couple HAAR like features (SuperDataScience Team, 2017)} \]

The features act like windows over images, calculating a single feature by subtracting the sum of pixels under the white region from that under the black region.

- **b. Integral Images**

'Integral Images' is a technique used in computer vision to speed up image processing. It involves creating a new image that stores the sum of all pixels' values in a rectangular area of the original image. This new image can then be used to quickly calculate the sum of pixel values within any rectangular region of the original image. This technique is often used in face detection algorithms to efficiently calculate HAAR-like features. Viola-Jones algorithm employs a base window size of 24x24 pixels, which leads to the computation of over 180,000 features within this window. By using the integral image, we can easily compute the sum of all the pixels under any rectangle, as we only need to know the four corner values of that rectangle as below Figure 7.
Using four array references, we can determine the sum of pixels within rectangle D in Figure 8. The integral image value at location 1 corresponds to the sum of pixels in rectangle A. Similarly, at location 2, it represents the sum of A and B, at location 3 it represents A and C, and at location 4, it signifies the sum of A, B, C, and D. To compute the sum within rectangle D specifically, we can apply the formula 4+1-(2+3).

**Data Acquisition**

The Labeled Faces in the Wild (LFW) dataset, featuring real time images, was utilized for face identification and recognition tests. Labeled Faces in the Wild (LFW) [21] as sample with up-to-date population is a collection of pictures of faces that was made to study the problem of recognizing faces in a free setting. The database has 13,233 pictures of 5,749 people that were found on the Internet and recognized by the Viola-Jones face detector [22] as sample with up-to-date population. Experiments were done on 1700 images, which include 1,602 pictures of person with two or more different photos in the dataset and 20 pictures taken from real time. BU3DFE [23] as sample with up-to-date population of the database that is used for the 3D face reconstruction studies.

**Table 1:**
The proposed model’s similarity function between labels and names is shown in the table below.

<table>
<thead>
<tr>
<th>Sl.No</th>
<th>Labels of the image to be recognized</th>
<th>Number of images</th>
<th>Name detected by the proposed model</th>
<th>Similiarity Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Abdoulaye_Wade</td>
<td>4</td>
<td>Abdoulaye_Wade</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>Bishal Ghimire</td>
<td>2</td>
<td>Bishal Ghimire</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>Pramod Kumar Soni</td>
<td>3</td>
<td>Pramod Kumar Soni</td>
<td>1</td>
</tr>
<tr>
<td>4</td>
<td>Sital</td>
<td>2</td>
<td>Sital</td>
<td>1</td>
</tr>
<tr>
<td>5</td>
<td>Abrial</td>
<td>3</td>
<td>Abrial</td>
<td>1</td>
</tr>
<tr>
<td>6</td>
<td>Messi</td>
<td>2</td>
<td>Messi</td>
<td>1</td>
</tr>
</tbody>
</table>

The formula for:

\[
\text{Similarity (T1,T2)} = 2 \times \frac{M}{T} \quad \text{........(eq. 1)}
\]

Where,

M represents the number of characters in both strings
T represents the total number of items in both sets.
T1 and T2 acquisition schemes were similar.

\[
\text{Accuracy in } \% = \frac{\text{Number of images correctly recognized (m)}}{\text{Total Number of images (t)}} \times 100
\]

\[\text{........ (eq. 2)}\]
Table 2:

Proposed model's accuracy on the LWF dataset

| Number of images correctly recognized (m) | 1602 |
| Total number of images (t) | 1700 |
| Accuracy in % | 94.23% |

Several sets of trained and test pictures were fed to the model during the experiments. There are different lighting conditions, backgrounds, and angles used to take the pictures. If the model finds a match between the test picture and the training image, it labels the face with the name of the image. Table 1 shows how Eq. (1) is used to figure out the Similarity Index between the image's label and the names that the model picked out after recognizing the face. Equation (2) is used to figure out how accurate the model is. There were 1700 images used in the experiments. These included 1,602 images of person with two or more different shots in the dataset and 20 images taken from real time. Table 1 shows that the model is correct 94.23% of the time.

Table 3:

The proposed method for face recognition was compared to other methods that are already out there.

<table>
<thead>
<tr>
<th>METHOD</th>
<th>Shape</th>
<th>Texture</th>
<th>ACCURACY</th>
</tr>
</thead>
<tbody>
<tr>
<td>3DMM(LFW)</td>
<td>N</td>
<td>N</td>
<td>63.34%-68.92%</td>
</tr>
<tr>
<td>3DDFA(LFW)</td>
<td>Y</td>
<td>Y</td>
<td>73.79%-76.07%</td>
</tr>
<tr>
<td>PROPOSED (LFW-IMAGES CAPTURED IN REALTIME)</td>
<td>Y</td>
<td>Y</td>
<td>64.42-69.54%</td>
</tr>
<tr>
<td></td>
<td>Y</td>
<td>Y</td>
<td>94.23%</td>
</tr>
</tbody>
</table>

Face Recognition Accuracy

In Table 3, you can see how well the suggested method (LFW dataset with Images captured from real time) recognizes faces compared to other methods on the LWF dataset. The accuracy of face recognition on the LWF dataset ranges from 63.34% to 68.92% when the 3DMM method is used alone. Similarly, there are differences in accuracy when either one or both of the factors are taken into account using different methods. The suggested method works well for images that aren't frontal and isn't affected by changes in lighting or pose. It's 94.23% accurate when both shape and texture are used as parameters. Computer vision is starting to do more and more work with face recognition. There are already a number of different ways to complete the job. Table 3 shows how accurate the suggested method is compared to other methods that are already used.

Results and discussions

The main rule is that video is captured by the camera and the captured video is turned into image which converted as viewing. In addition, it detected the know image for dataset, else the system is set as unknown image and marks image as missing. In each execution, while running the program it automatically starts camera which search an face and when face is detecting in frame, it display an name of person from dataset as shown below figure .8.
Fig 8: Result Face Detection

Conversely, the system recognizes the video face which shows name on the display screen, If the face detected is register. If the face isn't registered the program will not display name.

Conclusion

The discipline of computer vision is still with the problem of FR. Due to its numerous uses in different fields, it has drawn a lot of interest in recent years. Although there is a lot of research being done in this field, facial recognition algorithms are far from ideal and cannot function properly in all real-world scenarios. The paper provided a succinct overview of the problems, solutions, and applications related to facial recognition. There is still much work to be done to develop techniques that represent people faces and utilize the temporal evolution of a face's recognition. Face detection and recognition tests were done on 1700 pictures chosen from the LWF dataset as well as some images that were taken in real time. Face mesh is used to put together a full face using markers on the face. The LFW dataset is looked at from various views to figure out how accurate the face reconstruction is. It is possible to find and recognize non-frontal human faces of people of all ages and races. The efficiency of face recognition has been compared to other methods that are already in use. For the tests done with different restrictions, the model gets an accuracy of 94.23%.

Acknowledgment

The authors express gratitude to senior taxonomists Dr. Rajendra Adhikari, Naryan Dhamala, and Pradip Khatiworda for their support and suggestions during the study, and to friends Raju Poudel, Pramod Kumar Soni, Bishal Ghimire, and Pankaj Uprety for their assistance in collecting a real labeled image dataset. We would also like to acknowledge to R.M.C. and all the faculties & staff of Mechi Multiple Campus for providing all the facilities concern for this mini research.

References


