Survey Paper on Automatic Student Attendance System Using Deep Facial Recognition

Ms. Bharti Nana Durgade, Ms. Pranita Argade, Ms. Ashwini Joshi, Ayesha Mujawar
[1] *Assistant Professor, Dept. of Information Technology, Sinhgad Institute of Technology and Science, Narhe, Pune*
[2,3,4] *Student, Dept. of Information Technology, Sinhgad Institute of Technology and Science, Narhe, Pune*

***Abstract*—Both educators and school officials find that tracking student attendance is a tedious and time-consuming endeavour. We therefore considered using the most recent developments in machine learning to automate this operation. This article suggests a facial recognition and detection method for managing attendance. A camera takes pictures of the classroom all the time. Students' face traits are identified and extracted by a thorough examination of the taken pictures. After that, their identities are predicted by a pattern recognition model. The suggested architecture is validated by the experiment's outcomes. No human involvement is required in the process of recording the kids' attendance.**

***Index Terms*— facial recognition and detection, pattern recognition, etc**

I. INTRODUCTION

The field of computer vision known as facial recognition (FR) focuses on analyzing facial picture features to determine an individual's identity. It is among the most trustworthy biometric methods for determining an individual's identity. Numerous data analytic, security, and industrial domains employ this technique [1]. We are using this technology to automate student attendance management because of the advancements it has made in recent years. In classrooms with many of pupils, manually keeping track of attendance becomes a tedious and time-consuming operation. In order to do this, we have considered utilizing the FR to establish a dependable system for automating the management of attendance by identifying the faces of the pupils. A high-resolution digital camera is positioned at the top of the board in the suggested system to record the classroom audience. The camera's pictures are then forwarded to a comprehensive analysis unit, which will identify any faces in the pictures and extract their characteristics. Finally, a pattern recognition model will correctly identify the individuals in the picture.

Taking attendance is a vital but frequently difficult activity in educational institutions. In addition to being time-consuming, traditional techniques like calling names or utilizing card-based systems are also vulnerable to manipulation and mistakes. Inaccurate record-keeping and disturbance of the learning environment can result from these inefficiencies. We suggest an automatic attendance system that makes use of face recognition technology in order to overcome these difficulties.

II. LITERATURE REVIEW

Md. Abdur Rahim, Md. Najmul Hossain, Tanzillah Wahid & Md. Shafiul Azam, “Face Recognition using Local Binary Patterns”[2], The human face provides a wealth of information about a person's identity and emotional condition. A fascinating and difficult topic, face recognition affects many crucial applications, including personal identification, banking and security system access authentication, and law enforcement identification. Our research is primarily divided into three sections: classification, feature extraction, and face representation. The process of modeling a face is known as face representation, which also establishes the subsequent algorithms for detection and recognition. During the feature extraction step, the face image's most valuable and distinctive features are extracted. The facial image is compared to the database photographs throughout the categorization process. In our study, we empirically assess face recognition that uses Local Binary Patterns for person-independent face recognition, which takes into account both shape and texture information to describe face images. In order to extract Local Binary Patterns (LBP) and histograms and concatenate them into a single feature vector, the face area is first separated into small sections. This feature vector is used to quantify image similarity and creates an effective face representation.

Minze Li, Xiaoxia Li, Wei Sun, Xueyuan Wang, Shunli Wang, "Efficient convolutional neural network with multi-kernel enhancement features for real-time facial expression recognition" [3], Personal emotions are most directly expressed through facial expressions. There is a lower feature difference in face expressions than in other pattern recognition issues. The general approaches either have too many parameters to allow for real-time processing or have trouble accurately characterizing the feature difference. In order to address the speed and accuracy of real-time facial expression identification, this research suggests a multi-kernel feature facial expression recognition network with a lightweight mobile architecture. The first step is to create a multi-kernel convolution block by parallelizing three depth wise separable convolution kernels of varying sizes. Both the tiny and large kernels are capable of extracting facial expressions' edge contour information and local details, respectively. To better characterize the variations in facial expressions, the multi-channel data is then combined to create multi-kernel improvement features. Second, the input of the multi-kernel convolution block undergoes a "Channel Split" operation, which can minimize the number of parameters to one-third of the initial value and prevent the recurrent extraction of incorrect information. In order to further enhance the feature representation capability, a lightweight multi-kernel feature expression recognition network is created by alternating between depth wise separable convolutions and multi-kernel convolution blocks. The suggested network achieves high accuracy of 73.3 and 99.5% on the FER-2013 and CK+ datasets, respectively, according to experimental results. Additionally, with 640 x 480 video, it reaches a speed of 78 frames per second. Its speed and accuracy are better than those of other cutting-edge techniques.

R. Verma and R. Kaur, "Enhanced character recognition using surf feature and neural network technique" [4], One of the most intriguing and difficult study topics in image processing and pattern recognition in recent years has been handwriting recognition. Identifying documents in both handwritten and printed formats is the biggest problem in the field of image processing. Because handwritten characters can vary due to various noises or font sizes, it can be challenging to recognize them. Character recognition is one of the most popular methods for document and human authentication. Character recognition using the Neural Network Algorithm model is the main emphasis of the solutions suggested in this study. Its main purpose is to retrieve a character that has been stored in memory when a noisy or insufficient version of that character is displayed. We use the Matlab software's Image Processing Toolbox to implement this suggested job. The MatLab-created image pre-processing utility implements numerous brightness adjustments and local pre-processing techniques. Strong robustness performance and good feature point distinction are capabilities of the suggested SURF feature and neural network, which has also significantly increased computation speed. The work done in this study has resulted in an average success rate of 98.7753% for the recognition of all characters utilizing SURF features and neural networks. These techniques, which are based on input picture noise, yield encouraging PSNR and MSE results.

Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, "Deepface: Closing the gap to human-level performance in face verification" [5], The traditional pipeline for contemporary face recognition goes like this: detect => align => represent => classify. Using explicit 3D face modelling, we perform a piecewise affine transformation and derive a face representation from a nine-layer deep neural network, revisiting both the alignment and representation steps. Instead, then employing the conventional convolutional layers, this deep network uses many locally linked layers without weight sharing, including around 120 million parameters. As a result, we used the largest facial dataset available to train it—four million facial photos from over 4,000 identities—which was identity tagged. Even with a basic classifier, the learnt representations that couple the massive facial database with the precise model-based alignment translate very well to faces in uncontrolled situations. With an accuracy of 97.35% on the Labeled Faces in the Wild (LFW) dataset, our methodology approaches human-level performance by lowering the error of the state-of-the-art by over 27%.

I. Melekhov, J. Kannala, and E. Rahtu, "Siamese network features for image Matching"[6], Many computer vision applications, including structure-from-motion (SfM), multi-view 3D reconstruction, image retrieval, and image-based localization, depend on the ability to find matching images across huge datasets. In this paper, we propose to use neural network-based feature vectors, whose similarity is assessed by Euclidean distance, to describe images in order to identify matching and non-matching pairs. Using a contrastive loss function in a Siamese

network architecture, convolutional neural networks are used to learn the feature vectors from labelled examples of matching and non-matching image pairs. Siamese architecture has been used in the past for facial image verification and local image patch matching, but not yet for whole-image matching or generic image retrieval. According to our experimental findings, the suggested features outperform baseline features derived from networks trained for image classification tasks in terms of matching performance. The characteristics enhance picture matching for new landmarks that are not visible during training and have good generalization. This is true even when our training data has imperfect matching and non-matching pair labelling. In light of image retrieval applications, the results are encouraging, and using more training image pairs with more precise ground truth labelling should lead to even better outcomes.

O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition" [7], This study aims to recognize faces, either from a single image or from a group of faces that are tracked in a video. Two aspects have contributed to recent advancements in this field: (i) the use of convolutional neural networks (CNNs) for end-to-end learning of the job, and (ii) the availability of very large-scale training datasets. We contribute in two ways: first, we demonstrate how a combination of automation and human intervention can assemble a very large dataset (2.6M images, over 2.6K people) and discuss the trade-off between time and data purity; second, we navigate the intricacies of face recognition and deep network training to present techniques and protocols to attain similar state-of-the-art results on the standard LFW and YTF face benchmarks.

Y. Sun, Y. Chen, X. Wang, and X. Tang,

"Deep learning face representation by joint

identification-verification" [8], Creating efficient feature representations to decrease intra-personal variations and increase inter-personal differences is the main problem in face recognition. In this research, we demonstrate that deep learning can effectively solve it with supervision from both face identification and verification signals. Carefully crafted deep convolutional networks are used to learn the Deep Identification-verification features (DeepID2). Essential to face recognition, the face verification task decreases intra-personal variations by combining DeepID2 features extracted from the same identity, while the face identification task increases inter-personal variations by separating DeepID2 features extracted from different identities. New identities not present in the training data can be effectively represented by the DeepID2 features that have been learned. The accuracy of face verification is 99.15% on the difficult LFW dataset. The error rate has been greatly decreased by 67% when compared to the best prior deep learning result on LFW.

M. Wang and W. Deng, "Deep Face Recognition" [9], When opposed to typical facial traits, face recognition relies on a very small number of distinguishing features. This increases the significance of color information in identification and recognition processes. However, in certain situations, such as legacy photos, handling grayscale facial photographs is essential. In this work, we examine how face recognition is affected when color information is lost. Using CNN-based colorization prior to a CNN classifier, we present a unique approach. The LFW benchmark dataset is used to test the suggested framework. The evaluation findings demonstrate how well the suggested framework works to lessen the detrimental impact of removing color information on face recognition performance.

III. PROBLEM STATEMENT

Traditional classroom attendance recording techniques are laborious, prone to human error, and frequently ineffective. Attendance is usually manually marked by teachers, taking up valuable class time and perhaps resulting in inaccurate attendance statistics. Manual systems also make it difficult to track and evaluate attendance patterns over time, which may result in the failure to notice children who have a history of poor attendance.

IV. OBJECTIVE

1. To create an intuitive and easy-to-use interface for teachers and administrators.
2. To enable real-time monitoring and updating of attendance records.
3. To develop a system to automatically capture and process images.
4. To Implement a centralized database or spreadsheet system.

V. METHODOLOGIES

The methodology involves several key phases. First, data collection is conducted by gathering a dataset of student images with various poses and lighting, along with high-resolution images of the entire classroom.

For model training, algorithms like Haar Cascades or MTCNN are used for face detection, while models such as FaceNet or DeepFace are trained on the collected student images for face recognition. Image quality is enhanced through preprocessing techniques like adjusting lighting and removing blurriness. During system development, software is implemented to detect and recognize faces in the captured images, automatically mark attendance, and store data in a centralized spreadsheet or database, with monthly attendance reports generated for each student.

VII. CONCLUSION

In this paper, we are studied a facial recognition-based attendance management system. Convolutional neural network models MTCNN and FaceNet, respectively, are used for face detection and recognition. Based on the findings, it can be said that the suggested design offers a workable way to control student attendance in classes.

REFERENCES

[1] A. K. Jain and S. Z. Li, Handbook of face recognition, vol. 1. Springer, 2011.

[2] M. A. Rahim, M. S. Azam, N. Hossain, and M. R. Islam, “Face recognition using local binary patterns (LBP),” Glob. J. Comput. Sci. Technol., 2013.

[3] Y. Feng, X. An, and X. Liu, “The application of scale invariant feature transform fused with shape model in the human face recognition,” in 2016 IEEE Advanced Information Management, Communicates, Electronic and Automation Control Conference (IMCEC), 2016, pp. 1716–1720.

[4] R. Verma and R. Kaur, “Enhanced character recognition using surf feature and neural network technique,” Int. J. Comput. Sci. Inf.

[5] Y. Taigman, M. Yang, M. Ranzato, and L. Wolf, “Deepface: Closing the gap to human-level performance in face verification,” in Proceedings of the IEEE conference on computer vision and pattern recognition, 2014, pp. 1701–1708.

[6] I. Melekhov, J. Kannala, and E. Rahtu, “Siamese network features for image matching,” in 2016 23rd International Conference on Pattern Recognition (ICPR), 2016, pp. 378–383.

[7] O. M. Parkhi, A. Vedaldi, and A. Zisserman, “Deep face recognition,” 2015.

[8] Y. Sun, Y. Chen, X. Wang, and X. Tang, “Deep learning face representation by joint identification-verification,” in Advances in neural information processing systems, 2014, pp. 1988–1996.

[9] M. Wang and W. Deng, “Deep Face Recognition: {A} Survey,” CoRR, vol. abs/1804.0, 2018.