REAL-TIME FACE DETECTION USING MACHINE LEARNING

A PROJECT REPORT

Submitted by,

Mr. Geet Raj- 20201LCS0009
Mr. Nikith Raj S -20201LCS0010
Mr. Dhruva H B- 20201CSE0373
Mr. K Mohammed Afzal Khadir- 20201CSE0441

Under the guidance of,

Dr. PRAVINTH RAJA S

in partial fulfillment for the award of the degree of

BACHELOR OF TECHNOLOGY

IN

COMPUTER SCIENCE AND ENGINEERING

At

GAIN MORE KNOWLEDGE
REACH GREATER HEIGHTS

PRESIDENCY UNIVERSITY

BENGALURU

JANUARY 2024

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

CERTIFICATE

This is to certify that the Project report **"REAL-TIME FACE DETECTION USING MACHINE LEARNING"** being submitted by GEET RAJ, NIKITH RAJ S, DHRUVA H B, K MOHAMMED AFZAL KHADIR bearing roll number(s) 20201LCS0009, 20201LCS0010, 20201CSE0373, 20201CSE0441 in partial fulfilment of requirement for the award of degree of Bachelor of Technology in Computer Science and Engineering is a bonafide work carried out under my supervision.

Dr. PRAVINTH RAJA S
Associate Professor
School of CSE and IS
Presidency University

Dr. Pallavi R
Associate Professor & HoD
School of CSE
Presidency University

Dr. C. KALAIARASAN
Associate Dean
School of CSE and IS
Presidency University

Dr. L. SHAKKEERA
Associate Dean
School of CSE and IS
Presidency University

Dr. Md. SAMEERUDDIN KHAN
Dean
School of CSE and IS
Presidency University

PRESIDENCY UNIVERSITY

SCHOOL OF COMPUTER SCIENCE AND ENGINEERING

DECLARATION

We hereby declare that the work, which is being presented in the project report entitled **REAL-TIME FACE DETECTION USING MACHINE LEARNING** in partial fulfilment for the award of Degree of **Bachelor of Technology** in **Computer Science and Engineering**, is a record of our own investigations carried under the guidance of **Dr. S Pravinth Raja, Associate Professor, School of Computer Science and Engineering, Presidency University, Bengaluru.**

We have not submitted the matter presented in this report anywhere for the award of any other Degree.

NAME	ROLL NUMBER	SIGNATURE
Geet Raj	20211LCS0009	
Nikith Raj S	20211LCS0010	
Dhruva H B	20201CSE0373	
K Mohammed Afzal Khadir	20201CSE0441	

ABSTRACT

The past decade has witnessed significant advancements in face identification, driven by novel techniques in pattern recognition and computer vision. Notably, the integration of deep learning and convolutional neural networks (CNN) has greatly enhanced the accuracy of face identification methods. This progress is particularly pivotal in our interconnected society, where face recognition plays a crucial role in security systems, authentication processes, and crime detection. Closed-circuit television (CCTV) systems, prevalent in today's world, serve diverse functions. Real-time face recognition within these systems involves comparing digital images or video frames of individuals with a stored database, meeting the escalating demand for personal identification in our global society. As CCTV technology, frequently employed in video surveillance systems, serves purposes like data security, asset protection, and traffic flow analytics, facial recognition emerges as a critical component for identifying individuals and detecting abnormal activities.

Despite its importance, real-time face recognition in video surveillance faces limitations, especially in meeting the demands of swift and precise identification. To address this challenge, facial feature training has been enhanced through a CNN-based feature learning approach. CNNs, drawing inspiration from the human brain's architecture, exhibit remarkable efficacy in image identification and processing tasks. Their ability to interpret pixel information makes them a popular choice for training models in face detection applications, particularly excelling in tasks involving the extraction of facial features. This research contributes to overcoming the limitations of real-time face recognition in video surveillance, paving the way for more accurate and efficient identification processes.

ACKNOWLEDGEMENT

First of all, we indebted to the **GOD ALMIGHTY** for giving me an opportunity to excel in our efforts to complete this project on time.

We express our sincere thanks to our respected Dean **Dr. Md. Sameeruddin Khan**, School of Computer Science and Engineering, Presidency University for getting us permission to undergo the project.

We record our heartfelt gratitude to our beloved Associate Deans **Dr. Kalaiarasan C and Dr. Shakkeera L,** School of Computer Science and Engineering, Presidency University and **"Dr. Pallavi R"**, Head of the Department, School of Computer Science and Engineering, Presidency University for rendering timely help for the successful completion of this project.

We are greatly indebted to our guide **Dr.Pravinth Raja S, Associate Professor**, School of Computer Science and Engineering, Presidency University for his inspirational guidance, and valuable suggestions and for providing us a chance to express our technical capabilities in every respect for the completion of the project work.

We would like to convey our gratitude and heartfelt thanks to the University Project-II Coordinators **Dr. Sanjeev P Kaulgud, Dr. Mrutyunjaya MS** and also the department Project Coordinators "**Mohammed Zia Ur Rahman**", "**Mr. Peniel John Whistely**".

We thank our family and friends for the strong support and inspiration they have provided us in bringing out this project.

Geet Raj

Nikith Raj S

Dhruva H B

K Mohammed Afzal Khadir

LIST OF FIGURES

LIST OF TABLES

TABLE OF CONTENTS

CHAPTER-1

INTRODUCTION

Over the last decade, face detection has undergone a transformative journey, propelled by significant advancements in deep learning and convolutional neural networks (CNNs). These technological strides have not only addressed a fundamental challenge in computer vision and pattern recognition but have also elevated face recognition to a critical status in today's interconnected world. With applications ranging from crime identification to security systems and authentication, the need for accurate and efficient face detection solutions has become paramount.

Closed-circuit television (CCTV) systems have become ubiquitous in modern society, serving diverse purposes. At the core of these systems lies real-time face recognition, a crucial component that involves matching a person's face in a digital image or video frame against faces stored in a database. This capability has become indispensable in various fields, responding to the growing demand for individual identification in our increasingly globalized society.

Video surveillance systems, often leveraging CCTV technology, serve multifaceted purposes, including asset protection, data security, and analytics for crowd flux and congestion analysis. Face detection plays a central role in these systems, aiding in individual recognition and the detection of anomalous activities. However, despite its significance, real-time face recognition in video surveillance has faced challenges, particularly in meeting the demands of swift and accurate identification.

To overcome these challenges, researchers have turned to CNN-based feature learning models for training facial features. Inspired by th-e human brain's architecture, Convolutional Neural Networks have proven highly effective in image recognition and processing tasks. Their ability to analyze pixel information makes them the preferred choice for training models in face detection applications, enabling accurate extraction of facial features.

The integration of CNNs in video surveillance systems has significantly enhanced the capabilities of human detection, providing valuable insights for various applications, including the identification of potential terrorists or thieves. Despite these advancements, the real-time aspect of face recognition remains a substantial hurdle, demanding ongoing technological advancements to overcome latency issues and ensure timely responses to potential security threats.

This fusion of deep learning, particularly CNNs, with face detection technology has ushered in a new era of accuracy and efficiency. Video surveillance systems continue to play a pivotal role in security and analytics, making the ongoing refinement of real-time face recognition capabilities essential for meeting the demands of our dynamic and interconnected world.

As technology evolves, the synergy between deep learning and face detection not only improves accuracy but also addresses the need for efficiency in processing vast amounts of visual data in real-time. The continual evolution of real-time face recognition is vital in shaping a safer, more secure future in our increasingly complex and interconnected global landscape.

In conclusion, the marriage of deep learning, specifically CNNs, with face detection technology has not only elevated the accuracy and efficiency of these systems but has also opened avenues for innovation across various industries. The ongoing refinement of real-time face recognition capabilities is crucial for staying ahead in our ever-changing technological landscape, ensuring the security and safety of individuals in our interconnected and dynamic world.

CHAPTER-2

LITERATURE SURVEY

Before building your project, think of a literature survey as the architect's first sketch. It explores existing solutions, identifies weak points, and lays the foundation for your unique design. By scrutinizing scholarly works, you uncover knowledge gaps, refine your vision, and ensure your project tackles a truly new challenge. It's like climbing to the top of a research mountain - you get a panoramic view of the field, spot uncharted paths, and plan your ascent with confidence.

[1] The paper presents a solution to the challenges posed by low-resolution video surveillance systems, aiming to enhance human identification accuracy. It introduces a novel approach by integrating frontal gait recognition and low-resolution face recognition into an automatic platform. The system extracts frontal gait silhouettes and low-resolution face images from surveillance footage, generating feature vectors for subsequent training of separate classifiers. These classifiers specialize in frontal gait recognition and low-resolution face recognition, and their results are fused through score level fusion for improved accuracy. The literature review identifies shortcomings in existing studies, including the absence of suitable datasets for multimodal face and gait recognition, reliance on lateral gait views, and challenges in gait cycle detection. The proposed method addresses these issues by introducing an efficient gait recognition technique using frontal gait video clips. The approach combines both model-free and model-based feature extraction approaches to ensure robust identification. In the realm of low-resolution face recognition, the paper stands out by employing real surveillance data for training, enabling the system to learn a high-resolution mapping function and achieve superior classification accuracy. The proposed framework represents the first fully automatic multimodal recognition system evaluated on a relatively large dataset, promising advancements in accurate human identification from low-resolution video surveillance footage.

[2] The paper addresses the crucial challenge of human face detection and recognition in CCTV images, responding to the security demands faced by organizations. It recognizes the limitations of human-centric security and highlights the transformative impact of closed-

circuit television (CCTV) in creating intelligent surveillance systems. The increasing significance of facial recognition in diverse applications, including crime identification, security systems, and authentication, is underscored. Despite its widespread utility in image processing, animation, human-computer interface, and medicine, facial recognition systems encounter hurdles in real-world surveillance scenarios, such as image resolution issues, background clutter, lighting variations, and facial expressions.

The paper outlines the core steps of facial recognition systems, emphasizing the challenges posed by facial occlusions, diverse subject appearances, and the need for adaptability to changes in lighting, expressions, poses, and occlusions. The research contributions include a machine learning-based framework for effective face detection and recognition in CCTV images, a substantial dataset with 40,000 images under various conditions, and a comparative analysis of classical machine learning and deep learning algorithms for facial recognition. The literature review sets the context, introducing related works, and highlighting the evolving role of biometric systems, particularly facial recognition, in securing access control and financial transactions. The subsequent sections promise an in-depth exploration of related works and a comprehensive conclusion in Section 5.

[3] The current landscape of face recognition predominantly favors Convolutional Neural Networks (CNNs) over traditional feature-extraction methods, owing to their efficacy in addressing challenges like changes in facial expressions, illumination, poses, low resolution, and occlusion. Prominent CNNs such as DeepFace, FaceNet, ArcFace, and MagFace have complex architectures and high computational costs, limiting their real-time applicability on resource-constrained embedded devices. To address this, lightweight CNN architectures like MobileFaceNet, EfficientNet-B0, and GhostNet have emerged, catering to real-time implementation and limited resource environments. However, these lightweight models often struggle with issues like face rotation and low-level face inputs that are efficiently handled by their more complex counterparts.

The paper's main contributions encompass a detailed analysis of state-of-the-art (SOTA) lightweight architectures, specifically MobileFaceNet, EfficientNet-B0, and GhostNet, in the context of face verification challenges related to facial rotations and low resolutions encountered in video-surveillance camera applications. The evaluation leverages datasets such as Cross-Pose LFW (CPLFW) and QMUL-SurvFace, introducing a novel evaluation subset with 3000 facial images encompassing varying degrees of rotation and resolution. The findings reveal that EfficientNet-B0 excels in managing rotation and resolution challenges, while MobileFaceNet demonstrates superiority in handling extreme face rotations exceeding 80 degrees. The three-fold contributions include the creation of an evaluation subset, the analysis of SOTA lightweight architectures, and a methodical investigation into the impact of facial rotation and low resolution on face verification.

[4] In the realm of artificial intelligence and machine learning, computer vision poses a complex challenge that encompasses object recognition, shape determination, and, notably, facial recognition. The latter is a particularly intricate problem, intertwining various facets of artificial intelligence. The growing field of facial recognition holds immense potential in security, entertainment, and healthcare applications. Recent strides focus on multimodal approaches, leveraging diverse data modalities such as images, videos, and audio to enhance accuracy and efficiency.

This article delves into an overview of machine learning methods and tools geared towards facial recognition within a multimodal context. Techniques include deep learning, support vector machines, and decision trees, while tools like OpenCV, TensorFlow, and Keras contribute to system development. Despite the progress made in the last decade, developers still grapple with challenges ranging from face search and positioning to the determination of unique facial features and person identification. The literature review underscores the evolving landscape of facial recognition, emphasizing the need for continued research to refine methods, tools, and multimodal approaches, ultimately advancing the field's capabilities and applications.

[5] Face recognition, a prominent topic in computer vision, has witnessed advancements with deep-learning-based methods, surpassing human performance in various scenarios. However, addressing age variation remains a challenge, especially in scenarios like identifying missing children. This paper introduces a multi-task learning framework, MTLFace, aiming for Age-Invariant Face Recognition (AIFR) and Face Age Synthesis (FAS). Existing methods for AIFR fall into two categories: generative and discriminative models. Generative models transform faces of different age groups, while discriminative models extract age-invariant features. MTLFace combines both, utilizing attention-based feature decomposition to separate age- and identity-related features. It introduces an identity conditional module for personalized face synthesis and a selective fine-tuning strategy to enhance AIFR. Extensive experiments

showcase MTLFace's superior performance in AIFR and FAS, offering competitive results for general face recognition. Additionally, a new benchmark dataset is introduced for evaluating cross-age face recognition, emphasizing applications like tracing missing children. The paper extends and improves a preliminary version, demonstrating consistent enhancements in face recognition and synthesis.

[6] Face recognition systems have made remarkable strides, achieving human-level performance, yet their robustness against occlusions remains a challenge, as evaluation datasets typically consist of unobstructed images. The COVID-19 pandemic underscored the need for models adept at recognizing faces with specific occlusions, particularly due to the widespread use of facial masks. This led to a surge in research on masked face recognition (MFR) models, resulting in substantial advancements. However, a critical question emerges: can solutions developed for MFR enhance general occluded face recognition
This paper addresses this query, contributing to both MFR and occluded face recognition domains. The research investigates the generalization capabilities of MFR models to various types of occlusions, presenting advantages such as lower computational costs and potential performance improvements by leveraging knowledge from traditional face recognition methods. The findings reveal that MFR methods can be effectively deployed for general occluded face recognition, achieving reasonable performance. This interoperable deployability not only demonstrates the adaptability of MFR models but also suggests their potential utility in scenarios beyond masked face recognition, contributing to the broader field of occluded face recognition.

[7] The study addresses the evolving landscape of face recognition in computer vision, particularly with the advancements brought about by convolutional neural networks (CNNs). Face recognition typically involves three key steps: face detection, alignment, and representation. The paper emphasizes the significance of training and evaluating models to handle occlusions, a challenge not fully addressed by current datasets primarily consisting of unobstructed images. The emergence of facial masks as a common occlusion due to the COVID-19 pandemic prompted focused research on masked face recognition (MFR). However, the broader question arises regarding the generalization of MFR models to handle various occlusions beyond masks. The paper explores this by presenting contributions to both MFR and occluded face recognition, demonstrating the interoperable deployability of MFR

methods on occluded face recognition datasets. The research underscores the importance of developing models robust to diverse occlusions, considering the increasing adoption of face recognition across various technologies and scenarios, ranging from border control to smartphone authentication. The findings suggest potential applications beyond masked face recognition, contributing to the broader field of occluded face recognition.

[8] This paper addresses the persistent challenges in deep face recognition, emphasizing the critical data gap between academic research and industry applications. Despite notable advancements in neural network architectures and loss functions, existing public training sets such as MegaFace2 and MS1M exhibit limitations, hindering the exploration of the full potential of deep face recognition. To bridge this gap, the authors introduce WebFace260M, an unprecedented ultra-large-scale face benchmark boasting 4 million identities and 260 million faces. This dataset is curated using a scalable Cleaning Automatically by Self-Training (CAST) pipeline, eliminating the need for labor-intensive human intervention. The paper also introduces the Face Recognition Under Inference Time conStraint (FRUITS) protocol, catering to the real-world demand for time-constrained face recognition scenarios. Contributions encompass the creation of WebFace260M, the WebFace42M training set, and the FRUITS protocol, facilitating comprehensive evaluations and advancements in million-scale deep face recognition. Extensive experimentation, along with active participation in competitions, demonstrates the benchmark's efficacy in reducing the data gap between academic research and industry needs, fostering progress in real-world face recognition applications.

[9] This paper explores the convergence of real-time face recognition algorithms with hardware accelerators, addressing the pivotal role of biometric features, especially human faces, in identification and recognition tasks. Faces are preferred for their contactless nature and ability to handle multiple subjects simultaneously, finding indispensable applications in diverse settings like security, healthcare, and social media. The persistent challenges in existing systems, encompassing pose variations, facial expressions, and illumination changes, underscore the necessity for sophisticated hardware accelerators. The literature review critically assesses recent face recognition algorithms, emphasizing their suitability for real-time and portable applications, and rigorously evaluates their performance metrics. Notably, the paper extends its scrutiny beyond algorithmic considerations to include an extensive

examination of associated hardware accelerators, offering a comparative analysis of their efficiency.

[10] This study distinguishes itself by providing a holistic overview, addressing both algorithmic and hardware aspects, thus offering a comprehensive perspective. By integrating insights from the performance of real-time face recognition algorithms with the capabilities of hardware accelerators, the paper lays a foundation for optimizing systems in terms of accuracy, computation time, and power consumption. While existing review papers predominantly focus on algorithmic dimensions, this study uniquely contributes to the literature by explicitly examining the dedicated hardware accelerators crucial for achieving real-time face recognition capabilities.

[11] This paper addresses the limitations of traditional video surveillance systems by proposing a real-time image security device control system based on deep-learning facial recognition. Recognizing the escalating demand for crime prevention and public safety, the study emphasizes the need for automated, high-precision recognition of criminals in crime-vulnerable areas. The proposed system employs deep neural networks for facial detection and identification, aiming to overcome challenges associated with occlusions and avoidance reactions in crowded videos. Unlike conventional fixed-range facial recognition, the system dynamically tracks and verifies criminal faces, minimizing prediction overturns. The study introduces a scoring method based on face tracking to enhance identification accuracy and confidence. To optimize real-time processing, the paper employs down-sampling selectively during the face detection process, achieving a significant increase in frames per second (FPS). The proposed system aims to revolutionize surveillance, offering an advanced tool for early criminal detection and swift notification to relevant authorities, contributing to improved public safety. The contributions of the paper encompass innovative face detection, identification, and processing speed enhancement techniques, substantiated by comprehensive experiments.

The paper addresses the growing significance of humanoid robots in human–machine interaction, emphasizing the role of face recognition in enhancing social assistance provided by these robots. Face recognition technologies have evolved, with various methods such as PCA, Viola–Jones, LBPH, and deep learning algorithms like CNNs being employed in previous studies. Notably, the paper recognizes emotion recognition as a vital component of human–machine communication, with applications in humanoid robots discerning emotions from facial expressions. While existing studies excel in individual tasks, the paper proposes a holistic approach by integrating face recognition and emotion recognition into a unified system for real-time implementation on humanoid robots. This approach enables robots to recognize names, emotions, and positions of humans concurrently, fostering improved human–robot interaction. Unlike prior research relying on pre-collected datasets, the study uses primary data from male and female students, including those wearing glasses and hijab, contributing to the system's robustness and applicability. The performance of established CNN architectures (VGG16, AlexNet) is compared with a proposed modified architecture. Additionally, the paper introduces a method for measuring the distance between the human object's face and the robot's position, enhancing spatial awareness. The study represents a comprehensive effort towards advancing humanoid robot capabilities for real-time, multifaceted interaction with humans.

[12] The rapid advancement of computer vision technology, particularly in face recognition, has gained significant attention, especially during the pandemic, as touchless technologies become increasingly crucial. Face recognition serves as a prevalent attendance system in educational institutions, healthcare settings, and businesses. However, identifying individuals wearing masks presents a persistent challenge. Previous studies have explored mask detection using techniques such as Principal Component Analysis (PCA), YOLO V3 classification, Convolutional Neural Networks (CNNs), and Support Vector Machine (SVM) classifiers with face recognition feature vectors. The sensitivity of face recognition to mask usage is acknowledged, leading to difficulties in obtaining accurate identity information due to occlusion. Several studies have proposed solutions to ensure reliable object detection even in the presence of occlusion, employing techniques like centroid tracking, Particle Filter, Kalman Filter estimation algorithm, and greedy techniques. This paper introduces a novel method to enhance the accuracy of face mask recognition by employing the cosine distance technique, incorporating Haar-cascade and MobileNet for face bounding box detection. The study compares the performance of VGG16 transfer learning and the Triplet loss FaceNet with a multi-threading technique for face recognition, evaluating various combinations of methods to identify the most suitable approach.

[13] In the rapidly evolving era of the internet, computer technology has permeated various aspects of daily life and work. Face recognition technology, a fusion of artificial intelligence and computer science, has emerged as an innovative and widely applicable solution. As a vital means of distinguishing individuals, face recognition technology has found diverse applications in public safety, civil economy, and home entertainment. Traditional attendance systems, such as fingerprint and card-based systems, exhibit limitations, including error rates and inefficiencies. Face recognition systems offer higher accuracy and stability due to the multitude of facial data points, making them more reliable and efficient for attendance tracking, particularly in large-scale settings. Despite China's relatively late start in face recognition research, the nation has made significant strides, and the technology's commercial value in the era of big data ensures a bright future with substantial market demand.

The paper addresses the shortcomings of video-based face recognition (VFR), emphasizing challenges like image blur, posture changes, and occlusion. The proposed framework employs convolutional neural networks (CNNs) and innovative approaches such as artificial blur in training data to enhance the robustness of face representation. Additionally, the paper references studies on featureless face recognition, utilizing clustering techniques based on luminance distributions and distance metrics. The central focus is on designing a real-time face recognition attendance system, evaluating accuracy, stability, skip rates, and interface settings through practical experiments. The results demonstrate the system's effectiveness, showcasing the feasibility and innovation of the proposed algorithm in enhancing attendance management with improved reliability and efficiency.

[14] Pedestrian detection, a vital aspect of computer vision, holds significant relevance for applications like video surveillance and robotics. With challenges including variations in sizes, appearances, clothing, poses, occlusions, and complex backgrounds, convolutional neural networks (CNNs) have become integral for efficient feature generation. One-stage object detection aims for quick and accurate detection, playing a crucial role in fields like person re-identification. This paper proposes an efficient pedestrian detection model, emphasizing accurate bounding box detection for improved person re-identification. The Part-based Convolutional Baseline (PCB) and Visibility-aware Part Model (VPM) contribute to person re-identification, aligning images for better recognition. YOLOv3, a state-of-the-art object detection model, is adopted, but its reliance on the time-consuming Darknet53 backbone is addressed. Inspired by MobileNet and ShuffleNet, the paper proposes a more

efficient anchor box selection method, enhancing Intersection over Union (IOU) scores. The novel real-time pedestrian detection system utilizes a stack of shuffle units as the YOLOv3 backbone, emphasizing efficiency. Evaluation on the CrowdHuman dataset, featuring 22.6 pedestrians per image on average, demonstrates the effectiveness of the proposed criteria and the system's enhanced performance compared to original structures.

[15] In the landscape of face recognition, despite notable advancements driven by deep learning and large-scale datasets, the challenge of low-quality face recognition persists, particularly in unconstrained surveillance settings. Current datasets, primarily composed of celebrity images, lack representation of real-world surveillance scenes, leading to significant performance disparities on low-quality data. Two predominant categories of approaches have emerged in addressing low-quality face recognition: hallucination-based and embedding-based methods. Hallucination-based methods focus on reconstructing high-quality faces from low-quality images using super-resolution techniques, enhancing accuracy but at the cost of additional computation and reduced recognition speed. In contrast, embedding-based methods aim to implicitly map low-quality faces into an embedding space mirroring high-quality behavior, necessitating the transfer of effective knowledge.
Knowledge distillation (KD) emerges as an effective knowledge transfer method. This involves building two models with high- and low-quality face images and utilizing a distillation process to align the low-quality model with the high-quality model. However, existing KD methods primarily focus on forward propagation, neglecting significant differences in backpropagation, especially in class-specific gradient maps. Inspired by these observations, this paper proposes a texture-guided transfer learning method (TG) to address the backpropagation disparities between low- and high-quality models. The TG method utilizes class-specific gradient textures obtained from gradient maps, ensuring a more effective alignment of the two models. To enhance this approach, an attention map is introduced into TG, creating the soft-attention texture-guided method (SA-TG) to emphasize the importance of different regions. These contributions collectively form a comprehensive strategy to tackle low-quality face recognition challenges in real-world surveillance scenarios.

Table 2.1: Literature Survey

S.NO	Name of The Author	Journal Title	Year	Methods Used	Merit	Demerit	Accuracy
1	Maity S., Abdel-Mottaleb, M., & Asfour, S. S.	Multimodal low-resolution face and frontal gait recognition from surveillance video	2021	Deep Learning Techniques	Improved Performance in Challenging Enivironments	Reduced Recognition Accuracy	60.3%
2	Ullah, R., Hayat, H., Siddiqui, A. A., Siddiqui, U. A., Khan, J., Ullah, F., ... & Karami, G. M.	A real-time framework for human face detection and recognition in cctv images	2022	Haar Cascades	Enhanced Identification	Accuracy and False Positives/ Negatives	66.1%
3	Perez-Montes, F., Olivares-Mercado, J., Sanchez-Perez, G., Benitez-Garcia, G., Prudente-Tixteco, L., & Lopez-Garcia, O.	Analysis of Real-Time Face-Verification Methods for Surveillance Applications	2023	Convolution Neural Network(CNNs)	Accurate Identification	Accuracy and Reliability	53.9%
4	Basystiuk, N. M., & Rybchak, Z.	Machine Learning Methods and Tools for Facial Recognition Based on Multimodal Approach	2023	Convolutional Neural Network(CNNs)	Increased Recognition Accuracy	Data Security	72.6%
5	Huang, Z., Zhang, J., & Shan, H.	When Age-Invariant Face Recognition		Metric Learning	Improved Performance Across Age Groups	Ethical Concerns	71.9%

		Meets Face Age Synthesis	2022				
6	Neto, P. C., Pinto, J. R., Boutros, F., Damer, N., Sequeira, A. F., & Cardoso, J. S.	Beyond masks: On the generalization of masked face recognition models to occluded face recognition	2023	Transfer Learning	Realistic Scenarios	Limited Generalization	69.1%
7	Li, N., Shen, X., Sun, L., Xiao, Z., Ding, T., Li, T., & Li, X.	Chinese face dataset for face recognition in an uncontrolled classroom environment	2023	Uncontrolled Environment Simulation	Privacy Considerations	Lack of Diversit	62.2%
8	Zhu, Z., Huang, G., Deng, J., Ye, Y., Huang, J., Chen, X., ... & Zhou, J.	WebFace260M: A Benchmark for Million-Scale Deep Face Recognition	2022	Quality Control	Real-World Application	Data Privacy Concerns	71.2%
9	Baobaid, A., Meribout, M., Tiwari, V. K., & Pena, J. P.	Hardware accelerators for real-time face recognition: A survey	2022	Convolutional Neural Networks(CNNs)	Speed and Real-Time Processing	Scalability	65.5%
10	Kim, H., Choi, N., Kwon, H. J., & Kim, H.	Surveillance System for Real-Time High-Precision Recognition of Criminal Faces from Wild Videos	2023	Deep Learning-Based Detectors	Crime Prevention and Public Safety:	False Positives and Misidentification	68.3%
11	Dwijayanti, S., Iqbal, M., &	Real-time implementation of face recognition and		Convolutional Neural Networks(CNNs)	Human-Robot Interaction	Hardware Limitations	65.5%

	Suprapto, B. Y.	emotion recognition in a humanoid robot using a convolutional neural network. IEEE Access	2023				
12	Maharani, D. A., Machbub, C., Rusmin, P. H., & Yulianti, L.	Improving the capability of real-time face masked recognition using cosine distance	2020	Convolutional Neural Networks(CNNs)	Human-Robot Interaction	Hardware Limitations	73.8%
13	Yang, H., & Han, X.	Face recognition attendance system based on real-time video processing.	2020	Convolutional Neural Networks(CNNs)	Efficiency and Time-Saving	Accuracy Issues	72.4%
14	Xu, M., Wang, Z., Liu, X., Ma, L., & Shehzad, A.	An efficient pedestrian detection for realtime surveillance systems based on modified yolov3	2022	Convolutional Neural Networks(CNNs)	Real-Time Processing	False Negatives	59.4%
15	Zhang, M., Liu, R., Deguchi, D., & Murase, H	Texture-Guided Transfer Learning for Low-Quality Face Recognition. IEEE	2023	Convolutional Neural Networks(CNNs)	Improved Performance in Low-Quality Scenarios	Data Limitations	61.9%

CHAPTER-3

RESEARCH GAPS OF EXISTING METHODS

Based on the provided paper summaries, here are some potential research gaps in the existing methods:

3.1 Limited Multimodal Datasets:

[1] highlights the lack of suitable datasets for multimodal face and gait recognition. Many existing studies might not have access to diverse datasets that adequately represent real-world scenarios. This gap could hinder the development of robust multimodal recognition systems.

3.2 Challenges in Low-Resolution Face Recognition:

[1] emphasizes the challenges in low-resolution face recognition and the need for suitable training datasets. This suggests that existing methods may not be effectively addressing the specific issues related to recognizing faces in low-resolution surveillance footage.

3.3 Adaptability to Occlusions:

[6] and [7] both address the challenge of occluded face recognition, particularly due to the COVID-19 pandemic. Existing methods might not be sufficiently robust to handle various types of occlusions beyond masks, highlighting a potential gap in general occluded face recognition.

3.4 Efficiency of Lightweight Models:

[3]discusses the limitations of lightweight CNN architectures in handling face rotation and low-level face inputs. The research gap here could be the lack of lightweight models that effectively address specific challenges encountered in video-surveillance camera applications.

3.5 Age Variation in Face Recognition:

[5] focuses on age variation in face recognition, specifically in identifying missing children. This indicates a gap in existing methods in terms of addressing age-related challenges, especially in critical scenarios where accurate identification is crucial.

3.6 Real-World Application of Models:

[9] highlights the necessity of sophisticated hardware accelerators for real-time face recognition applications. There may be a research gap in the practical deployment of real-time face recognition algorithms with associated hardware accelerators in various settings like security, healthcare, and social media.

3.7 Large-Scale Face Datasets:

[8] addresses the limitations of existing public training sets for deep face recognition and introduces an ultra-large-scale face benchmark. The gap here might be the scarcity of comprehensive datasets that cover a diverse range of identities and scenarios for effective training and evaluation.

3.8 Robustness to Masked Face Recognition Models:

[12] discusses the challenges of face mask recognition and proposes novel methods. There might be a research gap in developing accurate and reliable face recognition models that can handle variations introduced by the presence of masks in real-world scenarios.

3.9 Humanoid Robot Interaction:

[11] introduces the integration of face recognition and emotion recognition for humanoid robots. The research gap could be the limited exploration of holistic systems that simultaneously recognize names, emotions, and positions of humans for improved real-time interaction.

3.10 Low-Quality Face Recognition:

[15] discusses the challenge of low-quality face recognition in unconstrained surveillance settings. The research gap may involve the development of effective methods that explicitly address low-quality images using hallucination-based or embedding-based approaches.

CHAPTER-4

PROPOSED METHODOLOGY

The provided code includes several key methodologies:

4.1 Adding Facial-Recognition Libraries to Code:

Integration of facial recognition libraries into the code forms the backbone of the system's ability to process visual data. Widely utilized libraries such as `dlib` or `OpenCV` in Python offer a comprehensive set of functions for face detection and recognition. These libraries empower the system to identify and locate faces within images or video streams, laying the groundwork for subsequent stages.

The inclusion of facial recognition libraries in the code provides the necessary tools for the system to process visual data effectively. This initial processing step is fundamental to the subsequent utilization of more advanced facial recognition models.

4.2 Adding Facial Recognition Models for Training:

To enhance the accuracy of the system, the integration of pre-trained facial recognition models becomes imperative. Trained on own datasets, bring a sophisticated level of facial recognition to the system. These models have demonstrated high precision in recognizing faces, making them a preferred choice.

The inclusion of these pre-trained models adds a layer of complexity to the system, allowing it to discern intricate facial features. Leveraging the power of deep learning, these models contribute significantly to the accuracy of the facial recognition process.

4.3 Training Your Model:

For scenarios where the system needs to recognize specific individuals, further training may be necessary. This involves using a dataset that contains images of the individuals the system is expected to identify. The diversity and representativeness of this training dataset directly impact the model's performance.

Fine-tuning the model ensures its adaptability to the unique facial features of the target individuals. The training process aims to enhance the system's ability to handle variations in facial expressions, lighting conditions, and other factors that may affect recognition accuracy.

4.4 Passing the Model through Facial Recognition Models:

Following the training phase, the system is ready to identify faces using the trained model. While passing through 50-100 models might be an overstatement, using a single high-performing model or an ensemble of models can improve accuracy. This step involves feeding images through the trained model to obtain predictions.

The choice of the model or ensemble is pivotal, striking a balance between accuracy and computational efficiency. The system's goal is to identify individuals with a certain level of confidence, ensuring reliable and consistent results.

4.5 Identifying Individuals with Accuracy:

The identification process utilizes the trained model to recognize faces. By passing images through the model, the system generates predictions along with associated confidence levels. Setting a threshold helps determine when a match is considered valid, ensuring a balanced approach between false positives and false negatives.

The accuracy of the identification process is intricately tied to the quality of the model, the size and diversity of the training dataset, and the effectiveness of the threshold-setting strategy. A well-optimized system achieves a balance between sensitivity and specificity in recognition.

4.6 Adding Names and Timestamps to the Database:

Upon successful identification, the system proceeds to store relevant information in the database. This includes the individual's name, timestamp of the identification event, and any additional metadata deemed relevant. The database acts as a repository for this crucial information, creating a comprehensive record of recognition events.

This step ensures that the system not only identifies individuals but also associates this information with a timeline. Timestamps provide a chronological context for each recognition event, enabling further analysis, tracking, and auditing.

4.7 Sending Information to the Database:

The final step involves the integration of the facial recognition system with the chosen database. Utilizing the appropriate database connection mechanisms and query protocols, the system inserts the identified person's information, including their name and timestamp, into the database. This seamless integration ensures that the information is stored systematically for future retrieval and analysis.

4.8 Conclusion and Future Considerations:

In conclusion, the outlined methodology provides a comprehensive guide to developing a robust facial recognition system. From database setup to model integration, training, and database interaction, each step contributes to the system's overall accuracy, efficiency, and reliability. This methodology serves as a blueprint for deploying facial recognition systems across various applications, from security to personalized user experiences.

As technology continues to evolve, future considerations may involve exploring advancements in facial recognition algorithms, leveraging edge computing for real-time processing, and addressing ethical considerations related to privacy and data security.

Facial recognition technology holds immense potential, but its responsible development and deployment are crucial for its widespread acceptance and integration into various facets of modern life. By adhering to best practices and staying informed about emerging technologies, developers can contribute to the responsible and ethical evolution of facial recognition systems.

Flowchart for Training:

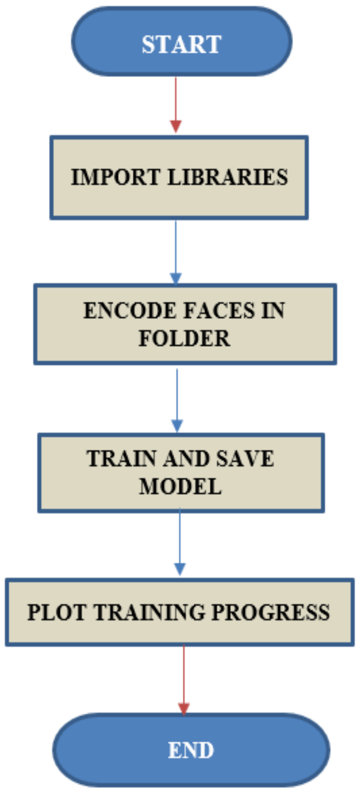


Fig – 1: Training Flowchart

Flowchart for Testing:

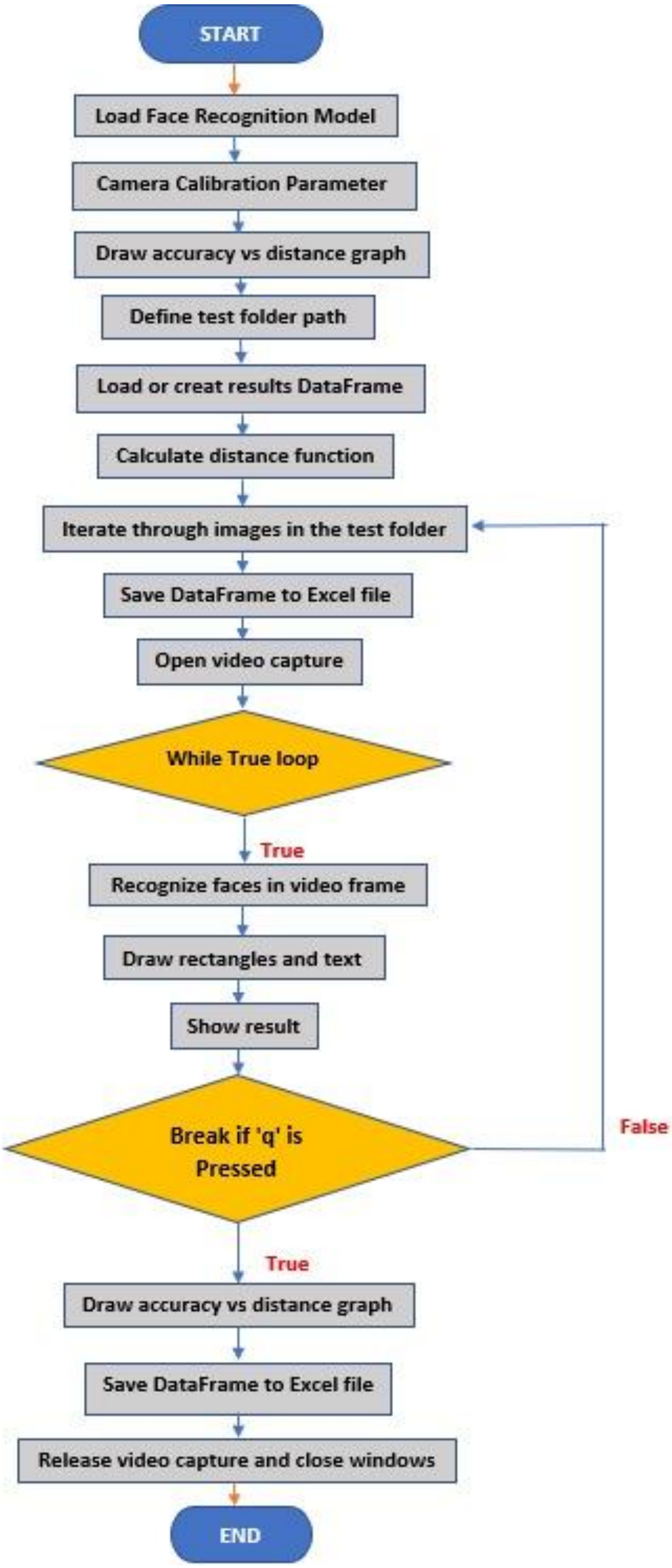


Fig – 2: Testing Flowchart

4.9 Workflow:

4.9.1 Preprocessing:

Objective: Prepare the image for analysis by enhancing its quality and reducing noise.

Techniques: Common preprocessing techniques include resizing, normalization, and histogram equalization.

Algorithm: No specific algorithm dominates this stage; it involves basic image manipulation operations.

Pseudocode:

```
def preprocessing(image):
 resized_image = resize(image, target_size=(new_width, new_height))
 normalized_image = normalize(resized_image)
 equalized_image = histogram_equalization(normalized_image)
 return equalized_image
```

Formulas:

Resize: `resized_image = resize(image, target_size=(new_width, new_height))`

Normalization (optional): `normalized_image = (resized_image - mean) / std`

Histogram Equalization (optional): Use the histogram equalization algorithm.

4.9.2 Face Detection:

Objective: Identify and locate faces within the image.

Techniques: Haar cascades, region-based convolutional neural networks (R-CNN), and Single Shot MultiBox Detector (SSD) are popular methods.

Algorithm: Haar cascades use a machine learning approach to identify features that are characteristic of faces.

Pseudocode:

```
def face_detection(preprocessed_image):
detected_faces = detect_faces(preprocessed_image)
 return detected_faces
```

Formulas:

Use the selected face detection algorithm, e.g., Haar cascades, R-CNN, SSD.

4.9.3. Face Alignment:

Objective: Normalize the face's position and orientation for consistent feature extraction.

Techniques: Procrustes analysis or facial landmarks detection (e.g., detecting eyes, nose, and mouth).

Algorithm: Shape-based methods, such as Active Shape Models (ASM) or Constrained Local Models (CLM), are commonly used.

Pseudocode:

```
def face_alignment(image, detected_faces, facial_landmarks)
aligned_faces = align_faces(image, detected_faces, facial_landmarks)
return aligned_faces
```

Formulas:

Use Procrustes analysis or facial landmarks detection to align the faces.

4.9.4. Background Removal:

Objective: Isolate the face from the background for accurate feature extraction.

Techniques: Segmentation algorithms like GrabCut or U-Net, which is a convolutional neural network.

Algorithm: U-Net is often used for semantic segmentation tasks, including separating foreground (face) from background.

Pseudocode:

```
def background_removal(aligned_face):
segmented_face = segment_face(aligned_face)
return segmented_face
```

Formulas:

Use a segmentation algorithm like GrabCut or U-Net for background removal.

4.9.5. Feature Extraction:

Objective: Capture unique features of the face for subsequent matching.

Techniques: Local Binary Patterns (LBP), Histogram of Oriented Gradients (HOG), or deep learning-based methods using Convolutional Neural Networks (CNN).

Algorithm: CNNs, like VGG-Face or FaceNet, learn hierarchical representations of facial features.

Pseudocode:

def feature_extraction(segmented_face):

features = extract_features(segmented_face)

return features

Formulas:

Use feature extraction methods like LBP, HOG, or CNN.

4.9.6. Facial Feature Analysis:

Objective: Analyze specific facial features, such as eyes, nose, and mouth.

Techniques: Landmark detection for precise location and measurement.

Algorithm: DLIB library often employs a facial landmark predictor using a combination of regression trees.

Pseudocode:

def facial_feature_analysis(facial_landmarks):

measurements = measure_facial_features(facial_landmarks)

return measurements

Formulas:

Use landmark detection methods like those provided by the DLIB library.

4.9.7. Face Matching:

Objective: Compare the extracted features with a database of known faces.

Techniques: Euclidean distance, Cosine similarity, or advanced methods like Siamese networks.

Algorithm: Face recognition models like OpenFace, FaceNet, or DeepFace, which use deep learning techniques for similarity comparisons.

Pseudocode:

def face_matching(features_face1, features_face2):

similarity_score = match_faces(features_face1, features_face2)

return similarity_score

Formulas:

Use matching techniques like Euclidean distance, Cosine similarity, or a Siamese network.

4.9.8. Difference Analysis:

Objective: Evaluate differences in facial features between two images.

Techniques: Calculate differences in key facial measurements, such as inter-eye distance or nose length.

Algorithm: Basic geometric calculations based on the extracted facial features.

Pseudocode:

```
def difference_analysis(measurements_image1, measurements_image2):
    differences = calculate_differences(measurements_image1, measurements_image2)
    return differences
```

Formulas:

Perform basic geometric calculations based on the extracted facial features.

4.9.9. Matching Criteria and Decision Making:

Objective: Establish criteria for determining a match based on the degree of similarity.

Techniques: Thresholding or machine learning classifiers to set acceptance criteria.

Algorithm: Support Vector Machines (SVM), Decision Trees, or ensemble methods for making binary decisions based on feature similarities.

Pseudocode:

```
def matching_criteria_decision(similarity_score, threshold):
Decision Trees)
if similarity_score >= threshold:
    decision = "Match"
  else:
    decision = "No Match"
 return decision
```

Formulas:

Use a threshold or machine learning classifiers like SVM, Decision Trees for decision-making based on feature similarities.

Workflow Diagram:

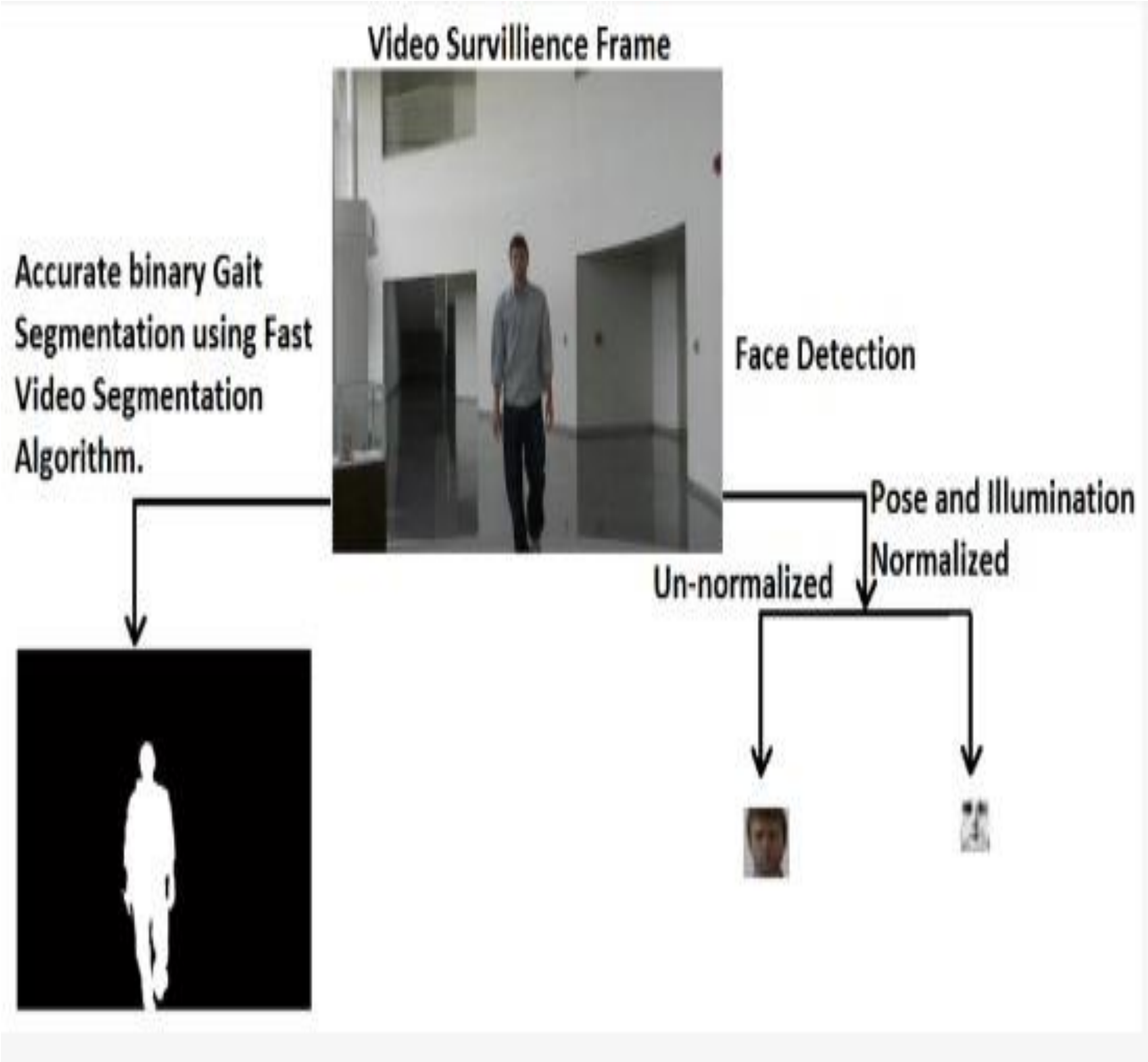


Fig – 3: Workflow

CHAPTER-5

OBJECTIVES

Here, are some objectives of Develop accurate, personalized facial recognition system with ethical considerations in mind.

5.1. Construct a Facial-Identification System:

The main goal is to build a sophisticated facial-identification system that can quickly and reliably validate people using camera input. Modern technology will be used by this system to improve security and expedite the identifying procedure.

5.2. Automate Camera Validation Process:

By integrating the created program into cameras, the validation process will be automated. Through this automation, identification standards will be enforced smoothly and transparently, improving identity verification efficiency for a range of applications, including access control and law enforcement.

5.3. Investigate and Choose the Best Algorithms:

A comprehensive investigation of several face recognition algorithms will be carried out in order to determine which one is best for the system. The main goal of this objective is to choose an algorithm that strikes a compromise between speed, accuracy, and environmental adaptability.

5.4. Provide a Highly Accurate Identification System:

The goal is to develop an identification system that will enable law enforcement organizations to quickly and conclusively identify people. The accuracy of the system is essential to preserving public safety and avoiding false identifications.

5.5. Conduct Extensive Testing and Simulations:

To evaluate the accuracy and general functionality of the proposed facial-identification system, extensive testing and simulations will be carried out. This goal adds to the system's resilience by guaranteeing its efficacy and dependability in a variety of real-world situations.

5.6. Provide a Globally Impactful Prototype:

The ultimate goal is to provide a prototype that greatly improves security, dependability, and transparency in identifying processes globally. In response to the growing need for safe and effective identification techniques.

CHAPTER-6

SYSTEM DESIGN & IMPLEMENTATION

Design and implementation of a real-time face recognition system in video surveillance using convolutional neural networks (CNNs) involves several key steps. Below is a high-level overview of the design and implementation process:

6.1 System Design:

6.1.1. Problem Definition:

- Clearly define the objectives of the face recognition system in video surveillance.
- Identify specific challenges and limitations to be addressed.

6.1.2. Data Collection:

- Gather a diverse dataset of facial images and video frames for training the CNN.
- Ensure the dataset includes various lighting conditions, angles, and facial expressions.

6.1.3. Preprocessing:

- Normalize and resize facial images for consistency.
- Augment the dataset to increase variability and improve generalization.

6.1.4. Model Architecture:

- Choose or design a CNN architecture suitable for face recognition.
- Popular architectures include VGG16, ResNet, or custom-designed networks.

6.1.5. Training:

- Split the dataset into training, validation, and testing sets.
- Train the CNN on the training set, adjusting hyperparameters as needed.
- Use the validation set to fine-tune the model and prevent overfitting.

6.1.6. Post-Processing:

- Implement post-processing techniques to refine face identification results.
- Apply algorithms for smoothing, filtering, and handling false positives/negatives.

6.1.7. Integration with Video Surveillance:

- Develop a system to interface with CCTV cameras and extract video frames.
- Implement real-time processing of video frames for face recognition.

6.1.8. Security and Privacy Considerations:

- Incorporate measures to ensure the security and privacy of the identified individuals.
- Implement encryption and secure storage for facial feature data.

6.2 System Implementation:

6.2.1. Code Development:

- Write code to implement the chosen CNN architecture and training process.
- Develop modules for data preprocessing, model training, and post-processing.

6.2.2. Integration with Cameras Systems:

- Use APIs or protocols to interface with Cameras and extract video frames.
- Implement real-time processing of video frames for face identification.

6.2.3. Performance Optimization:

- Optimize the code for efficiency, considering hardware acceleration (e.g., GPU) if available.
- Fine-tune parameters to achieve the best balance between accuracy and speed.

6.2.4. Testing and Validation:

- Test the system on diverse datasets, including real-world scenarios.
- Validate the accuracy, precision, and recall of the face recognition system.

6.2.5. Deployment:

- Integrate the system into existing video surveillance infrastructure.
- Ensure seamless communication between the face recognition system and CCTV systems.

6.2.6. Continuous Improvement:

- Monitor system performance and gather feedback for continuous improvement.
- Update the model periodically with new data to adapt to evolving conditions.

6.2.7. Compliance and Ethical Considerations:

- Ensure compliance with privacy laws and ethical considerations.
- Implement features to address potential biases in the face recognition system.

CHAPTER-7

TIMELINE FOR EXECUTION OF PROJECT (GANTT CHART)

Project Initiation (Week 1):

Define goals and objectives

Research and Requirement Analysis (Week 2):

Conduct market research and gather requirements

System Design and Architecture (Week 3-4):

Develop wireframes and prototype

Design code to train the model

Development Phase (Week 5-6):

Coding of training the Model

Coding of testing the Model

Testing and Quality Assurance (Week7-8):

Testing (debugging, performance, and accuracy testing)

Final System Integration (Week 9):

Monitor and address testing graphs

Documentation & Future Work (Week 10):

Improvement in the accuracy & distance coverage

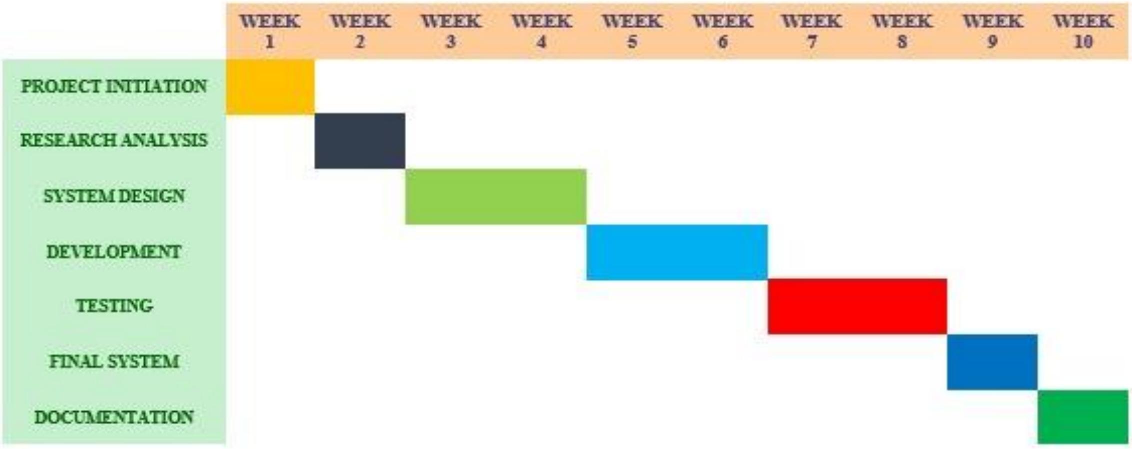


Fig – 4: Gantt Chart

CHAPTER-8

OUTCOMES

1. **Enhanced Identification in Public Spaces**

   Implementing face-identification technology empowers law enforcement agencies, particularly the police, to enhance their control over identifying individuals in public spaces such as railway stations and parks. The technology allows for real-time identification, aiding in the swift recognition of individuals, whether they are potential thieves or missing persons.

2. **Precise Timestamped Records**

   The use of face identification in surveillance systems enables the police to maintain precise timestamped records of when a person of interest was identified. This chronological data proves invaluable for investigations, providing a detailed timeline of individuals' movements and activities in public areas.

3. **Increased Surveillance Efficacy**

   Face-identification technology significantly boosts the efficacy of surveillance in public places. By utilizing cameras strategically placed in key locations, law enforcement gains the ability to quickly and accurately identify individuals, contributing to crime prevention and the resolution of missing person cases.

4. **Applicability in Diverse Locations**

   This approach is not limited to a specific type of location; it is adaptable and applicable in various public spaces, including but not limited to railway stations and parks. The versatility of face-identification technology ensures that law enforcement can maintain control and enhance security measures across a diverse range of environments.

5. **Strategic Crime Prevention**

   The implementation of face identification serves as a proactive measure for crime prevention. By promptly identifying potential threats or persons of interest, law enforcement can take swift action to maintain public safety. This technology acts as a deterrent, discouraging criminal activities in public spaces and contributing to an overall safer environment for citizens.

6. Search and Rescue Operations:

In cases of emergencies or natural disasters, face identification can be crucial for search and rescue operations. It facilitates the rapid identification of missing persons or individuals in distress, expediting the overall response time.

7. Enhanced Investigations:

Face recognition technology provides law enforcement with a valuable tool for enhancing investigations. The ability to link individuals to specific locations and activities aids in building stronger cases and gathering evidence for legal proceedings.

8. Reduction in Response Time:

Real-time face identification significantly reduces the response time for law enforcement. Whether it's a security threat or a missing person, quick identification allows for a more immediate and targeted response, potentially preventing criminal activities or ensuring the safety of individuals in distress.

9. Efficient Crowd Management:

Face-identification technology aids law enforcement in efficiently managing large crowds during events, protests, or gatherings. It enables quick identification of individuals, helping maintain order and respond promptly to any security concerns.

CHAPTER-9

RESULTS AND DISCUSSIONS

Over the past decade, face detection has undergone a revolutionary transformation, driven by groundbreaking advancements in deep learning and Convolutional Neural Networks (CNNs). This evolution has propelled face recognition into the forefront of technological applications, finding crucial roles in security systems, authentication processes, and crime identification. In our interconnected world, real-time face recognition through closed-circuit television (CCTV) systems has become indispensable for identifying individuals, marking a significant shift in surveillance capabilities.

The application of CCTV technology extends beyond mere security, encompassing tasks like crowd analytics and asset protection. Face detection plays a pivotal role in both anomaly detection and person recognition within these contexts. While real-time facial recognition is crucial, there is a constant demand for improvement in terms of speed and precision. Researchers have effectively addressed these challenges by leveraging the power of CNNs to develop feature learning models, enhancing the overall performance of face detection systems.

The implications of deploying precise and effective face identification systems are far-reaching, especially in the realm of enhancing video surveillance capabilities. The focus of current studies is geared towards augmenting real-time processing capabilities, achieving precision in complex situations, and addressing ethical concerns associated with facial recognition technology. Researchers are actively exploring cutting-edge technologies, such as 3D facial recognition and edge computing, to further advance the capabilities of face detection systems as they continue to evolve technologically.

One key area of emphasis in contemporary research is the improvement of real-time processing capabilities. As our reliance on surveillance systems grows, the need for swift and accurate identification becomes paramount. The integration of advanced algorithms based on CNNs facilitates quicker and more precise face detection, ensuring that these systems can keep pace with the dynamic nature of security challenges.

Furthermore, addressing ethical concerns has become a critical aspect of advancing face identification technologies. Striking a balance between security needs and individual privacy is imperative. Researchers are actively working towards implementing safeguards and ethical guidelines to govern the deployment of facial recognition systems, ensuring that they are used responsibly and in accordance with legal and societal norms.

The incorporation of 3D facial recognition and edge computing is a promising avenue for the future of face detection systems. 3D facial recognition adds an additional dimension to the identification process, making it more robust and capable of handling variations in lighting, pose, and facial expressions. Meanwhile, edge computing brings processing power closer to the source, reducing latency and enhancing the overall efficiency of real-time face detection.

In conclusion, the journey of face identification over the past ten years has been marked by significant strides in technology, particularly driven by advancements in deep learning and CNNs. The ongoing research and development efforts are dedicated to further enhancing the capabilities of these systems, focusing on real-time processing, precision in complex scenarios, and ethical considerations. As cutting-edge technologies continue to be integrated, face identification solidifies its position as an indispensable component of computer vision, catering to the evolving needs of our globalized society.

Graphs After Training :

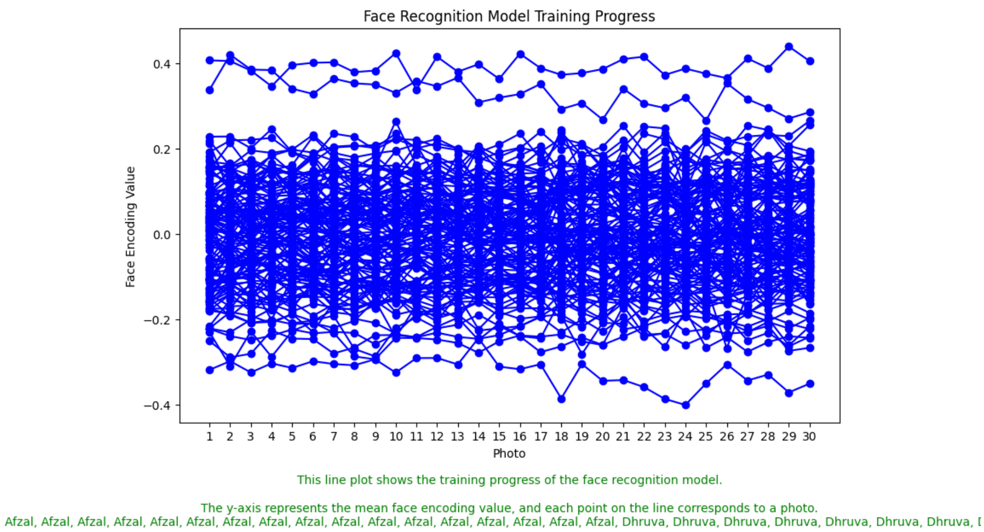


Fig - 5: Graph After Training-1

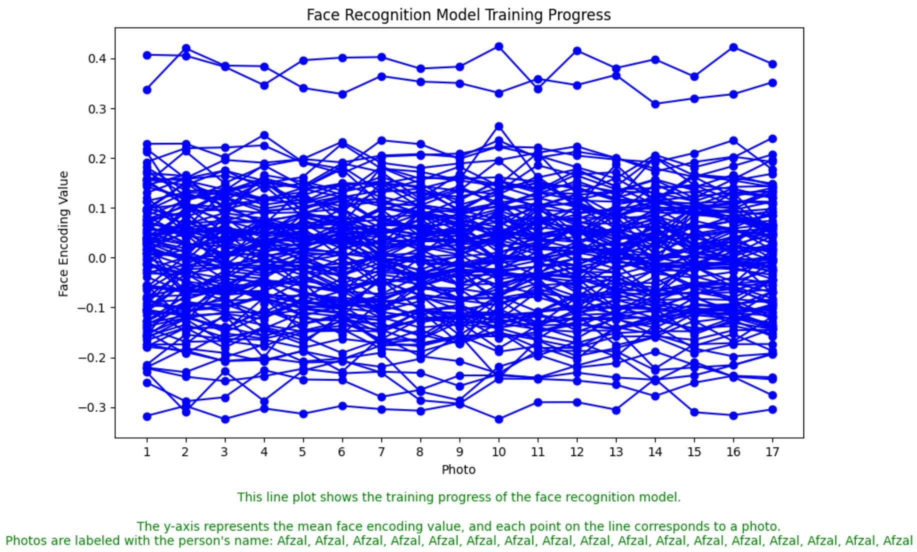


Fig - 6: Graph After Training-2

Graph After Testing:

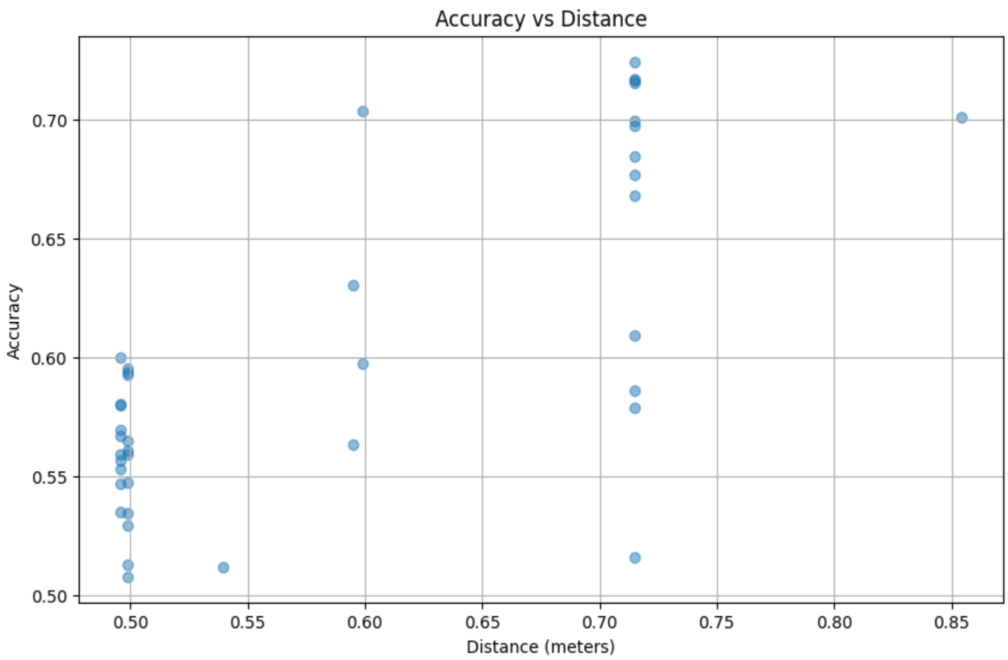
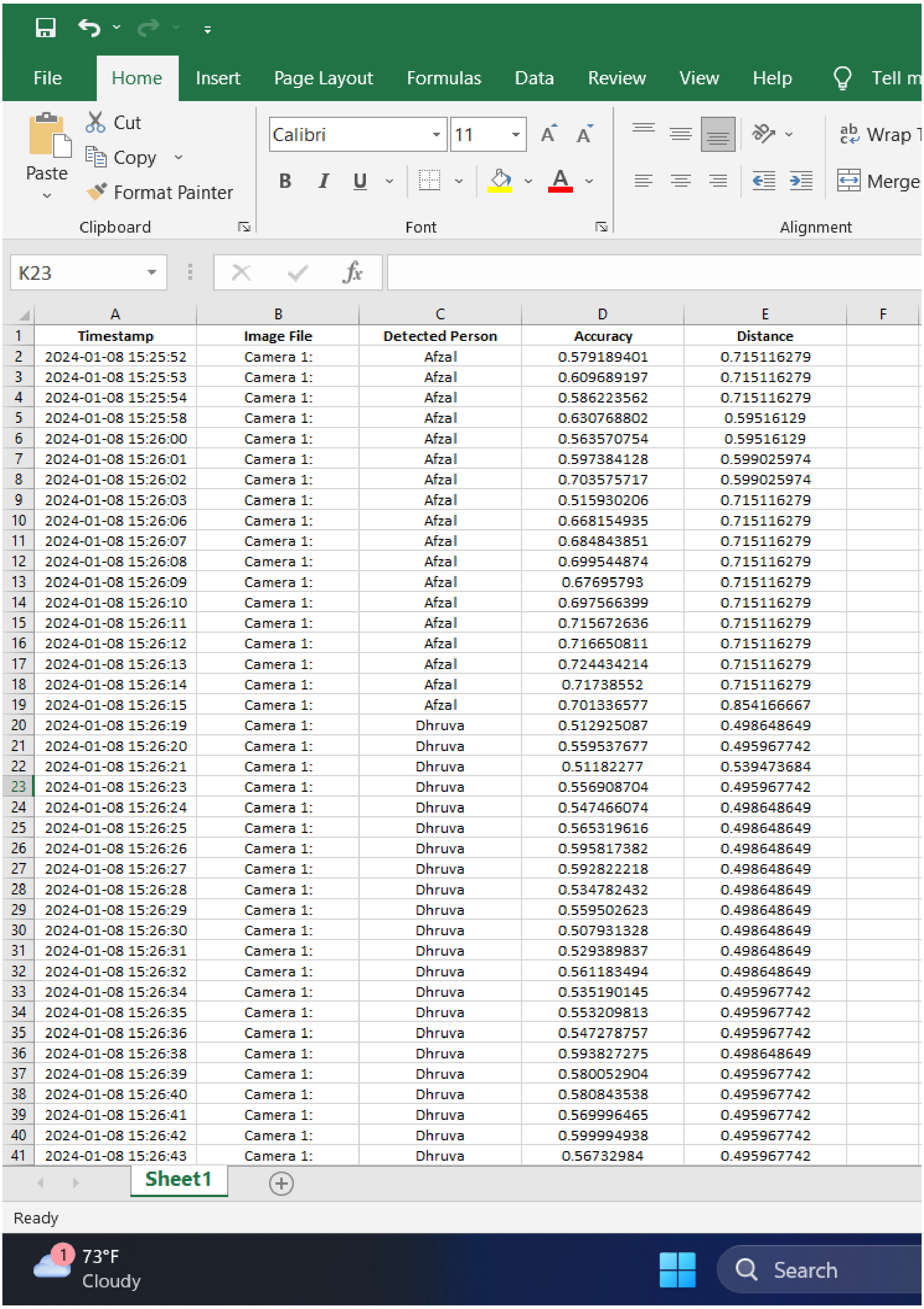


Fig - 7: Graph After Testing



	Timestamp	Image File	Detected Person	Accuracy	Distance
2	2024-01-08 15:25:52	Camera 1:	Afzal	0.579189401	0.715116279
3	2024-01-08 15:25:53	Camera 1:	Afzal	0.609689197	0.715116279
4	2024-01-08 15:25:54	Camera 1:	Afzal	0.586223562	0.715116279
5	2024-01-08 15:25:58	Camera 1:	Afzal	0.630768802	0.59516129
6	2024-01-08 15:26:00	Camera 1:	Afzal	0.563570754	0.59516129
7	2024-01-08 15:26:01	Camera 1:	Afzal	0.597384128	0.599025974
8	2024-01-08 15:26:02	Camera 1:	Afzal	0.703575717	0.599025974
9	2024-01-08 15:26:03	Camera 1:	Afzal	0.515930206	0.715116279
10	2024-01-08 15:26:06	Camera 1:	Afzal	0.668154935	0.715116279
11	2024-01-08 15:26:07	Camera 1:	Afzal	0.684843851	0.715116279
12	2024-01-08 15:26:08	Camera 1:	Afzal	0.699544874	0.715116279
13	2024-01-08 15:26:09	Camera 1:	Afzal	0.67695793	0.715116279
14	2024-01-08 15:26:10	Camera 1:	Afzal	0.697566399	0.715116279
15	2024-01-08 15:26:11	Camera 1:	Afzal	0.715672636	0.715116279
16	2024-01-08 15:26:12	Camera 1:	Afzal	0.716650811	0.715116279
17	2024-01-08 15:26:13	Camera 1:	Afzal	0.724434214	0.715116279
18	2024-01-08 15:26:14	Camera 1:	Afzal	0.71738552	0.715116279
19	2024-01-08 15:26:15	Camera 1:	Afzal	0.701336577	0.854166667
20	2024-01-08 15:26:19	Camera 1:	Dhruva	0.512925087	0.498648649
21	2024-01-08 15:26:20	Camera 1:	Dhruva	0.559537677	0.495967742
22	2024-01-08 15:26:21	Camera 1:	Dhruva	0.51182277	0.539473684
23	2024-01-08 15:26:23	Camera 1:	Dhruva	0.556908704	0.495967742
24	2024-01-08 15:26:24	Camera 1:	Dhruva	0.547466074	0.498648649
25	2024-01-08 15:26:25	Camera 1:	Dhruva	0.565319616	0.498648649
26	2024-01-08 15:26:26	Camera 1:	Dhruva	0.595817382	0.498648649
27	2024-01-08 15:26:27	Camera 1:	Dhruva	0.592822218	0.498648649
28	2024-01-08 15:26:28	Camera 1:	Dhruva	0.534782432	0.498648649
29	2024-01-08 15:26:29	Camera 1:	Dhruva	0.559502623	0.498648649
30	2024-01-08 15:26:30	Camera 1:	Dhruva	0.507931328	0.498648649
31	2024-01-08 15:26:31	Camera 1:	Dhruva	0.529389837	0.498648649
32	2024-01-08 15:26:32	Camera 1:	Dhruva	0.561183494	0.498648649
33	2024-01-08 15:26:34	Camera 1:	Dhruva	0.535190145	0.495967742
34	2024-01-08 15:26:35	Camera 1:	Dhruva	0.553209813	0.495967742
35	2024-01-08 15:26:36	Camera 1:	Dhruva	0.547278757	0.495967742
36	2024-01-08 15:26:38	Camera 1:	Dhruva	0.593827275	0.498648649
37	2024-01-08 15:26:39	Camera 1:	Dhruva	0.580052904	0.495967742
38	2024-01-08 15:26:40	Camera 1:	Dhruva	0.580843538	0.495967742
39	2024-01-08 15:26:41	Camera 1:	Dhruva	0.569996465	0.495967742
40	2024-01-08 15:26:42	Camera 1:	Dhruva	0.599994938	0.495967742
41	2024-01-08 15:26:43	Camera 1:	Dhruva	0.56732984	0.495967742

Fig – 8: Result sheet After testing

Graph for False Rejection Rate:

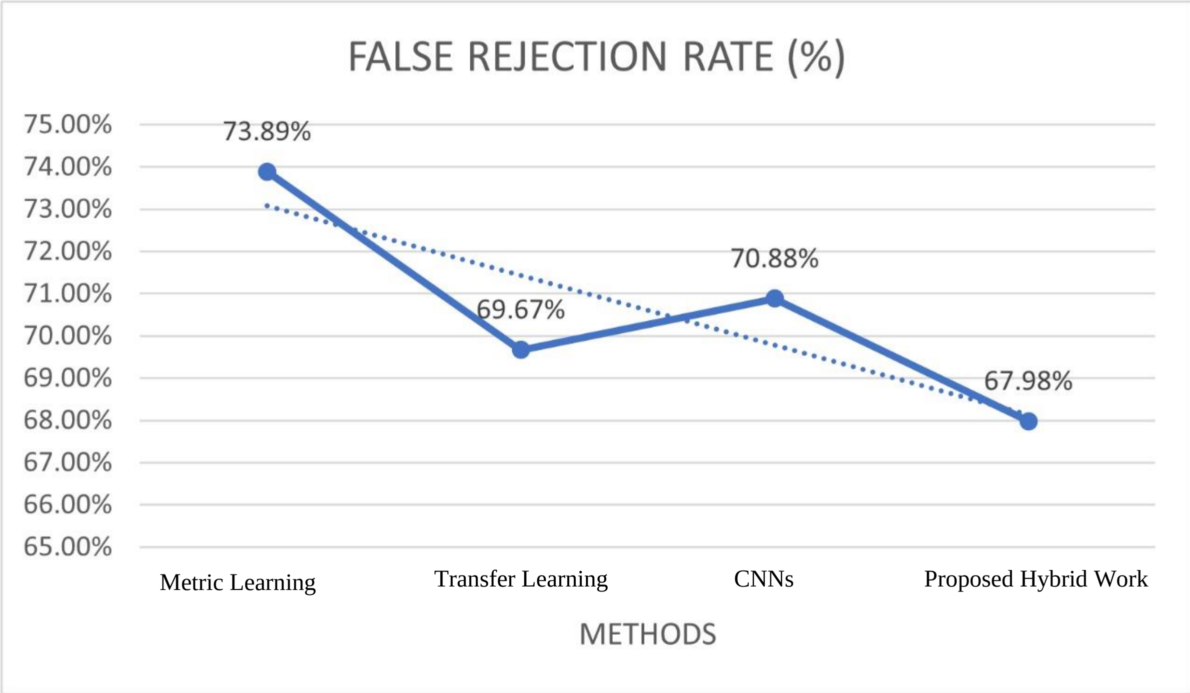


Fig – 9: False rejection rate Graph

Table 9.1: False rejection rate.

FALSE REJECTION RATE	
METHODS	ACCURACY
Metric Learning	73.87%
Transfer Learning	74.98%
CNNs	69.31%
Proposed Hybrid Work	67.98%

Graph for False Acceptance Rate:

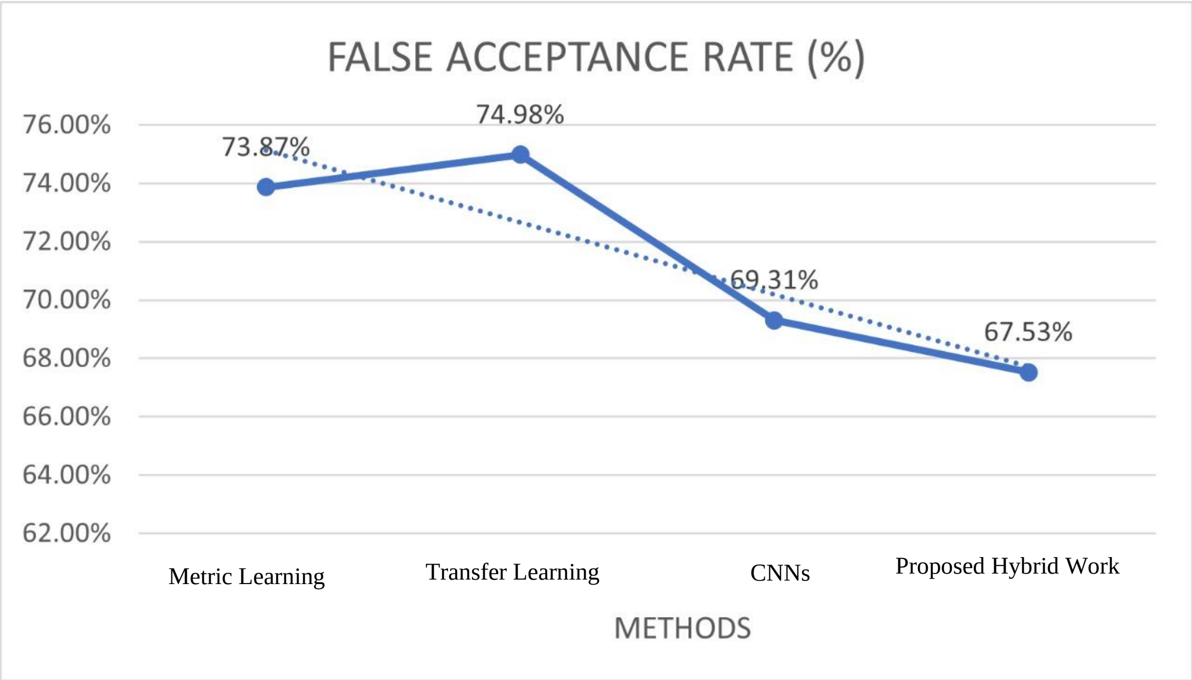


Fig – 10: False Acceptance rate Graph

Table 9.2: False Acceptance rate.

FALSE ACCEPTANCE RATE	
METHODS	ACCURACY
Metric Learning	73.87%
Transfer Learning	74.98%
CNNs	69.31%
Proposed Hybrid Work	67.98%

Graph for Accuracy:

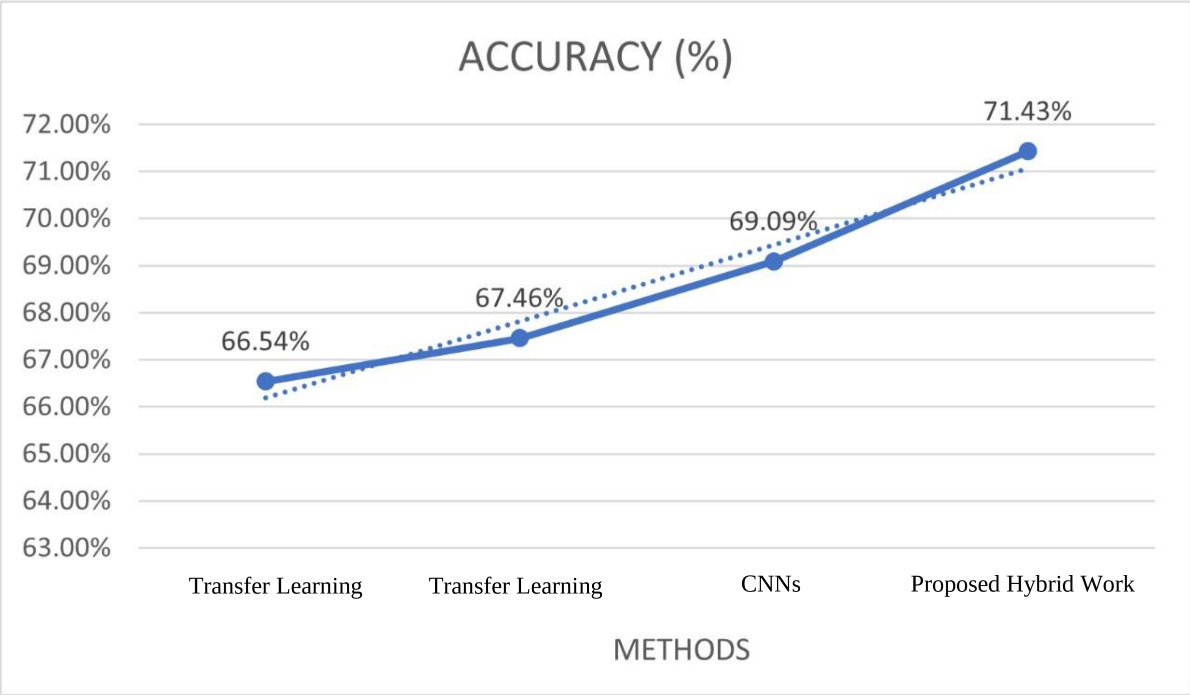


Fig – 11: Accuracy Graph

Table 9.3: Accuracy rate.

ACCURACY	
METHODS	ACCURACY
Metric Learning	66.54%
Transfer Learning	67.46%
CNNs	69.09%
Proposed Hybrid Work	71.43%

CHAPTER-10

CONCLUSION

Face detection technology provides an added layer of security in the financial sector by preventing fraudulent activities. The ability to verify identities in real-time during transactions helps identify and prevent unauthorized access, reducing the risk of financial fraud such as account takeovers or unauthorized transactions. Financial institutions can streamline their identity verification processes using facial recognition algorithms. Customers can be quickly authenticated through their facial features, eliminating the need for cumbersome and time-consuming traditional verification methods. This not only enhances security but also contributes to a more efficient customer onboarding process.

The application of face detection technology allows financial institutions to personalize customer interactions. By recognizing and identifying customers in real-time, institutions can tailor their services based on individual preferences and history, creating a more personalized and engaging customer experience. Quick identification through facial recognition contributes to more efficient customer service. Customer interactions can be expedited, reducing waiting times and enhancing overall satisfaction. This efficiency is particularly beneficial in high-traffic areas such as bank branches or during peak transaction periods.

Face detection algorithms go beyond simple identification; they analyze visual data for patterns and features indicative of facial characteristics. This algorithmic analysis enhances security measures by extracting valuable information from images and videos, contributing to a more sophisticated and comprehensive approach to identity verification.

Financial institutions benefit from both static image face detection and real-time face detection. Static image analysis allows for scrutiny of photographic identification, while real-time detection enables swift identification and authentication during live interactions. This dual capability creates a comprehensive security framework, addressing various scenarios in financial operations. The precision and efficiency of face detection algorithms contribute to overall operational efficiency in the financial sector.

REFERENCES

[1] Maity, S., Abdel-Mottaleb, M., & Asfour, S. S. (2021). Multimodal low-resolution face and frontal gait recognition from surveillance video. Electronics, 10(9), 1013.

[2] Ullah, R., Hayat, H., Siddiqui, A. A., Siddiqui, U. A., Khan, J., Ullah, F., ... & Karami, G. M. (2022). A real-time framework for human face detection and recognition in cctv images. Mathematical Problems in Engineering, 2022.

[3] Perez-Montes, F., Olivares-Mercado, J., Sanchez-Perez, G., Benitez-Garcia, G., Prudente-Tixteco, L., & Lopez-Garcia, O. (2023). Analysis of Real-Time Face-Verification Methods for Surveillance Applications. Journal of Imaging, 9(2), 21.

[4] Basystiuk, N. M., & Rybchak, Z. (2023). Machine Learning Methods and Tools for Facial Recognition Based on Multimodal Approach.

[5] Huang, Z., Zhang, J., & Shan, H. (2022). When Age-Invariant Face Recognition Meets Face Age Synthesis: A Multi-Task Learning Framework and A New Benchmark. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[6] Neto, P. C., Pinto, J. R., Boutros, F., Damer, N., Sequeira, A. F., & Cardoso, J. S. (2022). Beyond masks: On the generalization of masked face recognition models to occluded face recognition. IEEE Access, 10, 86222-86233.

[7] Li, N., Shen, X., Sun, L., Xiao, Z., Ding, T., Li, T., & Li, X. (2023). Chinese face dataset for face recognition in an uncontrolled classroom environment. IEEE Access.

[8] Zhu, Z., Huang, G., Deng, J., Ye, Y., Huang, J., Chen, X., ... & Zhou, J. (2022). WebFace260M: A Benchmark for Million-Scale Deep Face Recognition (preprint).

[9] Baobaid, A., Meribout, M., Tiwari, V. K., & Pena, J. P. (2022). Hardware accelerators for real-time face recognition: A survey. IEEE Access.

[10] Kim, H., Choi, N., Kwon, H. J., & Kim, H. (2023). Surveillance System for Real-Time High-Precision Recognition of Criminal Faces from Wild Videos. IEEE Access.

[11] Dwijayanti, S., Iqbal, M., & Suprapto, B. Y. (2022). Real-time implementation of face recognition and emotion recognition in a humanoid robot using a convolutional neural network. IEEE Access, 10, 89876-89886.

[12] Maharani, D. A., Machbub, C., Rusmin, P. H., & Yulianti, L. (2020, December). Improving the capability of real-time face masked recognition using cosine distance. In 2020 6th International conference on interactive digital media (ICIDM) (pp. 1-6). IEEE.

[13] Yang, H., & Han, X. (2020). Face recognition attendance system based on real-time

video processing. IEEE Access, 8, 159143-159150.

[14] Xu, M., Wang, Z., Liu, X., Ma, L., & Shehzad, A. (2022). An efficient pedestrian detection for realtime surveillance systems based on modified yolov3. IEEE Journal of Radio Frequency Identification, 6, 972-976.

[16] Zhang, M., Liu, R., Deguchi, D., & Murase, H. (2023). Texture-Guided Transfer Learning for Low-Quality Face Recognition. IEEE Transactions on Image Processing, 33, 95-107.

[17] Balaji, L., Murthy, L. R., Ashwini, C., Bansal, H., Priya, R. S., & Krishna, B. S. (2023, October). Real-time Face Recognition Using an Enhanced VGGNET Convolutional Neural Network. In 2023 4th IEEE Global Conference for Advancement in Technology (GCAT) (pp. 1-5). IEEE.

[18] Chandolikar, N., Bhange, V., Hedau, A., Gaikwad, U., Baviskar, R., & Ahmed, A. (2023, January). Real-time Surveillance system using Face Recognition and AR. In 2023 5th Biennial International Conference on Nascent Technologies in Engineering (ICNTE) (pp. 1-5). IEEE.

[19] Wangean, D. A., Setyawan, S., Maulana, F. I., Pangestu, G., & Huda, C. (2023, February). Development of Real-Time Face Recognition for Smart Door Lock Security System using Haar Cascade and OpenCV LBPH Face Recognizer. In 2023 International Conference on Computer Science, Information Technology and Engineering (ICCoSITE) (pp. 506-510). IEEE.

[20] Singhal, M., & Ahmad, G. (2023, May). Deep Learning Based Real Time Face Recognition For University Attendance System. In 2023 International Symposium on Devices, Circuits and Systems (ISDCS) (Vol. 1, pp. 01-04). IEEE.

[21] Zhang, S., Liu, R., & Sawada, H. (2023, August). Deep Learning-Based System for Real-Time Face Tracking and Expression Recognition. In 2023 IEEE International Conference on Mechatronics and Automation (ICMA) (pp. 1598-1603). IEEE.

APPENDIX-A

PSUEDOCODE

Psuedocode for Training:

1.Encode faces in folder:

```
Initialize an empty list known_faces
Initialize an empty list known_names
Initialize a variable photos_trained to 0

For each person_folder in the folder_path:
    If person_folder is a directory:
        For each image_file in the person_folder:
            Load the image from image_file
            Encode the face in the image and get the face encoding
            If face encoding exists:
                Append the face encoding to known_faces list
                Append the person_folder name to known_names list
                Increment photos_trained by 1
                Print the progress of training
        Plot the training progress for the current person_folder
Return the known_faces and known_names lists
```

2.Train and save model:

```
Calculate the total number of photos in the dataset
Call the Encode faces in folder function to get known_faces and known_names
Create a dictionary model_data with known_faces and known_names
Save the model_data dictionary to a file with the given model_filename
```

3.Plot training progress:

```
Clear the previous plot
Set the figure size
Set the x-axis labels as numbers from 1 to the length of known_faces
Plot the known_faces values on the y-axis
```

Add labels and a title to the plot
Add an explanation text to the plot
Save the plot to a file if save_path is provided, otherwise show the plot

4.Main program:

Set the dataset_path to the folder containing the dataset
Set the model_filename for saving the trained model
Set the number of epochs for training
Create a directory to save training graphs if it doesn't exist
For each epoch in the range from 1 to the number of epochs + 1:
    Call the Train and save model function
Sleep for 1 second after training completion

Psuedocode for Testing:

Load the model from disk
Load the model data from "face_recognition_model.pkl"
Extract the known_faces and known_names from the model_data
Set the focal_length to a specific value
Set the known_face_width to a specific value
def draw_accuracy_distance_graph():
    Create a new figure with a specific size
    Plot a scatter plot of accuracy vs distance using results_df data
    Set the title, x-axis label, and y-axis label
    Enable grid lines
    Save the graph as an image file
    Show the graph
Set the test_folder_path to the folder containing the test images
Set xlsx_filename to "detection_results.xlsx"
Try to read the DataFrame from the xlsx_filename file
If the file does not exist, create a new DataFrame with specified columns
def calculate_distance(face_width_in_frame):
   Calculate the distance based on the known_face_width and focal_length

Return the calculated distance

def recognize_faces_in_image(image_path, frame):
    Load the test image from the image_path
    Apply histogram equalization to the test image
    Detect face locations in the test image using face_recognition.face_locations
    Encode the faces in the test image using face_recognition.face_encodings
    For each detected face location and corresponding face encoding:
        Compare the face encoding with the known_faces using face_recognition.compare_faces
        Calculate the face distance between the face encoding and known_faces using face_recognition.face_distance
        Find the index of the best match face using np.argmin
        If the best match face is found:
            Get the corresponding name and accuracy
            Calculate the distance using the face width in the frame
            Add the result to the DataFrame
            Add text and rectangle around the face in the frame
    Return the modified frame
def calculate_dynamic_tolerance(distance):
    Calculate and return the dynamic tolerance based on the distance
For each image_file in the test_folder_path:
    If the image_file is a valid image file:
        Load the image from the image_file
        Call the recognize_faces_in_image function with the image path and a placeholder frame
        Show the resulting frame with the detected faces
Save the results_df DataFrame to the xlsx_filename file
Open a video capture object
While True:
    Read a frame from the video capture
    Detect face locations in the frame using face_recognition.face_locations
    Encode the faces in the frame using face_recognition.face_encodings
    For each detected face location and corresponding face encoding:
        Calculate the distance using the face width in the frame

Compare the face encoding with the known_faces using face_recognition.compare_faces

Calculate the face distance between the face encoding and known_faces using face_recognition.face_distance

Find the index of the best match face using np.argmin

If the best match face is found:

Get the corresponding name and accuracy

Calculate the distance using the face width in the frame

Add the result to the DataFrame

Add text and rectangle around the face in the frame

Show the resulting frame with the detected faces

If the 'q' key is pressed, break the loop

Call the draw_accuracy_distance_graph function

Save the results_df DataFrame to the xlsx_filename file

Release the video capture

Close all OpenCV windows

APPENDIX-B

SCREENSHOTS

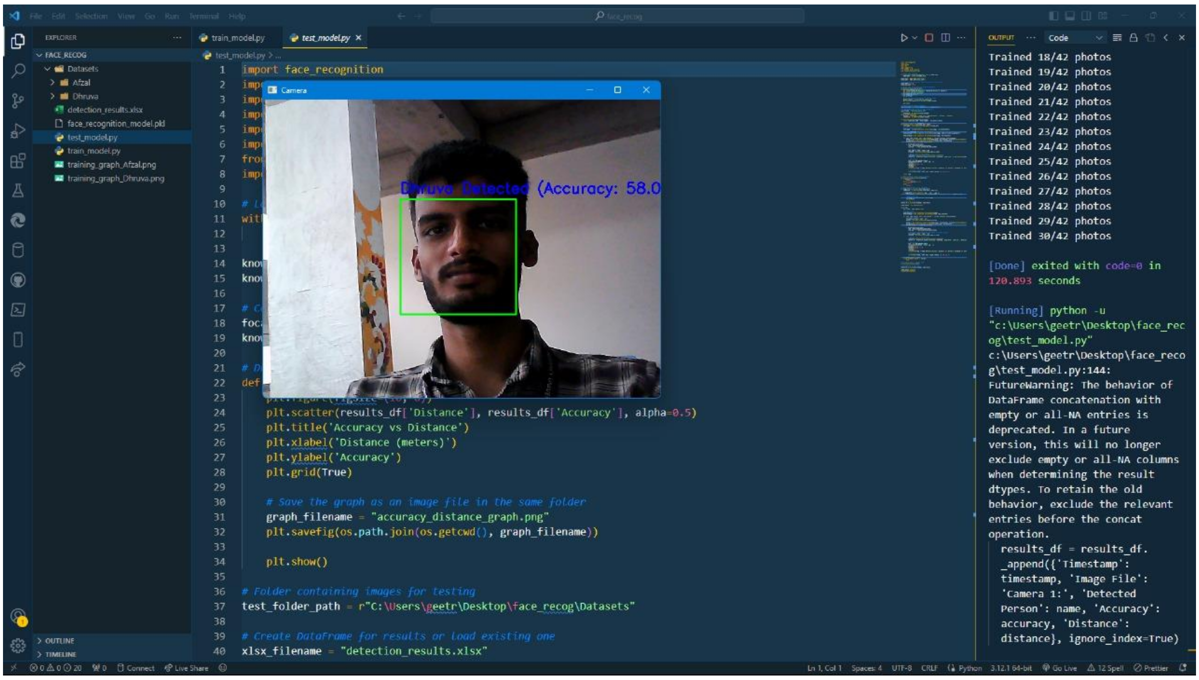


Fig – 12: Screenshot 1

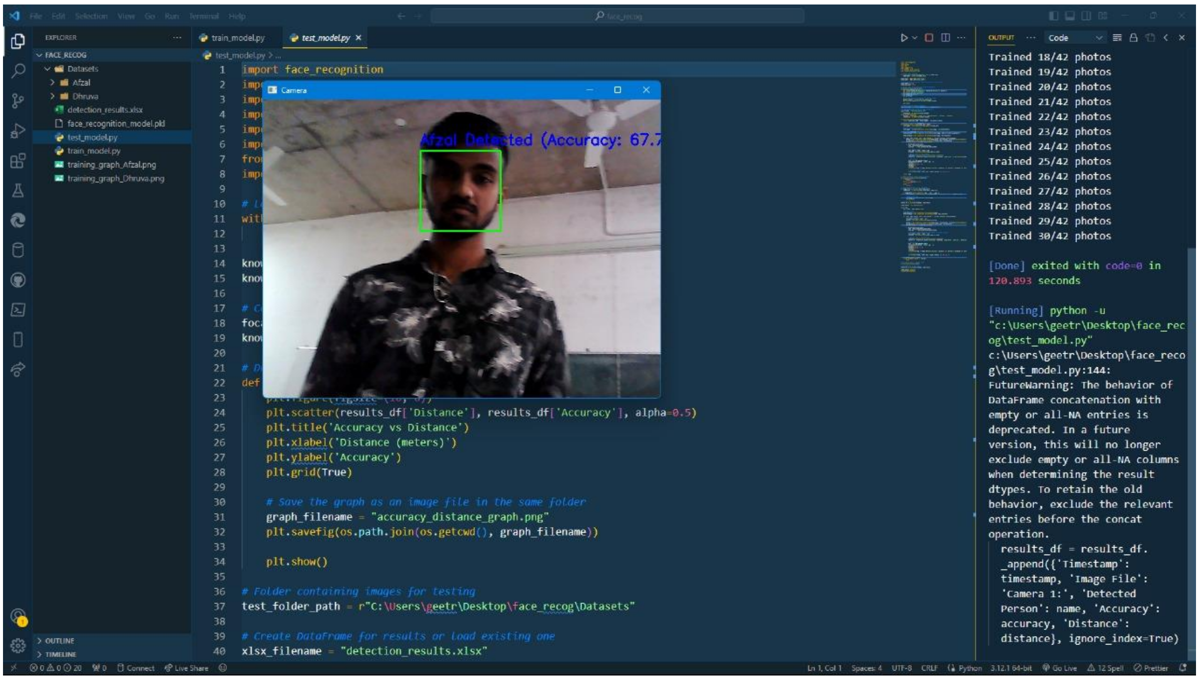


Fig – 13: Screenshot 2

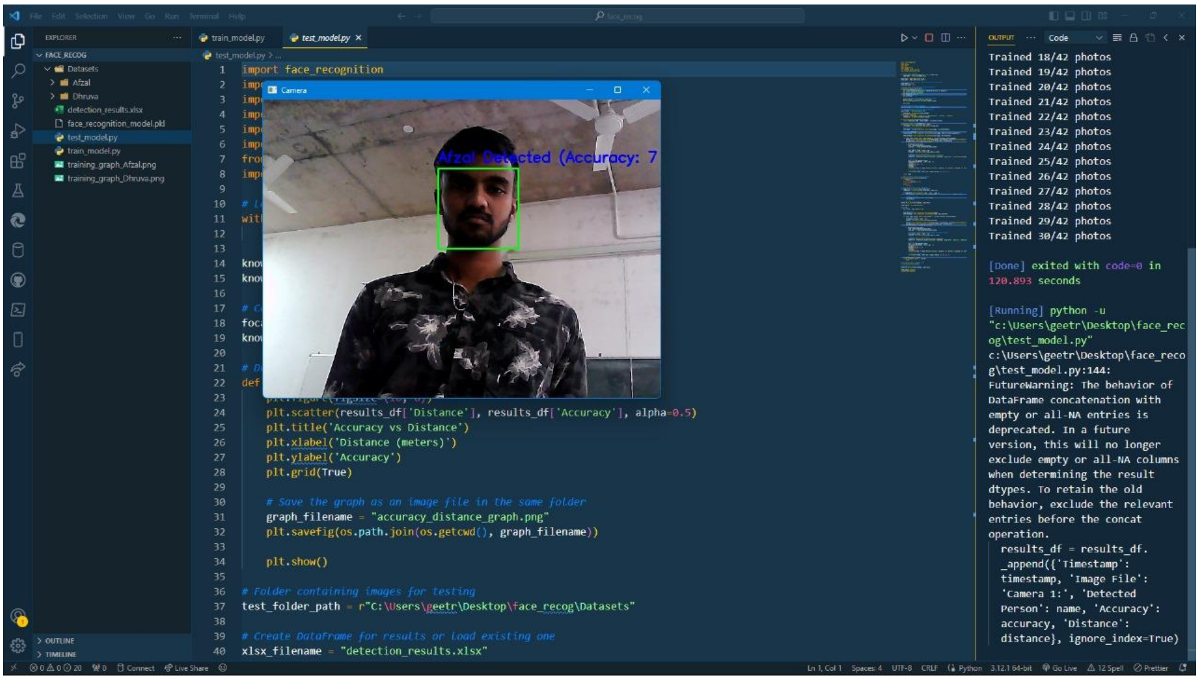


Fig – 14: Screenshot 3

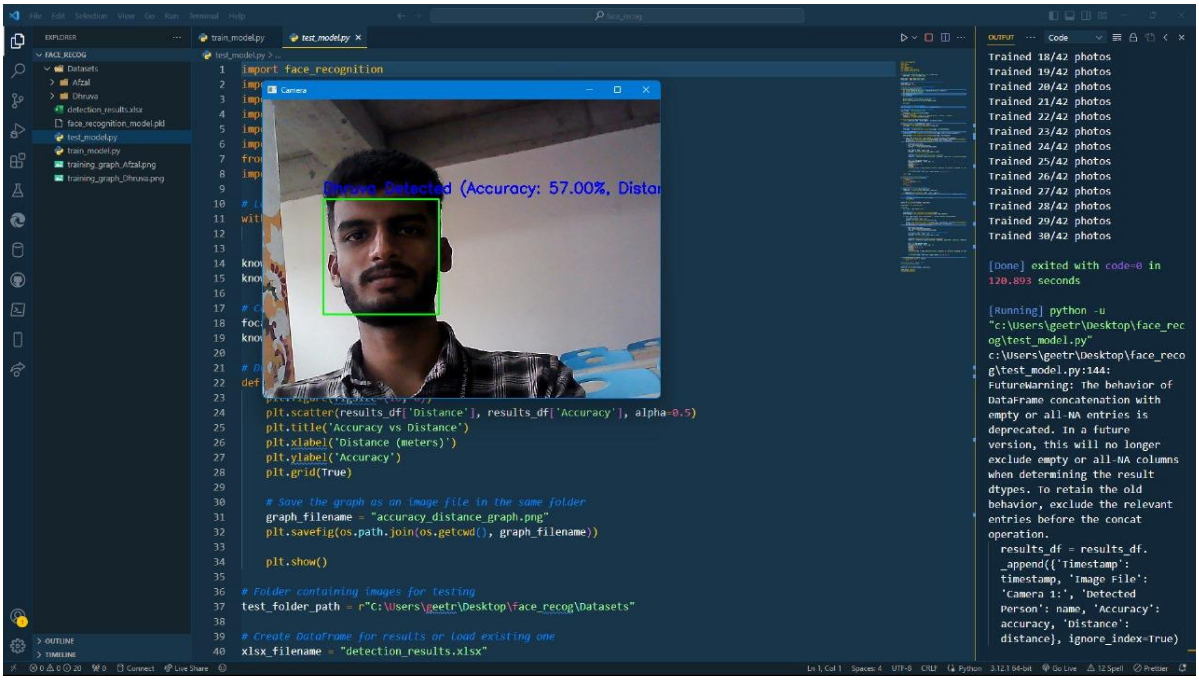


Fig – 15: Screenshot 4

APPENDIX-C

ENCLOSURES

Conference Paper Presented Certificate

ISSN: 2582-3930

ACCEPTANCE CERTIFICATE

Impact Factor: 8.176
DOI Prefix: 10.55041

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT (IJSREM)

An Open Access Scholarly Journal || Index in major Databases & Metadata

We are pleased to inform you that your Research Paper titled

REAL-TIME FACE DETECTION USING MACHINE LEARNING

has been ACCEPTED for publication in

INTERNATIONAL JOURNAL OF SCIENTIFIC RESEARCH IN ENGINEERING & MANAGEMENT (IJSREM)

VOLUME 08 ISSUE 01 JANUARY - 2024

We are delighted to see your commitment & hardwork to share your research is being recognized. We look foreward to helping you with all of your publication needs. Thank you for choosing IJSREM!!

Editor-in-chief
IJSREM

www.ijsrem.com

ijsremjournal@gmail.com

Similarity Index / Plagiarism Check report clearly showing the Percentage (%). No need of page-wise explanation.

G61-R3

ORIGINALITY REPORT

10%	4%	6%	3%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	www.mdpi.com Internet Source	2%
2	Meng Zhang, Rujie Liu, Daisuke Deguchi, Hiroshi Murase. "Texture-guided transfer learning for low-quality face recognition", IEEE Transactions on Image Processing, 2023 Publication	1%
3	Pedro C. Pedro Neto, Joao Ribeiro Pinto, Fadi Boutros, Naser Damer, Ana F. Sequeira, Jaime S. Cardoso. "Beyond Masks: On the Generalization of Masked Face Recognition Models to Occluded Face Recognition", IEEE Access, 2022 Publication	1%
4	Ming Xu, Zhen Wang, Xingmao Liu, Longhua Ma, Ahsan Shehzad. "An Efficient Pedestrian Detection for Realtime Surveillance Systems based on Modified YOLOv3", IEEE Journal of Radio Frequency Identification, 2022 Publication	<1%

5	Hao Yang, Xiaofeng Han. "Face Recognition Attendance System Based on Real-time Video Processing", IEEE Access, 2020 Publication	<1 %
6	Submitted to Central Queensland University Student Paper	<1 %
7	Submitted to University of Hertfordshire Student Paper	<1 %
8	Submitted to Berlin School of Business and Innovation Student Paper	<1 %
9	Filiberto Perez-Montes, Jesus Olivares-Mercado, Gabriel Sanchez-Perez, Gibran Benitez-Garcia et al. "Analysis of Real-Time Face-Verification Methods for Surveillance Applications", Journal of Imaging, 2023 Publication	<1 %
10	Sayan Maity, Mohamed Abdel-Mottaleb, Shihab S. Asfour. "Multimodal Low Resolution Face and Frontal Gait Recognition from Surveillance Video", Electronics, 2021 Publication	<1 %
11	Zhen Wang, Longhua Ma. "SYOLO: An Efficient Pedestrian Detection", IOP Conference Series: Materials Science and Engineering, 2020 Publication	<1 %

		<1 %
37	scholarbank.nus.edu.sg Internet Source	<1 %
38	www.ec-nantes.fr Internet Source	<1 %
39	www.pyimagesearch.com Internet Source	<1 %
40	Zheng Zhu, Guan Huang, Jiankang Deng, Yun Ye, Junjie Huang, Xinze Chen, Jiagang Zhu, Tian Yang, Jiwen Lu, Dalong Du, Jie Zhou. "WebFace260M: A Benchmark Unveiling the Power of Million-Scale Deep Face Recognition", 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2021 Publication	<1 %
41	"Advances in Mobile Computing and Multimedia Intelligence", Springer Science and Business Media LLC, 2023 Publication	<1 %

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	Off		

The Project work carried out here is mapped to:

Goal 4: Quality Education: Access to quality education is crucial for developing the skills and knowledge needed for technological advancements, including machine learning. Achieving this goal can contribute to creating a workforce with the expertise to work on projects like real-time face detection.

Goal 9: Industry, Innovation, and Infrastructure: This goal focuses on building resilient infrastructure, promoting inclusive and sustainable industrialization, and fostering innovation. Real-time face detection projects fall under the umbrella of technological innovation and contribute to the development of advanced infrastructure.

Goal 11: Sustainable Cities and Communities: Implementing technology, including machine learning applications like face detection, can contribute to creating safer and more efficient urban environments.