International Journal of Scientific Research in Computer Science, Engineering and Information Technology

ISSN : 2456-3307

OPEN ACCESS

Available Online at : www.ijsrcseit.com

doi : https://doi.org/10.32628/CSEIT2410273

AI and RFID Based Multipurpose Authentication and Surveillance System

S. I. Ranjitha[1], E. Deepankumar[2]

[1]Department of Computer Science and Engineering, Excel Engineering College, Komarapalayam Namakkal, Tamilnadu, India

[2]M.E., Department of Computer Science and Engineering, Excel Engineering College, Komarapalayam Namakkal, Tamilnadu, India

ARTICLEINFO

Article History:

Accepted: 10 April 2024

Published: 19 April 2024

Publication Issue

Volume 10, Issue 2

March-April-2024

Page Number

589-601

ABSTRACT

One of the main goals for face recognition in surveillance settings is recognising a person recorded on camera or in an image. This means that faces in still photos and video clips must match. High-quality still picture automatic face identification can perform satisfactorily, while video-based face recognition is difficult to do at a comparable level. Video sequences have a number of drawbacks over still image face recognition. First off, the majority of the time, CCTV cameras provides low-quality photos. There is more background noise, and moving objects or out-of-focus subjects might cause photos to become blurry. Second, video sequences often have lesser picture resolution. The resolution of the real facial image may be as low as 64 by 64 pixels if the subject is extremely far away from the camera. Lastly, in video sequences, differences in facial image, including lighting, emotion, position, occlusion, and motion, are more significant. By creating many "bridges" to link the still image and video frames, the method may effectively handle the uneven distributions between still photos and films. In order to match photos with videos and identify unknown matches, the Grassmann algorithm may be used in this research to develop a still-to-video matching strategy with RFID technologies. Matching feature vectors based on deep learning techniques and reading the feature vectors using the Grassmann method. Additionally, when an unknown face is detected, send out an SMS and email notice. After that, provide reports for the attendance system.

Keywords : Face Recognition, Deep Learning, Attendance System, Features Extraction, Notification

I. INTRODUCTION

Face recognition is a technique that uses a person's face to determine or confirm their identification. Face recognition frameworks can be used to recognise faces in images, videos, or real-time. Cell phones are another tool that law enforcement may use to identify

individuals during a police stop. However, face popularity records are susceptible to errors that might lead to the accusation of innocent persons in crimes. Facial recognition software is particularly bad at detecting African Americans and other ethnic minorities, girls, and young individuals. It often misidentifies or ignores them, which has a disproportionately negative influence on positive groups. Strictly speaking, multi-view face recognition refers to situations when many cameras record the problem (or scene) at the same time, and an algorithm employs all of the collected images/movies. On the other hand, the era has long been utilised to identify faces in different positions. This ambiguity has no effect on the enjoyment of the photos, though; in terms of pose variations, a collection of images showing individuals enjoying themselves with a single digital camera and others taking serious pictures with many cameras at different angles are comparable. When it comes to video records, however, the two circumstances differ. Even if a system with several digital cameras may capture multi-view records instantly, there is a significant chance that the same information might be captured with just one camera. These differences become important in non-cooperative reputation packages, coupled with monitoring. For the sake of clarity, we will refer to the monocular video collection taken when the difficulty modifications pose as a single-view video and the multi-view video collection collected with synchronised cameras as a multi-view video. As camera networks have become more widespread, multi-view surveillance videos have become more typical. However, the majority of multi-view video face recognition algorithms available today produce the majority of single-view films. When given two face photos to confirm, they show up within the set to "align" the look of the face portion in one photo to the same posture and lighting of the other. This method may even call for the positions and lighting conditions to be imagined for each face picture. The concept of "ordinary reference set" has also been used to the holistic matching algorithm, which documents the idea of matching degree through the grading of look-up outcomes. The standard configuration for face detection is seen in Figure 1.

II. RELATED WORK

M. Ayazoglu, B. Li, et.al,…[1] The suggested approach consists of the constraint that the target's 2D trajectories in the picture planes of each camera must develop in the same subspace under mild circumstances. This statement allows for the identification of a single (piecewise) linear version that describes all of the available 2D data at any instant in time. Conversely, this model might be applied to a modified particle filter out to forecast target locations for destiny. When a number of cameras are unable to capture the target, it is still possible to forecast the missing measurements by using the recordings that each camera has to fall into the subspace covered by earlier measurements and satisfy epipolar restrictions. Therefore, the suggested method can handle significant occlusion with robustness by taking use of both dynamic and geometric limitations. It does this without requiring acting on three-dimensional reconstruction, calibrating cameras, or imposing restrictions on sensor separation. Many real-world examples including goals that radically alter appearance and movement patterns while obscured to part of the cameras demonstrate the effectiveness of the suggested tracker.

D. Baltieri, et.al,…[2] executed 2D models as well as recently introduced non-articulated 3D representations of the human body. In contrast to movement capture or movement evaluation techniques, the models just need to be quick and loosely defined. On the other hand, model-based localization provides more accurate comparison of matching frame parts and more cogent and concise descriptions. It is also possible to reduce issues brought on by segmentation errors and occlusions. These

paintings achieved a breakthrough by showcasing a brand-new, innovative 3-D technique for re-identifying that is based on articulated frame styles. To map look descriptors to skeletal bones, a three-dimensional version is used. Utilised as input are the colour, depth, and skeleton streams generated by the Microsoft Kinect sensor and the OpenNi libraries. The skeleton is just as delicate when using a mastery technique that lets you create a collection of "bones." Because the obtained signature is precisely associated with the physical structure of the body, it also allows for the inclusion of a characteristic-based description, which is helpful in some packages. Furthermore, a calculated metric is used that serves as both a frame portion weighting and a characteristic choice at the same time.

D. Baltieri,et.al,…[3] developed the SARC3D system, a complete device for human re-identification based only on the mapping of appearance descriptors to three-dimensional frame fashions. Utilising 3-D frame styles for re-identification is relatively recent compared to other computer vision domains, such as motion capture and posture assessment. In order to achieve proper model-to-picture alignment, challenging scenarios involving 3-D models rely on the need to accurately identify, segment, and estimate the 3-D orientation by humans. The computation of global functions is smooth and doesn't require any alignment stages. Furthermore, they exhibit increased insensitivity to faults in segmentation or detection. But the lack of close skills results in misleading comparisons. Part-based matching is made possible by the partial issue mitigation provided by the use of 2D models. The precise comparison of upper and lower body parts increases the signature's specificity. However, issues specifically related to a certain orientation may also surface.

I. B. Barbosa, et.al,…[4] built a model of human appearance based on clothing, with the obvious problem that changing garments between camera acquisitions severely impairs reputation performance. There has been a significant advancement in this situation because rgb-d records provide significantly more records. However, current rgbd sensors are unable to operate at the same distance as standard surveillance cameras; as a result, rgb specialised work remains vital. This research presents a convolutional neural network-based re-identification method designed to address the aforementioned issues. The framework made a number of amazing features renowned. Firstly, compared to a siamese configuration, the network's topology is simpler. It serves as a function extractor and is based on the inception structure. Our preference is sold by quantitative experiments, which provide high scores in reputation terms.

Bedagkar-Gala, et.al,…[5] investigated several person re-identities across cameras in the setting of large-scale surveillance, treating matching as a project issue only based on a spatiotemporal appearance version. Although biometrics such as face or gait can be utilised for matching, they are typically difficult to recover due to camera resolution or body charge limitations. The most common applications for appearance models are in re-identification and the analysis of spatiotemporal interactions between cameras in the context of phoney matches. For assessment purposes, individual re-identification in the context of monitoring across multiple cameras is an open set matching problem. This means that the probe set varies dynamically for every digital FOV, the gallery changes over time, and not all of the probes inside a hard and fast are necessarily a subset of the gallery.

Y. Yan, et.al,…[6] suggested a unique lively pattern selection method (energetic learning) for each form of image employing net pix. Go-media modelling of different media kinds has been shown to be helpful for multimedia content evaluation in earlier studies. The online photos are frequently linked to elaborate written explanations (such as captions, surrounding

words, etc.). Show that textual content characteristics are helpful for learning robust classifiers, allowing for improved active learning performance of picture category, even while such text facts are not provided in testing images. Conventional active sampling techniques only handle a single media type and are unable to use many unique media types at once. This issue may be resolved by applying the new supervised mastering paradigm, more especially, mastering the usage of privileged records (LUPI). Within the training system in a LUPI scenario, privileged information is accessible in addition to critical functions. Tests cannot access privileged facts; they can only be used in training. In active learning, the most often employed approach is uncertainty sampling. In this study, we propose for treating textual material as privileged data and using both visual and text aspects for active pattern choosing. Teach SVMs about visible features and weak features related to text capabilities using LUPI. Give five methods to combine these two classifiers' uncertainty measures. Use the range measurement to verify that the selected samples are less similar to one another and more similar to each other as consultants. Create a ratio objective function that minimises the similarity of the chosen information and maximises cross-media uncertainty. Next, suggest gauging diversity and uncertainty in the selection of educational patterns. The suggested model is solved using a novel optimisation technique that automatically determines the best uncertainty to similarity ratio. By doing this, the exchange-off parameter between the two measurement techniques may be avoided.

Y. Yang, et.Al,…[7] implemented a new set of feature choice rules that improves characteristic choice performance by using the information from related pair of responsibilities. The observations have shown the following lessons: For supervised study, sharing records among related tasks is helpful. That being said, total performance isn't always improved if a few duties aren't connected. The benefits of multitask learning are typically more apparent as compared to single assignment mastery, with only a few training instances each mission. Adapting inter-task knowledge no longer always helps since intra-project information is sufficient for education when the large range of fine schooling data increases. Function selection does not always result in better overall performance. Still, it's quite helpful because it enhances performance. Furthermore, selecting a trait would improve our capacity to understand the skills. When using specialised classifiers, the feature selection process is improved in many ways. For example, following feature selection, SVM's performance gain is less than KNN's since linear SVM genuinely has the capacity to give unique weights to distinct features. The following is the prepared relaxation for this article. give the objective characteristic. The optimisation strategy is suggested together with evidence of its convergence. then present the findings of the experiment and complete the report. When certain classifiers are employed, characteristic selection's overall performance increase differs. Examine how different characteristic selection methods perform with different unique classifiers using the KTH dataset as an example. Use KNN as a substitute classifier in unique and the correctness of the six motion kinds' frequent appeal. By comparing Table and Figure, one can observe that KNN's accuracy is lower than SVM's when using all characteristics for motion recognition. However, KNN outperforms SVM in terms of accuracy following characteristic selection. A plausible explanation is that SVM can weigh exceptional characteristics, which means that feature selection yields less benefit.

X. Chang, et.Al,…[8] sought to recommend a compound rank-k projection algorithm for discriminant bilinear analysis and address the shortcomings of the current discriminant evaluation techniques for excessive-order facts. In contrast, our optimisation method's convergence is expressly guaranteed. try a few different orthogonal projection

styles to obtain additional discriminant projection guidance. To be more exact, find a low dimensional representation of the unique information using the h units of projection matrices. Each of the h projection matrices is orthogonal to the others. This allows for the conversion of the original statistics into an amazing orthogonal foundation and the receipt of statistics from several angles. Our method's main innovation is that it uses two projection models, which can be integrated and painted jointly. In this way, a larger search space is provided in order to identify the most satisfying answer and improve student performance. Give Compound Rank-ok Projection for Bilinear Analysis (CRP) as the name of the suggested set of guidelines. It is important to note that excessive-order tensor discriminate analysis may easily be performed using the same set of principles. The key accomplishment of our work is that CRP can deal with matrix representations directly, without having to transform them into vectors. The original statistics' spatial relationships can therefore be maintained. The processing complexity is reduced as compared to traditional approaches. When evaluating discriminate in two dimensions, CRP benefits from the trade-off between avoiding overbearing hassles and maintaining degree of freedom. Due to the among-class scatter matrix's singularity, the standard 2DLDA's iterative optimisation approach fails to converge even though it achieves optimal overall performance. In contrast, our algorithm's convergence is guaranteed explicitly. This paper's relaxation is structured as a summary of both 2DLDA and a traditional LDA review. A new compound rank-k projection is suggested for bilinear assessment. Demonstrate our experimental results on five outstanding datasets. We talk about the conclusion of our work.

J. Luo, et.Al,…[9] suggested a framework for the analysis and retrieval of multimedia material that includes two algorithms. A novel transductive rating method is first suggested, called ranking with Local Regression and Global Alignment (LRGA). In contrast to distance-based total ranking methods, LRGA makes use of the distribution of samples throughout the entire set of data. In contrast to inductive approaches, simplest requires an instance of the query. As an evaluation of the MR method, which computes the Laplacian matrix by directly adopting the Gaussian kernel, LRGA learns a Laplacian matrix for fact ranking. To anticipate the ranking scores of its adjoining points, rent a nearby linear regression version for each data point. Propose a single objective characteristic to globally align local linear regression models from all the information variables so that a surest ranking score may be assigned to each fact point. There is no ground reality to adjust the settings of rating algorithms such as MR in retrieval programmes. Therefore, expanding a novel method that acquires a gold standard has much greater significance. Statistics rating using Laplacian matrix. Second, suggest a long-term reinforcement learning approach that is semi-supervised. To save the historical RF statistics noted by each client, a device log is created. Using a statistical method, improve the vector representation of multimedia data so that it is consistent with the log facts. Convert the RF data into paired restrictions that are divided into two businesses in order to finish there. While the statistics pairs in the second institution are differed from one another, the information pairs in the first group are comparable to each other conceptually. Although these two types of data may be exploited using LDA, the useful information contained in the unlabelled data is not used. In order to improve the vector illustration, take into account the records RF facts and the multimedia records distribution of both labelled and unlabelled samples. This study proposes a semi-supervised learning set of rules. Examine how well our algorithms perform in content-based full move-media retrieval in this study, where the query example and the results of the retrieval may come from different media types. For instance, the user can use an instance picture or an instance audio record to search photos. likewise adhere to the suggested techniques for content-based whole picture retrieval and 3-dimensional motion/pose fact

retrieval. Numerous tests show that, in comparison to the current related research, our algorithms achieve superior retrieval performance.

W. Li, et.Al,…[10] presented a method for individual re-identification using a filter pairing neural network (FPNN). Comparing this deep learning technique to existing studies, it offers several noteworthy surprises and strengths. It works together to manage misalignment, geometric and photometric transformations, occlusions, and historical background clutter under a single deep neural network. All of the essential elements are simultaneously optimised during schooling. Every part works in tandem with the others to optimise its electricity. It automatically learns high-quality features for the task of person re-identification from records, together with the analysis of photometric and geometric transformations, in place of using handmade skills. To extract characteristics, two paired filters are used to distinguish across camera views. Photometric transformations are encoded by the filter pairs. Existing works consider pass-view transformations to be uni-modal; nevertheless, a mixture of complex transforms may be modelled thanks to the deep structure and its maxout grouping layer. Second, employ carefully thought-out training techniques, such as bootstrapping, data augmentation, dropout, and fact balancing, to acquaint the suggested neural community. These methods tackle the problems of overfitting, severe imbalance of positive and negative schooling samples, and incorrect patch correspondence identification in this assignment. Thirdly, re-evaluate the individual re-identification issue and create a large-scale dataset to compare the effects of automated pedestrian detection. The short duration of all the existing datasets makes it difficult for them to train a deep neural network. See evaluation for information on the 13,164 photos of 1,360 pedestrians in our collection. Current datasets rely on perfect detection in assessment processes and only provide manually cropped photos of pedestrians. Given that automated detection in exercises produces significant misalignment, the effectiveness of current strategies may be significantly impacted. For a thorough evaluation, our dataset offers both manually cropped images and automatically discovered bounding boxes using a state-of-the-art detector.

III. FACE DETECTION ALGORITHMS

Utilising computer algorithms, face recognition software recognises certain, distinctive traits on an individual's face. After being converted into a mathematical representation, these characteristics—like chin angle or eye distance—are then matched to information from other faces in a face recognition database. A face template is information about a particular face that is different from an image in that it is intended to contain just specified characteristics that may be utilised to distinguish one face from another.
Appearance-based methods

Depending on how many photos are utilised to build the human representation, appearance-based techniques for human re-identification may be categorised into two groups: single-shot methods and multiple-shot approaches. While multiple-shot approaches employ a collection of photos to calculate the feature vectors to represent the human, single-shot methods just need one image to do so. While single-shot approaches yield less information than multiple-shot methods, the latter are less compute-intensive and need more complicated algorithms. In situations where tracking data is lacking, single-shot techniques can be helpful. The two groups that may be formed from each category are learning-based approach and direct-based method. The goal of the direct-based method is to create a stable and reliable feature representation of the person under different circumstances. The similarity between the photos is ascertained by basic matching once the features have been extracted. The learning-based method looks for matching algorithms that decrease the gap between similar classes and maximise the distance between different classes using training

data. Deep learning is currently being used in human re-identification tasks in addition to these two categories since it combines feature extraction and classification tasks into a single integrated framework.

Single-shot Methods

Researchers like to use a variety of aspects, such as colour, texture, and form, into their works in order to address typical difficult problems. An appearance descriptor was created by feature extraction using colour and edge characteristics, and matching was done using a correlation-based similarity measure. A parts-based approach was put forth, in which the human picture was split into three horizontal sections, with each part's signature being extracted using colour histograms in the HSV colour space. Wang and colleagues introduced the shape and appearance context model. They divided the human vision into areas and recorded the spatial connection between colours in a co-occurrence matrix. However, the effectiveness of their suggested approach was limited to situations in which the human photos were taken from comparable angles by various cameras. And put out a model of appearance that was built via kernel estimation. In their work, they employed path-length and Colour Rank Feature.

Multi-shot Methods

In order to address the posture variation issue, it may split the human picture into 10 horizontal stripes. From these 10 stripes, they derived the color's median hue, saturation, and lightness (HSL). To minimise the dimensionality of the features, Linear Discriminant Analysis (LDA) was used. For low-resolution videos, like those seen in the majority of surveillance systems, their suggested approach is unreliable and suggested a unique method of creating spatiotemporal graphs by segmenting ten successive human frames. Colour and edge histograms were used to portray human appearance. The suggested method is only useful in situations where people's appearances across various cameras remain consistent. Their technique is not feasible for practical use in the real world due to this restriction. Additionally, their suggested approach is unreliable for position variations and occlusion.

Distance metric learning-based methods

Many metric learning and matching techniques were put out as convenient stand-ins for a potent procedure that allows for nearby man-or woman re-identification. These tactics were created with the intention of examining the quality measure between the equal person's look functions across all camera pairings. These metric analysis methods are classified as either global or local, or supervised vs unsupervised. The majority of Person Reid's efforts are classified as supervised international distance metric learning. Unsupervised techniques, in particular, focus on function design and feature extraction and do not necessitate the manual labelling of training samples; supervised tactics, on the other hand, often require the assistance of manually tagged schooling examples, leading to greater overall performance. Additionally, global metric learning strategies focus on learning the vectors of the same elegance to be closer together while pushing the vectors of different lessons farther apart. To improve the traditional kNN type made, the Large-Margin Nearest Neighbour metric (LMNN), which falls under the supervised local distance metric studying category, has been proposed. But using the k closest in-class samples for LMNN has grown to be a time-consuming procedure. A moderate improvement over the previously obtained results with the addition of a reject alternative for unexpected fits, known as LMNN-R, has been introduced as an LMNN version. Information Theoretic Metric Learning (ITML) and Logistic Discriminant Metric Learning (LDML) have also been used as process optimisation techniques to avoid the overfitting problems seen in LMNN. The Keep It Simple and Straightforward measure (KISSME) approach was designed to be a straightforward and uncomplicated way to learn a distance measure without equivalency limitations. Unlike many other approaches like LMNN, ITML, LDML, and KISSME,

this one does not rely on computationally expensive iterations and intricate optimisation. However, when applied to real-world situations, the performance of the device learning version is likely to deteriorate with time as freshly received records are likely to deviate from the initial educational data. Conventional approaches, which need a lot of time, are retrained in the batch mode to utilise both new and existing data concurrently. In order to address this kind of flaw, a number of well-known incremental learning algorithms, including Self-Organizing Map and Growing Neural Gas, were created using the theory that a neural network represents the topological shape of unlabeled records and serves as an average from which data can be grouped according to predetermined guidelines. Self-Organizing Incremental Neural Network (SOINN), which greatly outperforms the previously mentioned algorithms by allowing to learn the important wide variety of neural nodes and efficiently representing the topological shape of enter possibility density, was proposed as a more appropriate method for processing online information. This method's usage of Euclidean distance to measure the distance between the input data and nodes is a notable drawback. But considering the frequency of intra-elegance and inter-class variations, the Mahalanobis measure seems like a better option for addressing character Re-ID-related issues.

RGBD-Based Approaches

RGBD-based strategies like the RGB look-based man or woman for Person Re-ID Re-ID presents a novel idea that is entirely focused on depth and assumes that everyone wear the same clothes and utilise the most basic 2D information; Because intensity data is color-neutral and retains more invariant information for longer than RGB statistics, it can remain more invariant even in the face of harsh and inconsistent lighting. In actuality, it is not difficult to extract depth and skeletal data from an interior environment using depth cameras (such as Microsoft Kinect). Regardless of the colour or lighting of the item in indoor packaging, Kinect sensors use infrared to determine the intensity value (distance to the digicam) of each pixel. By using depth recordings, one may extract a person's skeleton and existence size point cloud, which provide shape and physiological details about their body. Furthermore, individuals may be more easily distinguished from their legacy thanks to the depth price of each pixel, meaning that the influence of history may be substantially reduced. There have been very few planned efforts on intensity-based person Re-ID. used a feature of the skeleton that was mostly based on the anthropometric size of joint distances and the geodesic distances on the body floor. created a factor cloud version for us all to see as a gallery by combining a quick and easy factor cloud from distinct perspectives, and then used Point Cloud Matching (PCM) to calculate the difference between samples. Eigen-depth feature and depth voxel covariance descriptor are proposed to explain body shape. The eigen-depth function is a fully featured covariance-based function that doesn't require point cloud alignment and is rotationally invariant in a given area. As an intensity shape descriptor, the Eigen-depth function eliminates the uncertainty that resulted from using anthropometric measurements of skeletons alone in the previous intensity modelling for re-identification. Some Re-ID techniques have evolved to incorporate RGB look cues with depth statistics since Kinect allows for the simultaneous acquisition of both RGB and depth facts, allowing for the extraction of additional discriminative feature illustrations. These tactics may be separated into two groups. Appearance-based methods, which combine look and intensity statistics, are the main type of strategy suggested a Re-ID method that was mostly based on bone data. Character signatures are obtained by concatenating feature descriptors that are derived from the skeleton joints of the individual. The geometric skills form the basis of the 2D sort of technique: Re-ID is achieved by using the given approach to match frame shapes in terms of full point clouds distorted to a fashionable position. For Re-ID, they use the anthropometric degree approach.

They compute the geometric functions, which include limb lengths and ratios, using the skeleton tracker-provided 3-dimensional position of body joints.

IV. HYBRID AUTHENTICATION SYSTEM FACE AND RFID ACCESS

Now a day's most educational institutions are concerned about student irregular attendance. Face based recognition of the people is very helpful to ascertain their identity. Many papers have been proposed related to RFID and face-based attendance system. Biometric-based techniques shows to be more auspicious techniques for identifying an individual, instead of authenticating people and providing them access to physical and virtual domains based on passwords, PINs, smart cards, tokens, keys and so forth. These methods inspect an individual's behavioral and physical characteristics in order to determine his or her unique identity. But the passwords and PINs are difficult to recall and can be predicted easily or stolen. So here person identification plays an important role. Amongst the person recognition methods, face recognition is known to be the most familiar ones, as the face modality is a method that uses to identify people in everyday lives. Although other approaches, such as fingerprint identification, can offer improved performance, still those are not suitable for natural smart interactions due to their extended nature. We are integrating the face recognition techniques and hence proposing a prototype which will not only be helpful for attendance recording and tracking but also it will enhance security. When a student tries to gain entry into the campus, he is asked to swipe his ID card. Web camera captures the live video & the frames from the live video are processed for the face detection first. As face detection is the first step towards the specific employee face recognition. Once face is detected, then the face of the employee captured is matched with the face of the concerned student face photo already present in the database which has been tagged for the particular ID number. Once all the parameters match then face is taken and it is also checked for authentication. If found matched then the Attendance is marked in the file with the log in details.

A common mathematical method for analysing and comparing facial characteristics in face recognition is the Grassmann algorithm. The following are the fundamental procedures for using the Grassmann algorithm to facial recognition:

Feature extraction: First, facial traits must be extracted from the photos of the faces that need to be recognised.

Face representation: After that, a high-dimensional space is used to represent the retrieved face characteristics as points. This area is frequently called a face space or a feature space.

Grassmann manifold: In a high-dimensional space, the Grassmann manifold is a mathematical representation of every feasible subspace of a given dimension. This manifold is used by the Grassmann method to compare the subspaces that represent the various faces' facial traits.

Subspace projection: In the feature space, every face is represented as a subspace. These subspaces are then projected onto the Grassmann manifold via the Grassmann method, yielding a collection of comparable points.

Distance computation: Lastly, a distance metric like the Grassmann distance is used to calculate the distance between the subspaces. The distance metric calculates the degree of similarity or dissimilarity between the subspaces while accounting for the geometry of the Grassmann manifold.

Classification: The faces are then categorised as belonging to known or unknown people using the calculated distances. In this stage, a distance metric threshold value is usually defined, above which a face is deemed unknown.

All things considered, the Grassmann algorithm is an effective face recognition method that can adapt to changes in illumination, posture, and facial emotions. Because it can accurately represent the diversity of face characteristics in a low-dimensional space, it is

especially helpful when there are few training examples. The suggested architecture diagram is shown in Fig. 1.

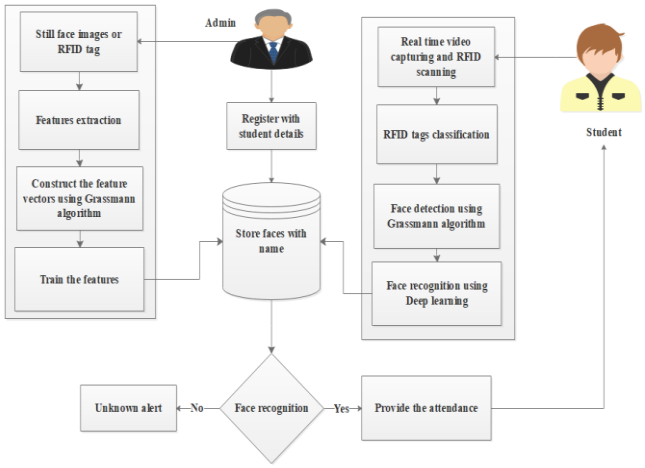

Fig 2: System Architecture

The proposed modules are shown below

FRAMEWORK CONSTRUCTION

Over the past 20 years, there have been several advancements in personal identification. Any technique that uses a person's bodily characteristics to identify them—a feature that can be seen and assessed depending on behavior—is referred to as biometrics. When used on huge data sets, traditional biometric techniques like fingerprints and handwritten signatures have shown to be unreliable. The admin and student interfaces are built in this module. All student information may be stored by the administrator for further verification.

FACIAL FEATURES EXTRACTION

In addition to being recorded in real time for registration purposes, student faces also allow the camera to remove background pixels from the face. Use preprocessing techniques in this module to determine which pixels are in the foreground. Take the foreground pixels out of the entire picture. To extract facial traits like skin tone, eyes, and other aspects, use the Grassmann algorithm. Feature vectors are used to build these features.

REGISTER THE FACES

Feature vectors are made for each student in this lesson. A vector that includes several details about an object is called a feature vector. Feature space may be created by combining object feature vectors. The characteristics might collectively represent a full picture or only one single pixel. The level of detail relies on what is being attempted to be understood or represented regarding the student object.

CLASSIFICATION

One of the most creative approaches to advancing facial recognition technology is deep learning. Extraction of face embeddings from face-containing photos is the notion. For various faces, these facial embeddings will be distinct. And the best method for completing this goal is to train a deep neural network. Use a deep learning method in this module to categorise the feature vectors. Neural networks are used to categorise several features at once. Using convolutional neural networks, we can extract a variety of information from pictures.

ATTENDANCE SYSTEM

This module allows to identify a known individual with RFID number. Provide features with their attendance based on a recognised individual. Additionally, provide notice of any unfamiliar faces. Lastly, a summary of the attendance details

V. EXPERIMENTAL RESULTS

In this simulation, provide real time face datasets in student attendance system using Python framework as front end and MySQL as back end.

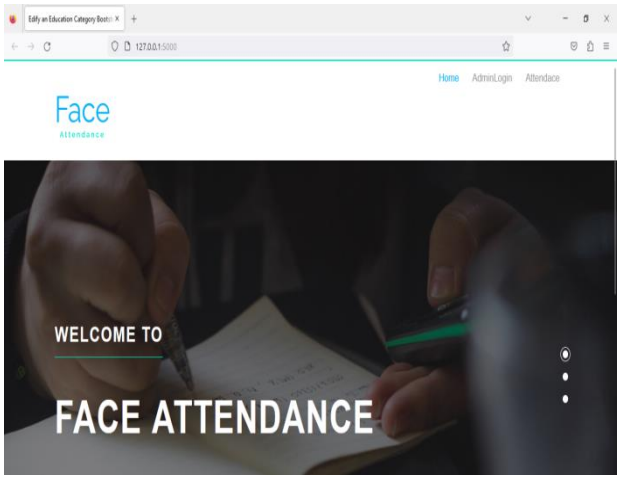

Fig 3: Home page

In this screen show the home page for student attendance system

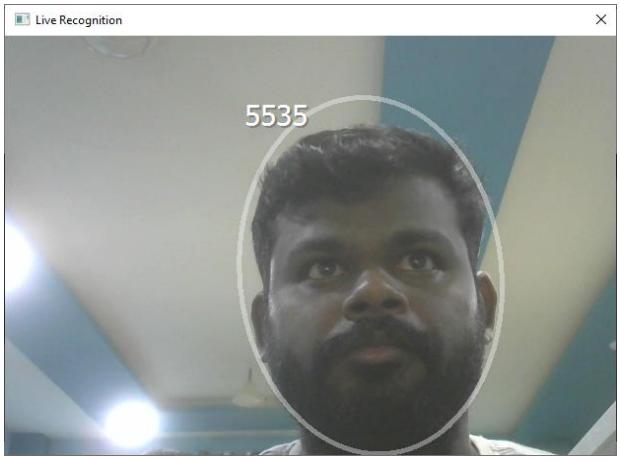

Fig 4: Attendance for real time face

Real time face datasets are used to train the person details in terms of feature vectors

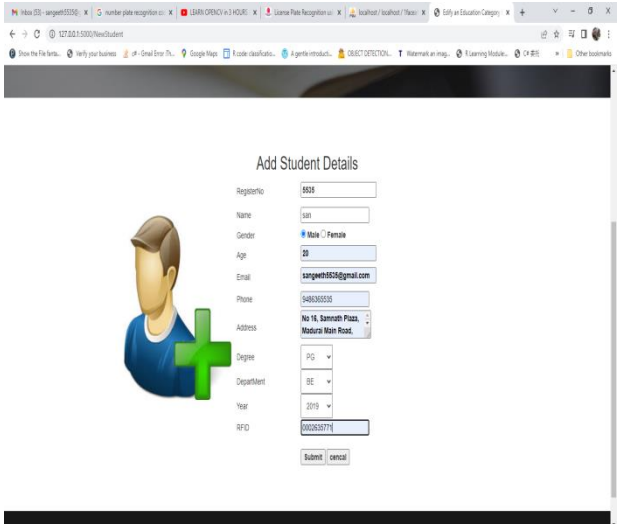

Fig 5 Add the RFID card

In this screen shows the RFID card details for each and every student

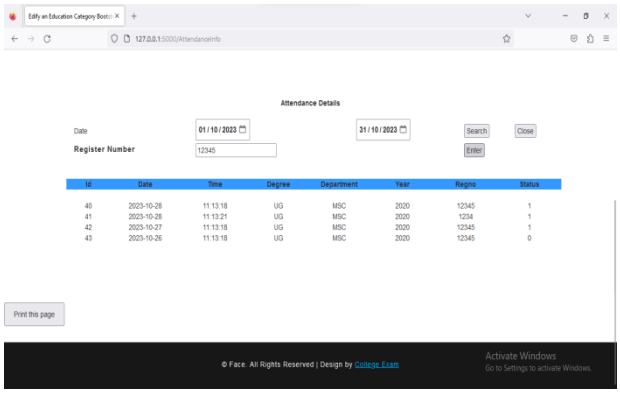

Fig 6: Attendance report

This shows the overall attendance report for each and every student

VI. CONCLUSION

The proposed system has been designed based on the principle of multi-layer security. We have demonstrated successfully RFID can be used integrated with face recognition techniques for attendance &also it's a low cost easily affordable. Face recognition is a well-developed technology used for person identification. Radio-frequency identification (RFID) tags are used for access control system which contains electronically stored information. These kinds of devices can be used in secure and sensitive places where security is of utmost importance. As the proposed attendance system is based on multilayer security, the system can not only be used for Attendance but also pre security checks we are doing to establish whether the student is authentic or not. With the proposed system secure establishments can enhance their internal security and hence can prevent possible intruder entry into the campus. Security officials can take up immediate actions whenever such a possible intrusion into campus occurs with the help of the proposed system. Thus, the proposed system not only works well as an attendance system but also as a security device.

VII. REFERENCES

[1]. M. Ayazoglu, B. Li, C. Dicle, M. Sznaier, and O. Camps. Dynamic subspace-based coordinated multicamera tracking. In 2011 IEEE International Conference on Computer Vision(ICCV), pages 2462–2469, Nov. 2011.

[2]. D. Baltieri, R. Vezzani, and R. Cucchiara. Learning articulated body models for people re-identification. In Proceedingsof the 21st ACM International Conference on Multimedia, MM '13, pages 557–560, New York, NY, USA, 2013. ACM.

[3]. D. Baltieri, R. Vezzani, and R. Cucchiara. Mapping appearance descriptors on 3d body models for people reidentification. International Journal of Computer Vision, 111(3):345–364, 2015.

[4]. I. B. Barbosa, M. Cristani, B. Caputo, A. Rognhaugen, and T. Theoharis. Looking beyond appearances: Synthetic training data for deep cnns in re-identification. arXiv preprintarXiv:1701.03153, 2017.

[5]. A. Bedagkar-Gala and S. Shah. Multiple person reidentification using part based spatio-temporal color appearance model. In Computer Vision Workshops (ICCVWorkshops), 2011 IEEE International Conference on, pages 1721–1728, Nov 2011.

[6]. 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

[7]. 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.

[8]. 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.

[9]. 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.

[10]. W. Li, R. Zhao, T. Xiao, and X. Wang, "Deepreid: Deep filter pairing neural network for person re-identification," in Proc. CVPR, 2014, pp. 152–159.

[11]. A. Bedagkar-Gala and S. K. Shah. Part-based spatiotemporal model for multi-person re-identification. PatternRecognition Letters, 33(14):1908 – 1915, 2012. Novel Pattern Recognition-Based Methods for Re-identification in Biometric Context.

[12]. J. Berclaz, F. Fleuret, E. T¨uretken, and P. Fua. Multiple object tracking using k-shortest paths optimization. IEEETransactions on Pattern Analysis and Machine Intelligence, 2011.

[13]. K. Bernardin and R. Stiefelhagen. Evaluating multiple object tracking performance: the CLEAR MOT metrics. EURASIPJournal on Image and Video Processing, (246309):1–10, 2008.

[14]. L. Beyer, S. Breuers, V. Kurin, and B. Leibe. Towards a principled integration of multi-camera re-identification and tracking through optimal bayes filters. CVPRWS, 2017.

[15]. M. Bredereck, X. Jiang, M. Korner, and J. Denzler. Data association for multi-object Tracking-by-Detection in multicamera networks. In 2012 Sixth International Conference onDistributed Smart Cameras (ICDSC), pages 1–6, Oct. 2012.

[16]. A. A. Butt and R. T. Collins. Multiple target tracking using frame triplets. In Computer Vision–ACCV 2012, pages 163– 176. Springer, 2013.

[17]. Y. Cai and G. Medioni. Exploring context information for inter-camera multiple target tracking. In 2014 IEEE WinterConference on Applications of Computer Vision (WACV), pages 761–768, Mar. 2014.

[18]. S. Calderara, R. Cucchiara, and A. Prati. Bayesiancompetitive consistent labeling for people surveillance. PatternAnalysis and Machine Intelligence, IEEE Transactionson, 30(2):354–360, Feb 2008.

[19]. L. Cao, W. Chen, X. Chen, S. Zheng, and K. Huang. An equalised global graphical model-based approach for multicamera object tracking. ArXiv:11502.03532 [cs], Feb. 2015.

[20]. Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh. Realtime multiperson 2d pose estimation using part affinity fields. In CVPR, 2017.