IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 6.252

Volume 2, Issue 3, April 2022

Automatic Attendance System through Facial Recognition using Deep Learning

Rohan Chakor[1], Nikhil Hulawale[2], Kamlesh Kohle[3], Pravin Bibave[4], Prof. Sanket N. Wawale[5]

Students, Department of Information Technology[1,2,3,4]

Faculty, Department of Information Technology[5]

Amrutvahini Polytechnic, Sangamner, Maharashtra, India

rohanchakor9145@gmail.com, nikhilhulawale2@gmail.com, kamleshkolhe2003@gmail.com, pravinbibave39@gmail.com, sanket.wawale@gmail.com

Abstract: *Due to a large number of students in the schools and colleges most of the time teachers are not able to monitor on the attendance of the students. This may create the problems to both students and teachers in the end of the year or semester. The manual handling of all these can take much more effort in making the students to attend the classes regularly and also to intimate their parents on the valid interval of time. So to make this system more efficient and thereby to increase the productivity of the college or the school proposed system uses the automatic attendance system, where the proposed model takes the input from the staff regarding the attendance details and facial images of the students, then the system trains a Convolutional Neural Networks that have been deployed specifically for the facial recognition purpose. This trained model is then utilized for the purpose of achieving the identification of the student's faces through the test image and the decision making approaches for the marking of the attendance. The methodology has been subjected to the evaluation for the accuracy of facial recognition which has resulted in highly accurate outcomes.*

Keywords: Open CV, Convolutional Neural Network, Haar Cascade, Decision Making

I. Introduction

The technological advancements have been making the lives of individuals easier and highly accessible in day to day tasks. The increased convenience offered by these platforms allows for a much more effective and useful implementation of the everyday activities. The technological advancements have been highly useful in the realization of a much better quality of life, which has been evident quite blatantly. The educational sector has been one of the fields that has been improved through the realization of technologies which have been helpful in understanding the enhancements in the uninterrupted learning of the students. The influence of technology on the education sector has been instrumental in the improvement of the learning ability of the students by a large margin.

The paradigm of image processing has been a fundamentally complex task that has been improving every single day. The computer vision methodologies have been instrumental in the efficient and reliable identification and recognition tasks. There have been numerous deployments of the image processing approaches that have been implemented to achieve effective and useful identification of different types of objects. The realization of the recognition performance through the use of an effective neural network for the purpose of achieving the identification has been crucial for the scope of the computer vision methodologies. The image processing tasks are extremely complex and have been a constant topic of research amongst the researchers.

The task for facial image recognition is one of the most challenging aspects of the image processing implementations. The complexity of the facial image recognition is unparalleled due to the uniqueness of facial features between individuals has been one of the driving factors. The realization of facial recognition systems with an increased accuracy is an ongoing task with the researchers in the computer vision approaches working extensively. The lack of a highly precise facial recognition has driven research in the computer vision technologies which have been tirelessly working to improve the accuracy of these implementations.

The facial recognition approaches have variety of implementations that can be applied to achieve an improvement in that particular field. The use of facial recognition for the purpose of attendance marking has been a suitable realization that can be useful in a variety of different scenarios. The facial recognition approaches have been realized through the effective combination of facial feature extraction and classification. The facial feature extraction approach extracts the relevant features of the input face which are then learnt by a neural network. These facial features have been instrumental in the recognition tasks as it allows the truly unique characteristics of a person's face.

The use of facial features as a biometric identification approach has been increasingly used for the purpose of achieving authentication in various laptops and smartphones nowadays. These characteristics can be easily circumvented through the use of an effective and accurate implementation of the neural networks. Therefore, this research paper outlines an effective and useful approach for attendance management through facial recognition. This approach utilizes the Convolutional Neural Networks and Decision making for the purpose of facial recognition of the students and achieving the attendance of the students. This approach has been tested extensively with common realization of the experimental methodologies which has resulted in the extremely satistfactory response through the outcomes.

The Literature Survey component of this research paper examines previous work. Section 3 delves into the approach in depth, while section 4 focuses on the outcomes evaluation. Finally, Section 5 brings this report to a close and gives some hints for future research.

II. Literature Survey

A new approach for recognizing facial expressions from depth films has been proposed by M. Z. Uddin et al. [1]. LDRHP, LDSP, KPCA, and GDA were utilized to extract prominent facial characteristics for each pixel in a depth picture. LDRHP takes into account edge strengths in all directions, LDSP concentrates on edge strengths in the two most prevalent directions, KPCA employs a nonlinear method for dimension reduction and GDA clusters depth pictures of faces in a nonlinear feature space. The suggested feature extraction approach is resistant to variations in light and hence extracts important characteristics from depth pictures of faces. In addition, for training and identification, the characteristics have been merged using CNN-based deep learning. On several datasets, the suggested method was compared to other traditional ways, and it was found to be superior to the others. Therefore, any consumer system might use powerful expression recognition technology to better human-machine interaction.

B. Yang et al. presents a WMDNN-based FER approach for processing face grayscale and LBP facial pictures at the same time. The authors claim that both image channels are complementary, that they can collect a lot of information from face photos, and that they can help with recognition. To fully use the characteristics acquired from different picture channels, a weighted fusion technique is presented. To automatically extract characteristics of face emotions from facial grayscale photos, a partial VGG16 network is built. With initial parameters collected from ImageNet, fine-tuning is utilized to train the network [2]. Because no successful pre-trained model is depending on LBP pictures, a shallow CNN is built to automatically extract characteristics of facial emotions from LBP facial images. Following that, a weighted fusion technique to fuse both aspects is presented to fully use complimentary face information. A "softmax" procedure is utilized to achieve recognition results depending on fused features.

Q. Shi et al. [3] developed an effective ship classification framework by learning high-level features using deep CNN and integrating multi-scales rotation invariance characteristics. Traditional CNN has several drawbacks, making it difficult to apply in ship classification tasks. The proposed approach compensated for deep CNN's inadequacies by using a Gabor filter to collect features in various directions and MS-CLBP to get local texture, spatial, and profile information from ship pictures. Furthermore, the fused level representation has been proved to be stronger and more comprehensive since both high-level and low-level parts complement each other.

H. Wu et al. present a real-time multi-task convolutional neural network cascade architecture for simultaneous face detection and posture estimation [4]. Face detection job gets a boost from posture estimation task thanks to multi-task learning. The authors give a solution to achieve real-time performance for handling these two jobs by using a cascade structure. Using feature fusion, they improve posture estimation performance even further. On two tasks, extensive trials on accessible unconstrained datasets show that their strategy beats most state-of-the-art methods.

From unconstrained "In The Wild" photos, G. Storey et al. demonstrate the suggested Integrated Deep Model for recognizing faces and performing landmark localization. The author's main contribution is a strategy for effectively combining features from two state-of-the-art deep learning architectures to make use of both of their strengths for more precise face identification [5]. On the AFW and FDDB test sets, they show that their technique is similar to other topperforming face detection systems. IDM, in particular, shows a large decrease in the amount of false-positive detections, enhancing accuracy while having a minor influence on recall.

Z. Wu et al. performed studies on clinical image identification of six common face skin disorders utilizing five standard CNN structures and created a data set mostly made up of facial skin illness pictures. The findings show that CNNs are capable of detecting face skin disorders. Based on their findings, the authors believe that different models should be utilized to identify illnesses in various body sections. Furthermore, the tests revealed that a more appropriate network topology might increase the model's performance [6]. In some disorders, the present network structure has delivered good results, but overall performance has to be enhanced.

M. Z. Khan et al. offers a Convolution Neural Networks (CNN)-based face detection and identification method that outperforms previous approaches. To test the efficiency of the proposed algorithm, an automatic attendance system has been proposed to reduce human mistakes that occur in traditional attendance-taking systems. The primary goal is to automate the system and develop a smart classroom that will benefit educational institutions. To attain state–of–the–art results, a faster region convolution neural network and Edge Computing approaches are used [7]. The algorithm correctly identified 30 of the 35 faces it recognized; however, the accuracy may be improved by acquiring a sharper photograph of the students.

M. Shi et al. presents a FER method that depends on an FCM clustering algorithm combined with CNN. To extract the expressive image features in the training and test sets, the FCM technique is used in the CNN's convolutional layer to generate a convolution kernel with a starting value [8]. It is predicted to improve the model's FE capacity while also reducing model training time. Depending on simulation results, the F-CNN technique presented in this research may improve the detection rate of facial emotions in complicated backdrops to varying degrees while also reducing the time it takes to train CNN models. Even though the algorithm in this research has made significant progress in processing FER in non-simple backgrounds, many gathered photos are not.

A. Song et al. established a strong face recognition approach that incorporates both internal and exterior facial characteristics. Their strategy assures that face matching accuracy is increased while also achieving a fine-grained recognition effect by merging the exterior and interior features. The solution increases the traditional face recognition model's face feature extraction efficacy while preserving the benefits of the original deep CNN model, which addresses the recognition accuracy problem produced by traditional methods when there are extremely identical face photos [9]. The major disadvantage of adding internal and external characteristics to aid fine-grained face recognition is that it adds more parameters to the model and makes training more challenging.

X. Sun et al. provide an innovative method for recognizing facial expressions. To aid the model learning, the authors incorporate three ROIs in the face: the mouth, nose, and eyes. They employ the attention mechanism in particular by calculating the distance connection between the feature points and their designated ROIs to determine the enhancement coefficient. The first layer of the model can extract more helpful features for face expression identification after the suggested features matrix. After the feature is recovered, the next layer applies CapsNet's vector technique in the second portion of the model to achieve the fusion of the semantic levels of the three components of interest. This technique is capable of retaining all feature information [10].

III. Proposed Methodology

The presented methodology for the automatic attendance management through facial recognition is elaborated with the stepwise process described below. The figure 1 above displays the system overview for the methodology.

- **Step 1: Preprocessing** - This is the first step of the approach where the student's facial images are captured through the use of the OpenCv library. The images of the particular student are captured through the Video Capture functionality of the cv2 library. The captured image is subjected to the facial region identification through the use of the Haar Cascade Classifier which is stored in the form of an xml file. Once the facial region in the image is identified, it is cropped out. The cropped facial image is then converted into grayscale using the cvtColor method. The grayscale cropped image is then resized to a size of 48x48 and then stored in a folder specific for that students. This procedure is performed extensively for all the students which results in the realization of the input training dataset.

- **Step 2: Image Segmentation** – The captured images need to be effectively utilized for the training purposes for which the training generator and the validation generator is being deployed. The training generator first initiates the target size of the image as 48x48, with the batch size as 64 and the grayscale as the color mode and the categorical class mode. The validation generator is also generated with the various attributes, such as the target size of image as 48x48, along with the use of 64 as the batch size and the color model set as grayscale with the class mode as categorical.

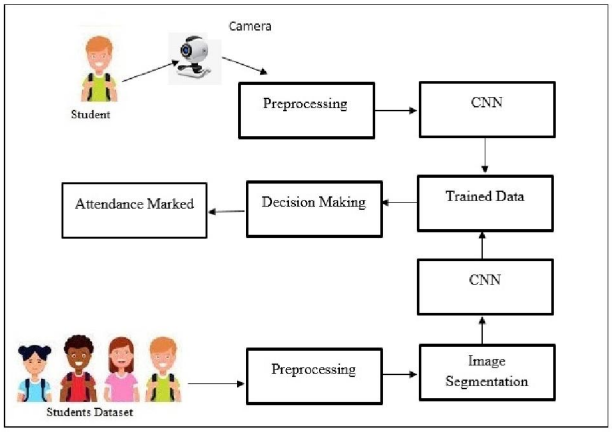

Figure 1: System Overview

- **Step 3: Convolution Neural Network –**This is the primary component of the proposed approach, and it is responsible for detecting and recognizing facial characteristics. The original picture are used as inputs in the convolutional Neural Network module. The input photos gathered, preprocessed, and segmented in the previous phases of the approach are used to train this model.

The training and testing image folders make up the input dataset. Each of the folders is then subdivided into distinct directories for each student's facial photos. These photos are fed into the CNN model as part of the training process. The supplied photos should be first downsized to a width and height of 48 x 48 pixels. On these photos, the model is trained for 500 epochs with 64 as the batch size with the dense as 7 as we are using 7 students for our demonstration. In the python environment, the TensorFlow and Keras libraries are utilized to enable the individual elements of the CNN model. The architecture is depicted in the diagram 2 below.

Layer	Activation
CONV 2D 32 X 3 X 3	Relu
CONV 2D 64 X 3 X 3	Relu
MaxPooling2D 2 X 2	
Dropout 0.25	
CONV 2D 128 X 3 X 3	Relu
MaxPooling2D 2 X 2	
CONV 2D 128 X 3 X 3	Relu
MaxPooling2D 2 X 2	
Dropout 0.25	
Flatten	
Dense 1024	Relu
Dropout 0.25	
Dense 7	Softmax
Adam Optimizer	

Figure 2: CNN network Architecture

The CNN model obtained with this architecture is then execute for 500 epochs to achieve a trained model file with an extension as .h5.

- **Step 4: Decision Making –** Once the trained model through CNN is achieved, the approach now can be tested for its facial recognition for the registration of the attendance. Through the use of the OpenCV platform the camera is initiated to get the face of the student cropped through the use of Haar Cascade. This cropped mage is converted into grayscale and resized to evaluate with the data of .h5 file. This process yields the matched student name with the dictionary, this is then sorted with all the matched faces at that instance. If the obtained count crosses a threshold value, then the student is marked with his attendance for the subject along with the date and time.

IV. Results and Discussion

The proposed methodology for achieving automatic attendance system through facial recognition has been deployed on python programming language using the Spyder IDE. The approach utilizes the Keras, TensorFlow and OpenCV libraries to achieve the desired goals. For the evaluation purposes the proposed model is trained for 50 epochs on 10 students and the outcomes for the same consisting of accuracy and model loss are provided in the figure 3 and 4 given below.

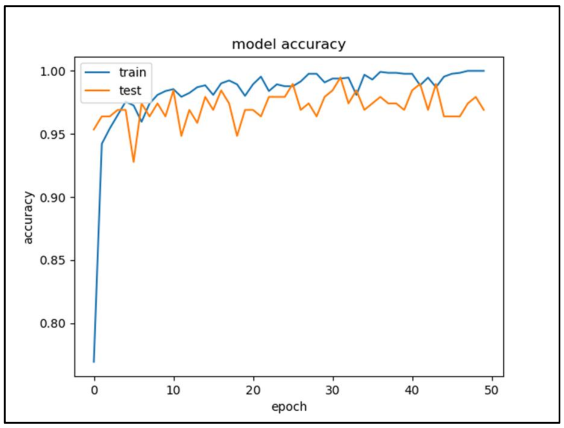

Figure 3: Model Accuracy

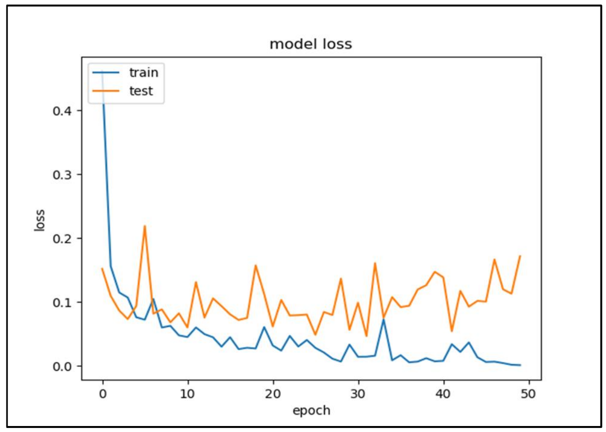

Figure 4: Model Loss

V. Conclusion And Future Scope

The presented approach for the automatic attendance through facial recognition has been proposed in this research article. The approach utilizes the Keras and TensorFlow libraries for python for the deployment of the neural networks. The approach is initiated with the collection of the student's facial images for the training of the neural network. The students facial images are collected and stored in specific directories based on their name. These images are effectively preprocessed through the deployment of the Haar Cascade approach effectively identifies the facial region in the image which is effectively cropped and the image is converted into grayscale. This image is then resized to 48x48 and stored in the respective folder as the training data. The images are then utilized for the training purposes to train the Convolutional Neural Networks. The CNN approach is defined to train the model to achieve a .h5 file. This file is then utilized to perform the facial recognition of the students. The testing is performed by initiating the video stream to capture the facial frames of the student. Once the facial images are detected the frames are captured and subjected to preprocessing through facial region detection through Haar Cascade, after which the image is cropped, converted to grayscale and resized. This image is then subjected to the .h5 model file to perform the identification. The Decision making approach performs this distinction based on a threshold value which is then utilized for the purpose of marking the attendance and storing it in the database. The proposed approach is evaluated for the model accuracy and model loss, which has achieved highly precise outcomes.

References

[1]. M. Z. Uddin, W. Khaksar and J. Torresen, "Facial Expression Recognition Using Salient Features and Convolutional Neural Network," in IEEE Access, vol. 5, pp. 26146-26161, 2017, DOI: 10.1109/ACCESS.2017.2777003.

[2]. B. Yang, J. Cao, R. Ni, and Y. Zhang, "Facial Expression Recognition Using Weighted Mixture Deep Neural Network Based on Double-Channel Facial Images," in IEEE Access, vol. 6, pp. 4630-4640, 2018, DOI: 10.1109/ACCESS.2017.2784096.

[3]. Q. Shi, W. Li, F. Zhang, W. Hu, X. Sun and L. Gao, "Deep CNN With Multi-Scale Rotation Invariance Features for Ship Classification," in IEEE Access, vol. 6, pp. 38656-38668, 2018, DOI: 10.1109/ACCESS.2018.2853620.

[4]. H. Wu, K. Zhang, and G. Tian, "Simultaneous Face Detection and Pose Estimation Using Convolutional Neural Network Cascade," in IEEE Access, vol. 6, pp. 49563-49575, 2018, DOI: 10.1109/ACCESS.2018.2869465.

[5]. G. Storey, A. Bouridane and R. Jiang, "Integrated Deep Model for Face Detection and Landmark Localization From “In The Wild” Images," in IEEE Access, vol. 6, pp. 74442-74452, 2018, DOI: 10.1109/ACCESS.2018.2882227.

[6]. Z. Wu et al., "Studies on Different CNN Algorithms for Face Skin Disease Classification Based on Clinical Images," in IEEE Access, vol. 7, pp. 66505-66511, 2019, DOI: 10.1109/ACCESS.2019.2918221.

[7]. M. Z. Khan, S. Harous, S. U. Hassan, M. U. Ghani Khan, R. Iqbal, and S. Mumtaz, "Deep Unified Model For Face Recognition Based on Convolution Neural Network and Edge Computing," in IEEE Access, vol. 7, pp. 72622-72633, 2019, DOI: 10.1109/ACCESS.2019.2918275.

[8]. M. Shi, L. Xu, and X. Chen, "A Novel Facial Expression Intelligent Recognition Method Using Improved Convolutional Neural Network," in IEEE Access, vol. 8, pp. 57606-57614, 2020, DOI: 10.1109/ACCESS.2020.2982286.

[9]. A. Song, Q. Hu, X. Ding, X. Di, and Z. Song, "Similar Face Recognition Using the IE-CNN Model," in IEEE Access, vol. 8, pp. 45244-45253, 2020, DOI: 10.1109/ACCESS.2020.2978938.

[10]. X. Sun, S. Zheng and H. Fu, "ROI-Attention Vectorized CNN Model for Static Facial Expression Recognition," in IEEE Access, vol. 8, pp. 7183-7194, 2020, DOI: 10.1109/ACCESS.2020.2964298.