IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Face Recognition Based Smart Attendance System

Vaggi Ramya[1], VVGK JAHNAVI[2], S ASHISH[3], S KARTHIK[4]

[1]Asst Professor, Dept of Computer Science and Engineering, GMR Institute of Technology, Rajam, Andhra Pradesh-532127

[2,3,4,5,6]Student, Dept of Computer Science and Engineering, GMR Institute of Technology, Rajam, Andhra Pradesh-532127

Abstract:

The majority of the time allotted to the professors for teaching is spent recording the attendance of the pupils who are present in the class. This is viewed as a problem because it deprives teachers of time they could use for teaching and interacting with students, as well as because it causes greater disruption and a loss of decorum in the classroom. The current manual attendance management system is further complicated by the existence of proxy attendance. These issues are addressed by an automatic attendance system, which monitors students' attendance in each class using a continuous stream of images from a video streaming device placed within the classroom. The suggested remedy might lessen the time spent by teachers monitoring attendance will also lessen the commotion in the classroom. For automated face identification utilising Deep Learning models like YOLO-v3 (You Only Look Once) and face recognition using Siamese Network, this piece of study offered a superior conclusion. The model has hardware that allows it to keep track of pupils' attendance in rooms with cameras. For the purpose of storing student information, a database is kept up to date. The data in the database can be changed, destroyed, or both. The framework has a connection to a dynamic web page that facilitates obtaining the attendance record directly. There are many different institutions, colleges, and universities that can employ this sophisticated face recognition-based attendance system.

Keywords: Smart Attendance, Face Recognition, YOLO-v3, Siamese Network, Database.

1. Introduction

The most significant use of image processing, which is employed in many other industries, is face recognition. Facial Recognition-Based Attendance System for Educational Institutions is the topic of this research. The primary goal of building an automatic attendance management system is to replace the traditional method of manually updating attendance by teachers with an automated system that uses face recognition. The accuracy of the data gathered is the main problem, according to the prior attendance management system. This is due to the possibility that the attendance may not be personally recorded by the original person; in other words, a specific person's attendance may be recorded by a third party without the institution's knowledge, which compromises the accuracy of the data For example, if student A is too lazy to attend a particular class, student B will assist him/her in signing for attendance, even though student A did not attend the class, but the system will overlook this because no enforcement is used. If the institution establishes enforcement, it may have to waste a lot of human resources and time, which is not practical at all. As a result, all of the previous system's recorded attendance is untrustworthy for analysis. The previous system's second flaw is that it is too time consuming. Assuming that a student signs his or her attendance on a 3-4 paged name list in about 1 minute. Only about 60 students can sign their attendance in an hour, which is obviously time-consuming and inefficient. The third problem is with the legitimate interested party's ability to access those data. As an illustration, the majority of parents are quite concerned about tracking their children's location to make sure they are actually enrolled in classes at school or college. The parents cannot obtain this information under the prior system, though. Consequently, evolution is required to be made to the old system to increase productivity, data accuracy, and information accessibility for the rightful parties.

2. Related Work

[1] The author proposed Different Scales Face Detector (DSFD), a new face detection network based on Faster R-CNN that improves detection of small faces. The author used the WIDER FACE dataset, which contains 32,203 images and 393,703 labelled human faces, and 25,000 images were chosen at random from the dataset for training. The proposed DSFD model is divided into two parts. The multitask RPN, which is based on the boosting cascade face detection constraint and parallel-type Fast R-CNN, is one example. To obtain the human face ROI, an efficient multitask region proposal network (RPN) combined with boosting face detection is developed first, followed by a parallel-type Fast R-CNN network proposed and assigned to three corresponding Fast R-CNN networks. With 130 ms to process an image frame, the proposed method achieves a recall rate of 96.69%, while the original Faster R-CNN works at 140 ms.

In [2], The authors proposed a cascade convolutional neural network based on separable residual convolution that outperforms in terms of speed and accuracy. The dataset used in this model is the WIDER FACE dataset, which contains 393,703 annotated faces with a wide range of scale and pose variations. The proposed cascade convolutional neural network model is divided into three stages, with the detection stage performing two tasks: face classification and bounding box regression. The first step is to secure the region proposal and confidence. The second stage network aims to further screen the misjudged face region region proposal. The final face detection results are output by the third-stage network. On a medium set of wider faces, the model achieves a true positive rate of the discrete ROC curve at 2000 false positives (96.1%) and an accuracy of 84.6%.

The author in [3] proposed YOMO, a real-time face detector composed of depth-wise separable convolutions and multiple feature fusion structures in the form of a topbottom pattern The dataset used in this method is the WIDER FACE dataset, which is divided into 61 event classes; each class is then randomly divided into 40%, 10%, and 50% for training, testing, and validation sets, respectively. YOMO has only 21 million parameters and achieves superior performance on a GPU, achieving 51 FPS for a 544 input image. The architecture relies on MobileNet as its backbone, which is truncated before the classification layers and supplemented with 6 depth-wise separable convolution modules that output 1024 channels. The YOMO-Semi soft sample has a recall rate of 97.7% in discrete and 75% in continuous, and a small face ratio of 56%, a medium face ratio of 30.81%, and a large face ratio of 13.16%.

A novel face detector [4], Deep Pyramid Single Shot Face Detector (DPSSD), which detects faces with large scale variations and is fast. For training, the Universe face dataset was used, which is a combination of UMD Faces images and UMD Faces video frames. The Universe dataset contains approximately 5.6 million images representing approximately 58,000 identities. The author used two networks for feature representation and performed fusion by averaging their similarity scores. These two networks are based on the ResNet-101 and the Inception ResNet-v2 protocols. The model was tested on four difficult face detection datasets, including WIDER Face, Unconstrained Face Detection Dataset (UFDD), Face Detection Dataset and Benchmark (FDDB), and Pascal Faces. On the easy and medium sets, the WIDER Face dataset achieves mean average precision (mAP) of 92.5% and 90.8%, respectively, while the UFDD dataset achieves mAP of 70.6%, the FDDB dataset achieves mAP of 96.9%, and the Pascal Faces dataset achieves mAP of 96.11%.

In [5] To achieve high performance, the author proposed RefineFace, a single-shot refinement face detector. WIDER Face contains 393, 703 annotated faces in 32, 203 images, with variations in pose, scale, facial expression, occlusion, and lighting condition. These images are divided into three subsets: training (40%), validation (10%), and testing (50%), with each set classified as easy, medium, or difficult. This model has five modules: Selective Two-step Regression (STR), Selective Two-step Classification (STC), Scale-aware Margin Loss (SML), Feature Supervision Module (FSM), and Receptive Field Enhancement (RFE) (RFE). WIDER has undergone extensive testing FACE, AFW, PA+SCAL Face, FDDB, and MAFA show that the proposed method produces cutting-edge results. The proposed method improves the AP score of current results by 1.55% on the AFW dataset and 99.5% aP on the PASCAL Face dataset, which is the highest score.

The work in [6] An automated attendance system built with OpenCV is designed to provide student attendance in any college or university. The primary goal of this system is to improve the traditional attendance system of various colleges and universities in order to reduce time and asset waste, as well as to improve the adaptability and performance of the attendance system procedure while reducing long time processes and work. They used Haarcascade features for face detection and Local Binary Pattern Histogram for face recognition in this study (LBPH). LBPH is used to compile the local structure of an image/frame with each pixel's neighbour frames. In machine learning, the K-Nearest Neighbor regressor is also used for regression and classified recognition terms. This framework can not only aid in the participation framework, but it can also improve an organization's altruism and significantly reduce time and paper waste.

The primary goal is to create a face recognition attendance system that is based on real-time video processing [7]. The author divides the task into two parts: face detection and face recognition. Face detection collects information to determine whether there is a human face image in the image, as well as its size and position, and converts the detected human face image into an adult face region. Face recognition technology extracts facial feature and image information to determine if it is in the repository and matches identity information. Otherwise, no recognition result is obtainedFace feature extraction is carried out here using linear discriminant analysis, which is used to find a set of linear transformations that minimise intra-class variance between each category while maximising inter-class variance. According to experimental data, the video face recognition system has an accuracy of up to 82%.

The methodology presented in [8] involves tracking an object with a drone-mounted camera. Object tracking applications are becoming increasingly important. This paper builds on the core Deep SORT algorithm and a combination of YOLOv3 and RetinaNet to generate detections in each frame for the implementation of multi-object tracking benchmarks. SORT (Simple Online and Realtime Tracking) is one such detection-based tracking algorithm. YOLOv3 658974

is a real-time object detection model that identifies specific objects in videos, live feeds, or images. When compared to many existing trackers, RetinaNet was created by two improvements over existing single-stage object detection models: Feature Pyramid Networks (FPN) and Focal Loss. The author used the py-motmetrics library that supports CLEAR-MOT and ID metrics to compare the implementation with other trackers. Py-motmetrics records all relevant events in a snapshot, e.g. B. Matches, crashes, false positives and changes. . The results in this article for the 2018 VisDrone dataset show competitive performance against many existing trackers.

In the modern world, the education system is developing day by day with the introduction of the concept of "smart classroom" [9]. The article describes the existence of an intelligent attendance system without human intervention. In many organisations, schools and universities, attendance is still a manual method hence common methods such as a teacher/instructor calling students by name to indicate attendance. 10 algorithms used in auto-attendancethe system is Viola & Jones algorithm for face detection and recognition. With Viola & Jones converts the model from an RGB image to a YCbCr image to detect facial features. The mean square error (MSE) is calculated from the live image and the reference image pixels when the MSE is higherA face is detected above the threshold. The result of the proposed system is the replacement of manual systems for static materials such as registers and bulky documents with fast, efficient, rapid and economical automatic systems.

The work proposed in [10] aims to design a personalized face recognition technology as it is used in many applications such as medicine, security, and robotic navigation, etc. The algorithm used by the author is the YOLO v3 algorithm becausegenerates faster face recognition output by comparing existing deep learning algorithms like R-CNN, Fast R-CNN, etc. The author divides the whole process into two stages: face recognition and face recognition. There are two stages in detection methods: single and double shot detection.Single-shot detection, face detection is achieved in one step. Double Shot Recognition There are two steps in face recognition, the first step is feature extraction and the second step is classification. YOLO v3 is a single-shot object detection and recognition algorithm used in many real-time applications. The single-shot algorithm performs feature extraction and classification in a single step.The result of this algorithm is fast facial recognition for continuous monitoring purposes to flag help, medical assistance, etc.

The author of the article [11] mainly focused on the recognition of faces in crowded places, pose changes, light changes, distortions, facial expression changes, fragments of blurry images and low-resolution images. Resolution. In this article, the author developed a technique called Normalized Pixel Difference (NPD). The sum of the pixel values is calculated from the values of the two pixels and then the difference of the pixel values is determined. In the end, the sum and difference result in the ratio caught. Three public datasets: FDDB, GENKI, and CMU-MIT are used by the author throughout the book. The author used a new technique called "Deep Squaretic Tree" to understand the features produced by a normalized pixel Difference (NPD) and their combinations. The author used a flexible waterfall classifier to recognize different faces and assign them the appropriate class. It has been observed to be 6x faster than Open CV. The author even used the AdaBoost algorithm to understand deep square trees based on NPD features.

The importance of face detection in various applications like face recognition, liveness detection and age estimation [12]. In these approaches, it is necessary to have boxes bounding the face and an extra Non-Maximum Suppression (NMS) which is done after processing. Author proposed a dual-branch center face detector that conducts face detection in a pure convolutional neural network. It simplifies the processing steps of face detection and improves the detection speed. To reduce the influence of large variations of scales, Author proposed feature pyramid aggregation (FPA) module and scale adaptive Gauss mask. Both of them can significantly improve the performance of face detection. The evaluation of the proposed method is done on popular face detection benchmarks, such as AFW, PASCAL face, FDDB and WIDER FACE. It demonstrates the capability and competitiveness of the proposed method with other approaches.

Face detection is a fundamental step in many faces related applications such as: face alignment, face recognition, face verification and face expression analysis. In order to detect 12 small faces better, this paper proposes a new scale-invariant face detector called Small Faces Attention (SFA) face detector [13]. SFA performs face detection in one step by scanning the entire image with a sliding window. It detects faces directly from the initial feature maps by classifying a predefined set of anchors and regressing them simultaneously. Author introduced a multi-branch face detection architecture that plays an important role in small face detection. Here, feature map fusion is applied by fusing the feature maps of adjacent branches and using large features to detect small hard surfaces. It simultaneously performs multi-scale training and testing concurrently to make the model robust at different scales. This method achieves promising recognition performance on many demanding face recognition benchmarks such as WIDER FACE and FDDB datasets at competitive inference speed.

The main aim at the design of Face Detection in security surveillance based on Artificial Intelligence video retrieval technology. The main purpose of the video retrieval is to locate the development process of the causes of events and their associations. It proposed a video oriented cascaded intelligent face detection algorithm, that builds a deep learning network by cascading multiple features, such as edge features, contour features, local features to semantic features etc. It collects the images and video stream containing face, and automatically detects and tracks the faces in image, and then carry out series of face related technologies for the detected face [14]. Experimental results show that the algorithm has good detection performance for single face and multi face images, and has strong robustness for rotating faces.

The primary goal is to create Face Detection in security surveillance using Artificial Intelligence video retrieval technology. The primary goal of video retrieval is to locate the development process of event causes and their associations. It proposed a video-oriented cascaded intelligent face detection algorithm that constructs a deep learning network by cascading multiple features, such as edge features, contour features, local features, and semantic features, among others. It collects images and video streams containing faces, detects and tracks the faces in the images, and then applies a series of face-related technologies to the detected face [15]. The results of the experiments show that the algorithm has good detection performance for single and multi-face images, as well as strong robustness for rotating faces.

Table 1: Comparison Table

Ref No.	Description	Limitations	Advantages	Performance Metrics
1	Face recognition network based on Faster RCNN, a multi-scale face detector (DSFD) that improves the detection of small faces.	Network parameters are difficult to adjusta large selection of object scales.	High-quality face proposals	Recall rate of 96.69% with 130 ms.
2	A cascading convolutional neural network based on a separable residue convolution has been proposed.	Complicated and variable situations in the face region during the process of training a large number of samples	Improve the accuracy of candidate window positioning	84.6% accuracy on medium set of wider faces.
3	YOMO Face detector is proposed.	Small faces cannot be handled well by singlestage face detectors.	Less Computation.	Recall rate of 97.7% and 75% in discrete and continuous.
4	A new Deep Pyramid Single Shot Face Detector (DPSSD) has been proposed that is fast and capable of detecting faces with high variability	Domain adaptation and dataset bias and Training CNNs.	The model is fast and detects faces with large scale variations.	The model is fast and detects faces with large scale variations.
5	Single-shot refinement face detector namely Refine Face to achieve high performance.	The LOC error is triggered by the lack of strong regression ability.	It better distinguish faces from the complex background across different scales	Mean average precision (mAP) of 92.5%.
6	The main reason why this system was proposed is to improve the traditional system of attending different universities, avoid misuse of time and resources, and improve adaptability and efficiency.	For making attendance to individual student some pretrained models are used and also input to be taken in gray scale mode.	This software solves the realworld issue with schools and universities in proving attendance	Accuracy – 99%.

7	Real-time video recognition and a basic face recognition algorithm via a video recognition system were used for the experimental setup.	Compared with the traditional check-in method, the face recognition attendance system can be reduced by 60%.	The face recognition attendance system with real-time video processing can be quickly complete the tasks of students in the time.	Accuracy rate – 82%.
8	Simple Online and Realtime Tracking (SORT) and YOLO v3 (You Only Look Once, Version 3) are used for constructing the models.	Simple Online and Realtime Tracking (SORT) and YOLO v3 (You Only Look Once, Version 3) are used for constructing the models.	Exhibits a competitive performance when compared to other trackers.	Accuracy – 80%.
9	The algorithms used in the automated attendance system such as Viola & Jones algorithm for face detection and face recognition.	The use of ordinary generic cameras for video replay attacks in a huge storage requirements, vulnerable detection, and potential privacy issues	It provides a better security, easy integration, and automated identification.	Accuracy rate – 88%.
10	In this method YOLO-v3 algorithm can be used for facial recognition which produces faster output.	It has low Mean Average Precision (MAP) value while compared to the CNN algorithms.	YOLO-v3 is having a high processing speed and it is in millisecond of range.	Accuracy rate – 63%
11	The technique used is called Normalized Pixel Difference (NPD). Calculate the value of two pixels, and then find the difference in pixel values. Finally, the sum/difference ratio is calculated.	Occlusions and blur are two big challenges for face detection and work need to be done to overcome it.	Achieve promising results for face detection with large pose variations and occlusions	Speed – 70.06 FPS
12	A central two-armed face detector that performs the face recognition process in a purely convolutional neural network, which simplifies the face recognition processing steps and improves the recognition speed.	The issue is that NMS cannot be easily parallelized and it may become a bottleneck of detection speed.	Pure convolutional neural method will obtain higher speedup ratio from the increase of computing units.	Accuracy – 90.52%
13	Face recognition for many face-related applications, such as B. face alignment, face detection, face verification and facial expression analysis for small faces	The main drawback of this small face attention face detector is it has poor detection performances.	Small faces attention face detector significantly improves face detection performance, especially on faces.	Accuracy – 86%
14	Face Detection in security surveillance based on Artificial Intelligence video retrieval	Drawback of video oriented cascaded intelligent face detection algorithm is	Facial recognition technology is used in public safety CCTV	Accuracy rate – 77%.

	technology.	it consumes a lot of manpower and material resources.	cameras, reducing the likelihood of illegal crime	
15	face alignment method for pose invariant face recognition by using Adaptive Pose Alignment (APA) method.	Drawback of Adaptive Pose Alignment method is when it processes the faces containing extreme poses, this alignment method is not conducive to face recognition	APA reduces the intra-class difference and corrects the noise caused by the traditional method in the alignment process.	Accuracy – 99.8%.

3. Methodology

3.1 Using R-CNN

a) Improving RPN The flowchart for the detection procedure is presented to help convey our proposed DSFD approach more intuitively. The first is the enhanced RPN procedure, while the second is the enhanced Fast R-CNN technique. Steps for Improving RPN:

1. In the first step, Input units are passes i.e sample images are passed.
2. For image pre processing, we used boosting cascade face detection and separated the face ROI into several groups based on their scales. They also presented a strategy for ROI convolution to reduce the area of each layer of convolution. Because convolution consumes 90% of the time for the entire network operation, such a technique can significantly improve detector speed.
3. The proposal is generated by the anchor box in the improved RPN. Modified version of the original RPN [20] is used to generate face proposals. A shared small network that slides over the conv5_3 feature map generates a sequence of anchor boxes.
4. If the center of the anchor box inside the Face ROI Mask, the sliding stride is set to 1.
5. If the center of the anchor box doesn't lie inside the Face ROI mask the sliding stride is set to
6. 6.Then the proposal is generated.

b) Improving R-CNN

1. If the proposal generated in RPN had a small face, then small face Fast R-CNN IS improved.
2. likewise, the first step if the proposal generated in RPN had a medium and large face then medium face and large face fast R-CNN is improved.
3. In third step it detects the results of small, medium, and large faces
4. In final step the detection results are merge

3.2 Using CNN

Face Detection Process

Our approach is entire pipeline consists of two processes: training and testing. The training image is fed into the algorithm, and it is initially resized to different scales in order to form an image pyramid [27]. Iteratively trains the convolutional neural network's input to obtain the final convolutional neural network. The testing procedure is to feed the test images into the convolutional neural network, which is then trained to produce the face detection results. Steps for Face Detection Process:

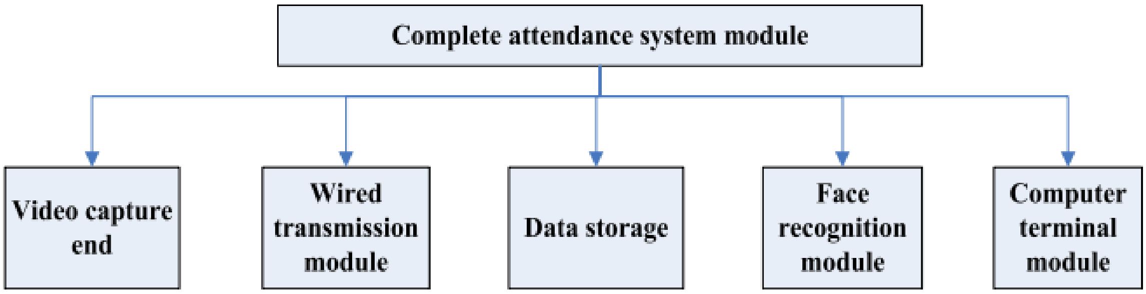
1. Input image I0(x, y), then use image pyramid [23] pre-processing to get set I=I1, I2 In.
2. The set I is fed into the first-stage network to generate face region proposals, and the bi in the face region proposals with confidence greater than the threshold T1 is outputted to form the set B.
3. Bounding box regression is applied to the face region proposal in set B, and the highly overlapping face region proposal is filtered using the NMS approach.
4. Resize the filtered face region proposal to a 2424 scale window as the input to the second stage network after mapping it to the original image I0.
5. With a confidence greater than the T2 threshold, output bi in the face area proposal and update the set B.
6. In the updated set B, the face region proposal is submitted to bounding box regression, and the highly overlapping face region proposal is filtered using the NMS approach. Step 7: The filtered face region suggestion is mapped to the original picture I0 and resized to 4848 pixels.

3.3 Using Local Binary Pattern Histogram

LBPH (Local Binary Pattern Histogram) in open computer vision. LBPH is used to fill the local image/image structure with the adjacent images of each pixel.Choose a specific pixel as the center and threshold of the others. Returns "1" if pixel quality is greater than equal to neighbor, otherwise returns "zero". So, Combinations of surrounding pixels produce local binary digits or local codes. In my experience, LBPH is a slightly better approach to face recognition than automatic faces and Fisher faces, which are used to process data in the form of vectors and provide high-dimensional spatial frames that are either false orNot identical, but in LBPH it offers a low-dimensional space that is identical and more recognizable. The LBPH concept used in the attendance system consists of recognizing the face of the entire class with an accuracy rate of 96% or keeping some of the distinguishing data for linear analysis, as is the case with fishermen Facial algorithm. So, to get good recognition rates we need at least 8 (+1) images per person and in our design, we aim for 20-30 images per person to avoid a drop in the recognition rate.

3.4 Using combination of Open CV and KNN algorithm

A complete attendance system consists of a combination of many modules, each fulfilling different functions. To reduce program complexity and facilitate code reuse, it supports maintenance and management of the entire system. The design of the face and presence recognition system in this system mainly includes multiple video capture terminal modules, a cable transmission module, data storage, a face recognition module and a computer terminal module



3.5 Using combination of YOLOv3 and RetinaNet

We used a sensor framework that is a combination of YOLOv3 and RetinaNet. YOLOv3 is fast and works well on common objects, but not well-suited to smaller, denser objects. On the other hand, RetinaNet works well, especially in cases where the objects are small and appear inclusters. This structure returns all detections of a given frame. Redundant detections were removed with non-maximum suppression (NMS), creating a set of all possible bounding boxes consisting of all newly detected objects in that box. the detections of each image were fed into a pre-trained CNN model on a Reid dataset. The MARS (Motion Analysis and Re-Identifying) dataset contains 1,261 identifiers with 200,000 track lets. six synchronized cameras were used to create this dataset, and each pedestrian captured by two cameras is added to the dataset. This model generates a deep association matrix associated with each detection it contains Characteristics of the appearance of the object. These aesthetic features were combined with information about the movement of objects detected in the array. This integration of well distinguishing features is useful When tracking objects after a short-term or long-term occlusion state, by assigning the same identity to an object after an occlusion.

3.6 Using linear discriminant analysis

Face recognition is a computer vision technique that examines data on facial features to identify individuals. Face detection and face recognition matching are the two components of face recognition. The process of extracting facial features makes use of linear discriminant analysis (LDA). The goal of the LDA algorithm is to identify a set of linear transformations that maximise interclass dispersion while minimising intraclass dispersion within each category. There are four main facial recognition techniques. As follows:

1) Method of Geometric Feature
2) Use of Subspace Analysis
3) Use of Neural Networks
4) Support Vector Machine Technique.

4. Conclusion:

To eliminate the inconvenience of traditional manual operation, the system implementation is based on an automatic attendance scheme based on facial recognition. It can be used by enrolled students atas an access control system Organization with their faces and subsequent recognition of students and attendance grade. Most existing approaches are time-consuming, cumbersome, and require manual work on the part of users. In this project we proposed deep learninga model that performs face recognition with YOLOv3 and face recognition with a Siamese network. In our work, we attempted to overcome these challenges by successfully demonstrating a face registration system. Electronic Participation The presence management system marks presence on the photos, after which the faces in the image are recognized and compared with the photos already taken. Finally, the student is marked as present or absent based on the results of the captured image.The proposed system is designed for economy and reliability. The accuracy of the system can be improved by training the system to correctly recognize faces with scarves, sunglasses, and beards.The proposed model promises no memory overload by removing the intermediate frames created by YOLO.

5. References:

[1] W. Wu, Y. Yin, X. Wang and D. Xu, "Face Detection With Different Scales Based on Faster RCNN," in IEEE Transactions on Cybernetics, vol. 49, no. 11, pp. 4017-4028, Nov. 2019, doi: 10.1109/TCYB.2018.2859482.

[2] R. Qi, R. -S. Jia, Q. -C. Mao, H. -M. Sun and L. -Q. Zuo, "Face Detection Method Based on Cascaded Convolutional Networks," in IEEE Access, vol. 7, pp. 110740-110748, 2019, doi: 10.1109/ACCESS.2019.2934563.

[3] J. Xu, Y. Tian, H. Wu, B. Luo and J. Guo, "You Only Move Once: An Efficient Convolutional Neural Network for Face Detection," in IEEE Access, vol. 7, pp. 169528-169536, 2019, doi: 10.1109/ACCESS.2019.2954936.

[4] Ranjan, Rajeev; Bansal, Ankan; Zheng, Jingxiao; Xu, Hongyu; Gleason, Joshua; Lu, Boyu; Nanduri, Anirudh; Chen, Jun-Cheng; Castillo, Carlos D.; Chellappa, Rama (2019). A Fast and Accurate System for Face Detection, Identification, and Verification. IEEE Transactions on Biometrics, Behavior, and Identity Science, () , 1–1. doi:10.1109/TBIOM.2019.2908436.

[5] Zhang, Shifeng; Chi, Cheng; Lei, Zhen; Li, Stan Z. (2020). RefineFace: Refinement Neural Network for High Performance Face Detection. IEEE Transactions on Pattern Analysis and Machine Intelligence, (), 1–1. doi:10.1109/TPAMI.2020.299745.

[6] Gupta, N., Sharma, P., Deep, V., & Shukla, V. K. (2020, June). Automated attendance system using OpenCV. In 2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO) (pp. 1226-1230). IEEE.

[7] Yang, H., & Han, X. (2020). Face recognition attendance system based on real-time video processing. IEEE Access, 8, 159143-159150.

[8] Kapania, S., Saini, D., Goyal, S., Thakur, N., Jain, R., & Nagrath, P. (2020, January). Multi object tracking with UAVs using deep SORT and YOLOv3 RetinaNet detection framework. In Proceedings of the 1st ACM Workshop on Autonomous and Intelligent Mobile Systems (pp. 1-6)

[9] Sharanya, T., Sucharith, P., Kasturi, U., & Mahesh, T. (2020). Online Attendance using Facial Recognition. Int. J. Eng. Res, 9(06), 202-207.

[10] Menon, S., Geroge, A., Aswathy, N., & James, J. (2021, May). Custom Face Recognition Using YOLO. V3. In 2021 3rd International Conference on Signal Processing and Communication (ICPSC) (pp. 454-458). IEEE.

[11] Liao, S., Jain, A. K., & Li, S. Z. (2015). A fast and accurate unconstrained face detector. IEEE transactions on pattern analysis and machine intelligence, 38(2), 211-223.

[12] Li, X., Lai, S., & Qian, X. (2021). Dbcface: Towards pure convolutional neural network face detection. IEEE Transactions on Circuits and Systems for Video Technology, 32(4), 1792-1804.

[13] Luo, S., Li, X., Zhu, R., & Zhang, X. (2019). SFA: Small faces attention face detector. IEEE Access, 7, 171609-171620.

[14] Dong, Z., Wei, J., Chen, X., & Zheng, P. (2020). Face detection in security monitoring based on artificial intelligence video retrieval technology. IEEE Access, 8, 63421-63433.

[15] An, Z., Deng, W., Hu, J., Zhong, Y., & Zhao, Y. (2019). APA: Adaptive pose alignment for pose-invariant face recognition. IEEE Access, 7, 14653-14670.