Deep Learning based Face Recognition Model
Final Report

Ramprakaash Kannan (rkanna14)
Rahavee Prabakaran (rprabak1)
CSE 598 – Statistical Learning Theory (2023 Spring)

I Introduction

Computer vision has come a long way in recent years, with advancements in deep learning and neural networks fundamentally changing how machines can interact with the world around them. From military applications to everyday convenience markets, computer vision is now being used more frequently than ever before.

At its core, computer vision is the ability of a computer to process visual input in a way that is similar to how humans process information. By breaking down images into individual components and analyzing them, machines can recognize objects, identify patterns, and make decisions based on visual data. This ability has a wide range of practical applications, from autonomous driving to medical imaging.

The advancement in computer vision has been largely influenced by the emergence of deep learning, which is a type of machine learning that employs artificial neural networks to detect patterns in information and generate predictions. These neural networks mimic the architecture of the human brain, featuring interconnected layers of neurons that analyze data and enable the network to learn and enhance its performance through experience.

Convolutional Neural Networks are popularly well-suited for image recognition and image processing applications. These networks are made up of several layers, each of which processes the image data in a different way. The first layer typically detects basic features like edges and corners, while later layers identify more complex patterns like shapes and textures. The final layer produces an output that indicates what object or objects are present in the image.

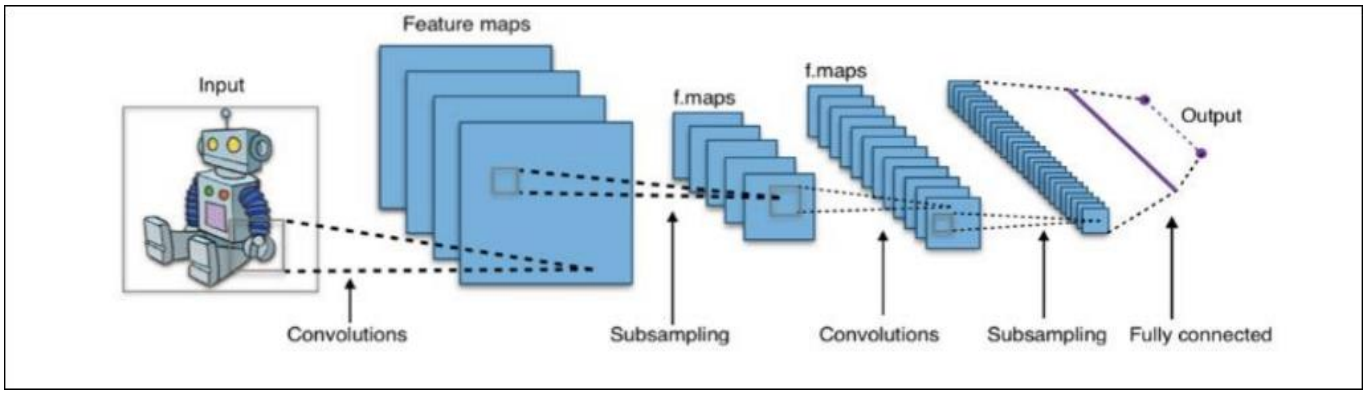
The main advantage of CNNs is their capability to reduce the dimensionality of image data without losing important features. This is achieved through a process called pooling, which aggregates information from adjacent pixels in the image to produce a smaller, more manageable representation of the data. By doing this, CNNs can process large amounts of visual data efficiently and accurately.

A large dataset of labeled images is typically used to train a CNN model. This dataset is used to "teach" the network what different objects look like, so that it can recognize them in new images. However, it is important to avoid overfitting, where the network becomes understand the data too well is nor able to generalize to new data. To prevent this, techniques like data augmentation and regularization are used to ensure that the network learns to recognize objects in a more general way.

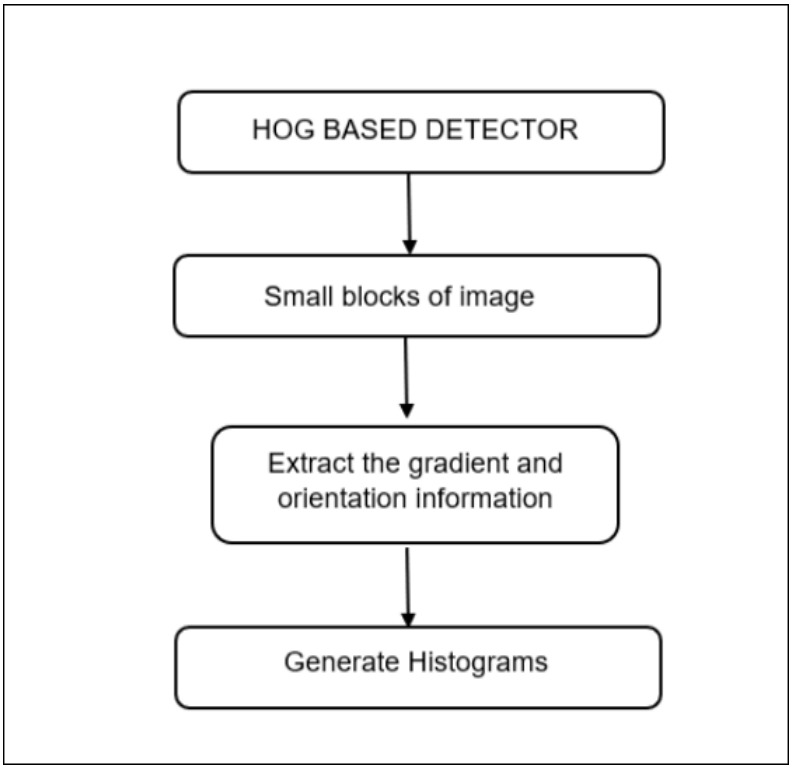
In the context of attendance tracking, CNNs can be used to identify visitors in a given area. By analyzing a video feed from a camera, the network can detect and recognize faces, and match them to a database of known individuals. This can be particularly useful in situations where it is important to track who is entering or leaving a particular space, such as in a workplace or school.

CV2 is a popular Python library that facilitates the implementation of computer vision techniques. It offers a broad range of features for image analysis and processing, and it supports various machine learning models. Notably, the library can assist with Convolutional Neural Networks (CNNs) and the Histogram of Oriented Gradients (HOG) algorithm to recognize images.

HOG is a feature descriptor that operates differently from CNNs. Rather than initializing data with weights like a neural network, HOG recovers image features by examining the exterior structure of the image. It computes the gradient and edge orientation while taking the image's structure into account. The image is divided into smaller chunks, and edge information is computed for each region. Histograms are then constructed for each of these smaller regions, and the image recognition process continues from there.



While HOG is not as widely used as CNNs in modern computer vision applications, it can still be useful in certain contexts. In particular, it can be effective for object detection in images where the object of interest has a distinctive texture or shape. For example, HOG has been used in applications like pedestrian detection



II Problem Description

Advancements in telecommunications technology have revolutionized various fields, and audio and video processing are no exception. In recent years, these technologies have become increasingly important for security and business applications. One of the most distinctive features of a human being is their face, making it an ideal identifier for various purposes. Due to this characteristic, facial recognition technology has emerged as a popular and favored choice for attendance management.

Traditionally, attendance management systems involved manual entry or scanning of identification cards, which were both time-consuming and prone to errors. This led to the need for more efficient strategies for managing attendance. One such strategy is the use of facial recognition technology, which has been gaining popularity due to its accuracy and convenience. The system involves the use of a face recognition module that scans the faces of students in a classroom and compares them with preset photographs. Each student's attendance is then automatically updated in a database, eliminating the need for manual data entry.

To recognize a face, an image of the face is captured and compared against a database of known images to determine the person's identity. Face recognition technology utilizes algorithms that extract various facial features, such as the distance between the eyes, the shape of the nose, and the curvature of the lips. These features are then compared to a database of known faces in order to identify the individual in the captured image.

Facial recognition technology offers multiple advantages over traditional attendance management methods. One of the primary benefits is that it eliminates the requirement for manual data entry, which can be a tedious and error-prone task. Additionally, the use of biometric information in facial recognition technology makes it highly reliable and accurate, as each individual's biometrics are unique to them. This feature reduces the likelihood of fraudulent attendance practices, such as having someone else mark attendance on behalf of a student. Another advantage of the system is its ability to track absences in real-time, which enables teachers to take prompt action such as contacting the student or their parents.

To develop a facial recognition-based attendance management system, a number of steps need to be taken. Initially, images of each student's face are captured by a camera. Subsequently, the captured images undergo processing to extract distinctive facial features, including the distance between the eyes and the shape of the nose. The extracted features are then matched against a pre-existing database of known faces in order to accurately identify the student. The system is trained to recognize the faces of all the students present in the classroom, ensuring precise attendance tracking.

To ensure the accuracy and reliability of the system, it is important to address potential issues such as lighting conditions, pose variations, and occlusions. Adequate lighting is essential to capture clear images of the face, while pose variations and occlusions such as glasses or hats can affect the accuracy of the system. These issues can be addressed by using advanced algorithms that are robust to such variations.

The use of facial recognition technology for attendance management is a novel and efficient strategy that has gained popularity in recent years. It offers several advantages over traditional methods such as accuracy, reliability, and convenience. Implementing such a system requires careful consideration of various factors such as lighting conditions, pose variations, and occlusions. With the advancements in telecommunications technology, facial recognition-based attendance management has very high scope for the future advancements.

III METHODOLOGY

The implementation of a proposed attendance system involves the use of facial recognition technology, specifically a Convolutional Neural Network. CNNs are a type of neural network designed to handle large image data and have different structures and properties than Artificial Neural Networks (ANNs). Both ANNs and CNNs have hidden layers, weights, biases, hidden neurons, activation functions, and utilize backpropagation. However, CNNs are more efficient in processing images because they can automatically recognize features for image/object classification.

To create an attendance system, the first step is to take a photograph of all the students in the class, which will be used to recognize the present students. The camera is placed at the center of the classroom to ensure that every individual is clearly visible in the picture. Once the photograph is taken, it is saved on a local server and analyzed with the help of a Python script.

In the proposed CNN, a deep learning approach is employed, comprising several layers including input, convolutional, max pooling, hidden, and output layers. This approach is chosen because traditional machine learning methods face challenges in handling high-dimensional data with large input and output sizes, which can lead to increased complexity and processing time. Therefore, deep learning is utilized as it provides a more efficient and effective solution.

The CNN is not limited to one convolutional layer and can include multiple such layers. The convolutional layer captures low-level features such as color and edges, while the max pooling layer reduces the size of convolved features, extracting dominant features that can be used for training the network. This process reduces computational time and improves the effectiveness of the system.

To create a convolutional network, the convolutional layer and pooling layer are combined. To improve the detection of low-level features while keeping computational demands in check, the number of convolutional layers can be increased. After the convolutional layer applies pooling to the input image, it is prepared for processing by a multilayer perceptron.

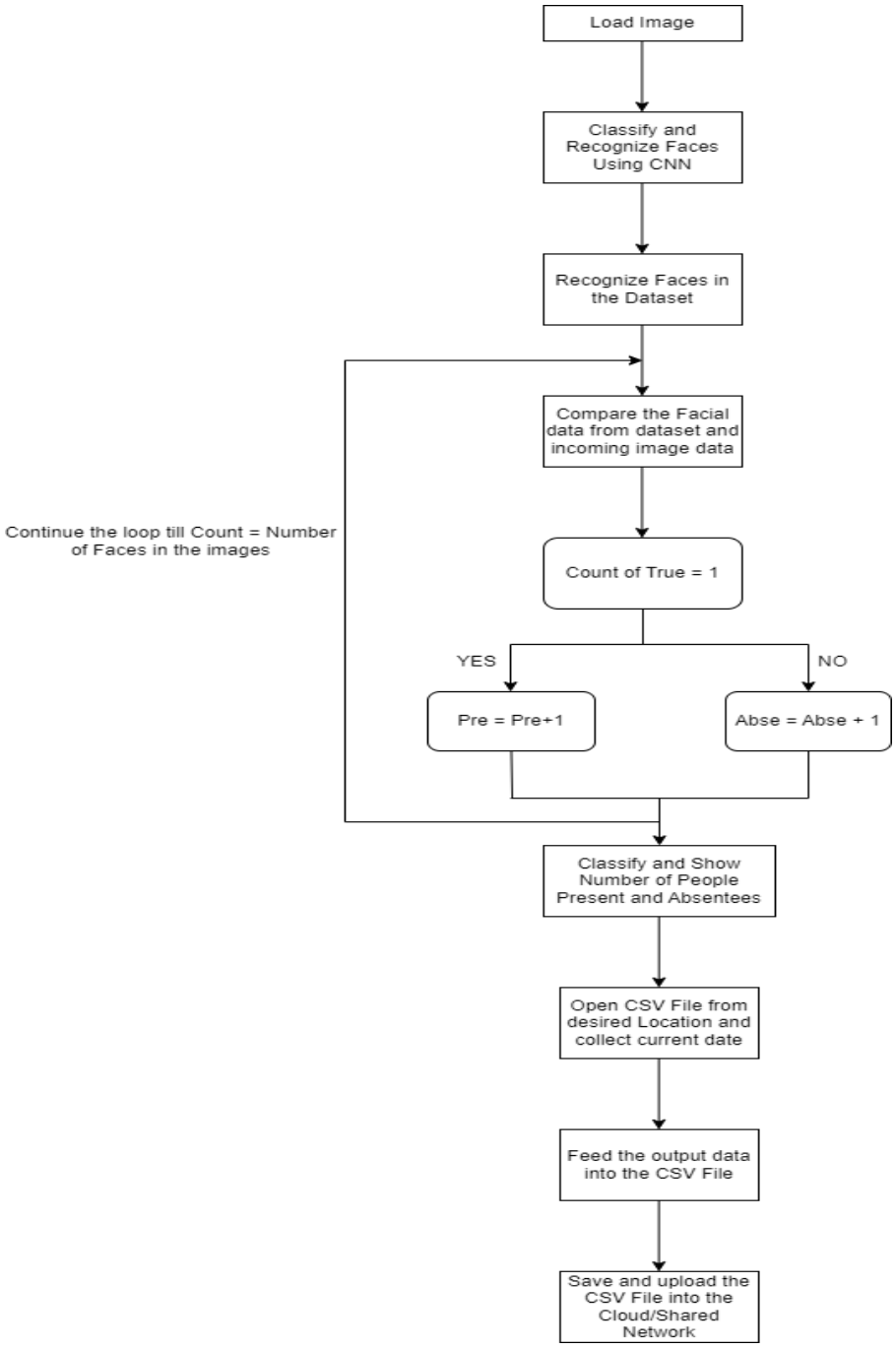
To prepare the pooled feature map for processing by a feedforward neural network with backpropagation, it must first be flattened into a column vector. As the training process progresses, the weights and biases of the network are fine-tuned to achieve optimal performance. This involves reducing the network error through backpropagation, whereby the error is propagated backwards and the weights and biases are adjusted accordingly.

The face recognition library is utilized to compare the faces detected in the image with the images of students in the student database. The attendance is marked for a student based on the result of this comparison. The attendance record is then converted to a CSV file to generate an attendance sheet, which is subsequently uploaded to the cloud for storage of attendance records.

The CNN approach to facial recognition has numerous advantages over traditional attendance management methods. Firstly, the system saves time and effort as it automatically updates the attendance database, eliminating the need for manual attendance tracking. Secondly, the system is more accurate in identifying students compared to manual attendance tracking, which can be prone to errors due to human factors such as confusion or incorrect recording. Finally, the system provides real-time attendance tracking, enabling teachers to immediately identify absent students and take appropriate action.

However, there are also concerns with facial recognition technology, including privacy concerns, bias in the recognition algorithm, and the potential for misuse of the technology. Privacy concerns arise due to the system capturing images of individuals without their explicit consent, and the possibility of these

images being used for other purposes. Bias in the recognition algorithm is another concern, as the system may have difficulty recognizing individuals with certain physical attributes such as skin color or facial features. Finally, the potential for misuse of the technology is a significant concern, with the possibility of the technology being used for nefarious purposes such as tracking individuals without their knowledge or consent.



The proposed smart attendance management system utilizing facial recognition technology has the potential to revolutionize attendance tracking in educational institutions and businesses. However, it is important to address concerns regarding privacy, bias, and potential misuse of the technology. Proper regulations and ethical considerations should be implemented to ensure the responsible use of facial recognition technology.

IV RESULTS

The proposed system utilizes a Convolutional Neural Network (CNN) to detect and recognize the faces of students in a classroom.

Figure 1 shows the pictures of student faces that are captured by a camera placed in the center of the classroom. These images are then processed through a Python script to analyze the images and extract key features. CNNs are designed to handle large image data and have specific structures and properties compared to Artificial Neural Networks (ANNs). These structures and properties include hidden layers, weights, biases, hidden neurons, activation functions, and backpropagation.

The CNN used in the proposed system includes an input layer, convolutional layer, max pooling layer, hidden layer, and output layer. Deep learning is used in this context because traditional machine learning approaches struggle to handle high-dimensional data with large input and output sizes. Deep learning models automatically extract features for image/object classification, making it a more efficient approach for facial recognition in attendance systems.

The convolutional layer is responsible for capturing low-level features such as color and edges, while the max pooling layer is used to reduce the size of convolved features after convolution on the images. This process extracts dominant features that can be used for training the network, reducing computational time and improving effectiveness. The CNN can have multiple convolutional layers to better capture low-level features while minimizing computational power.

Once the convolutional layer has performed pooling on the input image, the resulting feature map is transformed into a suitable format for the multilayer perceptron. This is done by flattening the pooled feature map into a column vector and passing it through a feedforward neural network that uses backpropagation. To achieve optimal results, the weights and biases of the network are adjusted during the training process. Backpropagation is utilized to reduce network error by propagating the error in reverse and adjusting the weights and biases accordingly.

Figure 2 shows how the faces detected in the images are compared to the images of students in the student database to determine each student's attendance status. The face recognition library is utilized to make this comparison, and the attendance is marked for a student based on the result of this comparison. Figure 3 and 4 shows the attendance record is then converted to a CSV file to generate an attendance sheet, which is subsequently uploaded to the cloud for storage of attendance records.

The proposed system has a face recognition rate of 93% to 95%, making it incredibly efficient and the best choice for usage in sophisticated attendance tracking systems. However, it is important to note that facial recognition technology is not perfect and can have accuracy issues. False positives and false negatives can occur, and the system may have difficulty recognizing faces in poor lighting or if the individual's face is partially obstructed.

By leveraging facial recognition technology, the attendance management system presented here offers a precise and efficient way of tracking attendance. Deep learning and a convolutional neural network (CNN) are employed to achieve accurate facial recognition, while the system's ability to transfer attendance records to the cloud provides a secure and hassle-free method of accessing attendance data. Nevertheless, it is crucial to acknowledge the potential drawbacks of facial recognition technology and implement necessary measures to mitigate any associated risks.

Figure 1

Figure 2

```
/content/gdrive/My Drive/Attendance/student_images/18225760007_7.jpg is absent
/content/gdrive/My Drive/Attendance/student_images/18225760013_12.jpg is Present
/content/gdrive/My Drive/Attendance/student_images/18225760013_11.jpg is Present
/content/gdrive/My Drive/Attendance/student_images/18225760013_10.jpg is absent
/content/gdrive/My Drive/Attendance/student_images/18225760013_8.jpg is Present
/content/gdrive/My Drive/Attendance/student_images/18225760011_9.jpg is absent
/content/gdrive/My Drive/Attendance/student_images/18225760001_1.jpg is Present
/content/gdrive/My Drive/Attendance/student_images/18225760002_2.jpg is Present
/content/gdrive/My Drive/Attendance/student_images/18225760003_3.jpg is Present
/content/gdrive/My Drive/Attendance/student_images/18225760004_4.jpg is absent
/content/gdrive/My Drive/Attendance/student_images/18225760011_5.jpg is absent
/content/gdrive/My Drive/Attendance/student_images/18225760007_6.jpg is absent
Total Number of Students Present =  6
Total Number of Students Absent =  6
```

Figure 3

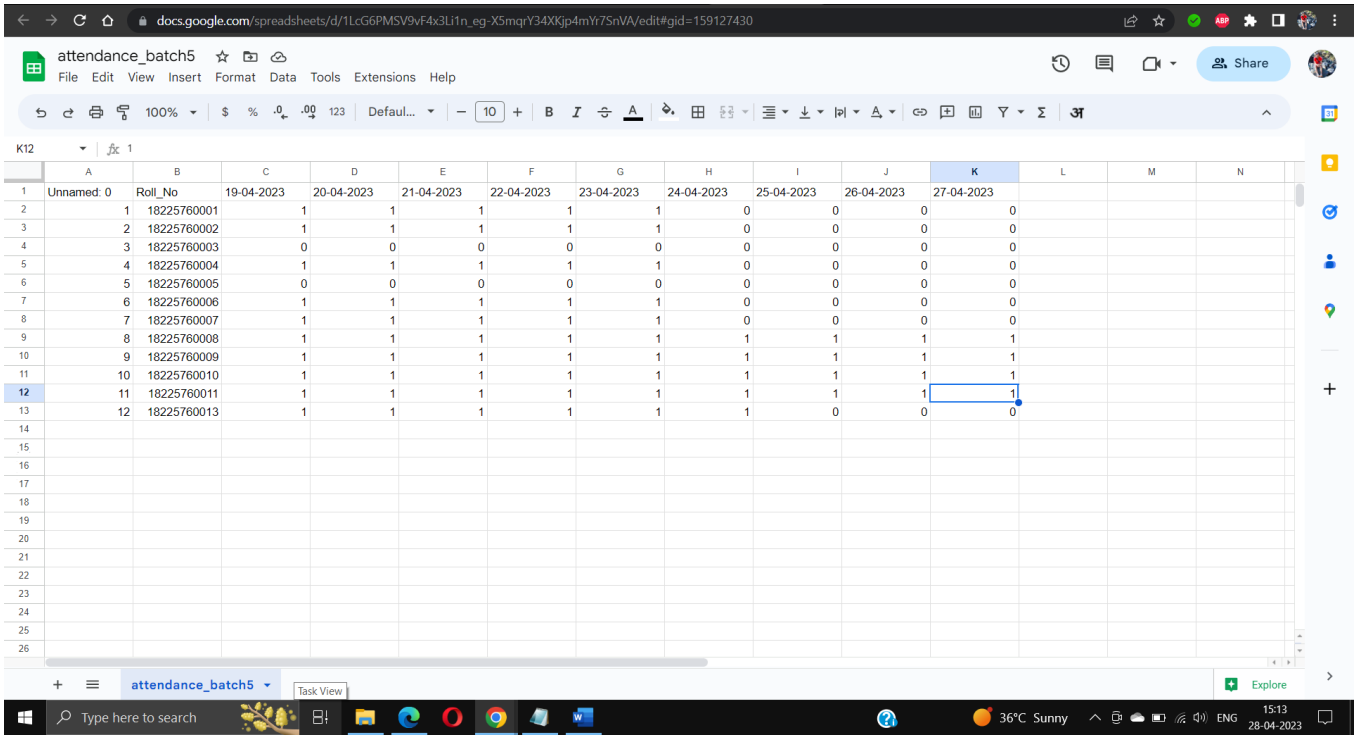


Figure 4

V Conclusion and Future Work

Attendance management is an essential aspect of any educational institution. It not only helps in keeping track of students' attendance but also enables educational institutions to maintain accurate data that can be used for various purposes such as grading, disciplinary actions, and monitoring student progress. However, the traditional method of attendance management has been proven to be time-consuming and prone to errors. Therefore, there is a need for a smarter and more efficient attendance management system that can eliminate these drawbacks.

The suggested attendance management system aims to overcome the limitations of the conventional approach. It offers a budget-friendly alternative that prevents any falsification of data and guarantees more precise and reliable attendance logs. The key goal of the proposed system is to simplify the attendance management procedure, resulting in time and resource savings.

The proposed attendance management system is easy to use and can be implemented in classrooms using LAN or the Cloud. At predetermined intervals, a trigger prompts the camera to take a picture of the class. The system utilizes facial recognition technology to identify students and mark their attendance. The system's facial recognition technology is highly accurate, with a recognition rate of 93% to 95%, making it a highly efficient and reliable method of managing attendance.

In addition to monitoring attendance, the proposed system also includes a teacher presence indicator. This feature allows the system to track the teacher's presence in the classroom, ensuring that teachers are present and engaged during class sessions. Moreover, the system has been improved to track student behavior more effectively. This enhancement makes it possible to detect fraudulent activities such as paper chasing during exams, ensuring the integrity of the examination process.

The attendance management system under consideration comes with a new attribute called the "academic indicator." It serves the purpose of identifying high-performing students academically. The academic indicator adds value by providing an extra layer of data that enables tracking and monitoring of students' progress. Moreover, the system generates attendance and academic performance reports that teachers and administrators can use to pinpoint areas that need improvement and devise strategies to enhance the learning experience.

Overall, the proposed attendance management system offers a smart and efficient solution to traditional attendance management. It is a cost-effective and reliable method of managing attendance that saves time and resources. The facial recognition technology used in the system ensures high accuracy, making it a trustworthy method of managing attendance. Additionally, the system's additional features such as the teacher presence indicator, behavior tracker, and academic indicator provide additional layers of information that can be used to enhance student learning and progress.

VI REFERENCES

[1] T. Manikandan, J.S. Jayashwanth, S. Harish and K.N. Harshith Sivatej, “Self-driving car”, International journal of psychosocial rehabilitation, vol. 24(5), pp. 380-388, 2020. DOI: 10.37200/IJPR/V24I5/PR201704.

[2] T. Manikandan, V. Nandalal, J.L. Mazher Iqbal, S. Joshua Kumaresan and S. Karthikeyan, “Protection of plantation from wild animals using wireless sensor network”, International journal of psychosocial rehabilitation, vol. 24(5), pp. 2719-2723, 2020. DOI: 10.37200/I JPR/V24I5/PR201974.

[3] T. Manikandan, D. Tamil Selvan, R. Vaidhyanathan, B. Vigneshvaran and V. Nandalal, “Home groceries management system using IOT”, International journal of psychosocial rehabilitation, vol. 24(5), pp. 2731- 2746, 2020. DOI: 10.37200/I JPR/V24I5/PR201977. Fig. 4: Attendance database of students International Journal of Psychosocial Rehabilitation, Vol. 24, Issue 02, 2020 ISSN: 1475-7192

[4] Kouser.R, Manikandan.T and Kumar.V, “Heart disease prediction system using artificial neural network, radial basis function and case based reasoning”, Journal of computational and theoretical nanoscience, vol. 15, pp. 2810-2817, 2018.

[5] H.A Ali Sharif Razavian et. al, "CNN Features off-the-shelf: an Astounding Baseline for Recognition”, 2014. Assessed at https://arxiv.org/pdf/1403.6382.pdf.

[6] T. Manikandan, J.L. Mazher Iqbal, V. Nandalal, A. Muruganandham and S. Joshua Kumaresan, “An automated computer-aided diagnosis system for malignant tumor localization from lung CT images for surgical planning”, International journal of psychosocial rehabilitation, vol. 24(5), pp. 2725-2729, 2020. DOI: 10.37200/I JPR/V24I5/PR201975A.

[7] T. Manikandan and N. Bharathi, “Lung cancer detection using fuzzy auto-seed cluster means morphological segmentation and SVM classifier”, Journal of medical systems, vol. 40(7), pp. 1-9, 2016.

[8] T. Manikandan and N. Bharathi, “A novel semi-automated 3-D CAD visualization system as an aid for surgical planning of lung cancer”, ARPN journal of Engineering and applied sciences, vol. 10(4), pp. 1872-1878, 2015.

[9] T. Manikandan and N. Bharathi, “Lung cancer detection by automatic region growing with morphological masking and neural network classifier’, in Asian journal of information technology, vol. 15 (21), pp. 4189-4194, 2016.

[10] M.D Zeiler and R. Fergus, “Visualizing and Understanding Convolutional Networks”, Lecture Notes in Computer Science, 978-3-319-10590-1, Vol 8689. Springer, Cham. DOI: https://doi.org/10.1007/978-3-319-10590-1_53.

[11] J. Schmidhuber, “Deep Learning in Neural Networks: An Overview”, Neural Networks, Vol 61, pp 85- 117, 2015. DOI: 10.1016/j.neunet.2014.09.003

[12] A. Purohit et. al, “Smart attendance”, proceedings of 2017 International conference of Electronics, Communication and Aerospace Technology, ICECA 2017, Coimbatore, Vol. 1, pp. 415-418, 2017.

[13] O.T Arulogun et. al, “RFID-Based Students Attendance Management System”, International Journal of Scientific and Engineering Research, Vol. 4, Issue 2, 2013.

[14] S. Kardy and M. Smaili, “Wireless attendance management system based on iris recognition”, Scientific research and essays, Vol. 5, issue. 12, pp. 1428-1435, 2010

[15] K. He, X. Zhang, S. Ren, and J. Sun, "Deep Residual Learning for Image Recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016.

[16] A. Krizhevsky, I. Sutskever, and G. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," in Advances in Neural Information Processing Systems (NIPS), 2012.

[17] L. Zhang, L. Wang, X. Liang, and T. Li, "Categorical Regularization with GANs," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2019.

[18] H. Zhao, J. Shi, X. Qi, X. Wang, and J. Jia, "Pyramid Scene Parsing Network," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[19] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image Quality Assessment: From Error Visibility to Structural Similarity," IEEE Transactions on Image Processing, vol. 13, no. 4, pp. 600-612, 2004.

[20] S. Xie, R. Girshick, P. Dollár, Z. Tu, and K. He, "Aggregated Residual Transformations for Deep Neural Networks," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.

[21] T. Hassner, S. Harel, E. Paz and R. Enbar, "Effective Face Frontalization in Unconstrained Images," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[22] F. Schroff, D. Kalenichenko and J. Philbin, "FaceNet: A Unified Embedding for Face Recognition and Clustering," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2015.

[23] H. Zhang, M. Cisse, Y. N. Dauphin and D. Lopez-Paz, "mixup: Beyond Empirical Risk Minimization," in International Conference on Learning Representations (ICLR), 2018.

[24] Y. Wen, K. Zhang, Z. Li and Y. Qiao, "A Discriminative Feature Learning Approach for Deep Face Recognition," in European Conference on Computer Vision (ECCV), 2016.

[25] S. Chopra, R. Hadsell and Y. LeCun, "Learning a Similarity Metric Discriminatively, with Application to Face Verification," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2005.

[26] Y. Sun, X. Wang and X. Tang, "Deep Learning Face Representation by Joint Identification-Verification," in Advances in Neural Information Processing Systems (NIPS), 2014.

[27] 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 (CVPR), 2014.

[28] W. Liu, Y. Wen, Z. Yu, M. Li, B. Raj and L. Song, "SphereFace: Deep Hypersphere Embedding for Face Recognition," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.