Attendance Tracking System Utilizing Facial Recognition

Abstract: Face recognition is an important part of image processing, especially in fields like technology. Recognizing faces is crucial for confirming identities, especially for student attendance. Using biostatistics based on face features, a face recognition attendance system replaces the traditional method of calling names and keeping manual records. Current attendance systems are cumbersome and prone to errors through manual recording, and traditional biometric methods can be tricked. This paper proposes a solution to these issues by employing various technologies such as Haar classifiers, KNN, CNN and SVM. After recognizing faces, attendance reports are generated and stored in Excel. The system is rigorously tested under different conditions like lighting and student-camera distance variations to ensure its accuracy and robustness. Ultimately, the proposed system provides an efficient and reliable way to manage classroom attendance without manual effort and is cost-effective with easy installation.

***Keywords* –** KNN, SVM, CNN, Neural Networks.

I. INTRODUCTION

Attendance tracking is an essential administrative task, but it can be arduous and prone to inaccuracies when done manually. Traditional roll call methods are limited, especially with large student numbers, as it is difficult to call out names accurately and maintain proper records. Organizations have adopted various approaches, such as document-based systems, biometric fingerprinting, or card swiping. However, these methods can be time-consuming, requiring students to wait in queues, and may fail if a student forgets their ID card.

Evolving technologies have enabled advancements in intelligent attendance systems, such as biometric face recognition. While promising, existing face recognition techniques struggle with challenges like scaling, pose variations, illumination changes, rotation, and occlusions. The proposed framework aims to address these limitations of current systems.

The vital steps in face recognition involve face detection, feature extraction, and recognition. The system employs multiple cameras installed on the classroom ceiling to capture images of the entire area. Generative Adversarial Networks (GANs) can be used to enhance blurred images caused by student movement, improving efficacy. The processed images then undergo face detection, feature extraction using Gabor filters, and face recognition via algorithms like K-Nearest Neighbors (KNN), Convolutional Neural Networks (CNNs), and Support Vector Machines (SVMs), with comparative studies.

Upon successful recognition, the system generates the names and identification numbers of students present in the image. Attendance is then marked in an Excel format, including the date and subject details. This system requires minimal hardware resources, making it cost-effective for institutions.

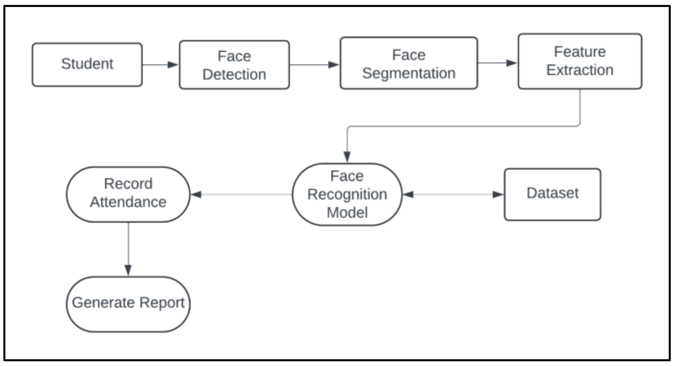


Figure 1 Architecture Diagram

This diagram illustrates the process of a face recognition-based attendance system. The process starts with the input from the student, which is likely an image or video feed. The face detection step identifies and locates faces present in the input. The face segmentation step separates the detected faces from the background or other elements in the image. The feature extraction step analyzes the segmented faces and extracts relevant features or characteristics that can be used for recognition.

The extracted features are then compared against a dataset, which likely contains pre-recorded facial data of enrolled students or authorized individuals. This comparison is performed by the face recognition model, which uses algorithms to match the input facial features with the stored data in the dataset.

If a match is found, the system records the attendance for the recognized individual. Finally, the attendance data can be used to generate reports, which could include information such as attendance records, absentee lists, or attendance statistics.

Overall, this diagram depicts a systematic approach to automating attendance tracking using facial recognition technology, where various steps are involved to process the input data, identify individuals, and maintain attendance records..

II. LITERATURE REVIEW

In [1] described by Bhise, Khichi, Korde, and Lokare in 2015, utilizes Near Field Communication (NFC) technology along with an embedded camera on a mobile device. NFC enables short-distance wireless communication between an active and a passive device, typically achieved through inductor coils responding to electromagnetic induction. In this system, students are provided with NFC tags containing unique identification numbers. During class, students swipe their NFC tags near an NFC reader, such as the lecturer's phone, to register their attendance. However, this system is susceptible to impersonation, where one individual could swipe in for another.

In [2], proposed a real-time computer vision algorithm for automating attendance management. The system utilized non-intrusive cameras installed in classrooms to capture images for comparison with stored faces. Machine learning algorithms commonly applied in computer vision were utilized, along with HAAR classifiers for image training. Captured images were converted to grayscale, subjected to subtraction, and then sent for storage on a server for later processing. In 2012, N. Kar presented an automated attendance management system employing face recognition techniques, particularly Principal Component Analysis (PCA). OpenCV was used for computer vision functionalities, while FLTK (Fast Light Toolkit) facilitated interface design. The system consisted of two main functions: Request Matching and Adding New Faces to the Database. During Request Matching, the camera captured images, frontal faces were extracted, recognized using training data, projected onto PCA, and the closest match was displayed. Adding a new face involved capturing and extracting frontal face images, applying the Haar cascade method for face detection, employing the PCA algorithm, and storing the information in an XML file. This system primarily aimed to enhance face detection algorithms for images or videos acquired in real-time.

In [3] Jyotshana Kanti introduced a smart attendance marking system that integrates Principal Component Analysis (PCA) and Artificial Neural Network (ANN) algorithms. The author's aim was to address issues with traditional attendance marking systems and reduce time consumption. The system employs PCA for facial feature extraction and identification of similarities within the face database using acquired images. ANN is utilized to handle input data and learn from it, along with expected values. The author implemented a backpropagation algorithm in conjunction with mathematical functions within the system. According to the author's findings, the system demonstrated effectiveness in recognizing individuals across various environments.

III. PROPOSED SYSTEM

A. Architecture

The proposed system offers simplicity, ease, and manageability with clear-cut operations. It incorporates a database containing students' facial profiles alongside pertinent details such as their names, enrollment numbers, and courses. Depending on classroom size and requirements, two or more cameras will be mounted on the ceiling to comprehensively cover the entire area. These cameras will capture images multiple times throughout a lecture, enhancing system efficiency by ensuring that even if certain students are not within the view of one camera, they will likely be captured by others.

Given the diverse range of expressions and poses students may adopt, the system accounts for potential challenges in face detection due to unfavorable angles or poses. In such cases, subsequent image captures provide additional opportunities for detection. Once image acquisition commences, initiated by the teacher clicking the start button, the system proceeds with face detection across all camera feeds and instances. Detected faces are then cross-referenced with stored student images in the database. Upon successful matching, attendance is recorded in an Excel format alongside the corresponding enrollment number and name.

B. Methodology

Developing an intelligent attendance management system, some steps need to be followed to achieve this Successful task. The steps are definable as follows:

Database creation

The database will be built in the first phase when students are enrolled. General student information such as name, ID number, course, and semester subjects will be stored in the database. In addition, the system is supposed to record the student's image in order to train the suggested framework. For training purposes, this system records a single image for each pupil.

Facial recognition for every student in a lecture is made possible by the database containing all of the images the student has uploaded. It is possible to achieve.

Image amelioration

The camera may catch a blurry image due to a pupil moving about the classroom. With the use of Generative Adversarial Networks, the image can be improved. GANs are renowned for their capacity to preserve textural information in photos, produce results that resemble the real range of characteristics, and have a believable appearance.

$$I_B = k(M) * I_s + N$$

where In this scenario, IB represents a distorted image, with k(M) representing unknown blur kernels detected through the motion field M. IL stands for the clear latent image, and the symbol * denotes convolution, while N signifies additional noise..

Face detection

"Facial detection relies on the identification of 70 key facial landmarks. Haar classifiers, a method based on machine learning, are utilized for this task. This involves training a cascade function with a diverse set of positive and negative images to recognize faces. The approach involves assessing pixel values in black and white regions, with the difference informing the classifiers. Handling 6000 characteristics per window frame posed a challenge, leading to the adoption of classifier cascades. AdaBoost is applied to streamline feature selection, focusing on essential aspects termed weak classifiers. These weak classifiers are then combined in a weighted manner to form a robust classifier, leveraging AdaBoost for optimization."

$$P(Y) = \sum (S_i * p_i(Y))$$

In this context, P(Y) represents a robust classifier, while s_i denotes the respective weights assigned to each weak classifier $p_i(T)$.

Feature extraction

Gabor filters are used to capture facial features that are angled at different angles in order to extract features. This is a crucial stage because it's thought that a feature extractor that works well chooses a function that is resistant to lighting, pose variance, occlusion, and context. Spatial distortions brought on by variations in lighting and position are corrected with 2D Gabor filters.For feature extraction, Gabor filters are used to lay hold on facial features inclined at various angles. It is a very critical step since it is believed that a successful feature extractor

selects a function that is not prone to occlusion, lighting, context, and pose variance. 2D Gabor filters are used to resolve spatial distortions caused by position and lighting variances.

$$W(x, y, \theta, \lambda, \varphi, \sigma, \gamma) = e^{\left(-\frac{x^2+\gamma^2 y^2}{2\sigma^2}\right)} \cos\left(2 \prod \frac{x^2}{\lambda} + \varphi\right)$$

$$x' = x \cos\theta + y \sin\theta$$

$$y' = -x \sin\theta + y \cos\theta$$

here (x, y) defines the situation of a light impulse and μ, ϕ, γ, λ, σ are parameters of the sinusoidal wavelet.

Face recognition

Face recognition employs various algorithms such as the K-nearest neighbor, convolutional neural networks (CNNs), and support vector machines (SVMs). A comparison among these algorithms is conducted based on criteria including accuracy, robustness, and time complexity..

A. K-nearest neighbor algorithm

KNN is often referred to as lazy learning because it retains the information from the training examples without building a specific model. The Euclidean distance metric is commonly chosen to establish the positions of data points in KNN. In this method, an object's classification is determined by the majority vote of its neighbors, with the object assigned to the most frequent class among its nearest neighbors (where k is a positive integer). When k equals 1, the object is allocated to the class of itsclosestneighbor.

$$\boxed{d(x.y) = \sqrt{(x_1 - y_1)^2 \ldots \ldots \ldots (x_n - y_n)^2}}$$

d(x.y) is the euclidean distance which is by default used by KNN to find the nearest class.

B. Convolutional neural networks

Convolutional Neural Networks (CNNs) enable the extraction of diverse features from images. This principle extends to facial recognition, where CNNs utilize 70 facial landmarks to create 128-dimensional encodings representing facial features in RGB format. These encodings are compared to identify matching faces, with the level of comparison stringency adjustable via a tolerance parameter Redundancy removal

As the system encompasses multiple cameras. there might be a possibility of the presence of the face of a single student in different images. redundant faces will be removed and single faces will be considered to mark single attendance for a student during a lecture.

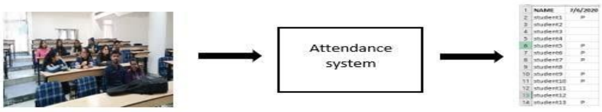
Report generation

Trailing face recognition reports are generated by marking present in front of the student name and enrollment number in excel format during a lecture

IV. RESULTS

Three distinct algorithms were used to evaluate the system; with a 98% accuracy rate, the KNN method turned out to be the best. The system was evaluated under a variety of circumstances, including lighting, head motions, facial expressions, and the distance between the students and the camera. Even in cases where the image includes both bearded and bald faces wearing spectacles, the system performs as expected. The suggested method shown remarkable ability to identify faces with a two-year age difference. When tested under these circumstances, KNN outperformed the others, obtaining an overall accuracy of 97%. CNN and SVM both obtained 95% and 88% overall accuracy when tested under the aforementioned parameters. CNN revealed to have minimal time complexity while looking at the time complexity aspect. Of the three algorithms mentioned, SVM was found to have the highest time complexity. 200 real-time photos of a classroom with a maximum capacity of 70 students are used to test the suggested approach. The suggested solution is capable of handling 70 pupils in a classroom.

The figure below shows the result of our proposed system



V. COMPARISON

The provided table compares the accuracy of three machine learning models (KNN, CNN, and SVM) under different testing conditions. These conditions include normal circumstances, scenarios with a 2-year age difference, scenarios involving individuals with beards, and scenarios with different facial expressions. The accuracy percentages for each algorithm are given in the table. For instance, under normal conditions, KNN achieved an accuracy of 98%, CNN had 96%, and SVM had 88% algorithms have been compared on the grounds of time complexity accuracy in various conditions.

Conditions under testing	Accuracy		
	KNN	CNN	SVM
Normal	99.27	95.54	89.15
With age difference (2yrs)	97.90	95.00	86.75
With beard	90.00	90.00	80.00
With different expressions	92.00	90.00	78.00

A. Head movements

In a classroom head movement at different angles is possible during a lecture. Head movements can be categorized into three categories measured in angles which are pitch, yaw, and roll at the respective x-axis and y-axis.

Pitch: This refers to the movement of the head up or down, akin to nodding. It is measured along the x-axis. When the head moves downwards, it registers as a negative value on the x-axis, while an upward movement results in a positive value.

Yaw: Yaw involves the rotation of the head from side to side, resembling a "no" gesture. It is also measured along the x-axis. When the head moves towards the left, it's represented by a negative value on the x-axis, and conversely, a movement to the right is indicated by a positive value.

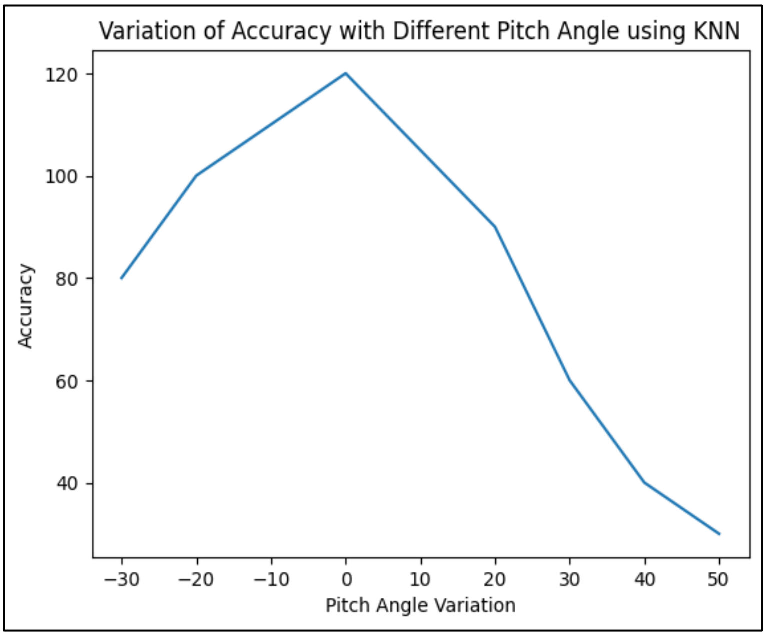
Roll: Roll refers to the tilting motion of the head from one side to the other, similar to tilting one's head to look at something from a different angle. This movement is measured along the y-axis. A tilt towards the left is denoted by a negative value, while a tilt towards the right is shown as a positive value on the y-axis.

The mentioned algorithms underwent testing with various head movements, and the plots below depict how their accuracy varies with different angles.

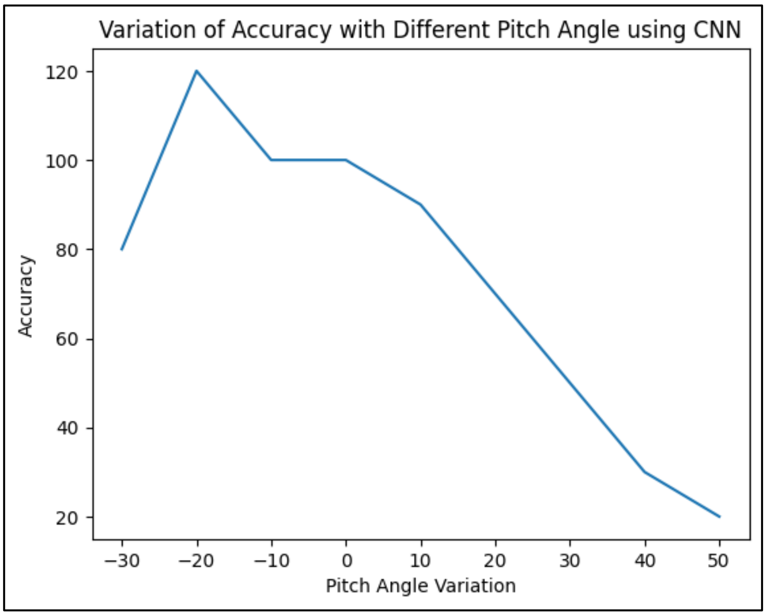
Regarding Roll and Yaw angles, negative values on the x-axis represent head movement to the left, while positive values signify movement to the right.

Concerning the Pitch angle, negative values on the x-axis represent downward head movement, whereas positive values indicate upward movement.

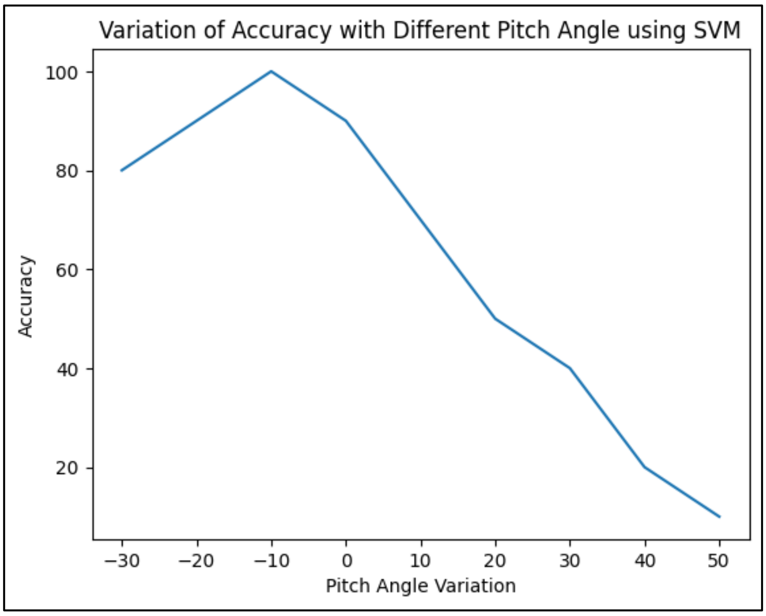
1) Pitch angle variation



Plot 1 The plot depicts the relationship between accuracy and pitch angle variation when using a K-Nearest Neighbors (KNN) algorithm. The accuracy is highest at a pitch angle of 0 degrees, indicating that the KNN model performs optimally when the object or image is aligned with the standard orientation. As the pitch angle deviates from 0 degrees, either positively or negatively, the accuracy decreases symmetrically. This suggests that the KNN model's performance deteriorates equally for positive and negative pitch angle variations.



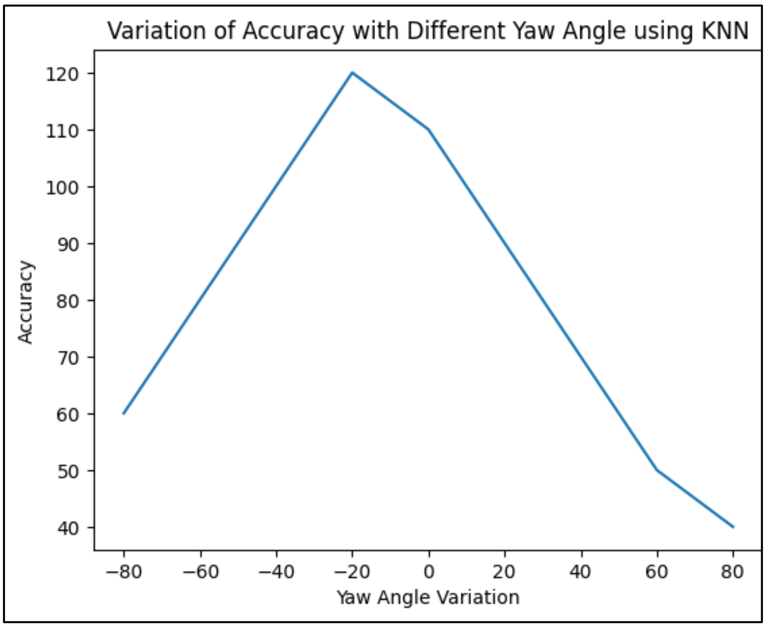
Plot 2 This plot illustrates the variation of accuracy with different pitch angles when employing a Convolutional Neural Network (CNN) algorithm. Similar to the KNN model, the accuracy peaks at a pitch angle of 0 degrees and declines as the pitch angle deviates from this standard orientation. However, the curve exhibits an asymmetric shape, with a steeper drop in accuracy for positive pitch angle variations compared to negative variations. This asymmetry indicates that the CNN model may be more sensitive to positive pitch angle changes.



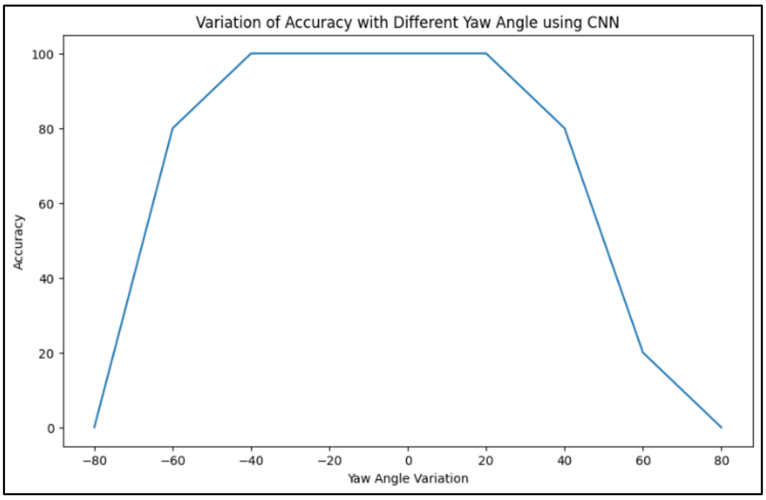
Plot 3 The plot represents the variation of accuracy with different pitch angles when using a Support Vector Machine (SVM) algorithm. The overall trend is comparable to the KNN and CNN plots, with the highest accuracy achieved at a pitch angle of 0 degrees and a decrease in accuracy as the pitch angle deviates from this point. However, the drop in accuracy appears more gradual for the SVM model, suggesting that it may be more robust to pitch angle variations compared to the KNN and CNN models

.

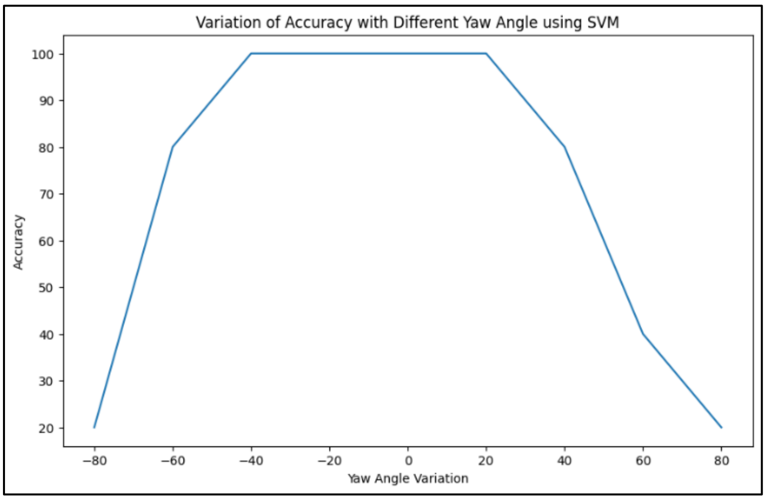
2) Yaw angle variation



Plot 4 This plot illustrates the relationship between accuracy and yaw angle variation when using a KNN algorithm. The accuracy is highest at a yaw angle of 0 degrees and decreases symmetrically as the yaw angle increases or decreases from this point. This behavior is consistent with the pitch angle variation observed for the KNN model, indicating that the KNN algorithm is sensitive to deviations from the standard orientation in both pitch and yaw angles.

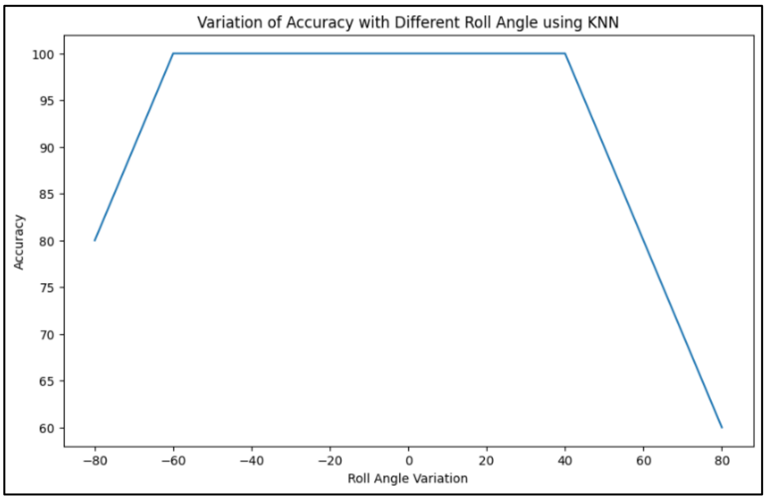


Plot 5 The plot shows the variation of accuracy with different yaw angles when using a CNN algorithm. The curve exhibits a similar shape to the pitch angle variation for the CNN model, with the highest accuracy at a yaw angle of 0 degrees and a steeper drop in accuracy for positive yaw angle variations compared to negative variations. This asymmetric behavior suggests that the CNN model may be more sensitive to positive yaw angle changes, similar to its sensitivity to positive pitch angle changes observed in the previous plot.

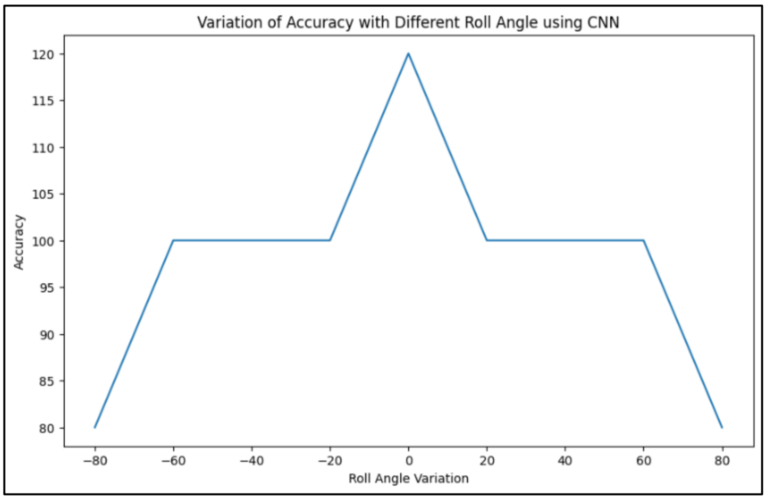


Plot 6 The plot shows the accuracy of an SVM (Support Vector Machine) system varies with changes in the yaw angle. Accuracy peaks around 0 degrees yaw angle and decreases symmetrically on either side, forming a U-shaped curve.

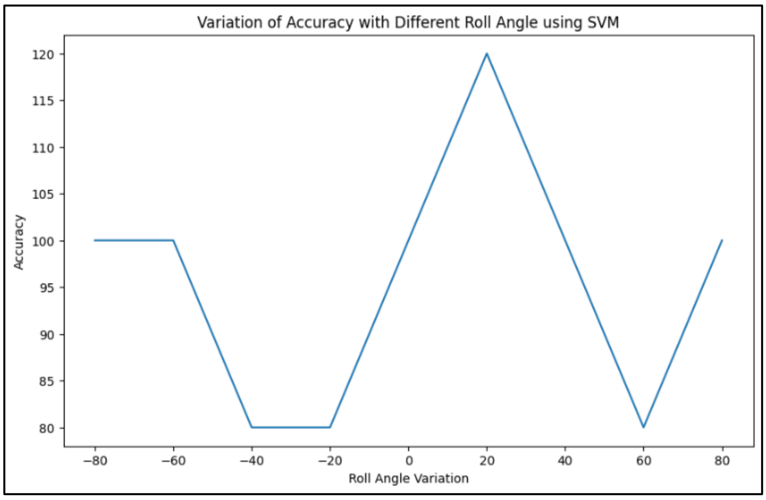
3) Roll angle variation



Plot 7 This plot illustrates how the accuracy of a KNN (K-Nearest Neighbors) system changes with roll angle variation. Accuracy is highest near 0 degrees roll angle and decreases gradually as the roll angle increases or decreases, creating a relatively flat U-shape..

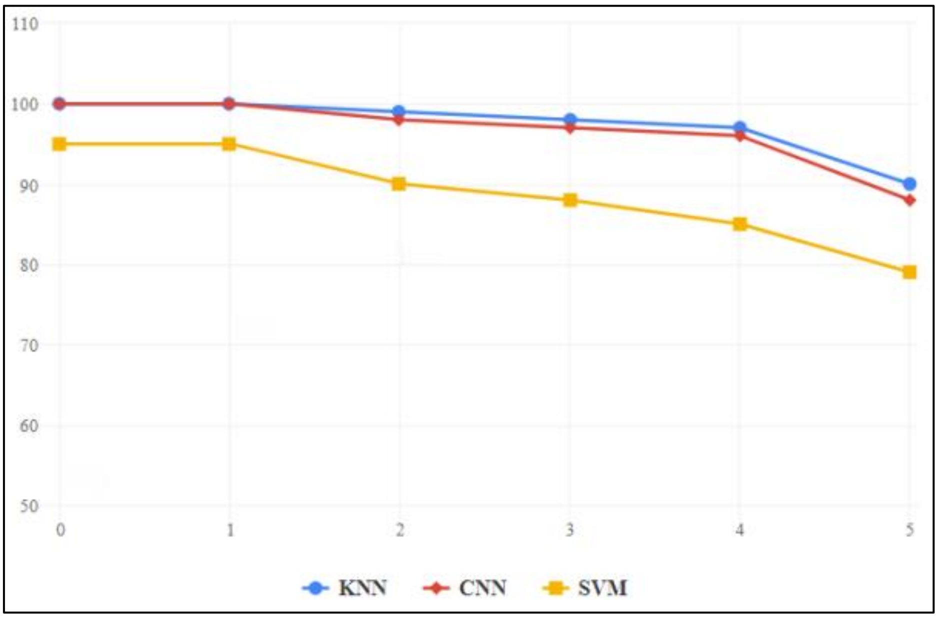


Plot 8 The plot displays the accuracy of a CNN (Convolutional Neural Network) system varying with roll angle changes. Accuracy is maximum at around 0 degrees roll angle and drops sharply in a symmetric V-shape as the roll angle increases or decreases..



Plot 9 This plot shows an irregular curve for the accuracy of an SVM system as the roll angle varies. There are multiple peaks and valleys, with the highest accuracy around 0 degrees roll angle and lower accuracies at larger positive or negative roll angles

Even though cameras are intended to be mounted on the ceilings of classrooms, there's a chance of varying distances between students and cameras due to students sitting in different seats. A system employing three distinct algorithms was evaluated under conditions of fluctuating distances between students and cameras. The plot below illustrates how the accuracy of these algorithms changes in response to the aforementioned scenario.



The plot compares three systems - KNN, CNN, and SVM - as the roll angle changes from 0 to 5 (arbitrary units). KNN has the highest accuracy at lower roll angles, followed by CNN and SVM. All three show a gradual decline in accuracy as the roll angle increases.

B. Overall result

Taking into account all the above-mentioned conditions and situations overall accuracy, precision, recall, F1 score, and time complexity of the algorithm are calculated.

The table listed below describes the above statement.

Algorithm	KNN	CNN	SVM
Overall accuracy	99.27	98.54	80.15
Overall time complexity	124 seconds	120 seconds	480 seconds
Precision	0.99	0.98	0.78
Recall	0.98	0.97	0.75
F1 score	0.984	0.974	0.764

VI. CONCLUSION

The proposed system successfully fulfills the goal of attaining high precision while maintaining low computational complexity. It's cost-effective and requires minimal manual intervention. The incorporation of Gabor filters significantly enhances accuracy. In face recognition, three algorithms were employed: K-nearest neighbor, convolutional neural networks, and support vector machine. Among these, the KNN algorithm demonstrated the highest accuracy at 98%. Convolutional neural networks showed lower computational complexity, while the SVM algorithm exhibited comparatively lesser efficiency.

VII. REFERENCES

[1] Yohei Kawaguchi, , “Face Recognition-based Lecture Attendance System”, “3rd AERU…”, 2005.

[2] B. Kavinmathi, , “Attendance System for Face Recognition using GSM module”, 5th International Conference on Signal Processing Integrated Networks”, 2018.

[3] Ketan N. Mahajan, ,” Classroom attendance system using surveillance camera”, International Conference on Computer Systems, Electronics and Control”,2017.

[4] Shubhobrata Bhattacharya, , Aurobinda Routray “Smart Attendance Monitoring System (SAMS): A Face Recognition Attendance System for Classroom Environment”, IEEE 18th International Conference on Advanced Learning Technologies, 2018.

[5] E.Varadharajan, a,., “Automatic attendance system using face detection”, 2017.

[6] Guo, Jing-Ming, “Complexity reduced face detection using probability based face mask prefiltering and pixel-based hierarchical-feature Ada boosting”, Signal Processing Letters, IEEE 2011. .

[7] K.Senthamil Selvi1, , Jenitha S, “face recognition based attendance marking system”, International Journal of Computer Science and Mobile Computing, 2014.

[8] Chen, Joy Iong Zong. "Smart Security System for Suspicious Activity Detection in Volatile Areas." Journal of Information Technology 2, 2020. “Attendance System for Face Recognition using GSM module”, 4th International Conference on Signal Processing and Integrated Networks”, 2018.

[9] Rajat Kumar Chauhan, Lokanath M, “ Smart Attendance System Using CNN”, International Journal of Pure and Applied Mathematics,2018.

[10] Mayank Yadav, , “ Motion based attendance system in real time environment for multimedia application”, 2018.