Automatic Attendance System using Deep Learning

Sanskar Patel, *Student, IIITV-ICD,* Ashish Gupta, *Student, IIITV-ICD,* Raj Tejaswee, *Student, IIITV-ICD,*

***Abstract*—This paper presents a new automatic attendance system using machine learning and deep learning algorithms for real-time face recognition. The system is integrated with a university management system and is designed to save time for both students and teachers by eliminating the need for manual attendance taking. The Deep Learning Assisted Attendance System (DPAAS) uses a stream of pictures from a video device in the classroom to keep track of student attendance. The DPAAS method overcomes the problems present in traditional attendance systems, such as multi-person identification, occlusion, and different lighting conditions. The proposed system has been tested and found to have an accuracy of 94.66%, which is higher than existing methods and does not require any interaction from the user.**

***Index Terms*—automatic attendance system, deep learning, real-time face detection, multi-person identification, deep learning algorithms**

I. Introduction

Attendance taking is an important part of the educational process, as it helps to keep track of student participation and engagement in class. However, traditional methods of manual attendance taking can be time-consuming and prone to errors, such as proxy attendance and incorrect record keeping. In order to address these issues, there has been a growing interest in the use of automatic attendance systems, which use advanced technologies such as machine learning and deep learning to improve accuracy and efficiency.

This paper focuses on the development of an automatic attendance system using deep learning algorithms. The system is designed to detect the faces of students in real-time and integrate with an existing university management system. The use of deep learning algorithms enables the system to overcome the challenges faced by traditional attendance systems, such as multi-person identification and varying lighting conditions. The proposed system is expected to provide a more accurate and efficient solution to the problem of attendance taking, while also freeing up valuable time for both students and teachers.

The aim of this research is to explore the potential of deep learning algorithms for the development of an effective automatic attendance system, and to evaluate its performance in a real-world scenario. The results of this study will be of interest to educators, researchers, and practitioners in the field of automatic attendance systems and related technologies.

II. Related Work

Deep Learning is a subfield of Machine Learning where data passes through multiple layers of non-linear mappings to allow the algorithm to learn a representation. Deep Learning Models are capable of learning almost any function based on the input data. One of the challenging and widely researched topics in Deep Learning is Face Detection and Face Recognition. The field has been hindered by the lack of a clean and labeled training dataset. However, with the advent of Graphical Processing Units like Nvidia Graphics cards, which enable fast matrix operations, the state-of-the-art solutions in this domain are constantly being improved. Before the rise of Deep Learning, Face Recognition relied on techniques like Viola Jones, Clustering, and HOG methods, which were limited by factors such as noise in the data and occlusion. Convolutional Neural Networks have now revolutionized Image and Vision Computing in Deep Learning. By using Data Augmentation techniques during training, Deep Learning can combat data noise and incorrect predictions caused by occlusion and luminosity variance.

Face Recognition is a two-part problem, consisting of detecting a face in an image and correctly categorizing the image. This can be seen as an Object Detection and Object Classification problem. Object Detection frameworks can be viewed as an issue of image localization, where the neural network needs to determine where in the image the object is located and predict its class with confidence. There are two main approaches to solving this problem, Branch Networks and Single Network. Branch Networks use multiple Neural Network architectures to perform different subtasks, while Single Network uses a single architecture. In the context of Object Detection networks, one branch of the network is used to determine the most promising region in the image that might contain the desired object, while the other branch is used to classify the detected object.

This paper introduces a new approach to the automatic attendance system using deep learning. Unlike the traditional branch network architecture, the proposed method uses a single network to achieve faster and more accurate results even in conditions with varying lighting and partial occlusion. This solution overcomes the difficulties that branch networks face in reaching a near-optimal solution. The proposed DPAAS (Deep Learning Based Automatic Attendance System) is capable of accurately identifying multiple objects in various conditions.

III. Implementation

A. Dataset

The dataset of Automatic Attendance System using Deep Learning contains the close face images of students. Dataset may contains the side view or front view of face. The dataset will be used to facial detection for pre-trained model, which can automatically identify and track the attendance of individual students based on their images. The training images are clear and contains clearly visible faces So that, The model can detect face easily. There is only one image per student in dataset.

B. Model

The model is trained based on dataset, contains the images of students. We are using Webcam of device to detect the face of student and give output of student ID. We are training the model by two algorithms. First algorithm deals with the problem that faces turned different directions look totally different to a computer. Second algorithm matches the input image with all the images of dataset with certain parameters.

*1) **Face Landmark Estimation**:* In the dataset, we have different looks of images like some are close views, some are side views etc. So, It may cause problem in detection because some of the parameters may be null. Ex. If any input image is side view, then the values of parameters of 2nd eye will be null. Now, If the input image of same person have some values in eye parameters, then it will cause problem in detection.

We are using the **face landmark estimation** algorithm to address the aforementioned problem. Face landmark estimation is the process of detecting and locating key points on a human face, such as the corners of the eyes, the tip of the nose, and the corners of the mouth. This is typically achieved using computer vision algorithms and machine learning techniques. The resulting set of facial landmarks can be used for a variety of applications, including face detection, facial expression analysis, and virtual makeup or accessory try-on.

Essentially, we aim to identify 68 specific points or "landmarks" on each face, such as the top of the chin, the outer edge of each eye, and the inner edge of each eyebrow. To achieve this, we will train a machine learning algorithm to detect these landmarks on any face. Once these landmarks are identified, we will adjust the image by rotating, scaling, and shearing it to ensure that the eyes and mouth are as centered as possible. We will only use basic image transformations that preserve parallel lines, such as rotation and scaling. This will allow us to center the eyes and mouth in a consistent position, regardless of how the face is turned, thereby improving the accuracy of the system.

Face detection estimation uses BlazeFace architecture. Google has created a machine learning model known as BlazeFace that can quickly and accurately identify the location and key points of human faces. This model can simultaneously detect the position of the face and the six crucial facial features, including eyes, nose, ears, and mouth. Additionally, BlazeFace is capable of detecting multiple faces within a single image. The benefits of BlazeFace extend beyond its speed and accuracy. The model's ability to detect multiple faces simultaneously makes it useful for applications such as group photography or video conferencing. The key point detection feature is also beneficial for various facial analysis applications, including emotion recognition and facial expression tracking.

*2) **Face Encoding**:* The most straightforward way of face detection involves comparing an unknown face with all the tagged pictures of individuals. By extracting basic measurements from each face, we can compare the measurements of the unknown face with the tagged faces and find the one with the closest measurements. Though it may seem simple, it requires careful consideration of the parameters used to compare the faces. We need to extract a few basic measurements from each face. Then we could measure our unknown face the same way and find the known face with the closest measurements. The examples of parameters are the size of each ear, the spacing between the eyes, the length of the nose, etc.

In order to generate 128 measurements for each face, the system employs a training process that involves examining sets of three face images simultaneously. The first two images in each set are of the same person, while the third image is of a different individual. This training process involves loading a training image of a known person, followed by another image of the same person, and then loading a third image of an entirely different person.

The process of generating 128 measurements for each individual's face is known as face encoding, and it is a critical step in the facial recognition process. These measurements are unique to each person and capture specific facial features and characteristics that distinguish them from others. By analyzing these features, the neural network can determine whether two faces belong to the same individual or not.

The encoding process used by the neural network is highly efficient and accurate, making it ideal for a range of applications. For example, in the security industry, facial recognition technology can be used to verify the identity of individuals entering a restricted area or accessing sensitive information. Similarly, in the retail industry, facial recognition can be used to identify repeat customers and provide them with personalized experiences.

Moreover, the facial recognition technology based on face encoding has the potential to improve safety in a variety of settings. For instance, it can be used to locate missing individuals, monitor public spaces for suspicious activity, and assist in law enforcement investigations. Despite the potential benefits, there are also concerns regarding privacy and security when it comes to facial recognition technology. Therefore, it is important to ensure that the use of such technology is regulated and transparent, with appropriate measures in place to protect individuals' rights and prevent abuse of the technology.

In conclusion, the process of encoding facial features to generate 128 unique measurements for each individual is a crucial step in the facial recognition process. With its high accuracy and efficiency, this technology has a wide range of potential applications, but it is important to ensure that it is used responsibly and ethically to protect individuals' rights and privacy.

*3) **Finding the student from the encoding**:* Once the system has generated the 128 measurements for a given face, the next step is to compare these measurements to the set of known individuals in the system's database. This involves finding the person in the database who has the closest measurements to the test image. To accomplish this, the system uses a linear SVM classifier to compare the output of encoding with the known individuals in the database. In this case, the SVM classifier is used to find the known person whose set of 128 measurements is closest to the test image's set of measurements. If the closest match found by the SVM classifier is above a certain threshold, the system will output the Student ID of the matched student. However, if the closest match is below the threshold, it means

that no image in the database is a close enough match to the test image. In this case, the system will output "Unregistered" to indicate that the individual is not recognized in the system's database.

The facial recognition system utilizes the closest distance method to determine the identity of a detected face. This method involves calculating the closest measurements of the detected face parameters to the stored dataset face parameters. By comparing these measurements, the system can determine whether the detected face matches a face in the dataset. To achieve this, the system applies a threshold value to discard measurements that are greater than a certain threshold. This helps to ensure that only the most accurate measurements are used in the comparison process, thereby improving the accuracy of the facial recognition process. Once the comparison process is complete, the system returns true for the smallest measurement value, indicating a match with a face in the dataset. It is important to note that while the closest distance method is highly accurate, it is not infallible. There are instances where facial recognition technology may fail, such as when dealing with low-quality images or in scenarios where the lighting conditions are poor.

Overall, this process allows the system to accurately and efficiently recognize individuals based on their facial characteristics, making it useful for a variety of applications such as security, access control, and identity verification.

*4) **Marking the attendance**:* The system is designed to detect a face and match it with the dataset of registered students. To mark attendance using the attendance system, a student must place their face within the webcam frame. Once the system detects a face, a counter will begin counting. During the first 10 counts, the system will focus on detecting the face, while from counts 10 to 20, the system will compare the detected face with the stored dataset of faces. When a match is found, the system will record the attendance for the identified student on that particular day in the database. The system will then display the student's name, photo, student ID, batch, and the total number of attendance records recorded for that particular student on the screen. Finally, the system will display a message indicating that the attendance has been marked successfully.

To register a new student in the attendance system, the student needs to complete a registration form which includes their name, batch, student ID, and a real-time image captured from a webcam. This information is stored in a database for future reference. The database contains real-time data related to registered students, including their name, student ID, batch, attendance percentage, total attendance, and total number of lectures attended. This information is regularly updated based on the attendance records captured by the system. To provide an accurate representation of attendance records, the system also records the date and time of the last recorded attendance for each student. This allows for the tracking of attendance patterns over time and can help to identify any potential issues or concerns related to attendance. Overall, the database plays a crucial role in the functioning of the facial recognition system, providing a centralized location for storing and accessing important data related to registered students. By keeping accurate records of attendance and other relevant information, the system can help to improve the efficiency and effectiveness of attendance tracking, while also providing valuable insights into student attendance patterns and trends.

However, there may be instances where the attendance marking process fails, which can lead to discrepancies in attendance records. When a student successfully marks their attendance and the corresponding data is updated in the database, the facial recognition system will display a message indicating that the attendance has been marked successfully. This message serves as confirmation to the student and helps to ensure that attendance records are accurate and up-to-date. This allows the user to take corrective action in case of any errors or issues with the attendance marking process. Another potential issue is the possibility of multiple attendance records being marked for the same person on the same day. In order to prevent students from marking their attendance multiple times, the facial recognition system is designed to only accept attendance records after a certain period of time has elapsed since the last recorded attendance. Specifically, if a student attempts to mark their attendance less than 12 hours after their last recorded attendance time, the system will not accept the record and will display a message indicating that the attendance has already been marked. This feature ensures that attendance records are accurate and prevents any potential issues or errors that could arise from multiple attendance records being marked for the same student.

Overall, the facial recognition system provides an efficient and reliable way to mark attendance in a classroom setting. By automating the attendance marking process, the system saves time and reduces the potential for errors or inaccuracies in attendance records. However, it is important to ensure that the system is properly maintained and calibrated to ensure accurate attendance marking and prevent any potential issues or errors.

IV. References

1) J Joshan Athanesious, Vanitha, S. Adithya, C. Anirudh Bhardwaj, Jugat Singh Lamba, A.V. Vaidehi, Deep Learning Based Automated Attendance System, Procedia Computer Science, Volume 165, 2019, Pages 307-313, ISSN 1877-0509.
2) https://towardsdatascience.com/face-landmark-detection-using-python-1964cb620837