Deep Learning-Based Facial Recognition for Surveillance Systems

Masapu Venkata. Pranav Sahith | Dr. K. Kavitha (Sr. Assistant Prof) | GMRIT



# Abstract:

The deep learning technique in artificial intelligence is providing remarkably effective solutions in the video surveillance sector, by providing real-time detection and recognition capabilities through improved data analysis. IP Camera Surveillance is a method that integrates IP (internet protocol) cameras with deep learning algorithms to improve video surveillance systems by analyzing real-time video footage for advanced monitoring, detection, and recognition features. This article reviews the use of an IP camera surveillance system with deep learning based face detection and recognition system in detail. The system employs convolutional neural networks (CNN) for the purposes of object detection and recognition in motion video clips in a real time mode. The design facilitates accurate detection and identification of complex objects in the video streams. The face prints of the persons will be stored inside the database with relevant statistics and does the face recognition. When any unknown face is recognized then alarm will ring so one can alert the security systems and in addition actions will be taken. The system testes how deep learning can transform traditional IP camera surveillance into a user-friendly, intelligent monitoring tool, providing enhanced security and efficiency for various applications.

# Keywords:

*Convolutional Neural Networks, IP Camera Surveillance, Deep Learning, internet protocol, face detection, face recognition.*

# Introduction:

People are developing to be intensely troubled in their surroundings they don't have guarantee that they are staying securely. Surveillance systems, particularly those using face recognition technologies, are now a significant part of security in various public and private spaces. The vigilance of people is setting in, thanks to the necessity for security surveillance. With this in mind, CCTV monitoring fixtures or spy cameras are put in every sector/activity where convenience or security is of great concern. A system for surveillance includes tracking a number of interactions and includes the use of monitoring device and face detection and recognition. As one of the technologies used in biometric identification systems, facial recognition is preferred because of easy information gathering and high level of correctness. Thus this technology is embraced in information management, country defense, and even road traffic control. Face recognition refers to a biometric software program that performs algorithms that maps out a person’s face and keeps it in its database as a faceprint. For the reasons of effective monitoring of subjects, while the surveillance system captures the image of the subject, the image also needs to be compared with the images of known profiles to identify the subject. Face recognition techniques are essential in these cases. The underlying technology for this system is entirely based on a Convolutional Neural Network (CNN), which belongs to the deep learning algorithms, which is one of the most sought after algorithm for dealing with images. CNNs are excellent classifiers, segmenters and recognizers of images making them suitable for any face recognition jobs. Once the system has been

trained with facial characteristics, it is able to recognize the registered individuals and raise the alarm upon detection of an unrecognized individual. The deep learning system as well as enhances the face detection and recognition system even further with the passage of time. Another model described by this far for accurate measures of faces is a FaceNet image compression to a 128 dimensional image compression level, which makes it easy to compare a live face with the one in the database. This feature allows it to prevent the border violation attempts by other individuals with very low chances of any errors by the security agents thereby improving the security of the area under surveillance. In deep learning, the term Convolutional Neural Network (CNN) is applied successfully because recent advances in CNN technology are allowing for the development of applications that support a very large amount of training datasets.

# Literature Review:

1. International Journal of Open Information Technologies ISSN: 2307-8162 vol. 12, no. 3, (2024) " Face detection and recognition in video surveillance systems” A.B. Mudrich, K.V. Ezhova.
   1. B. Mudrich, K. V. Ezhova proposed "Detection and Recognition of Faces in CCTV Systems" Using a deep learning approach, and the methods used here were CNN and RNN and they also presented a face detection and recognition system using a Haar cascade classifier and a support vector machine (SVM). proposed a dataset for face detection and recognition called the Flickr-Faces-HQ (FFHQ) dataset. The dataset contains 70,000 images of faces. The system achieved an accuracy of 95% on a dataset of 1000 images.
2. Deep Learning- Based Surveillance System using Face Recognition January 2020 ITM Web of Conferences 32:03011 32:03011 DOI[:10.1051/itmconf/20203203011](http://dx.doi.org/10.1051/itmconf/20203203011)

Divya Kapil, Aishwarya Kamtam, Akhil Kedare, Smita Bharne proposed "Deep Learning-Based Surveillance System using Face Recognition" using deep learning techniques for automated biometric surveillance used CNN, a pre-trained FaceNet model, which converts a facial image into a 128-dimensional vector encoding, Triplet Loss Function. system was trained using a dataset where 50 frames per user were captured and fed into the model. In total, 2000 images. Achieved high accuracy face recognition model.

1. Gómez-Bautista, Andrés David, and Francisco Carlos Calderón-Bocanegra. "Enhancing facial recognition in surveillance systems through embedded super-resolution." Revista Facultad de Ingeniería Universidad de Antioquia 112 (2024): 98-110.

Andrés David and Francisco Carlos propoed "Enhancing Facial Recognition in Surveillance Systems through Embedded Super-Resolution", presents a method to improve facial recognition by enhancing image resolution using a super-resolution technique. Used Efficient Sub-Pixel Convolutional Neural Network (ESPCN) and NVIDIA Jetson TX2, for enhancing the resolution of facial images, The training used a dataset of 22,000 synthetic images generated by Generative Adversarial Networks (GANs). Image Quality Improvement, Face Distance Metric are achieved.

1. Personal Identification Based on Deep Learning Technique Using Facial Images for Intelligent Surveillance SystemsAugust 2019 9(4):465-470 DOI[:10.18178/ijmlc.2019.9.4.827](http://dx.doi.org/10.18178/ijmlc.2019.9.4.827)

Van-Dung Hoang proposed "Personal Identification Based on Deep Learning Technique Using Facial Images for Intelligent Surveillance Systems" Face Detection, Convolutional Neural Network (CNN), Data Augmentation, trained the augmented dataset using CNN and performed both feature extraction and classification. As input consists of low- resolution, 192x192x3 facial images collected from surveillance cameras.

1. Deep Learning Architectures for Face Recognition in Video Surveillance July 2019 DOI[:10.1007/978-981-10-](http://dx.doi.org/10.1007/978-981-10-5152-4_6) [5152-4\_6](http://dx.doi.org/10.1007/978-981-10-5152-4_6)

Saman Bashbaghi and Eric Granger proposed "Deep Learning Architectures for Face Recognition in Video Surveillance" used Cross-Correlation Matching CNN (CCM-CNN), Trunk-Branch Ensemble CNN (TBE-CNN), HaarNet, Canonical Face Representation CNN (CFR-CNN) methods as input high-quality still images and lower- quality video frames captured under uncontrolled conditions.

1. Deep Learning Face Detection and Recognition June 2019 Conference: International Conference on Automation Science and Engineering (ICASE-2019), At: Giza, Egypt <https://www.researchgate.net/publication/336472655_Deep_Learning_Face_Detection_and_Recognition>
   1. T. Khalil, and A. S. Mohra proposed "Deep Learning Face Detection and Recognition"

The proposed method for face recognition in this paper is to use Convolutional Neural Networks (CNNs) and deep neural network face recognition classifications. They performed several experiments on thr Face96 database and their proposed method got 99.67% of accuracy. Input consists of 1000 images with a resolution of 112x92 pixels, The database was divided into a 70% training set and a 30% validation set.

1. International Journal of Open Information Technologies ISSN: 2307-8162 vol. 12, no. 3, Singh, A., Bhatt, S., Nayak,

V. et al. Deep Learning based Intelligent Surveillance System sep 2020 <https://doi.org/10.1051/itmconf/20203203011> Rashid Amin and Hamza Aldabbas proposed "Deep Learning based Intelligent Surveillance System", It pre- processing techniques, like feature extraction. The system is trained using SVM classifier and validated using four datasets. aims to achieve high accuracy in classifying human activities, taking 1000 datasets under gone by pre- processing techniques achieved accuracy and efficiency in recognizing human activities

1. Deep Learning Based Real-Time Face Recognition System June 2022 NeuroQuantology 20(6):7355-7366 20(6):7355-7366 DOI:[10.14704/nq.2022.20.6.NQ22737](http://dx.doi.org/10.14704/nq.2022.20.6.NQ22737)

Mohammed Y. Shakor and Ahmed Abdalla proposed "Deep Learning Based Real-Time Face Recognition System" used methods are Histogram of Oriented Gradients (HOG), facial landmark estimation technique, pre-trained deep learning model based on CNN, Support Vector Machine (SVM). As for input they took 40 images of 8 different people (5 per person) and trained. The system achieved an accuracy of 96.88% on both still images and video frames

1. Face Detection for Video Surveillance-based Security System March 2020 IEEE Access PP(99):1-1 PP(99):1-1 Olena Yakovleva proposed "Face Detection for Video Surveillance-based Security System" used methods are MTCNN (Multi-Task Cascaded Convolutional Networks), FaceBoxes, DSFD (Dual Shot Face Detector), RetinaFace (with ResNet50 and MobileNet0.25 as backbone networks), CenterFace, SCRFD-500MF (Single-Stage Cascade

Residual Face Detector). As per inputs images included face sizes from 20x20 to 310x310 pixels and head rotations from -90 to +90 degrees in both the X and Y axes. The system achieved an accuracy of 97.88% on both still images and video frames

1. Singh, A., Bhatt, S., Nayak, V. et al. Automation of surveillance systems using deep learning and facial recognition. Int J Syst Assur Eng Manag 14 (Suppl 1), 236–245 (2023).

Arpit Singh and Vishal Nayak proposed "Automation of surveillance systems using deep learning and facial recognition" used methods are Convolutional Neural Networks (CNN) Specifically, they used the VGGFace, Single Shot Detector (SSD) For face detection, Support Vector Machines (SVM). self-made dataset of 7,500 images created by converting video frames into images, Real-time video footage from surveillance cameras. facial recognition results in real time, with recognition accuracy ranging from 78.54% to 100%, averaging 96% accuracy.

1. Jie xu proposed "A deep learning approach to building an intelligent video surveillance system" September (2022). Jie xu proposed "A deep learning approach to building an intelligent video surveillance system" spatial pyramid pooling networks (SPPnets), and region-based convolutional neural networks (R-CNNs) for object detection tasks, 3- dimensional convolutional networks (3D ConvNets) and video salient object detection (VSOD), As per input took own dataset of images. The system archives key results, including timestamps, categories, and confidence scores for detected objects, and timestamps with recognition scores for unidentified faces
2. Smart Surveillance System using Deep Learning Dayana R, Suganya M, Balaji P, Mohamed Thahir A & Arunkumar P September (2022).

Dayana R, Suganya M, Balaji proposed "Smart Surveillance System using Deep Learning" methods used are Convolutional Neural Networks (CNNs) to detect and recognize faces from video feeds, Principal Component Analysis (PCA) used for extracting facial features and Data Augmentation. Datasets like the Mega Face Challenge and Labeled Faces in the Wild (LFW) are used for training and testing the model. The system achieved accurate classification of detected faces.

1. A Deep Learning based Human Detection and Tracking for Security Surveillance Systems March 2023 Applied and Computational Engineering 2(1):569-577 2(1):569-577 DOI[:10.54254/2755-2721/2/20220606](http://dx.doi.org/10.54254/2755-2721/2/20220606)

Muhammad Asif and Arfa Hassan proposed "A Deep Learning based Human Detection and Tracking for Security Surveillance Systems" used methods are Multi-Task Cascaded Convolutional Neural Networks (MTCNN), Proposal- Net (P-Net), Refine-Net (R-Net), Output-Net (O-Net) and FaceNet Model. system used 50 test video sequences captured in indoor and outdoor environments via CCTV and mobile cameras. The system achieved 97% accuracy, with outputs displayed as bounding boxes around the detected individuals along with their assigned IDs

1. Face recognition in real-world surveillance videos with deep learning method June 2019 DOI[:10.1109/ICIVC.2017.7984553](http://dx.doi.org/10.1109/ICIVC.2017.7984553) Conference: 2017 2nd International Conference on Image, Vision and Computing (ICIVC).

Ya Wang and Ming Zhu proposed "Face Recognition in Real-world Surveillance Videos with Deep Learning Method" used method are Face Detection and Tracking, Graph Clustering, Fine-tuning VGG Face Model. Used Surveillance Video Frames and VGG Face Model Features as inputs. The fine-tuned VGG model achieved an accuracy of 92.1% on the real-world surveillance dataset, an improvement from 83.6% before fine-tuning.

1. Intelligent Video Surveillance System using Deep Learning Neha Kardile, Rutuja Deshmukh, Vaibhav Kalhapure, Project Guide: Prof. Devidas Jaybhay International Research Journal of Engineering and Technology (IRJET) May 2022.

Neha Kardile, Rutuja Deshmukh, Vaibhav Kalhapure proposed "Intelligent Video Surveillance System using Deep Learning" Probabilistic Neural Network (PNN) and Convolutional Neural Network (CNN) for activity recognition are used methods. As per input they used video streams of surveillance system and achieved

# Methodology:

1. **The FaceNet architecture:**

The face recognition system employs primarily this technique. This is a deep learning face representation model developed by Google for face recognition, verification as well as clustering. What it does, in fact, is to convert in input face images into a 128-dimensional vector which assists in telling apart varying faces by their normal features.

# Minimizing Triplet Loss:

The Facenet model is trained with this triplet loss function. It ensures that the model germane to the tasks performed distinguishes between like and unalike faces with the ease of mapping them in an embedding space.

**Anchor:** A picture that shows the person to focus on.

**Positive:** Another photograph of the same person.

**Negative:** A photo of another individual.

The main of this model is to reduce the distance between the anchor and positive images and increase the distance between the anchor and negative images.

# One-Shot Learning:

One Shot Learning model is a machine learning model, what this model will do is that it will make the model Learn to generalizefrom few examples. It differs from the normal machine learning method in such a way that it will not equire so many inputs in order to reach the goal state, a few specified number of inputs will be Provided and the most optimum model will be Given.

# CNN:

The Convolutional Neural Network (CNN) is a type of deep learning model that is predominantly used in the field of image processing. In this instance, it is employed along with the facenet model for the purpose of identifying and

categorizing faces. The role of the CNN is to help in processing the already extracted features and searching for face prints in the available database.

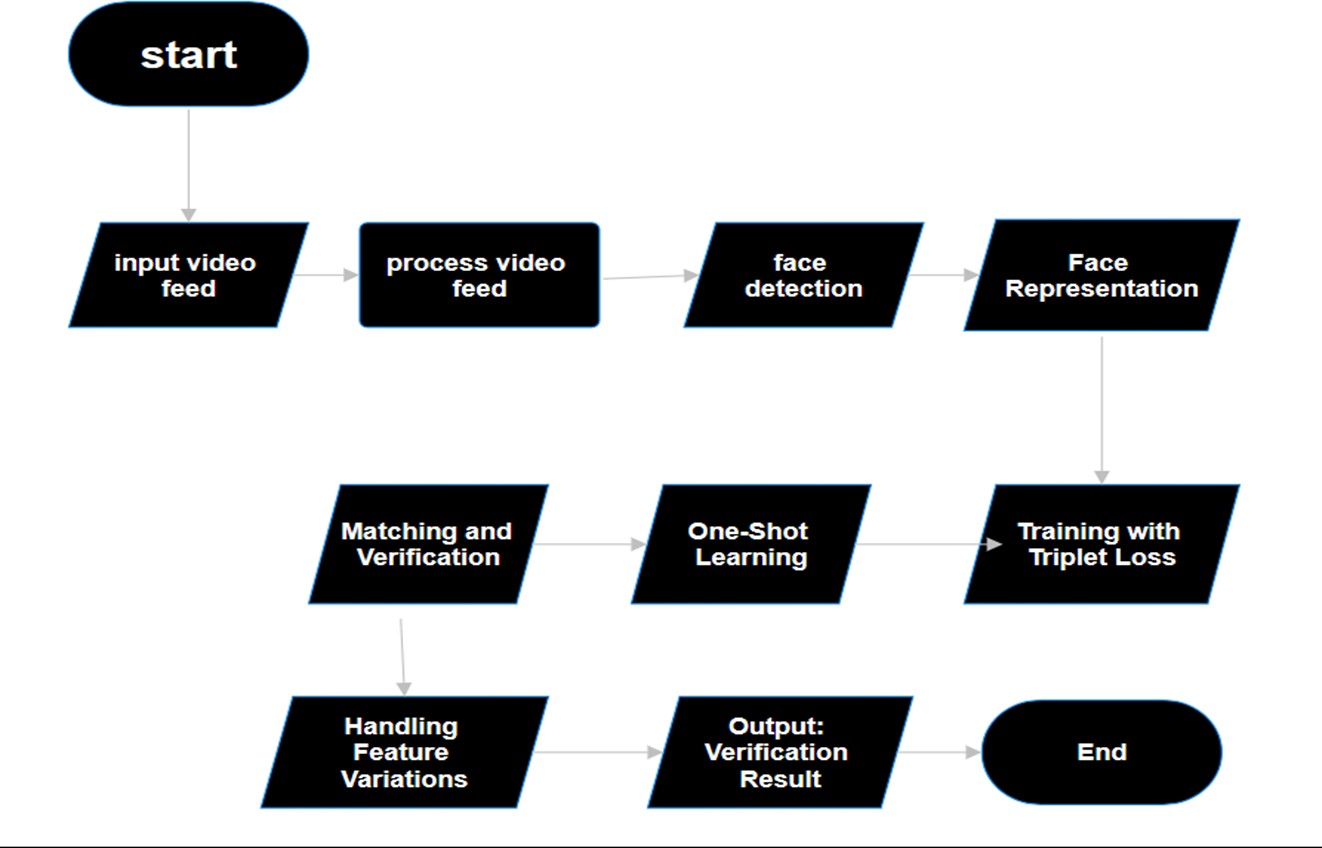
# Real-Time Implementation:

The system operates in real time, capturing images from surveillance cameras then processing facial images for comparison with a database of stored images for identity verification.

# Handling Variations in Features:

This is where the system is assessed for other different aspects like movements of the face such as wearing glasses and not.

# Work flow:



First we will start from the input video feed which is the input that we will give like image or video footage, persons in it and after that we will process the video feed and now here will use our CNN and detect the wheather it is a person or not and by here we will get clarity that it is a person or not and after that we will use FaceNet model by which a 128-dimensional vector will drawn around person’s face now we will use triplet loss function and compare the face of same and different person, by now we will get the result means the face detection will be occured but inorder to get more precission we will train the model with one shot learning and by doing this we can train the model with less number of input and get accurate and presiced output. And we will repeat this process this is how the system works.

# Face detection models:

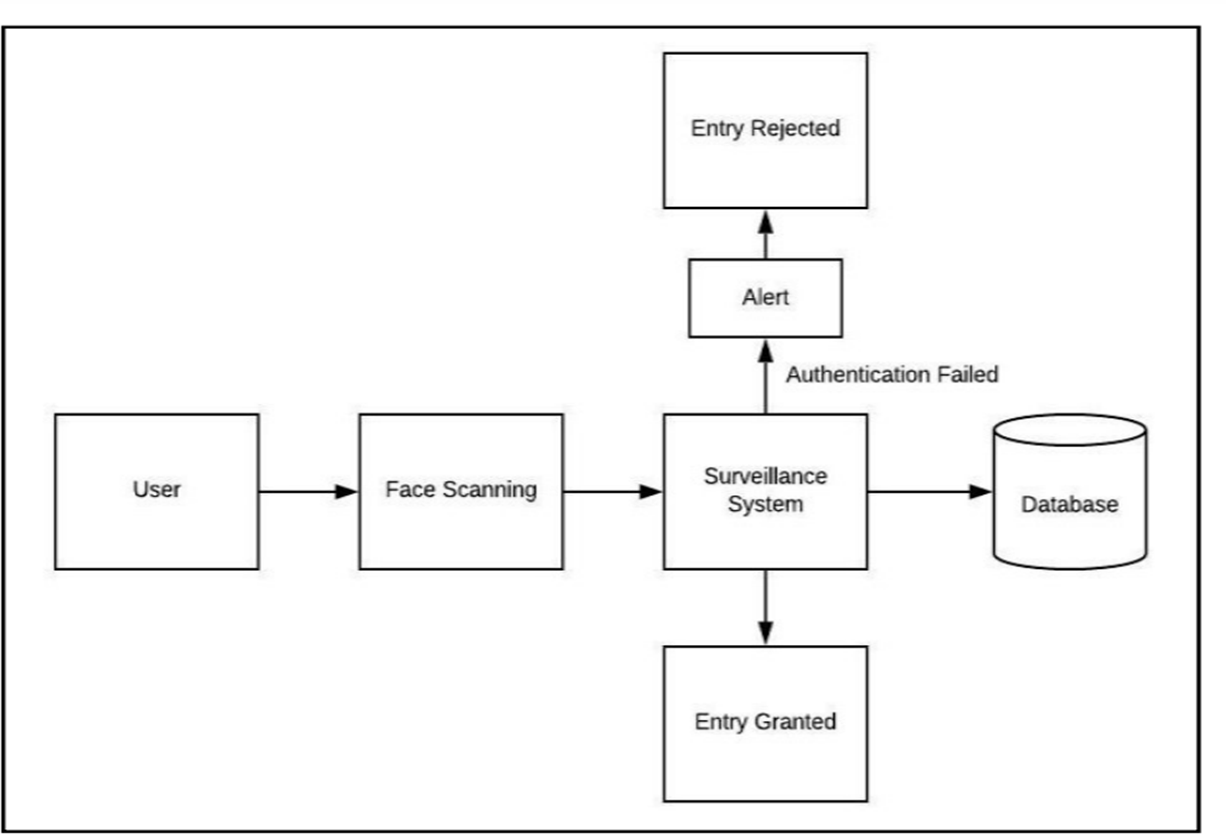
In this research, many face detection methods such as cnn, facenet model, openCV, yolo can be applied. Nevertheless the main emphasis is given on Facenet model due to its focus on face recognition, verification and clustering tasks with remarkable accuracy and efficiency.

# Facenet model:

Google’s FaceNet is a system for face recognition, face verification and face clustering based on deep learning. It implements a Convolutional Neural Network to generate a 128-dimensional embedding from an image of a face, this embedding contains the information about the unique characteristics of the face. The main idea of the system is a Triplet Loss function, which tunes the embeddings by pulling together pictures of the same persons and pushing apart photographs of various people. FaceNet is also fast, and it performs well even under different lighting and facial expression conditions hence it finds applications in security, biometric systems, and social networks.

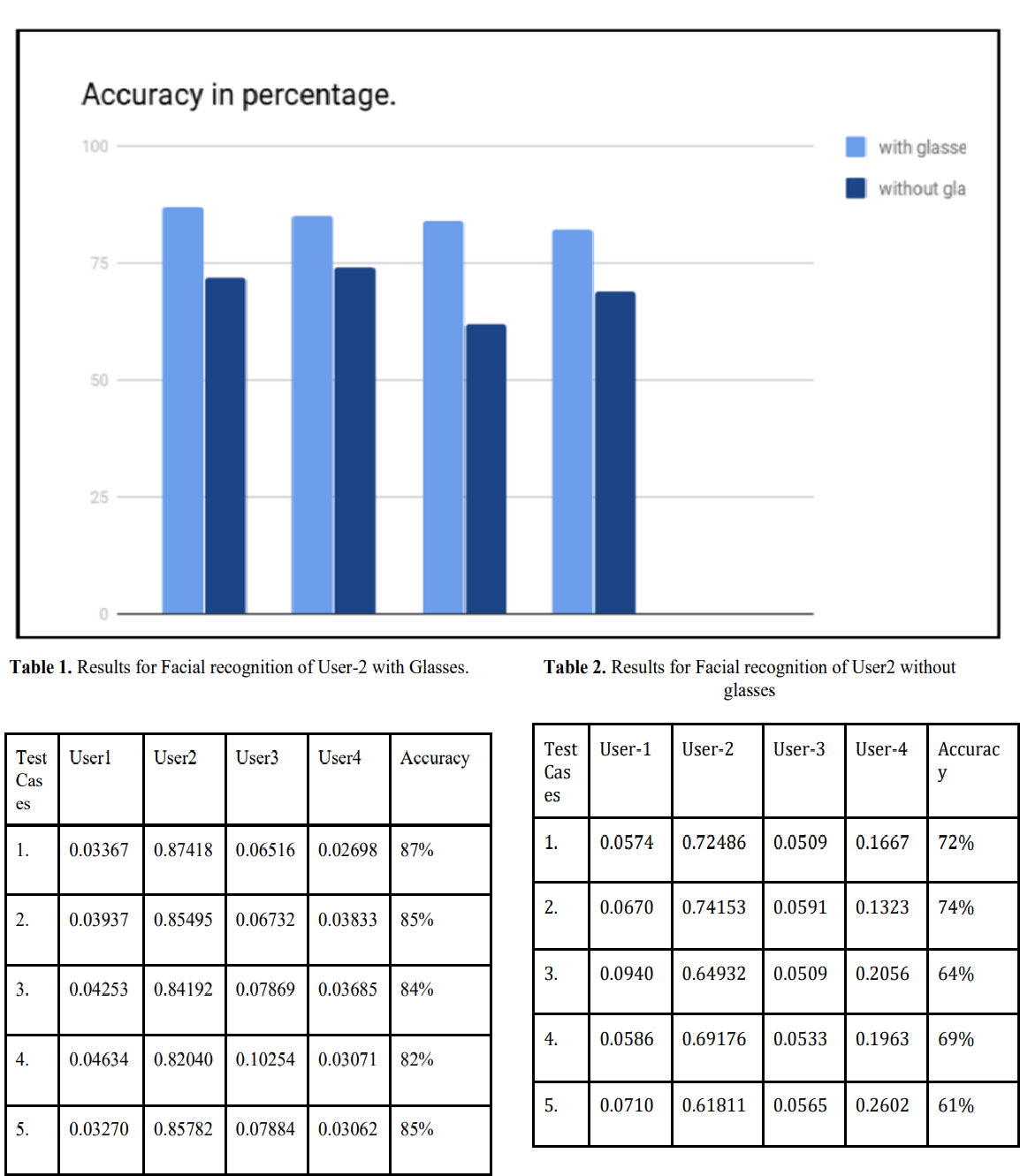
# Facenet Architecture:

**System Design:**



The user will scan the face and the surveillance system that we did will perform it’s job by applying the methodology and will grant the person if his face is stored in the data base and if not it will give use an alert that this face is not registered and for future purpose it will store the data of the unrecognized person face if he entered again. This is how the face recognisation system were developed and work, we can use this system in many ways we can use it in malls for cctv’s and we use for our house for more guranteed security or we can also use it for anamoly detection.

# Results:



**Conclusion:**

Utilities integrating graceful learning algorithms and especially the FaceNet architecture with Convolutional Neural Networks (CNN) has revealed evident enhancement in the real-time face recognition systems for IP camera-based surveillance. The two-dimensional facial images are mapped to 128-dimensional embeddings and a triplet loss function is applied, making it difficult for the identification system to mix between known and unknown faces effectively and efficiently. Real time characteristics mean that there is an immediate action like alarm ringing without

detection of the faces that are recognized which is an additional security and surveillance measure to it. It however also fits in well with changing face features and conditions making it effective when used in face monitoring systems that are mobile. Additionally, the application of one-shot learning improves the performance of the system by allowing it to learn from few examples. This strategy improves the efficiency of such surveillance systems as well as their responsiveness, and there is a clear scope for improvement of this technology in the context of intelligent video surveillance systems.

# References:

1. International Journal of Open Information Technologies ISSN: 2307-8162 vol. 12, no. 3, (2024) " Face detection and recognition in video surveillance systems” A.B. Mudrich, K.V. Ezhova.
2. Singh, A., Bhatt, S., Nayak, V. et al. Automation of surveillance systems using deep learning and facial recognition. Int J Syst Assur Eng Manag 14 (Suppl 1), 236–245 (2023).
3. Gómez-Bautista, Andrés David, and Francisco Carlos Calderón-Bocanegra. "Enhancing facial recognition in surveillance systems through embedded super-resolution." Revista Facultad de Ingeniería Universidad de Antioquia 112 (2024): 98-110.
4. Smart Surveillance System using Deep Learning Dayana R, Suganya M, Balaji P, Mohamed Thahir A & Arunkumar P September (2022).
5. Automation of surveillance systems using deep learning and facial recognition Arpit Singh, Saumya Bhatt, Vishal Nayak, Manan ShahInternational Journal of System Assurance Engineering and Management 14 (Suppl 1), 236-245, (2022).
6. Deep Learning Based Real-Time Face Recognition System June 2022 NeuroQuantology 20(6):7355-7366 20(6):7355-7366 DOI:[10.14704/nq.2022.20.6.NQ22737](http://dx.doi.org/10.14704/nq.2022.20.6.NQ22737)
7. Personal Identification Based on Deep Learning Technique Using Facial Images for Intelligent Surveillance SystemsAugust 2019 9(4):465-470 DOI[:10.18178/ijmlc.2019.9.4.827](http://dx.doi.org/10.18178/ijmlc.2019.9.4.827)
8. Deep Learning Architectures for Face Recognition in Video Surveillance July 2019 DOI:[10.1007/978-981-10-](http://dx.doi.org/10.1007/978-981-10-5152-4_6) [5152-4\_6](http://dx.doi.org/10.1007/978-981-10-5152-4_6)
9. Deep Learning Face Detection and Recognition June 2019 Conference: International Conference on Automation Science and Engineering (ICASE-2019), At: Giza, Egypt <https://www.researchgate.net/publication/336472655_Deep_Learning_Face_Detection_and_Recognition>
10. International Journal of Open Information Technologies ISSN: 2307-8162 vol. 12, no. 3, Singh, A., Bhatt, S., Nayak, V. et al. Deep Learning based Intelligent

Surveillance System sep 2020 <https://doi.org/10.1051/itmconf/20203203011>

1. Deep Learning based Intelligent Surveillance System May 2020 International Journal of Advanced Computer Science and Applications 11(4) 11(4) DOI[:10.14569/IJACSA.2020.0110479](http://dx.doi.org/10.14569/IJACSA.2020.0110479)
2. Smart Surveillance System using Deep Learning May 2020 International Journal of Recent Technology and Engineering (IJRTE) 9(1):1151-1155 9(1):1151-1155 DOI[:10.35940/ijrte.A1464.059120](http://dx.doi.org/10.35940/ijrte.A1464.059120)

[13]A Deep Learning based Human Detection and Tracking for Security Surveillance Systems March 2023 Applied and Computational Engineering 2(1):569-577 2(1):569-577 DOI[:10.54254/2755-2721/2/20220606](http://dx.doi.org/10.54254/2755-2721/2/20220606)

1. Face recognition in real-world surveillance videos with deep learning method June 2019 DOI[:10.1109/ICIVC.2017.7984553](http://dx.doi.org/10.1109/ICIVC.2017.7984553) Conference: 2017 2nd International Conference on Image, Vision and Computing (ICIVC)
2. Intelligent Video Surveillance System using Deep Learning Neha Kardile, Rutuja Deshmukh, Vaibhav Kalhapure, Project Guide: Prof. Devidas Jaybhay International Research Journal of Engineering and Technology (IRJET) May 2022