National College of Ireland

# Improving the Face Recognition capabilities using DenseNet architecture for Attendance management system in Cloud Computing Environment

MSc Research Project
Cloud Computing


Rajiv
Student ID: x17165822

School of Computing
National College of Ireland


Supervisor: Aqeel Kazmi

**National College of Ireland**
**Project Submission Sheet**
**School of Computing**

| | |
|---|---|
| **Student Name:** | Rajiv |
| **Student ID:** | x17165822 |
| **Programme:** | Cloud Computing |
| **Year:** | 2021 |
| **Module:** | MSc Research Project |
| **Supervisor:** | Aqeel Kazmi |
| **Submission Due Date:** | 31/01/2022 |
| **Project Title:** | Improving the Face Recognition capabilities using DenseNet architecture for Attendance management system in Cloud Computing Environment |
| **Word Count:** | 7235 |
| **Page Count:** | 19 |

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

**ALL** internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

| | |
|---|---|
| **Signature:** | Rajiv |
| **Date:** | 31st January 2022 |

**PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:**

| | |
|---|---|
| Attach a completed copy of this sheet to each project (including multiple copies). | ☐ |
| **Attach a Moodle submission receipt of the online project submission**, to each project (including multiple copies). | ☐ |
| **You must ensure that you retain a HARD COPY of the project**, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer. | ☐ |

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

| **Office Use Only** | |
|---|---|
| Signature: | |
| Date: | |
| Penalty Applied (if applicable): | |

# Improving the Face Recognition capabilities using DenseNet architecture for Attendance management system in Cloud Computing Environment

Rajiv
x17165822

**Abstract**

In the educational institutions, managing the attendance of student and staff is a tedious task. The traditional approach of attendance system is manual, time consuming and may contain fallacies in the record due to human mistake. Therefore, in order to resolve such issues in this work we have proposed a framework for attendance management system. which can automatically detect the person based on captured face data. As a onetime process, student needs to register themselves by providing their name and capture their image. The captured image will be trained on the deep learning model and next time the model will be able to detect the trained face. In this work we have used 3 different deep learning architectures that are VGG-19, ResNet-50 and DenseNet model. After comparative analysis, we have obtained better results using DenseNet architecture where Facenet and MTCNN has been used to localize the face from video and extract the face features.

# 1 Introduction

In recent times, technological advances have mimicked the human intellect. The growing pace has combined computer and human interfaces. With the introduction of Artificial Intelligence (AI), it has been actively being utilized in various domains. From futuristic innovations to the human assistive system, AI which is the sub-field of Deep Neural Network (DNN) has contributed to this. For decades the researchers have been working to replicate human intelligence to the computer which was known as Computer Vision. The human brain is capacious to sense and develop reasoning under the environmental presence. Similarly, the Computer Vision process is hand-to-hand to human vision. Computer Vision digitally recognizes and grasps the graphical representation of the environment and interprets it through computational algorithms. The specific instance employed through computer vision is object detection and recognition which is further explored. Considering our paper, the detection and recognition of face are taken. Nowadays, facial recognition has widespread adoption in various domains such as security verifications, authentication, and surveillance. This system is being utilized and incorporated in programs from simplified face recognition in social media like Facebook to facial authentication in the payroll system. Apart from these, other applications in the domain are facial-based point of sale (POS) transactions, missing person detection, assisting blind people and many more. Although the system of face detection and recognition has a wide range of application out of which security authentication and verification is most sought upon.

This system works on two processes which are detection and recognition. In the primary process, the face is being utilized which is based on the fundamentals of the object's detection module. Multiple features of the face are computed in the model such as shapes, size, similarity indexes for face detection. Then comes the part of face recognition. Here, the standardized face images are computed among different variance and graphical compositions such as lighting, contrast, background effect, etc. The challenge of face recognition relies on pattern recognition standards. Over time various conventional approaches were brought up such as elastic map match, fisher face, and eigenface which relied on the performance modelling. Although in the later stage, the researchers considered the approach of deep neural network for face recognition system providing better enhancement and efficiency. With DNN implementation, the space model is primarily formed to interpret the different angles of the face. Further with the construction of the space model, the characteristics of images are extracted. After the extraction of various dominant characteristics of the images, the DNN algorithms are computed to implement in the model. In the final stage, the composed framework is implemented in a different system to process the unique data such as video, images, or any other graphical representations.

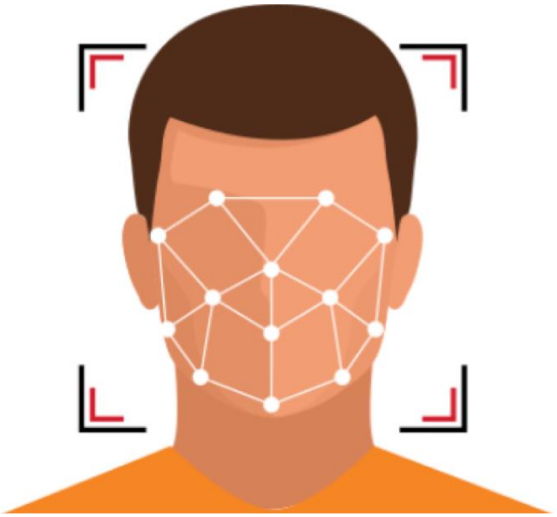


Figure 1: Face Detection

In this paper, we will study the Attendance Management System which relies on the authentication and verification of the human impression. Attendance management had been a longing issue for schools, colleges and large companies. In previous times, themost common approach was to take attendance manually. Although, it had a considerable number of fallacies. In the educational campuses, this approach was very time-consuming and exhausting to some extent. The period utilized on checking the attendance could be utilized in other productivity. On the other hand, for large companies, a manual approach can be a gruesome task and also not economically feasible. Although some other mechanisms such as semi-automated approaches were implemented to overcome the challenges. But these systems such as password-based authentications, identification cards and various other biometric systems had their fallacies. The worst-case scenarios

such as forgotten passwords, misplaced id cards, fraudulent attendance can be a great hurdle for all. Therefore, to overcome all these challenges and obtain efficient authentication and verification, face detection and recognition is the best approach. This system didn't obligate for the manual approach which is time-consuming and gruesome or the semi-automated such as passwords or id cards. It could autonomously identify a face and recognize the person by comparing it to the database, once trained. This way it could make the process economically feasible, reliable, and convenient.

For this purpose, various machine learning and deep learning algorithms have been utilized by researchers. Due to its multiple advantages over the conventional approaches, the attendance management system utilizing the face recognition framework is being widely adopted among various sectors. Generally, in the framework, a wide set of image or video input data is considered to train the model. In this research, the Five-Celebrity Data set has been collected from kaggle *5 Celebrity Faces Dataset* (n.d.), which contains the differ-ent images of 5 celebrities. After acquiring the dataset, these are converted into the RGB (Red-Green-Blue) format for which either of the processing frameworks is implemented such as DLIB (library for machine learning) or MTCNN (Multi-cascaded Convolutional Neural Network). This enhances the model for face classification. Thereafter the featuresof these images are extracted through the approach of feature engineering. Finally, the algorithms are implemented to the processed dataset and the model is evaluated withthe implementations of the module. In this research, various pre-trained algorithms such as VGG-19, Resnet-50, Dense Net have been utilized in the model for face recognition and detection in the attendance management system. Then the classification model is evaluated through multiple metrics such as Accuracy, Precision, Recall, F-1 score and Loss.

# 2 Literature Review

In this section, we will discuss the study by the researchers on the various related domain. This section is further divided into a conventional approach, machine learning approach and deep learning approach.

## 2.1 Conventional Approach

In a research paper Adamu (2019), an approach for an attendance management system was proposed utilizing the fingerprint scanner and iris scanner. This system was imple- mented in the educational areas to keep a check on the attendance of the student. The author stated that the system consists of two sets namely, hardware and software units.In the latter one, it consisted of a fingerprint scanner, iris scanner and a computer modulewhereas in the next one it consisted of a user interface and student database. This system worked in the sequential process wherein the initial stage the data had to be fed for the further process. Then comes the stage of identification where the data is being identified under the database. At the final stage, the inputs are verified with the database and in case not being verified the system shows an error. This system should have an overall accuracy of 75% over a year which reduced some challenges to an extent. Although,the study needs to be improved for increasing the overall performance. Similarly, Bansal (2018) also studied the presence of the fingerprint system for the attendance management system in educational institutions and large organizations. This study also demonstrated

the challenges that occurred in the mid-way due to the mishandling of the database and system. Moreover, to this, the author stated some of the fallacies in this system such as the utilization of duplicate copies of the fingerprint which is nowadays a feasible task. Also, the system may not detect the biometrics of biological twin siblings which may be a great hurdle. In this system, during the time of enrolment of biometric the features are extracted and stored in the database. Thereafter at the time of authentication the features are once more extracted and compared to the database for the verification of the person.

In another study Desai (2018), an attendance management system using the biometric system with SMS alert was proposed by the author in the paper. This study showed a novel approach by incorporating SMS alerts after using the biometric system which could be helpful for the periodical report generation in institutions and corporations. For the institutions, this system would initially authenticate and verify the student from the database and thereafter send an SMS to the parents. It also generates an attendance report of a student. This system was built using Arduino with a GSM module to send SMS. The study helps in overcoming the challenges due to the traditional approach. Fur- thermore, the author also described the future scope of the study which could be iris and facial recognition. Vibin Vinod et al. (2021) proposed an attendance system relying ona low-cost convenient smart card. This system of smart cards is built using the combin- ation of RFID (Radio Frequency Identification) with the fingerprint system. Through this system, the institutions and corporations could save time, manpower and money. The system was built on the framework of Raspberry PI which was connected to the RFID reader and fingerprint scanner. A user interface was designed for easy accessibility through the desktop or mobile phones which further had a robust security level by im- plementing certain access permission. This model was efficient, feasible and convenient for real-time implementation in the educational campuses.

Ravi Supunya (2018) recommended the utilization of IoT-based recognition for the attendance management system. As the manual procedure for attendance management system is not feasible in the current era, the author stated the use of an automated sys- tem. But the use of the known automated system such as fingerprint scanners, RFID readers or face recognition is not convenient and accurate when used independently in the educational campuses among the students. Therefore, the author proposed combiningmore than one system to get better performance of all. When this is connected with the web application, it enhances accessibility and convinces. The author also suggested the future task to integrate with cloud service and build an efficient model to track students' location and behaviour. In another paper, Elaskari et al. (2021) suggested an approachto track students' attendance and assets through the barcode system. Through a bar-code scanner, students can enter the campuses conveniently by automated attendance. This will also help the institution staff to save time, and generate an error-free report of the students. Also, through the dedicated barcode of the students, the institutions can connect the assets such as degrees and certificates of students for verification in higher education. This system can be easily configurated by anyone as the database has to be just connected with the system to work. It will act as a feasible solution to common challenges.

## 2.2 Machine Learning Approach

Mohan and Prasad (2020) surveyed the machine learning classifier model for the attendance management system. A robust model was researched to overcome the challenges of the conventional approach. The object detection framework was implemented in the model, for which the Haar Cascade classifier was implemented. Initially, the database of the model is constructed by capturing the images of each person. Then the facial features are extracted from the images feed in the database. Here in the system, an administrator is dedicated who can update or changing the database. To detect the facial features, libraries such as Dlib and OpenCV. Then for verification and identification, face recognition algorithms could be utilized such as eigenfaces, local binary pattern histograms, fisher faces, etc. Although in this paper, local binary pattern histogram algorithms were util- ized. Furthermore, histograms such as HOG are utilized for better efficiency. In the final stage, the model achieved an accuracy of 87% and a precision of 89%. In another paper by Alankar Patil (2020), similar work was proposed with better tuning and performance. The author utilized the Viola-Jones algorithms for the face detection framework. This algorithm relies on four object detection frameworks namely, Haar features, Integral image, Adaptive boost algorithm, and Cascading. The detection using this algorithm is fast but the training time is a disappointment. This study also utilized the local binary pattern histogram algorithm for the face recognition protocol. Overall, this system is ought to be less complex as under hardware the model can work with just a computer and camera. It is a reliable, accurate, convenient model to replace the conventional approach.

R.S et al. (2020) proposed smart student attendance through face recognition utilizing machine learning algorithms. The video or image dataset sequence was obtained from the classroom which will show the faces of each student present. Once obtained, Haar Cascade and MTCNN model algorithms were utilized to detect and recognize the facial features and boundaries. After this process, the attendance of the students was marked using the Decision Tree algorithm. When the model was implemented in the real-time environment, it showed an enhanced performance with higher accuracy and precision when compared to the previous approaches. Similarly, Khan (2020) utilized machine learning algorithms for computer vision modules. The author studied the applications and implementations of computer vision among different countries and domains. The techniques of computer vision are further split into supervised, semi-supervised and unsupervised learning approaches. Generally, the most utilized machine learning algorithms are support vector machine, k-nearest neighbour, etc. Nowadays, major entities have adopted the models to incorporate a convenient feature. The libraries implemented in the studies are TensorFlow and OpenPose which are usually detected for computer visionand object detection. The paper also evaluated the applications of object detection in various instances. This system was majorly implemented in the field of biological scienceand miscellaneous human activities which stood at 19% whereas in the other fields such as face recognition, medical science, weather forecasting and agriculture stood at 6%. Furthermore, this system was also utilized in the fields such as traffic level detection, professional sports, industrial tasks, etc. The author also suggested the scope of future task which would assess the performance of the machine learning framework.

In another paper, machine learning algorithms were reviewed for the applications of computer vision by Ayub Khan et al. (2021). Various approaches to machine learning

framework were discussed for incorporating the image processing. Aspects of the specific tools, applications, methodologies and datasets were touched upon. Challenges that occurred due to underfit or overfit of the data were also discussed. Some of the most prominent space research in this domain by the researchers are object or vehicle detection, activity recognition and human pose estimation. With the further enhancement, many companies have implemented the neural network to their model. The paper stated multiple instances such as recommendation systems in Amazon, translator in Microsoft, facial detection and recognition in Facebook, spam filter in Google's Gmail and more. Shekhar et al. (2020) surveyed different algorithms of machine learning framework which is implemented in the field of computer vision. The author considered the philosophyof the famous scientist named Charles Darwin's law of evolution in the growing paceof technological advances in machine learning. The implementation of the models in various fields and disciplines such as a computer or mobile programs, web applications, etc. Majorly utilized machine learning frameworks such as Decision Tree, Na¨ıve Bayes, Random Forest Distribution algorithms were discussed in the paper. Furthermore, the author here discussed the development that occurred in the field of medical science due to the implementation of machine learning in computer vision.

## 2.3 Deep Learning Approach

J et al. (2019) proposed a face recognition module for security systems utilizing the deep learning framework. In the paper, an artificial neural network (ANN) is being utilized in the model. The author has proposed a novel approach that can be implemented in schools and colleges in real-time. The proposed approach also could combat the challenges that occurred in the conventional approach. For the hardware, the Raspberry PI system was utilized which could be developed with different modules. Moreover, a user-friendly and convenient GUI (graphical user interface) was developed. During the real-time application, the model achieved an accuracy of 74%. According to a study by Tapyou et al. (2021), multiple entities are interested in adopting the facial analysis method for protection agencies and combating violence. Furthermore, the goal is for robots to replicate human proficiency in remembering appearances as precisely as possible. To evaluate the quality of computer vision applications, many assessment measures are used. Failed identification has to be inside an appropriate spectrum to meet the precision criteria. To match the particular requirements, the findings are calculated and evaluated. A facial identification system, which is reliant upon this Internet of Things (IoT), is also used for automated enrollment. Pictures are captured via cellphones and tablets, computers and then processed by the identification program.

In another paper Arsenovic et al. (2017), a convolutional neural network (CNN) algorithm was utilized for the face recognition system. The author used the process of image augmentation to extend the dataset as the DNN framework achieves higher accur- acy with the larger dataset. CNN cascading was implemented for face detection. For face embedding, the pre-trained model of Face Net was implemented. Finally, for the classification SVM algorithm was implemented. This proposed model achieved an accuracy of 95.02%. In another investigation by .A et al. (2018), two distinct pattern identification and characterization approaches, called Scale Invariant Feature Transformation (SIFT)

and Speeded Up Robust Features (SURF) were utilized. Employing the bag of words approach, the segmentation methods are picked to build a universal classification approach. SVM, Extreme Learning Machine (ELM), and Extremely Randomized Trees (ERT) are used to evaluate the categorization. These would be subsequently sent into the polling algorithm as forecasting models. The researchers of the paper Nyein and Oo (2019) created an academic attendance management system using FaceNet for image retrieval and SVM for identification. The face photographs are placed in a 128-dimension space, and the triple error mechanism aids in load optimization. In a multi-faced environment, the study proved capable of reaching an impressive performance.

Similarly, Kuang and Baul (2020) utilized a pre-trained framework called Haar Cascade algorithm to efficiently detect the faces through the given image or video sequence. Furthermore, the framework of Face Net is also utilized in the model which also helps encompass the error function. The system initially captures facial images through the web camera. Then the face recognition algorithm is implemented by comparing it to the predetermined database. Islam et al. (2017)'s other research, which evaluated ANN for object detection using BoW, HOG, and image pixel as component selection methods,saw almost 96%, 99%, and 99.50% efficiency scores, correspondingly. The picture pixel resolution methodology with ANN beat the other rival techniques, SVM and CNN, which used similar attribute retrieval methods. On other hand, Sarkar et al. (2019) introduced a novel model for the automated attendance system utilizing deep learning mechanisms. Primarily, the challenges were discussed in the previously built model and the blueprint was constructed for the new model. In the paper, shallower face verification architecture was implemented but the performance was enhanced in the overall output. The model was able to achieve 98.67% in the learning set and 100% in the validation set.

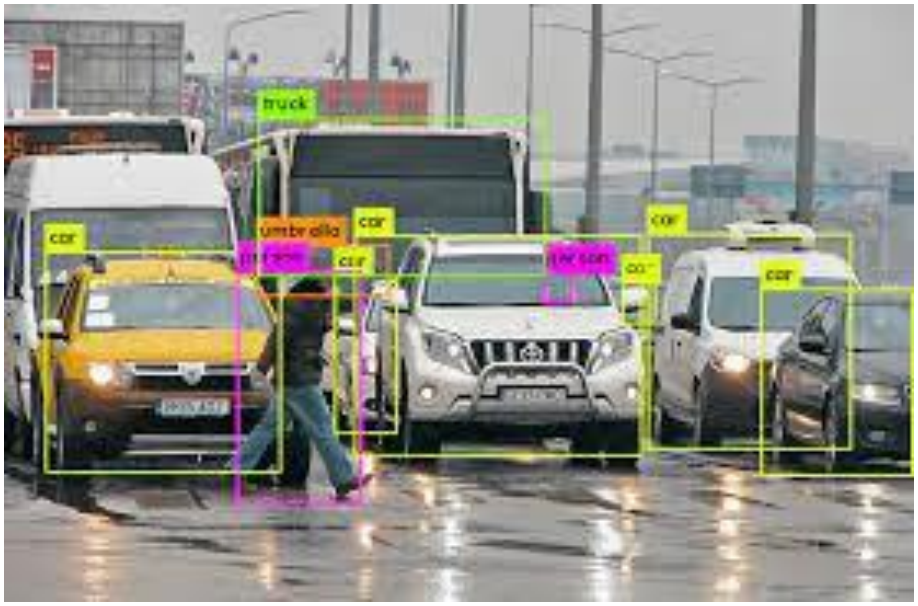



Figure 2: Detection of Object

# 3 Methodology

The need for a robust approach for attaining the attendance of a person in workplaces, educational campuses, etc has been increasing with the increasing technological innovations. In this section, we will discuss the methodology for our model. Our model primarily aims to detect the face and recognize the person for the entry of the attendance. For this, various advanced deep neural network algorithms have been utilized and the best performing approach has been considered for further implementations. In addition to this, the model system is incorporated using the cloud architecture thereby aligning with the convenience and demand of the current ecosystem.

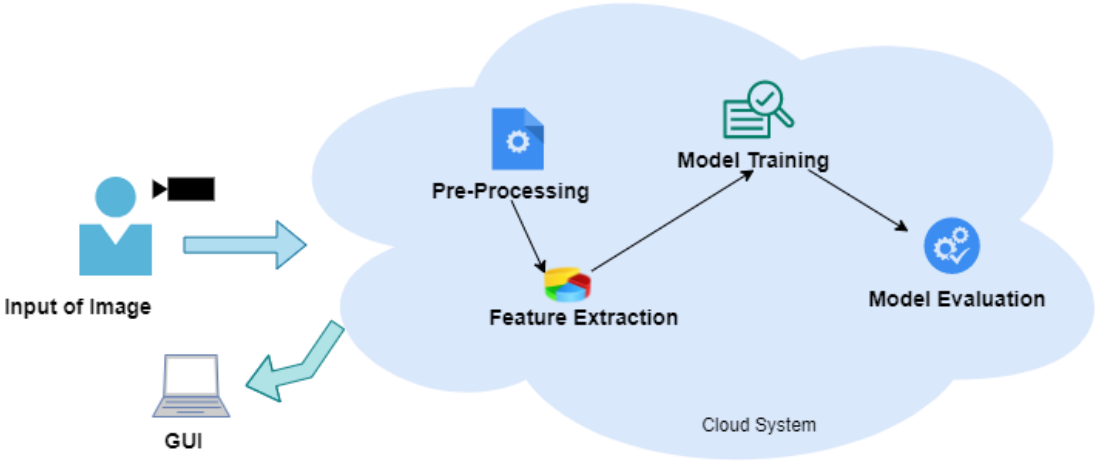



Figure 3: Model Methodology

## 3.1 Acquisition of Dataset

To implement and train the model, we have sourced the dataset from the Kaggle repositories. The sourced dataset consists of five celebrity images with image samples of different variants such as postures, angles, lightings, etc. In addition to this, the pixel dimensions for each image sample in the dataset vary. This obtained dataset is named as 5 Celebrity Faces Dataset taken from kaggle *5 Celebrity Faces Dataset* (n.d.). It consists of 118 image samples for which the training set and validation set are compiled before. The dataset had samples for five celebrities in which the training samples consists of 14-20 images foreach celebrity and validation samples consisted of 5 images for each celebrity.

## 3.2 Pre-Processing of the Dataset

In this phase of pre-processing, the samples obtained are balanced and standardized for the efficacy of the model implementations. Initially, the image samples are converted intothe RGB format. Once the completion of the conversion, the crucial pre-processing part is implemented by the utilization of Multi-Task Cascaded Convolutional Neural Network (MTCNN). The algorithm is an upgraded and robust alternative to the previously util- ized Viola-Jones Detector. Through the MTCNN algorithm, the facial presence by the standardized features and vectors is obtained from the image's samples. It works in the

three detailed procedures which are Proposal Network (P-Net), Refine Net (R-Net), and Output Net (O-Net). Here, P-Net is a primary stage of facial detection from the image samples. Then in the second stage which is R-Net, the false positive is filtered out and the boundary region over the image sample is constructed. Then the final stage which is O-Net, the detected facial regions and boundary regions are validated. Ultimately, only the main features of the face are considered such as nose, ear, eyes, lips, etc out of which the unnecessary pixels are trimmed out.

## 3.3 Extraction of the feature

After the phase of pre-processing, the features of the image samples are extracted. In our model, we have utilized the Face Net framework relying upon the deep learning architecture to extract the dominant features of the facial image samples. Although there are two types of approaches to extract the features of the image samples. In the first take, the DNN algorithms such as VGG-19 and ResNet-50 can be implemented which will both extract features and train the model. The merit of this approach is that the model will provide the output in a single step. Considering the second take, the FaceNet framework is implemented for feature extraction. The image samples are fed into the framework and the feature vectors are obtained with a vector dimension of 128. These vectors obtained are known as embedding. The figure depicts the image representation of the Face Net framework. Also, the label of the obtained input samples was categorical for which the method of one-hot encoding was implemented to avail a binary snippet.

## 3.4 Training of the Model

For our model of the attendance management system, we have considered various DNN algorithms. We have implemented three pre-trained DNN algorithms for our model which are VGG-19, ResNet-50, and Dense Net. Each algorithm was implemented and evaluated to implement in the model. VGG-19 is a pre-trained convolutional neural network with pre-defined weights of 19 layers. This framework is trained with the implementation of multiple images samples from the ImageNet dataset. Similarly, the ResNet-50 is alsoa pre-trained convolutional neural network trained using the ImageNet database. This framework consists of 48 convolutional layers with 1 Max Pool layer and 1 Average Pool layer. On the other hand, Dense Net is a pre-trained densely connected convolutional network. It uses dense blocks over the ResNet-50 architecture. The advantage of this algorithm is that it can counter the challenges of vanishing gradients which affects the accuracy of the model.

## 3.5 Evaluation of the Model

After the implementation of the algorithms, these are assessed by utilizing the appropriate evaluation metrics. The metrics used in our study are Accuracy, Loss, Precision, Recall, and F-1 score. Using these metrics, each algorithm is evaluated, and the best performing algorithm is considered for further implementation. The accuracy of the model defines the number of correct predictions as the total number of predictions. The higher is the accuracy of the model, the more effective the model is. On the other hand, the loss is the number of incorrect predictions which depicts the weakness of the model. Accuracy and Loss are anti-proportionate in comparison. Considering the confusion matrix which

defines the metrics such as precision, recall, and f-1 score. Here, the precision depicts the number of true positives to the number of classified positives whereas the recall depicts the number of true positives to the number of actual positives. These metrics have an overall impact on the efficacy of the trained model. On the other hand, the f-1 score is the weighted average of precision and recall.

# 4 Design Specification

In this section, we will discuss the architecture of each algorithm implemented in our model. Three DNN algorithms are implemented which are VGG-19, ResNet-50, and Dense Net.

## 4.1 VGG-19

The first algorithm implemented in our model is VGG-19. This is a pre-trained algorithm that relies on the convolutional neural network with pre-defined weights of 19 layers. It is trained by feeding the ImageNet database which contains a voluminous set of image samples. The 19 layers are further distributed into 16 convolutional layers, 3 fully connected layers, 5 max pool layers and 1 SoftMax layer. The input image dimension for his algorithm is 224x224 pixels. In every layer, the ReLu activation function has been established with the dropout function. Here, the dropout function counters the chal- lenges of overfitting in the model. The image representation of the VGG-19 architecture is depicted in the figure 4.

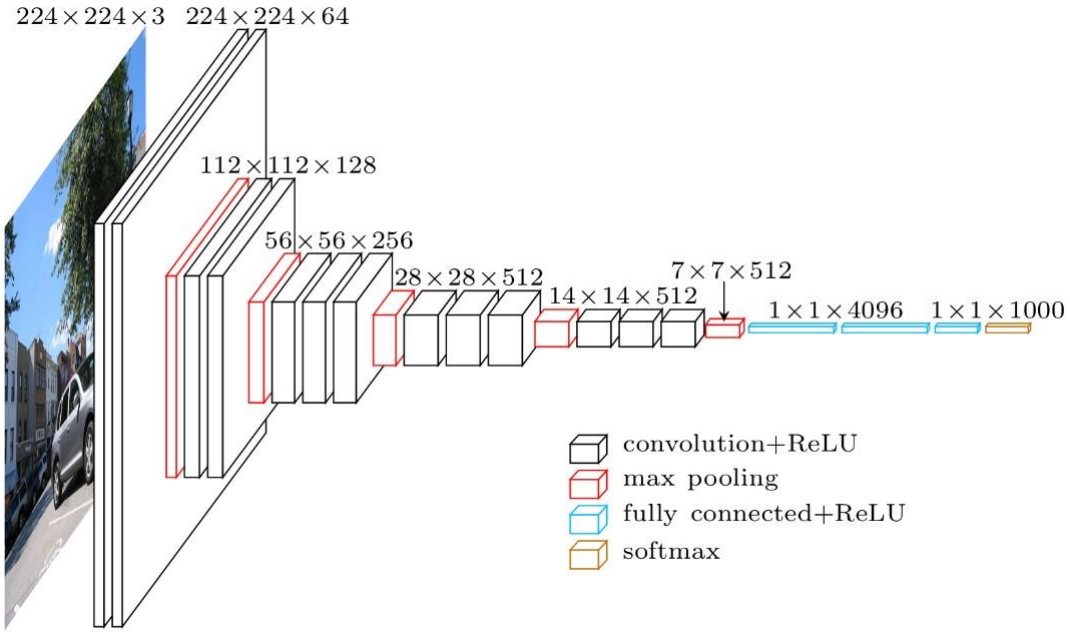



Figure 4: VGG-19

## 4.2 ResNet-50

Like the VGG-19, the second algorithm which is ResNet-50 is a pre-trained convolutional neural network. It is also known as the residual network. Although, unlike the first algorithm, it has pre-defined weights with 50 layers. It is also trained by feeding the ImageNet repositories to the database. The 50 layers in this model are distributed in 48 convolutional layers (with 34 residual layers), 1 max pool layer and 1 average pool layer. The following figure 5 represents the ResNet-50 architecture in image representation. Through its novel technique of skip connections, it can overcome the challenge of vanishing gradient by skipping the internal layers and directly connecting to the output layer. This algorithm is called under the library of Keras.

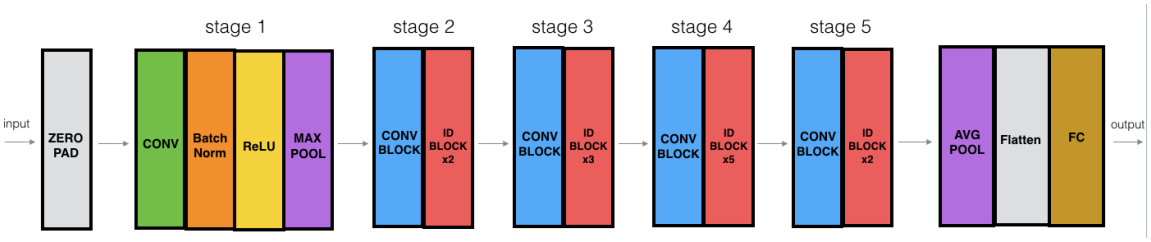



Figure 5: ResNet-50

## 4.3 Dense Net

On the other hand, the Dense Net algorithm is a robust and upgraded model of the ResNet algorithm. It is also a pre-trained convolutional neural network with a dense block layer in addition. This algorithm has higher computational capacity but for this, it also requires a higher GPU requirement. Only the major uphold here is that it connects the convolutional layer densely. Also, this algorithm overcomes the challenge of vanishing gradient. In addition to these, this algorithm is also trained using the ImageNet Database for further implementations. The figure 6 represents the graphical representation of the Dense Net architecture.

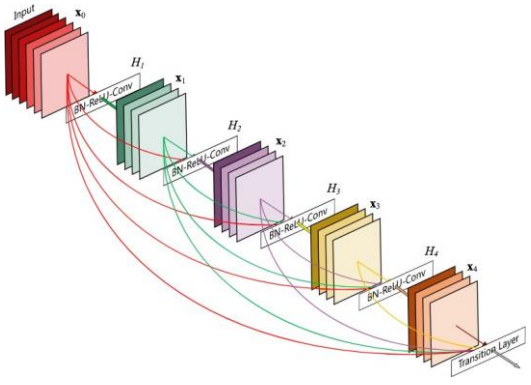



Figure 6: DenseNet

# 5 Implementations

In our study, the preliminary aim is to develop an effective model for attendance management systems utilizing the face recognition module. To develop and implement a robust model, we have utilized the three pre-trained DNN algorithms which are VGG- 19, ResNet-50, and Dense Net. Each algorithm is implemented and evaluated using the evaluation metrics such as Accuracy, Loss, Precision, Recall, and F-1 score. For the implementation of the model, we have obtained the dataset from the Kaggle repositories of the 5 Celebrity Face Dataset *5 Celebrity Faces Dataset* (n.d.). Then using the MTCNN framework the obtained samples are pre-processed. After the phase of pre-processing, the Face Net framework is utilized in our case for the implementation of feature extraction. Primarily for the Graphical User Interface (GUI), the Tkinter and various python libraries are utilized for the development. The prominent library in our model is the OpenCV library which guides the detects and recognized the facial presence in the provided samples. For the effective implementation of the model, we have utilized multiple libraries and modules in our model such as NumPy, Pandas, Matplotlib, Keras, MTCNN, os, cv2, etc. In addition to this, Google Collab is utilized for the implementation interface which provides the provision of GPU at a certain threshold with a hard disk space of 15GB. Primarily, in the user interface the face of a person is recognized, and entry is noted for attendance management. This can be utilized for student attendance. Initially, when the face is detected, the model recognizes whether the student is a new or existing one as observed in the Figure 7.

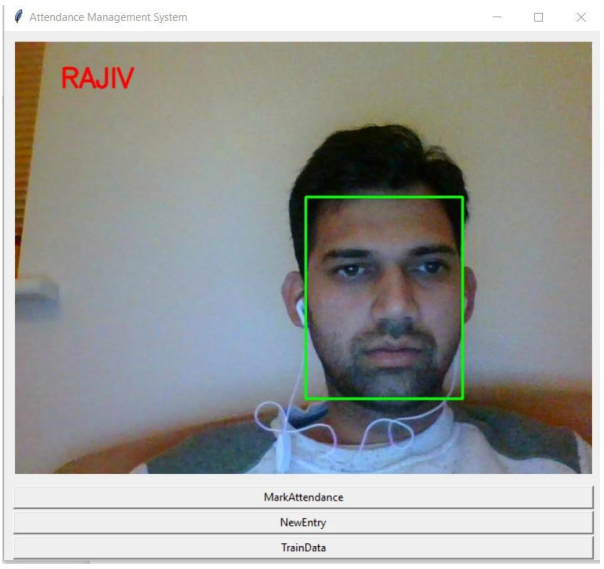



Figure 7: Application Recognizing the Trained Face

To add the facial data for the new attendance, the labelled new entry is selected which processes to recognize the person. Here, around 50 image samples are calculated where captured images will be splitted into train and test set. During the previous implementation, the obtained optimal model is considered to train the final output. Once the face is trained and stored in the database, the student can make their attendance through the system. When new student comes infront of the camera our system mark them as new student as shown in Figure 8. The trained model and the data of attendance also will be stored in the cloud. The student needs to click the mark attendance label

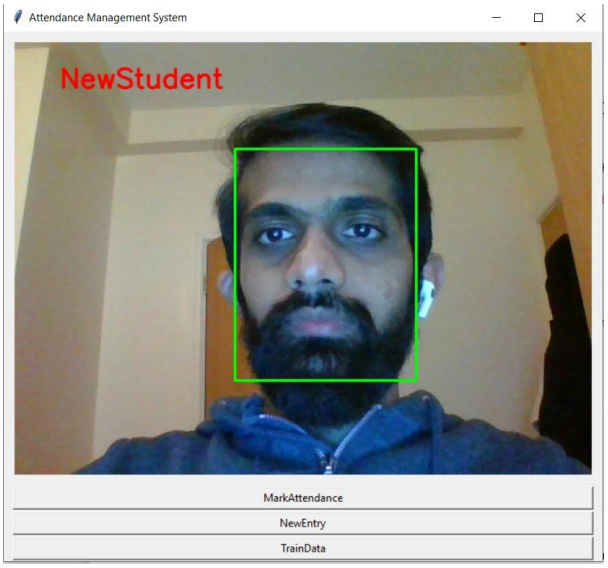
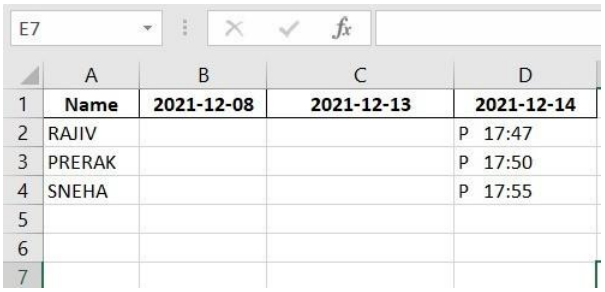



Figure 8: Application Recognizing the New Face

which will capture the attendance of the student and store the Enter the attendance in an excel sheet automatically. In the sheet considering the figure, the detailed report ofa student can be obtained such as the date and timing with the attendance marking as shown in Figure 9



| | A | B | C | D |
|---|---|---|---|---|
| 1 | Name | 2021-12-08 | 2021-12-13 | 2021-12-14 |
| 2 | RAJIV | | | P 17:47 |
| 3 | PRERAK | | | P 17:50 |
| 4 | SNEHA | | | P 17:55 |
| 5 | | | | |
| 6 | | | | |
| 7 | | | | |

Figure 9: Attendance management in Excel Sheet

The system with following configuration is required to run the machine and train the deep learning architecture. However, for this work mainly google colab has been utilized to train the deep learning model, as google colab provide the GPU without any additional cost.

| | |
|---|---|
| Operating System | Windows, Linux |
| GPU | Provided by Google Collab |
| RAM | 16 GB |
| Hard Disk | 108 GB |
| Programming Language | Python |
| Application Interface | Google Collab |
| Library Implemented | Tkinter, OpenCV, pandas, numpy, matplotlib, Keras |

Table 2 : System Specification

# 6 Evaluation

In this section, each implemented algorithms which is VGG-19, ResNet-50, and Dense Net are assessed utilizing the evaluation metrics. For the study, we implemented the metrics such as Accuracy, Loss, Precision, Recall, and F-1 score. As discussed, accuracy is the number of correct predictions over the total predictions whereas the loss is the number of incorrect predictions over the total predictions. On the other hand, Precision is the ratio of true positives over the classified positives and Recall is the ratio of true positives over the actual positives, whereas the F-1 score is the weighted average of the precision and recall. Each outcome of the algorithms through metrics is discussed and the optimal model is considered here.

## 6.1 Experiment 1 / Evaluation Based on Accuracy

In this section, the accuracy of the three pre-trained DNN algorithms which are VGG- 19, ResNet-50, and Dense Net is assessed and evaluated to find the optimal algorithmto implement in the model. For every algorithm implemented, it is observed that the input size of the image is set at 160x160 pixels. For the activation function, both the ReLu and SoftMax function was considered. Here, the SoftMax function standardizes the input for the multi-dimension output whereas the ReLu function rectifies the linear unit function. Also, the Adam optimizer is implemented in the model. To counter the chances of overfitting, the dropout function was set at 0.5. The accuracy for both the training and testing set was considered. During the processing of the model, the epoch was 25 with a batch size of 4. For the VGG-19, the highest training accuracy achieved was 80.8% and the highest testing accuracy achieved was 84% whereas for the ResNet-50,the highest training accuracy achieved stood at 77.75% and the highest testing accuracy stood at 94%. In the Dense Net algorithms, the highest training accuracy stood at 94.88% and the highest testing accuracy stood at 96%. Considering the figure, it was observed that the highest training accuracy was achieved by Dense Net and the lowest training accuracy was by ResNet-50. On the other hand, the highest testing accuracy was achieved by Dense Net and the lowest was by VGG-19. This showed that in termsof accuracy the Dense Net algorithm was optimal for the implementation in the model.

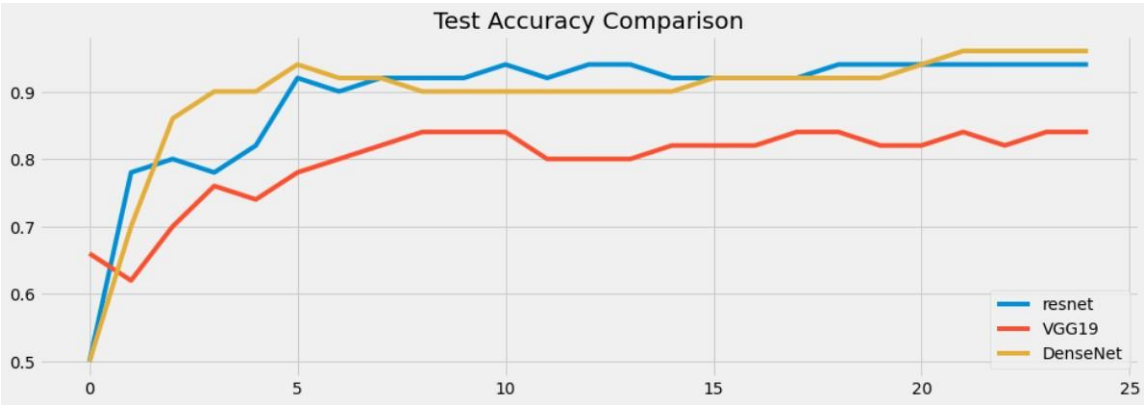



Figure 10: Accuracy Comparison

## 6.2 Experiment 2 / Evaluation Based on PRF Score

After evaluating the other metrics, the metrics such as precision, recall and f-1 score are assessed. These metrics depend on the outcomes of the confusion metrics such as true positive, true negative, false positive, and false negative. Here precision is the ratio of true positives over the total number of classified positives. Considering the figure, the VGG-19 algorithm achieved a training precision of 1 and a testing precision of 0.4615 whereas the ResNet-50 algorithm achieved a training precision of 0.9998 and a testing precision of 0.443. On the other hand, both the training and testing precision for the Dense Net algorithm stood at 1. Through this, it was observed that the Dense Net hadthe highest testing precision whereas the ResNet-50 achieved the lowest testing precision.

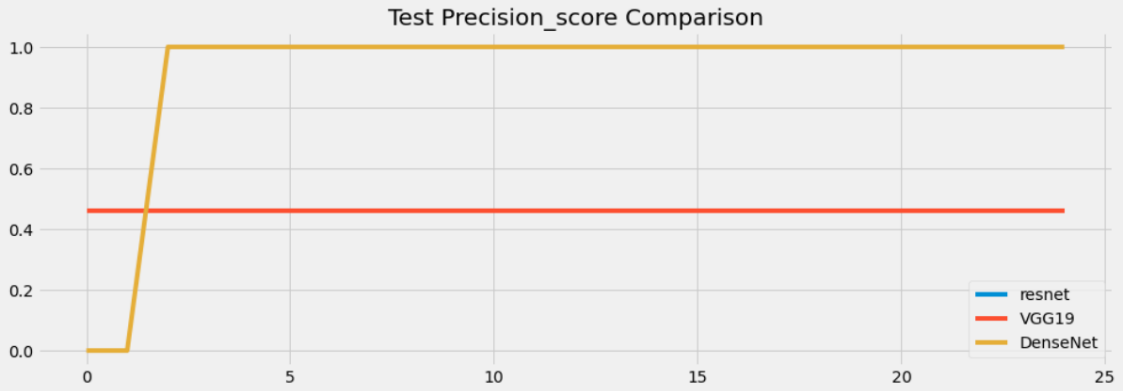



Figure 11: Precision Comparison

After this metric, recall is considered which is the ratio of the true positive over the total number of actual positive. As in the figure, the training and testing recall achieved by VGG-19 are 0.9969 and 0.443, respectively whereas for the ResNet-50 algorithm it stood at 1 and 0.4615, respectively. For the Dense Net algorithm, the training recall stood at 0.9968 and the testing recall stood at 0.9545. It was observed that the highest testing recall was achieved by the Dense Net algorithm whereas the lowest testing recall was achieved by the VGG-19 algorithm.

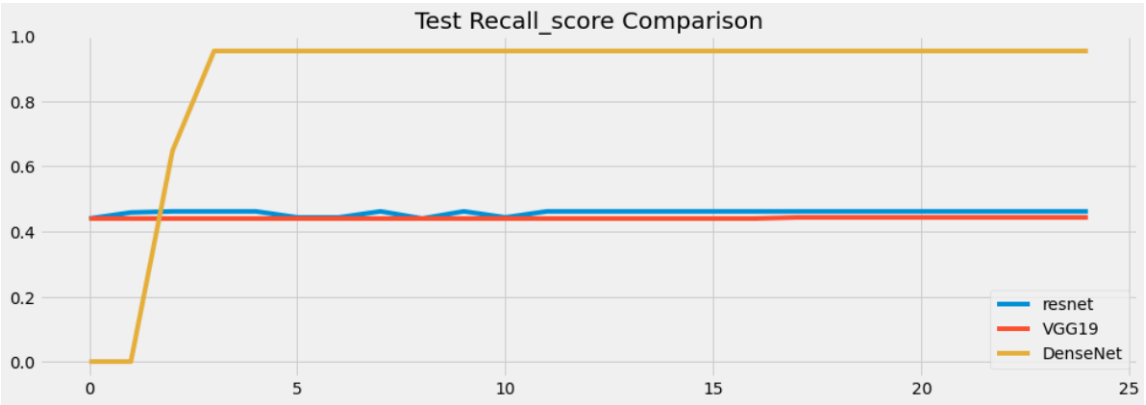



Figure 12: Recall Comparison

The final metric is the f-1 score that is the weighted average of the precision and recall. A detailed comparison was shown for this in the figure. The VGG-19 algorithm scored an

f-1 score of 0.9855 and 0.452 for the training and test set, respectively. Whereas for the ResNet-50, the training f-1 score stood at 0.9974 and the testing f-1 score stood at 0.4615. Considering the Dense Net algorithm, the training f-1 score was 0.9935 and the testing f-1 score was 0.9762. Like other metrics, here also the highest recall was achieved by the Dense Net algorithm whereas the lowest recall was achieved by VGG-19. The following metrics showed that the Dense Net is an optimal algorithm whereas the VGG-19 is a weak algorithm, and the ResNet-50 was considered a moderate algorithm.

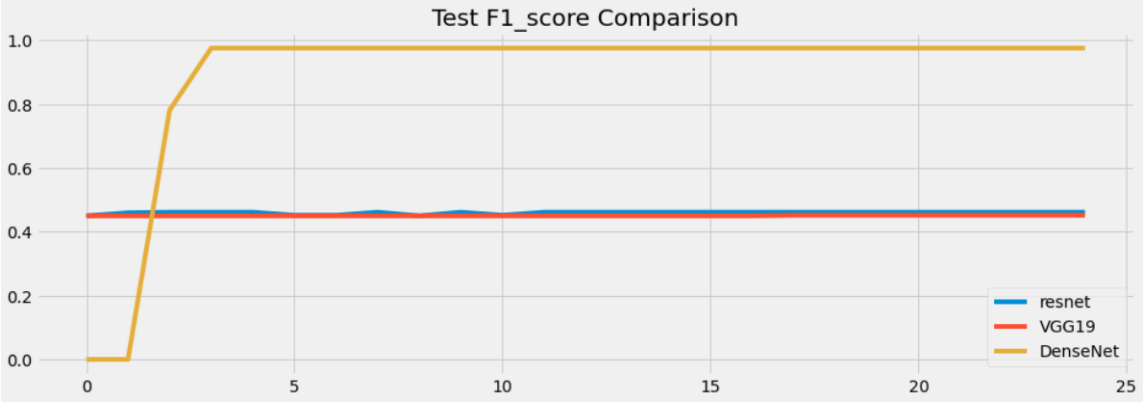



Figure 13: F-1 Score Comparison

## 6.3 Experiment 3 / Evaluation Based on the Loss

After the evaluation of accuracy, the metric comes as the loss value. Loss is the number of incorrect predictions over the total prediction. This metrics is generally inverse tothe accuracy. The lower is the loss value, the higher is the accuracy of the model. The model is optimal only when the accuracy is higher, and the loss value is the lowest. The training loss value of VGG-19 is 0.6392 and the testing loss value is 0.5373 whereas for the ResNet-50, the training loss value is 0.7272 and the testing loss value is 0.2549. On the other hand, the Dense Net loss value for the training set is 0.2199 and the testing setis 0.173. With the overall outcome in the figure, it is observed that the Dense Net had the lowest testing loss value whereas the VGG-19 had the highest loss value.

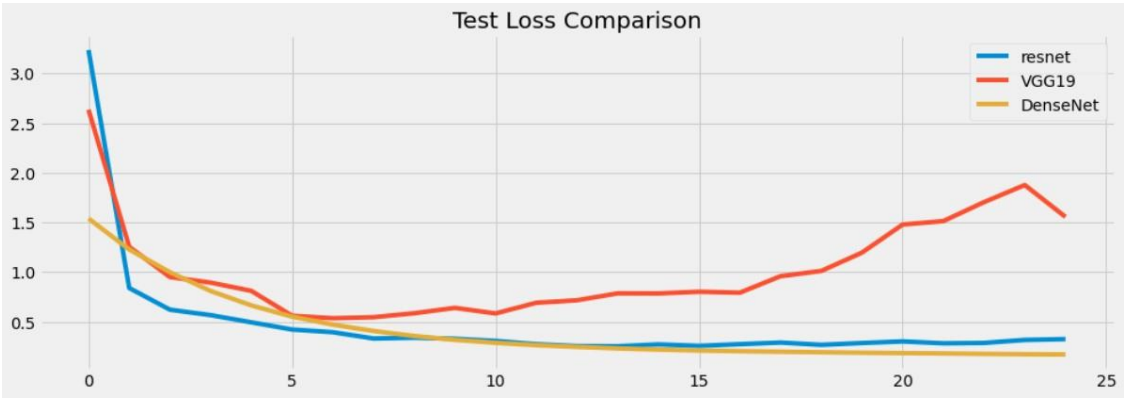



Figure 14: Loss Value

## 6.4 Discussion

Through our study, we aimed to develop a robust model for the attendance management system. Our models detect the face and recognize the student with existing data avail- able in the database. For the effective output, three pre-trained DNN algorithms were utilized which are VGG-19, ResNet-50, and Dense Net. Here, the VGG-19 and ResNet-50 algorithms could extract the features from the given dataset and the model was trained instantly. Whereas for the Dense Net algorithm, the features were extracted through the Face net framework and then the algorithm was implemented. The algorithms were evaluated using the evaluation metrics such as accuracy, loss, precision, recall, and f-1 score to obtain an optimal algorithm. The best-performing and trained algorithms were stored in the system and considered for further implementations. With final comparison, it is depicted that the Dense Net algorithm outperforms the other algorithms which are VGG-19 and ResNet-50. Our studied model showed an effective output for the scope of real-world application. In addition to this, the user interface was developed using the Tkinter library for the convenience to use the program. Furthermore, to overcome the large requirement of GPU computation, the Google Collab was utilized which provideda free range of GPU computation.

# 7 Conclusion

The growing technology in the domain of the workplace or educational environment has urged the need to adopt novel approaches. In our study, we have introduced an attendance management system for students using an effective facial recognition system. Three pre-trained DNN algorithms which are VGG-19, ResNet-50, and Dense Net was implemented in the model and then the best-performing algorithm was evaluated using the appropriate metrics. Each algorithm was evaluated using the metrics such as accuracy, loss, precision, recall, and f-1 score. With the experiments in our study, it was concluded that the most optimal algorithm for our model is Dense Net. In our model, for pre-processing of the dataset we have utilized the MTCCN framework whereas, for the feature extraction of the dataset, the Face Net framework was utilized. Through the implementation of these frameworks, our model was much more effective during the implementations. Our model could efficaciously recognize a person with existing data and direct the person to upload its facial images if the person using the model the first time. The trained model and the user database were stored in the cloud system which was transferred to the local machine on request. The trained model was effective and convenient enough to implement in a real-time environment. Although, the model could be enhanced for a better outcome using various approaches such as tuning of parameters, using multiple advanced parameters and algorithms, etc in the future scope of work. This model could be implemented in various other real-time applications soon.

# References

*5 Celebrity Faces Dataset* (n.d.).
**URL:** *https://kaggle.com/dansbecker/5-celebrity-faces-dataset*

.A, V., Gupta, A., Bharadwaj, A., Srinivasan, A., Balasubramanya Murthy, K. and Sub-

ramanyam, N. (2018). Deep learning on binary patterns for face recognition, *Procedia Computer Science* **132**: 76–83.

Adamu, A. (2019). Attendance management. . . attendance management system using fingerprint and iris biometric, pp. 427–433.

Alankar Patil, Priya K. P, P. M. A. J. A. R. K. (2020). *Attendance Monitoring using Face Recognition and Machine Learning*, pp. 94–102.

Arsenovic, M., Sladojevic, S., Anderla, A. and Stefanović, D. (2017). Facetime—deep learning based face recognition attendance system.

Ayub Khan, A., Laghari, A., Awan, S., Lyari and Karachi, P. (2021). Machine learning in computer vision: A review, *ICST Transactions on Scalable Information Systems* .

Bansal, A. (2018). Attendance management system through fingerprint, *International Journal for Research in Applied Science and Engineering Technology* **6**: 2140–2148.

Desai, V. (2018). Attendance management using biometrics with sms alerts, *International Journal for Research in Applied Science and Engineering Technology* **6**: 3468–3471.

Elaskari, S., Elaskri, A. and Almasoudi, A. (2021). Using barcode to track student attendance and assets in higher education institutions, *Procedia Computer Science* **184**: 226–233.

Islam, K. T., Raj, R. and Al-Murad, A. (2017). Performance of svm, cnn, and ann with bow, hog, and image pixels in face recognition.

J, H., Sudarsan, A., Sadashiv, A. and Ajai, A. (2019). Vision-face recognition attendance monitoring system for surveillance using deep learning technology and computer vision, pp. 1–5.

Khan, A. (2020). Machine learning in computer vision, *Procedia Computer Science* **167**.

Kuang, W. and Baul, A. (2020). *A real-time attendance system using deep learning face recognition.*

Mohan, G. and Prasad, D. (2020). Machine learning classifier model for attendance management system, pp. 1034–1039.

Nyein, T. and Oo, A. (2019). University classroom attendance system using facenet and support vector machine, pp. 171–176.

R.S, D., S, A., P, A., Joshua, P. and G, E. (2020). Smart attendance system using face recognition, *Journal of Advanced Research in Dynamical and Control Systems* **12**: 1079–1084.

Sarkar, P., Mishra, D. and Sai Subrahmanyam, G. (2019). *Automatic Attendance System Using Deep Learning Framework*, pp. 335–346.

Shekhar, H., Seal, S., Kedia, S. and Guha, A. (2020). *Survey on Applications of Machine Learning in the Field of Computer Vision*, pp. 667–678.

Supunya, R. (2018). *IOT enabled Recognition based attendance Management System*, PhD thesis.

Tapyou, K., Chaisil, P. and Muangprathub, J. (2021). Smart school attendance system using face recognition with near optimal imaging, pp. 1–5.

Vinod, V., Murugesan, G., Mekala, V., Thokaiandal, S., Vishnudevi, M. and Siddharth, S. (2021). A low-cost portable smart card based attendance system, *IOP Conference Series: Materials Science and Engineering* **1012**: 012046.