Indian Journal of Computer Science and Technology
https://www.doi.org/10.59256/indjcst.20240302021
Volume 3, Issue2 (May-August 2024), PP: 140-147.
www.indjcst.com
ISSN No: 2583-5300

A Recurrent Neural Network on Video-Based Face Identification Suspicious and Also Wanted People Recognition in ATM

R. Vasugi[1], Joiceline J[2], Monika N[3], Piyarijon F[4], Priyanka S[5]
[1]*Assistant Professor, Department of Information Technology, Er. Perumal Manimekalai College of Engineering, Hosur, Tamilnadu, India.*
[2,3,4,5] *Department of Information Technology, Er. Perumal Manimekalai College of Engineering, Hosur, Tamilnadu, India.*

To Cite this Article: *R. Vasugi[1], Joiceline J[2], Monika N[3], Piyarijon F[4], Priyanka S[5], "A Recurrent Neural Network on Video-Based Face Identification Suspicious and Also Wanted People Recognition in ATM", Indian Journal of Computer Science and Technology, Volume 03, Issue 02 (May-August 2024), PP: 140-147.*

Abstract: *The need for an intelligent recognition system is still increasing. Traditional approaches based on access to restricted places like ATM or suspected actions as theft, scam, and loitering. They are insufficient to identify suspects. These actions do not represent a real key of suspects. This project is motivated not only by the limits of the traditional approaches but also by the complexity of Intelligent Algorithms (RNN). In this project, we propose an approach for the automatic comparative labeling of facial soft biometric. The three main categories wanted persons, regular customers, and new persons serve distinct purposes. For wanted people recognition, the system is trained to identify individuals with outstanding or flagged in law enforcement databases. Regular ATM customers are recognized based on their registered facial features. New persons, on the other hand, are those not previously registered, and the system is designed to capture their facial data for future reference. Using a subset from the RNN-network datasets, recurrent neural network our experiments show the efficacy of the automatic generation of comparative facial labels, highlighting the potential extensibility of the approach to other face recognition scenarios and larger ranges of attributes. Recurrent Neural Network (RNN) system can be implement on CCTV cameras and it will be alerted particular officers. The application is easy to use and it is equipped with pyttsx3 so that the face detected is communicated to the people as voice alert.*

Key Words: *ATM, CCTV, Recurrent Neural Network (RNN) System, Intelligent Algorithms, Scam, Loitering and Theft.*

I.INTRODUCTION

1.1 Overview

In response to the evolving challenges in traditional recognition systems, particularly those relying on restricted access and predefined actions, this project introduces an innovative approach leveraging intelligent algorithms, specifically Recurrent Neural Networks (RNN).

The traditional methods, often limited in recognizing suspects or responding to dynamic situations such as virus outbreaks, necessitate a more nuanced and adaptive solution. The proposed system focuses on the automatic comparative labeling of facial database, categorizing individuals into three main groups: wanted persons flagged in law enforcement databases, regular customers identified by their registered facial features, and new persons not previously recorded. The use of RNN-network datasets demonstrates the effectiveness of automatically generating comparative facial labels, the potential extension of this approach to diverse face recognition scenarios and a broader spectrum of attributes. By implementing the Recurrent Neural Network system on web cameras, the project aims to enhance surveillance capabilities, automatically alerting designated officers to the presence of wanted individuals and providing a robust foundation for future facial data reference.

1.2 Objective

- The main objective of this project is to improve the identification of suspects by implementing a face recognition system based on Recurrent Neural Networks (RNN). By training the system to identify wanted person in law enforcement databases, the project aims to enhance the accuracy and efficiency of suspect recognition and identification.
- This project aims to overcome the limitations of traditional approaches by leveraging advanced algorithms, contributing to more accurate, efficient, and adaptive facial recognition capabilities in various security and surveillance scenarios.

II.LITERATURE SURVEY

1. A. Swapna, S. Bikash, and P. M. Dipti, "Recognizing emotions through facial expression", The Visual Computer,

The first group methods attempt to identify feeling from basic emotions as anger, disgust, fear, joy, sorrow, and surprise from face's expression. Authors in use Support Vector Machine (SVM) classifier to extract eye movements from facial expression then compared with CK+ and MUG datasets. Results related to fear sensing is often misclassified (only 42 % of expression are

well recognized as fear feeling).

2. G. Zhao, and M. Pietikinen, "Dynamic texture recognition using local binary patterns with an application to facial expressions"

Authors propose Local Binary Patterns on Three Orthogonal Planes (LBP-TOP) method to identify fear from face expression. The true recognition rate (about 79 %) in this case is better than previous work but suffers from the increased response time.

3. M. Suk, and B. Prabhakaran, "Real-time mobile facial expression recognition system a case study", CVPR, Computer Vision Foundation

The authors propose to identify the basic emotions from facial expression according to mouth status in this case. SVM classifier is performed. The accuracy is lower than LBP-TOP method (about 71%) and the time response does not respect the real-time.

4. P. Glauner, J. A. Meira, P. Valtchev, R. State, and F. Bettinger, ''The challenge of non-technical loss detection using artificial intelligence: A survey,'' Int. J. Comput. Intell. Syst., vol. 10, no. 1, pp. 760–775, 2020.

Face recognition technology has been an important subject in the field of computer vision for a long time, and it was mainly used in the field of public security in the early stage. With the popularity of face attendance, face passing and other applications in recent years, face recognition is widely used in intelligent transportation, intelligent medical treatment, building intercom, financial education, safe city construction and other fields;

5. Q. W. Wang and Z. L. Ying, ''A face detection algorithm based on Haarlike t features,'' Pattern Recognit. Artif. Intell., vol. 28, no. 1, pp. 35–41, 2019.

Intelligent video retrieval technology has become an increasingly indispensable part of video monitoring system. The international anti-terrorism situation is becoming more and more serious, and the social demand for security is even stronger. Airport, customs, border defense and other places, which are necessary for entry and exit due to their wide distribution and large flow of people, have always been the areas with frequent emergencies, which determines the requirements of high security level control at checkpoints.

6. K. Patel, H. Han, and A. K. Jain, ''Secure face unlock: Spoof detection on smartphones,'' IEEE Trans. Inf. Forensics Security, vol. 11, no. 10, pp. 2268–2283, Oct. 2020.

Using face recognition, we can take the initiative to extract and record the faces in the surveillance video in real time, to give an alarm in time and use the stored information to effectively retrieve the video data, to effectively assist the security personnel in dealing with the crisis, to minimize the false alarm and missing alarm, and to effectively help the checkpoint deal with the attack and emergencies

7. J. Liu, J. Chen, and J. Zhang, ''Geographic information dynamic monitoring in ntelligent Era,'' Wuhan Daxue Xuebao (Xinxi Kexue Ban)/Geomatics Inf. Sci. Wuhan Univ., vol. 44, no. 1, pp. 92–96, 2019.

Face detection is to determine whether there is a face in a static or dynamic image after certain processing and analysis. If there is a face, the location and size information of each face is recorded and identified in the image. Although the research of face detection has been decades, and has made some achievements, but at present, the more mature face detection technology is basically for the detection of some ideal situations, and still faces many challenges in practical applications.

III.EXISTING SYSTEM & PROPOSED SYSTEM

3.1 Existing System

Abnormal behavior based on the tracking method results in a poor recognition rate and is not really related to a suspect action. Existing surveillance systems suffer from the following shortcomings manual/visual detection of suspicious behavior is untrustworthy, systems save only what has already happened. Systems related to specific case, non-real-time systems, and systems violate the privacy of citizens.

3.1.1 Drawbacks

- Facial recognition technology may produce inaccuracies, leading to false positives and negatives. Factors like lighting conditions and facial obstructions can affect its reliability.
- Relying on email alerts assumes a constant internet connection; poor connectivity or network issues may hinder timely alerts, impacting the system's effectiveness.
- Susceptibility to false alarms can lead to user annoyance, decreased confidence, and a disregard for alarms over time.

3.2 Proposed System

The proposed system for the intelligent face Recognition project aims to leverage Recurrent Neural Networks (RNN) for automatic comparative labeling of face in database, enhancing traditional approaches and addressing the limitations associated with them. The system will be trained to identify individuals who are outstanding or flagged in databases. Utilizing the capabilities of RNN, the system will analyze facial features and patterns to automatically and recognize wanted persons. Regular customers will be identified based on their registered facial features stored in the system's database. The RNN system will compare incoming

facial data with the registered features to recognize and categorize individuals as regular customers. New persons are those not previously registered in the system. The system is designed to capture and analyze facial data of unknown individuals. The RNN model adapts and learns from new data, continuously improving its recognition capabilities.

3.2.1 Features

- It is useful for many fields, especially in the surveillance system.
- Recurrent neural network (RNN) system will be implemented on cameras, and it will be alerted to particular officers.

IV.SYSTEM FUNCTION

4.1 Architecture Design

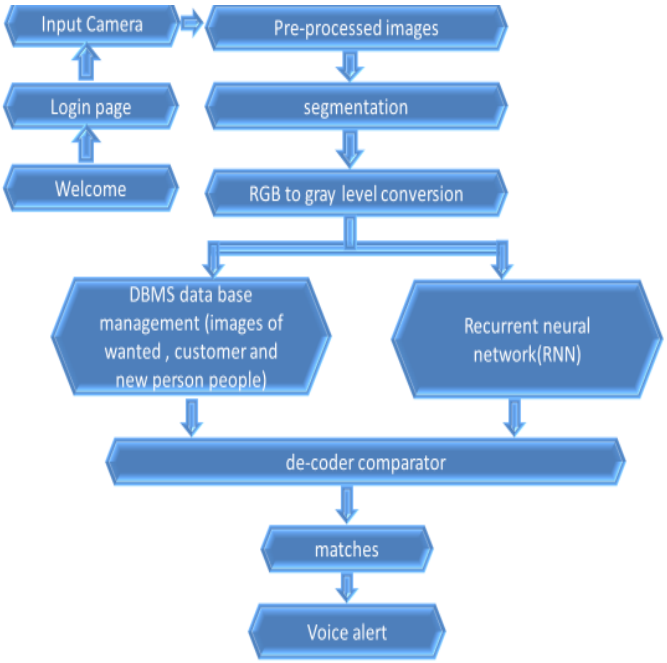

Figure 1 Proposed Block Diagram

V.METHODOLOGY

5.1 Recurrent Neural Network (RNN)

Recurrent Neural Network (RNN) are a type of Neural Network where the output from previous step is fed as input to the current step. In traditional neural networks, all the inputs and outputs are independent of each other, but in cases like when it is required to predict the next word of a sentence, the previous words are required and hence there is a need to remember the previous words. Thus RNN came into existence, which solved this issue with the help of a Hidden Layer. The main and most important feature of RNN is Hidden state, which remembers some information about a sequence. RNN have a "memory" which remembers all information about what has been calculated. It uses the same parameters for each input as it performs the same task on all the inputs or hidden layers to produce the output. This reduces the complexity of parameters, unlike other neural networks.

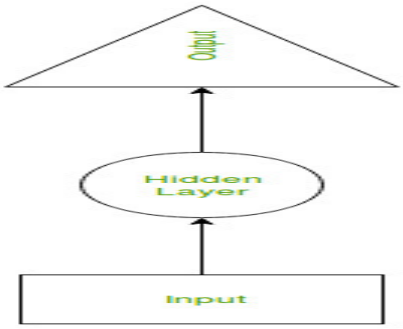

Figure 2 RNN Parameters

Now the RNN will do the following:

- RNN converts the independent activations into dependent activations by providing the same weights and biases to all the layers, thus reducing the complexity of increasing parameters and memorizing each previous outputs by giving each output as input

to the next hidden layer.

Hence these three layers can be joined together such that the weights and bias of all the hidden layers is the same, into a single recurrent layer.

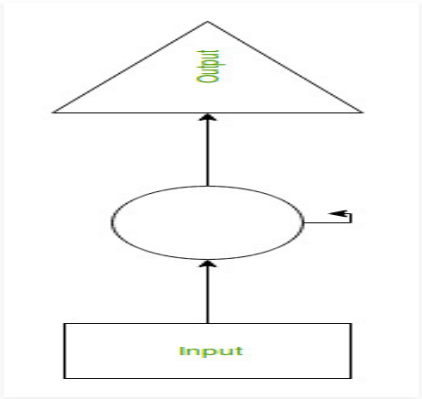

Figure 3 Hidden layers are the same, into a single recurrent layer.

Formula for Calculating Current State:

$$h_t = f(h_{t-1}, x_t) \quad (1)$$

Formula for applying Activation function (tanh):

$$h_t = \tanh (W_{hh}h_{t-1} + W_{xh}x_t) \quad (2)$$

Formula for calculating output:

$$y_t = W_{hy}h_t \quad (3)$$

5.2 Training through RNN

1. A single time step of the input is provided to the network.
2. Then calculate its current state using set of current input and the previous state.
3. The current ht becomes ht-1 for the next time step.
4. One can go as many time steps according to the problem and join the information from all the previous states.
5. Once all the time steps are completed the final current state is used to calculate the output.
6. The output is then compared to the actual output i.e the target output and the error is generated.
7. The error is then back-propagated to the network to update the weights and hence the network (RNN) is trained.

Advantages of Recurrent Neural Network

1. An RNN remembers each and every information through time. It is useful in time series prediction only because of the feature to remember previous inputs as well. This is called Long Short Term Memory.
2. Recurrent neural network is even used with convolutional layers to extend the effective pixel neighborhood.

Disadvantages of Recurrent Neural Network

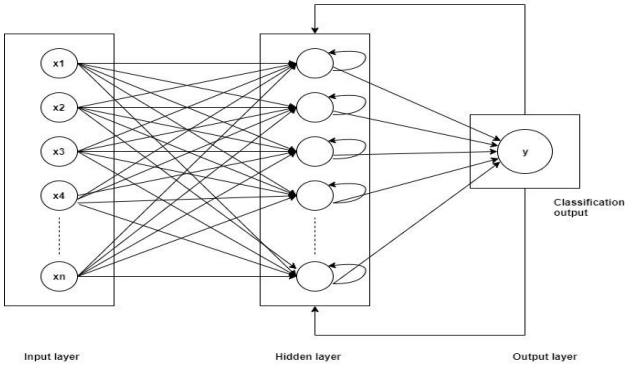

Figure 4 Recurrent Neural Network

1. Gradient vanishing and exploding problems.
2. Training an RNN is a very difficult task.

It cannot process very long sequences if using tanh or relu as an activation function.

5.3. Back propagation Through Time

To understand the concept of back propagation through time we'll need to understand the concepts of forward and back propagation first. We could spend an entire article discussing these concepts, so I will attempt to provide as simple a definition as possible. In neural networks, we basically do forward-propagation to get the output of our model and check if this output is correct or incorrect, to get the error. Back propagation is nothing but going backwards through our neural network to find the partial derivatives of the error with respect to the weights, which enables we to subtract this value from the weights. Those derivatives are then used by gradient descent, an algorithm that can iteratively minimize a given function. Then it adjusts the weights up or down, depending on which decreases the error. That is exactly how a neural network learns during the training process. So, with back propagation we basically try to tweak the weights of our model while training. The image below illustrates the concept of forward propagation and back propagation in a feed-forward neural network:

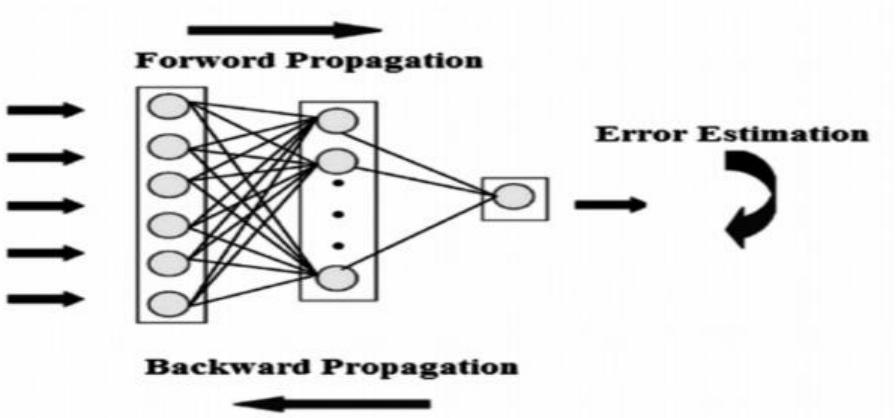

Figure 5. Neural Network

BPTT is basically just a fancy buzz word for doing back propagation on an unrolled RNN. Unrolling is visualization and conceptual tool, which helps we understand what's going on within the network. Most of the time when implementing a recurrent neural network in the common programming frameworks, back propagation is automatically taken care of, but we need to understand how it works to troubleshoot problems that may arise during the development process. We can view a RNN as a sequence of neural networks that we train one after another with backpropagation. The image below illustrates an unrolled RNN. On the left, the RNN is unrolled after the equal sign. Note there is no cycle after the equal sign since the different time steps are visualized and information is passed from one-time step to the next. This illustration also shows why a RNN can be seen as a sequence of neural networks.

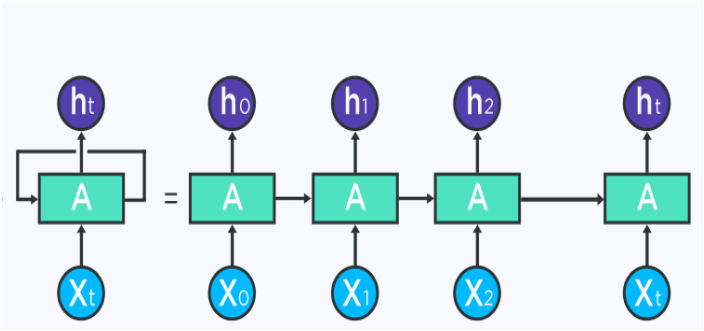

Figure 6 An unrolled version of RNN

If we do BPTT, the conceptualization of unrolling is required since the error of a given timestep depends on the previous time step. Within BPTT the error is back propagated from the last to the first time step, while unrolling all the time steps. This allows calculating the error for each time step, which allows updating the weights. Note that BPTT can be computationally expensive when we have a high number of time steps.

5.4 RGB Color Model

- The RGB color model is one of the most widely used color representation method in computer graphics. It uses a color coordinate system with three primary colors.
- R-RED; G-GREEN; B-BLUE

- Each primary color can take an intensity value ranging from 0(lowest) to 1(highest). Mixing these three primary colors at different intensity levels produces a variety of colors.
- The collection of all the colors obtained by such a linear combination of red, green and blue forms the cube shaped RGB color space.

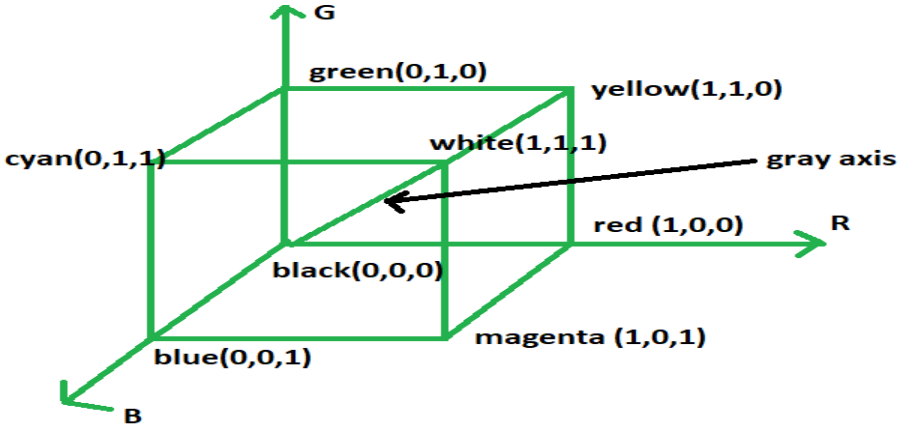

Figure 7 Cube shaped RGB color space

The RGB color model is an additive color model in which the red, green, and blue primary colors of light are added together in various ways to reproduce a broad array of colors. The name of the model comes from the initials of the three additive primary colors, red, green, and blue. The main purpose of the RGB color model is for the sensing, representation, and display of images in electronic systems, such as televisions and computers, though it has also been used in conventional photography. Before the electronic age, the RGB color model already had a solid theory behind it, based in human perception of colors.

RGB is a device-dependent color model: different devices detect or reproduce a given RGB value differently, since the color elements (such as phosphors or dyes) and their response to the individual red, green, and blue levels vary from manufacturer to manufacturer, or even in the same device over time. Thus an RGB value does not define the same color across devices without some kind of color management. Typical RGB input devices are color TV and video cameras, image scanners, and digital cameras. Typical RGB output devices are TV sets of various technologies (CRT, LCD, plasma, OLED, quantum dots, etc.), computer and mobile phone displays, video projectors, multicolor LED displays and large screens such as the Jumbo Tron. Color printers, on the other hand are not RGB devices, but subtractive color devices typically using the CMYK color model.

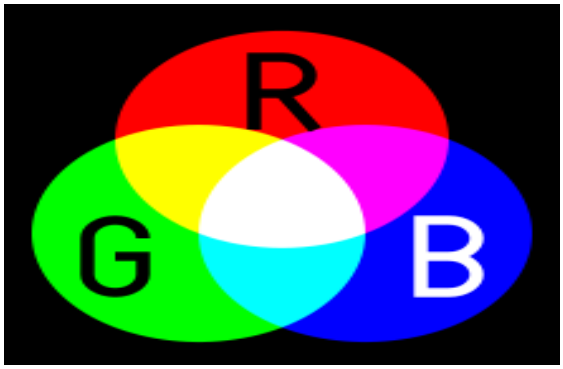

Figure 8 RGB Image

5.4.1 Gray Scale Image

Image formation using sensor and other image acquisition equipment denote the brightness or intensity I of the light of an image as two dimensional continuous function F(x, y) where (x, y) denotes the spatial coordinates when only the brightness of light is considered. Sometimes three-dimensional spatial coordinate is used. Image involving only intensity are called gray scale images.

1. Gray Levels

Gray levels represent the interval number of quantization in gray scale image processing. At present, the most commonly used storage method is 8-bit storage. There are 256 gray levels in an 8 bit gray scale image, and the intensity of each pixel can have from 0 to 255, with 0 being black and 255 being white we. Another commonly used storage method is 1-bit storage. There are two gray levels, with 0 being black and 1 being white a binary image, which, is frequently used in medical images, is being referred to as binary image. As binary images are easy to operate, other storage format images are often converted into binary images when they are used for enhancement or edge detection.

The Pseudo code

To get started, we need to import the cv2 module, which will make available the functionalities needed to read the original image and to convert it to gray scale.

```
import cv2
```

To read the original image, simply call the imread function of the cv2 module, passing as input the path to the image, as a string.

For simplicity, we are assuming the file exists and everything loads fine, so we will not be doing any error check. Nonetheless, for a robust code, you should handle these type of situations.

As additional note, which will be important for the conversion to gray scale, the imread function will have the channels stored in BGR (Blue, Green and Red) order by default.

```
image = cv2.imread('C:/Users/N/Desktop/Test.jpg')
```

Next, we need to convert the image to gray scale. To do it, we need to call the cvt Color function, which allows to convert the image from a color space to another.

As first input, this function receives the original image. As second input, it receives the color space conversion code. Since we want to convert our original image from the BGR color space to gray, we use the code COLOR_BGR2GRAY.

```
gray = cv2.cvtColor(image, cv2.COLOR_BGR2GRAY)
```

Now, to display the images, we simply need to call the imshow function of the cv2 module. This function receives as first input a string with the name to assign to the window, and as second argument the image to show.

We will display both images so we can compare the converted image with the original one.

```
cv2.imshow('Original image', image)
cv2.imshow('Gray image', gray)
```

Finally, we will call the waitKey function, which will wait for a keyboard event. This function receives as input a delay, specified in milliseconds. Nonetheless, if we pass the value 0, then it will wait indefinitely until a key event occurs.

Finally, once the user pressed a key, we call the destroy All Windows function, which will destroy the previously created windows.

```
cv2.waitKey(0)
cv2.destroyAllWindows()
```

VI.FACE RECOGNITION

The input design for the proposed intelligent recognition system involves capturing and processing various types of data to facilitate facial soft biometric analysis. The inputs encompass different sources of information necessary for effective recognition and categorization of individuals into three main groups: wanted persons, regular customers, and new persons. The key components of the input design are:

1. Facial Image Data:

- Captures high-resolution facial images from surveillance cameras or input devices.
- Includes images of individuals engaged in actions such as theft, scam, loitering, or those affected by a pandemic.

2. Database Inputs:

- Integrates data from law enforcement databases containing information about wanted persons.
- Utilizes flagged or outstanding records to train the system for accurate suspect recognition.

3. Registered Facial Features for Regular Customers:

- Gathers facial features of regular customers from a pre-established database.
- Includes unique identifiers and patterns associated with known individuals.

4. Real-time Facial Data for New Persons:

- Captures facial data for individuals not previously registered in the system.
- Enables the system to dynamically update its database with information about new persons.

5. Subset from RNN-network Datasets:

- Utilizes a subset of facial data from RNN-network datasets for training and validation.
- Incorporates diverse attributes and scenarios to enhance the system's adaptability.

6. Voice Alert Integration (pyttsx3):

- Implements the pyttsx3 library to provide voice alerts.
- Communicates detected faces or recognized categories through audible alerts for immediate awareness.

The input design ensures a comprehensive and diverse dataset for training, validation, and real-time recognition, enabling the system to effectively categorize individuals based on their facial soft biometric characteristics and actions.

VII.RESULT

This project introduces an innovative approach using recurrent neural networks (RNN) for automatic comparative labeling of facial soft biometrics, addressing the limitations of traditional recognition systems. By categorizing individuals into wanted

persons, regular customers, and new persons, it offers tailored recognition for diverse scenarios. Through experiments utilizing RNN-network datasets, the system demonstrates high efficacy in generating comparative facial labels, showcasing its potential for broader face recognition applications. Implemented on CCTV cameras, the system alerts designated officers efficiently, enhancing security measures. Its user-friendly interface, coupled with voice alerts via pyttsx3, ensures seamless integration and usability in various environments.

VIII.CONCLUSION

In the last 5years, facial recognition technology has come a long way. Today can check identity information automatically with regard to safe transactions, tracking and Security purposes and buildings access control. Such systems normally work in controlled environments and algorithms of recognition may manipulate environmental constraints to achieve high accuracy of recognition. Yet face-recognition technologies of next generation will be commonly used in smart settings where computers and machines are more like supportive helpers. In our project suspect like wanted people or disease affected people, when they are came in bank places means our project will give sudden detection of them, so we can easily identify the persons.

REFERENCES

1. *H. D. Flowe, "Do Characteristics of Faces That Convey Trustworthiness and Dominance Underlie Perceptions of Criminality?", PLOS One, Vol.7, pp.1-7, 2020.*
2. *N.D. Thomson, M. Aboutanos, K.A. Kiehl, C. Neumann, C. Galusha and K.A. Fanti, "Physiological reactivity in response to a fear-induced virtual reality experience: Associations with psychopathic traits", Psychophysiology, Vol. 56, pp. 158-164, 2019.*
3. *N. N. Oosterhof, A. and Todorov, "The functional basis of face evaluation", Proceedings of the National Academy of Sciences of the United States of America, Vol. 105, pp. 11087–11092, 2020.*
4. *C. P. Todorov, S. D. Said, and N. N. Oosterhof, "Understanding evaluation of faces on social dimensions", Trends in Cognitive Sciences, Vol. 12, pp. 455–460, 2019.*
5. *S. Subarna, S. Suman, and B. Abinash, "Human Behavior Prediction using Facial Expression Analysis" IEEE International Conference on Computing, Communication and Automation (ICCCA), pp. 399-404, 2021.*
6. *T. Chandan, H. Madasu, and V. Shantaram, "Suspicious Face Detection based on Eye and other facial features Movement Monitoring", IEEE Applied Imagery Pattern Recognition Workshop (AIPR), pp. 1-8, 2019.*
7. *G. Zhao, and M. Pietikinen, "Dynamic texture recognition using local binary patterns with an application to facial expressions", IEEE Trans. Pattern*
8. *Anal. Mach. Intell, Vol. 29, pp. 915–928, 2020.*
9. *M. Suk, and B. Prabhakaran, "Real-time mobile facial expression recognition system—a case study", CVPR, Computer Vision Foundation, pp. 132–137, 2021.*
10. *U. N. Mahesh, and R. Hanumantha, "Hybrid Approach for Facial Expression Recognition using HJDLBP and LBP Histogram in Video Sequences", Image, Graphics and Signal Processing, Vol. 2, pp. 1-9, 2019.*
11. *T. Pursche, J. Krajewski, and R. Moeller, "Video-based Heart Rate Measurement from Human Faces", IEEE International Conference on Consumer Electronics (ICCE), pp. 544-545, 2022.*
12. *M. Rapczynski, P. Werner, and A. Al-Hamadi, "Continuous Low Latency Heart Rate Estimation from Painful Faces in Real Time", 23rd International Conference on Pattern Recognition (ICPR), pp. 1165-1170, 2019.*
13. *K. Lin, D. Chen, and W. Tsai, " Face-Based Heart Rate Signal Decomposition and Evaluation Using Multiple Linear Regression", IEEE Sensors Journal, Vol. 16, pp. 1351-1360, 2020.*
14. *S. Fallet, V. Moser, F. Braun, and J. Vesin, "Imaging Photoplethysmography: What are the Best Locations on the Face to Estimate Heart Rate?", Computing in Cardiology, Vol. 43, pp. 341-344, 2021.*