

CAM BASED HUMAN ACTIVITY IDENTIFICATION WITH MACHINE LEARNING

Submitted in partial fulfilment of the requirements for the award of Bachelor of Engineering Degree in Electronics and Communication Engineering

BY

KALI ANANTH NAG(40130091)

KATTUNGA JAYARAM(40130110)



**DEPARTMENT OF ELECTRONICS AND COMMUNICATION
ENGINEERING**

SCHOOL OF ELECTRICAL AND ELECTRONICS

SATHYABAMA

**INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)**

**Accredited with Grade “A++” by NAAC JEPPIAAR NAGAR, RAJIV GANDHI
SALAI, CHENNAI - 600 119**

APRIL – 2024



SATHYABAMA

INSTITUTE OF SCIENCE AND TECHNOLOGY
(DEEMED TO BE UNIVERSITY)

CATEGORY-I UNIVERSITY BY UGC

Accredited "A++" Grade by NAAC | 12B Status by UGC | Approved by AICTE

www.sathyabama.ac.in

**DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING
BONAFIDE CERTIFICATE**

This is to certify that this Project Report is the bonafide work of **K.Ananth Nag(40130091)** ,**K.Jayaram(40130110)** who carried out the project entitled "**CAM BASED HUMAN ACTIVITY IDENTIFICATION WITH MACHINE LEARNING**" under our supervision from November 2023 to April 2024.

Internal Guide

Dr. A.ARANGANATHAN. M.Tech., Ph.D.,

Head of the Department

Dr. T. RAVI. M.E., Ph.D.,

Submitted for Viva voce Examination held on 17.04.2024.

Internal Examiner

External Examiner

DECLARATION

I, K Ananth Nag hereby declare that the Project Report entitled “ **CAM BASED HUMAN ACTIVITY IDENTIFICATION WITH MACHINE LEARNING**” done by us under the guidance of Dr.A.Aranganathan M.Tech.,Ph.D., is submitted in partial fulfillment of the requirements for the Bachelor of Engineering degree in Electronics and Communication.

DATE: 18-04-2024

PLACE : Chennai

SIGNATURE OF THE CANDIDATE

1. *K. Ananth Nag*

DECLARATION

I, K Jayaram hereby declare that the Project Report entitled " **CAM BASED HUMAN ACTIVITY IDENTIFICATION WITH MACHINE LEARNING**" done by us under the guidance of Dr.A.Aranganathan M.Tech.,Ph.D., is submitted in partial fulfillment of the requirements for the Bachelor of Engineering degree in Electronics and Communication.

DATE: 18/4/24

PLACE : Chennai

SIGNATURE OF THE CANDIDATE

1. K. Jayaram

ACKNOWLEDGEMENT

We are pleased to acknowledge our sincere thanks to **Board of Management of SATHYABAMA INSTITUTE OF SCIENCE AND TECHNOLOGY** for their kind encouragement in doing this project and for completing it successfully. We are grateful to them.

We convey our thanks to **Dr. N. M. NANDHITHA, M.E.,Ph.D., Professor & Dean, School of Electrical and Electronics** and **Dr. T. RAVI, M.E., Ph.D., Professor & Head, Department of Electronics and Communication Engineering** for providing us necessary support and details at the right time during the progressive reviews.

We would like to express our sincere and deep sense of gratitude to our **INTERNAL GUIDE Dr.A.ARANGANATHAN M.Tech.,Ph.D.,**for his valuable guidance ,suggestions and constant encouragement paved way for the successful completion of our project work.

We wish to express our thanks to all Teaching and Non-teaching staff members of the Department of Electronics and Communication Engineering who were helpful in many ways for the completion of the project.

ABSTRACT

Abnormal event detection, human behavior recognition, and object recognition are critical elements in the development of modern CCTV systems. These systems are designed to go beyond mere surveillance, offering proactive capabilities to identify and respond to potential threats or irregularities in real-time. Leveraging the advancements in Machine Vision and Machine Learning, CCTV systems can analyze video feeds with remarkable accuracy, enabling the detection of anomalies, aberrant behaviors, and specific objects of interest. CCTV systems relied on frame-by-frame processing and Supervised Learning methods for training. However, the inherent limitations of supervised approaches, such as the need for extensive labeled data and the inability to account for all possible anomalies, have led to the adoption of more sophisticated techniques. Unsupervised and semi-supervised learning methods have emerged as viable alternatives, allowing systems to learn and adapt autonomously to evolving situations without the need for human intervention. This evolution in training methods not only accommodates the variability of anomalies but also reduces the dependence on labeled data, thereby mitigating the need for extensive manual effort in detection and alert generation. By employing unsupervised and semi-supervised learning, these systems become more adept at autonomously identifying anomalies, thus minimizing the human workload. The implementation of distributed abnormality classification systems represents a significant advancement in CCTV technology. By decentralizing the classification process and distributing it across multiple dedicated systems, these architectures offer scalability, resilience, and improved response times. Each dedicated system can specialize in analyzing specific types of anomalies, allowing for more accurate and efficient classification.

CONTENTS

CHAPTER NO:	TITLE	PAGE NO:
1	INTRODUCTION	1
2	LITERATURE SURVEY	5
3	REQUIREMENTS ANALYSIS	13
	3.1 HARDWARE REQUIREMENTS	13
	3.2 SOFTWARE REQUIREMENTS	16
4	SYSTEM ANALYSIS	18
	4.1 EXISTING SYSTEM	18
	4.2 DIS EXISTING SYSTEM	19
	4.3 PROPOSED SYSTEM	20
	4.4 FLESIBILITY STUDY	21
	4.5 CONSTRAINTS OF THE SYSTEM	23
	4.6 TRADE OFF	24
	4.7 STANDARDS	25
5	METHODOLOGY	26
	5.1 DATA COLLECTION	27
	5.2 PRE-PROCESSING	28
	5.3 SEGMENTATION	29
	5.4 CLASSIFICATIONS	30
6	SYSTEM TESTING	31
	6.1 WHITE BOX TESTING	32
	6.2 BLACK BOX TESTING	33
7	ARCHITECTURE	35

8	RESULTS AND DISCUSSION	38
	8.1 RESULT	38
	8.2 DISCUSSION	41
9	CONCLUSION	43
	REFERENCES	45

LIST OF FIGURES

FIGURE NO:	FIGURE NAME	PAGE NO:
3.1	ARDUINO UNO	13
3.2	GSM MODULE	14
3.3	BUZZER	15
3.4	ARDUINO IDE	16
3.5	MATLAB	17
6.3	SYSTEM TESTING DIAGRAM	34
7.1	ARCHITECTURE DIAGRAM	36
8.1	FRAMES	39
8.2	WORKING MODEL	40
8.3	MESSAGE ALERT	40
8.4	ARDUINO CODE	41

CHAPTER 1

INTRODUCTION

In the contemporary technological landscape, the fusion of machine vision and machine learning algorithms with closed-circuit television (CCTV) systems signifies a monumental leap forward, fundamentally reshaping the traditional role of surveillance cameras. These systems, once relegated to passive observation, now emerge as dynamic and proactive entities, capable of real-time monitoring, analysis, and response. This transformative shift heralds a paradigmatic evolution in how we perceive and utilize CCTV technology, transcending its conventional boundaries to become integral components of modern security infrastructure.

The integration of machine vision and machine learning imbues CCTV systems with unprecedented capabilities, empowering them to not merely record events but actively interpret and respond to dynamic environments. Machine vision endows these systems with the ability to process and interpret visual data, allowing them to discern objects, detect patterns, and recognize anomalies in real-time. Meanwhile, machine learning algorithms enable CCTV systems to learn from past experiences, refine their analytical capabilities, and adapt to evolving circumstances autonomously.

Within this context, the concept of CCTV cameras as passive observers gives way to a new paradigm characterized by active surveillance and intervention. Rather than serving merely as archival tools for post-event analysis, these advanced CCTV systems function as proactive guardians of public safety, constantly scanning their surroundings, identifying potential threats, and alerting authorities to imminent dangers. This transformative shift not only enhances the efficacy of security measures but also fosters a sense of reassurance and confidence within communities, knowing that their safety is actively monitored and protected.

Moreover, the integration of machine vision and machine learning into CCTV systems holds promise for a myriad of applications beyond traditional security concerns. By leveraging the analytical capabilities of these technologies, CCTV systems can contribute to a wide array of endeavors, ranging from public health monitoring and environmental conservation to urban planning and transportation optimization. This multifaceted approach underscores the versatility and adaptability of modern surveillance technology, positioning it as a potent tool for addressing complex challenges in contemporary society.

In essence, the fusion of machine vision and machine learning with CCTV systems heralds a new era of surveillance characterized by proactive intervention, real-time analysis, and adaptive response capabilities. This transformative shift not only enhances the effectiveness of security measures but also opens up a plethora of possibilities for leveraging surveillance technology to address a diverse array of societal challenges. As these technologies continue to evolve and proliferate, their potential to reshape our world and enhance our collective well-being remains boundless.

In today's rapidly evolving technological landscape, the integration of machine vision and machine learning algorithms into closed-circuit television (CCTV) systems represents a significant advancement with profound implications across diverse domains. At its core, this integration revolutionizes the traditional role of CCTV cameras from passive observers to active participants in real-time monitoring, analysis, and response. This paradigm shift holds immense potential for enhancing public safety, urban management, and various aspects of daily life.

- *Public Health Monitoring:* Amidst the ongoing global health challenges, such as the COVID-19 pandemic, CCTV systems equipped with machine vision can play a pivotal role in public health monitoring. By analyzing video feeds, these systems can monitor compliance with health measures like mask-wearing and

social distancing. Moreover, they can detect symptoms such as fever or coughing, aiding in early detection of potential outbreaks and enabling timely intervention by health authorities.

- *Environmental Monitoring:*Beyond human activity, CCTV systems can contribute to environmental monitoring efforts by analyzing video feeds to track changes in air quality, pollution levels, and wildlife habitats. This data can provide valuable insights into the impact of human activities on the environment, guiding conservation efforts and informing sustainable development initiatives.
- *Emergency Response Optimization:*In emergency situations, such as natural disasters or accidents, CCTV systems enhanced with machine vision can facilitate more efficient emergency response efforts. By automatically detecting signs of distress, locating survivors, and assessing the extent of damage, these systems provide critical information to first responders, enabling them to allocate resources effectively and expedite rescue operations.
- *Smart Transportation:*In the realm of transportation, advanced CCTV systems can contribute to the development of smart transportation systems. By analyzing traffic patterns, pedestrian movements, and public transit usage, these systems can optimize route planning, improve public transit efficiency, and reduce congestion, thereby enhancing urban mobility and sustainability.
- *Privacy and Ethical Considerations:*However, the widespread adoption of CCTV systems raises important privacy and ethical considerations. Striking the right balance between public safety and individual privacy rights requires clear regulations, transparent policies, and robust data protection measures. It is essential to ensure that surveillance technologies are used responsibly and ethically, safeguarding personal privacy while upholding public safety.

- *Collaborative Security Networks:*Moreover, as surveillance systems become increasingly interconnected, CCTV cameras can be integrated into collaborative security networks. By sharing real-time data and coordinating responses, these networks enhance situational awareness and response capabilities, fostering greater resilience against security threats.

In conclusion, the integration of machine vision and machine learning into CCTV systems heralds a new era of surveillance and security, with far-reaching implications for public safety, urban management, and societal well-being. However, realizing the full potential of these technologies requires careful consideration of ethical, privacy, and regulatory issues, alongside ongoing innovation and collaboration among stakeholders.

CHAPTER 2

LITERATURE REVIEW

2.1. R. Shahbazian and I. Trubitsyna, "Human Sensing by Using Radio Frequency Signals: A Survey on Occupancy and Activity Detection," in IEEE,2023

Applications for human sensing, also known as (human) occupancy detection, include energy management systems for intelligent buildings, intruder detection, e-health systems, the identification of everyday activity, and the monitoring of vital signs. These applications require intelligent decision-making that relies on human sensing. Multiple technologies based on vision, sensors, or radio signals can be used to detect occupancy. Vision-based systems use a multitude of cameras to recognize the human presence, but they are restricted by light availability, line-of-sight coverage, expensive equipment, and privacy concerns. Sensor-based techniques refer to a prospective method that employs various combinations of sensors. These solutions are static and necessitate costly equipment installation and maintenance. Due to technical advancements, radio-based signals, such as WiFi, have been integrated into various forms of infrastructure, including homes, offices, and constructions. Due to how human body movements affect wireless signal propagation, it is possible to detect human motions by analyzing the received wireless signals (such as reflection, diffraction, and scattering). Due to its low cost and non-intrusive nature, wireless-based human activity detection has received substantial attention and become a key topic of study. This article reviews the underlying principles, methodologies, and system architectures of radio-frequency-based occupancy detection systems. We classify the reviewed research studies based on the technical measures and applications they employ. In addition to focusing on the security aspects of occupancy detection and discussing future trends and difficulties, we also discuss practical considerations.

2.2. M. M. Sandhu, S. Khalifa, K. Geissdoerfer, R. Jurdak, M. Portmann and B. Kusy, "FusedAR: Energy-Positive Human Activity Recognition Using Kinetic and Solar Signal Fusion," in IEEE, June 1, 2023

Today's wearable Internet of Things (IoT) devices, which have numerous practical applications, suffer from the limited lifetime of batteries due to the high power consumption of conventional inertial activity sensors. Recently, kinetic energy harvesters have been employed as a source of energy as well as context information to replace conventional activity sensors. However, the harvested power from human movements using miniaturized kinetic transducers may not be sufficient to enable a perpetual and self-powered activity recognition system. In this article, we propose a novel mechanism of fused signal-based human activity recognition (FusedAR), which employs miniaturized wearable solar and kinetic energy harvesters simultaneously as an energy source as well as an activity sensor. As human activities engender distinct movement patterns and interact and interfere with the ambient light differently, the kinetic and solar energy harvesting (SEH) signals incorporate unique information about the underlying activities while generating sufficient power. After detailed experiments, we find that the FusedAR, which employs both solar and kinetic energy signals, achieves superior activity recognition performance by up to 10%, particularly in outdoor and night-time contexts, and can recognize not only activities but also contexts, compared with the individual energy harvesting signals. Furthermore, our analysis demonstrates that FusedAR, in addition to significant energy generation, consumes up to 22% less power than the highly optimized conventional three-axis accelerometer-based mechanisms, achieving energy-positive human activity recognition (HAR) leading toward perpetual, uninterrupted, and autonomous operation of wearable IoT devices.

2.3. Ullmann, R. G. Guendel, N. C. Kruse, F. Fioranelli and A. Yarovoy, "A Survey on Radar-Based Continuous Human Activity Recognition," in IEEE, July 2023

Radar-based human motion and activity recognition is currently a topic of great research interest, as the aging population increases and older individuals prefer an independent lifestyle. This technology has a wide range of applications, such as fall

detection in assisted living, gesture recognition for human-machine interfaces, and many more. Numerous studies exist on various approaches for radar-based activity capture and classification. However, most of these employ rather artificial data, often obtained in laboratory environments, and typically collected under particular conditions. Specifically, most research so far has aimed at distinguishing a predefined set of single activities with a defined start, stop and duration. This paper aims at drawing the attention to a so far less researched issue, one that will be of vital importance for future real-world application of radar-based human activity recognition: continuous activity recognition, i.e. recognizing specific activities in a stream of several sequential activities with unknown duration and arbitrary transitions between different classes of activities. A review on the current state of the art in this relatively new topic is given, followed by a discussion on future research directions.

2.4.Hafeez, S. S. Alotaibi, A. Alazeb, N. A. Mudawi and W. Kim, "Multi-Sensor-Based Action Monitoring and Recognition via Hybrid Descriptors and Logistic Regression," in IEEE,2023

In the fields of body-worn sensors and computer vision, current research is being done to track and detect falls and activities of daily living using the automatic recognition of human actions. In the area of human-machine communication, different combinations of sensors and communication technologies are often used to capture human action. Many researchers have also worked with artificial intelligent systems to detect actions, understand scenes, and implement systems that are more efficient in human action recognition. Although effective approaches are needed to detect outdoor activities with the combination of human actions, feature extraction can be quite a complicated task in a human activity recognition system development. Thus, this paper proposed a solution to detect human activities via hybrid descriptors based on robust features and accurate results. In this study, complex backgrounds, including multiple humans in video frames, were detected. First, inertial signal and video frames are pre-processed using denoising techniques, after which the frames are used to remove the background by detecting human motions and extracting the silhouettes. Then, these silhouettes are further used to extract the human body key points to make the human skeleton. Then the time and frequency domain features are extracted for inertial

signals, and geometric features are extracted for the skeleton body points. Finally, multiple feature sets are combined and fed into a zero order optimization model, after which logistic regression is utilized to recognize each action. The proposed system has been evaluated on three benchmark datasets, including, the UP Fall dataset, the University of Rzeszow Fall dataset, and the SisFall dataset and proved its significance by achieving accuracy of 91.51%, 92.98%, and 90.23%, on the aforementioned datasets respectively.

2.5.-T. Shen et al., "Archangel: A Hybrid UAV-Based Human Detection Benchmark With Position and Pose Metadata," in IEEE,2023.

Learning to detect objects, such as humans, in imagery captured by an unmanned aerial vehicle (UAV) usually suffers from tremendous variations caused by the UAV's position towards the objects. In addition, existing UAV-based benchmark datasets do not provide adequate dataset metadata, which is essential for precise model diagnosis and learning features invariant to those variations. In this paper, we introduce Archangel, the first UAV-based object detection dataset composed of real and synthetic subsets captured with similar imagining conditions and UAV position and object pose metadata. A series of experiments are carefully designed with a state-of-the-art object detector to demonstrate the benefits of leveraging the metadata during model evaluation. Moreover, several crucial insights involving both real and synthetic data during model optimization are presented. In the end, we discuss the advantages, limitations, and future directions regarding Archangel to highlight its distinct value for the broader machine learning community.

2.6. J. Kim, H. Kang, J. Yang, H. Jung, S. Lee and J. Lee, "Multitask Deep Learning for Human Activity, Speed, and Body Weight Estimation Using Commercial Smart Insoles," in IEEE,Sept.15, 2023

Healthcare professionals and individual users use wearable devices equipped with various sensors for healthcare management. Recently, the joint usage of artificial intelligence and these wearable sensors has played an essential role in healthcare management by providing a wide range of applications, such as fitness tracking, gym activity monitoring, patient rehabilitation monitoring, and disease detection. These

tasks eventually aim to enhance personal well-being and better manage the user's physical health by monitoring different activity types and body weight changes. Here, we present an efficient multitask learning (MTL) framework based on commercial smart insoles that can solve three tasks related to physical health management: 1) activity classification; 2) speed estimation; and 3) body weight estimation. Our multitask framework converts the sensor data from the smart insole to a recurrence plot, which shows significant performance improvement compared to processing the raw time-series data. In addition, we utilized a modified MobileNetV2 as our backbone network, which has a total parameter of less than 100K and a computational budget of 0.34G of multiply-accumulate operations. Furthermore, we collected a vast data set from 72 users carrying out 16 experiments, which contains the largest number of people for MTL purposes using smart insoles. Extensive experiments show that the proposed MTL framework is extremely efficient while outperforming or leading to comparable performance against single-task models.

2.7. Kalabakov et al., "Federated Learning for Activity Recognition: A System Level Perspective," in IEEE, 2023

The past decade has seen substantial growth in the prevalence and capabilities of wearable devices. For instance, recent human activity recognition (HAR) research has explored using wearable devices in applications such as remote monitoring of patients, detection of gait abnormalities, and cognitive disease identification. However, data collection poses a major challenge in developing HAR systems, especially because of the need to store data at a central location. This raises privacy concerns and makes continuous data collection difficult and expensive due to the high cost of transferring data from a user's wearable device to a central repository. Considering this, we explore the adoption of federated learning (FL) as a potential solution to address the privacy and cost issues associated with data collection in HAR. More specifically, we investigate the performance and behavioral differences between FL and deep learning (DL) HAR models, under various conditions relevant to real-world deployments. Namely, we explore the differences between the two types of models when (i) using data from different sensor placements, (ii) having access to users with data from heterogeneous sensor placements, (iii) considering bandwidth efficiency,

and (iv) dealing with data with incorrect labels. Our results show that FL models suffer from a consistent performance deficit in comparison to their DL counterparts, but achieve these results with much better bandwidth efficiency. Furthermore, we observe that FL models exhibit very similar responses to those of DL models when exposed to data from heterogeneous sensor placements. Finally, we show that the FL models are more robust to data with incorrect labels than their centralized DL counterparts.

2.8. M. M. Sandhu, S. Khalifa, K. Geissdoerfer, R. Jurdak, M. Portmann and B. Kusy, "FusedAR: Energy-Positive Human Activity Recognition Using Kinetic and Solar Signal Fusion," in IEEE, June1, 2023

Today's wearable Internet of Things (IoT) devices, which have numerous practical applications, suffer from the limited lifetime of batteries due to the high power consumption of conventional inertial activity sensors. Recently, kinetic energy harvesters have been employed as a source of energy as well as context information to replace conventional activity sensors. However, the harvested power from human movements using miniaturized kinetic transducers may not be sufficient to enable a perpetual and self-powered activity recognition system. In this article, we propose a novel mechanism of fused signal-based human activity recognition (FusedAR), which employs miniaturized wearable solar and kinetic energy harvesters simultaneously as an energy source as well as an activity sensor. As human activities engender distinct movement patterns and interact and interfere with the ambient light differently, the kinetic and solar energy harvesting (SEH) signals incorporate unique information about the underlying activities while generating sufficient power. After detailed experiments, we find that the FusedAR, which employs both solar and kinetic energy signals, achieves superior activity recognition performance by up to 10%, particularly in outdoor and night-time contexts, and can recognize not only activities but also contexts, compared with the individual energy harvesting signals. Furthermore, our analysis demonstrates that FusedAR, in addition to significant energy generation, consumes up to 22% less power than the highly optimized conventional three-axis accelerometer-based mechanisms, achieving energy-positive human activity recognition (HAR) leading toward perpetual, uninterrupted, and autonomous operation of wearable IoT devices.

2.9. T. Shen et al., "Archangel: A Hybrid UAV-Based Human Detection Benchmark With Position and Pose Metadata," in IEEE,2023

Learning to detect objects, such as humans, in imagery captured by an unmanned aerial vehicle (UAV) usually suffers from tremendous variations caused by the UAV's position towards the objects. In addition, existing UAV-based benchmark datasets do not provide adequate dataset metadata, which is essential for precise model diagnosis and learning features invariant to those variations. In this paper, we introduce Archangel, the first UAV-based object detection dataset composed of real and synthetic subsets captured with similar imaging conditions and UAV position and object pose metadata. A series of experiments are carefully designed with a state-of-the-art object detector to demonstrate the benefits of leveraging the metadata during model evaluation. Moreover, several crucial insights involving both real and synthetic data during model optimization are presented. In the end, we discuss the advantages, limitations, and future directions regarding Archangel to highlight its distinct value for the broader machine learning community.

2.10. J. Kim, H. Kang, J. Yang, H. Jung, S. Lee and J. Lee, "Multitask Deep Learning for Human Activity, Speed, and Body Weight Estimation Using Commercial Smart Insoles," in IEEE,Sept.15, 2023

Healthcare professionals and individual users use wearable devices equipped with various sensors for healthcare management. Recently, the joint usage of artificial intelligence and these wearable sensors has played an essential role in healthcare management by providing a wide range of applications, such as fitness tracking, gym activity monitoring, patient rehabilitation monitoring, and disease detection. These tasks eventually aim to enhance personal well-being and better manage the user's physical health by monitoring different activity types and body weight changes. Here, we present an efficient multitask learning (MTL) framework based on commercial smart insoles that can solve three tasks related to physical health management: 1) activity classification; 2) speed estimation; and 3) body weight estimation. Our multitask framework converts the sensor data from the smart insole to a recurrence plot, which

shows significant performance improvement compared to processing the raw time-series data. In addition, we utilized a modified MobileNetV2 as our backbone network, which has a total parameter of less than 100K and a computational budget of 0.34G of multiply-accumulate operations. Furthermore, we collected a vast data set from 72 users carrying out 16 experiments, which contains the largest number of people for MTL purposes using smart insoles. Extensive experiments show that the proposed MTL framework is extremely efficient while outperforming or leading to comparable performance against single-task models.

CHAPTER 3

REQUIREMENT ANALYSIS

3.1 HARDWARE REQUIREMENTS:

In this project we are using Arduino UNO,GSM Module,GPS Module,Buzzer.

3.1.1 :*Arduino UNO:*



Fig 3.1:Arduino uno

From fig 3.1 the Arduino Uno is a compact and versatile microcontroller board equipped with the ATmega328P chip, providing 14 digital input/output pins and 6 analog input pins for interfacing with a wide range of sensors, actuators, and other electronic components. It features USB connectivity for easy programming and serial communication, enabling users to upload and execute code effortlessly. With its user-friendly Arduino IDE, programming the Uno becomes a straightforward process, supported by a vast online community and extensive documentation. Its affordability and simplicity make it an accessible platform for beginners and experienced users alike, facilitating rapid prototyping and experimentation in electronics projects. Additionally, the Uno's compatibility with various shields allows for further expansion

and customization, making it an ideal choice for hobbyists, educators, and professionals seeking to explore the world of embedded systems and robotics.

3.1.2 :GSM Module:



Fig 3.2 : GSM Module

As shown in fig 3.2 we can see a GSM (Global System for Mobile communication) module is a hardware component that enables devices to communicate via cellular networks. Typically, it integrates a GSM modem along with other necessary components such as a SIM card slot and antenna. These modules support GSM networks worldwide, allowing devices to send and receive SMS messages, make voice calls, and access mobile data services. The module communicates with a microcontroller or computer via serial communication protocols such as UART, enabling seamless integration into various electronic systems. GSM modules are widely used in applications such as remote monitoring, telemetry, security systems, and IoT devices, where access to cellular networks is essential. They offer reliable wireless connectivity, making them suitable for applications in remote or mobile environments where traditional wired communication is impractical. Additionally, GSM modules come in different form factors and variants, including 2G, 3G, and 4G/LTE, providing flexibility and compatibility with different network technologies and frequencies. The integration of GSM modules into electronic projects enables remote control and monitoring capabilities, empowering users to stay connected and informed

regardless of their location. Moreover, these modules often come with additional features such as GPS functionality, enabling location tracking and asset management applications. Overall, GSM modules serve as indispensable components in modern communication systems, facilitating reliable and efficient wireless communication over cellular networks.

3.1.3: Buzzer:



Fig 3.3 :Buzzer

As shown in fig 3.3 a buzzer is an electronic component that produces sound when an electrical signal is applied to it. It consists of a piezoelectric element or an electromagnetic coil attached to a diaphragm. When activated, the piezoelectric element vibrates, producing audible sound waves, while the electromagnetic coil creates sound by causing the diaphragm to vibrate. Buzzers are commonly used in electronic devices and projects for generating audible alerts, notifications, or alarms. They come in various types, including passive and active buzzers. Passive buzzers require an external signal to produce sound, while active buzzers have an internal oscillator circuit that generates a sound when power is applied. Buzzers are widely used in applications such as alarms, timers, doorbells, and electronic games due to their simplicity, low cost, and ease of integration.

3.2 SOFTWARE REQUIREMENTS:

Arduino IDE , Matlab.

3.2.1 Arduino IDE:

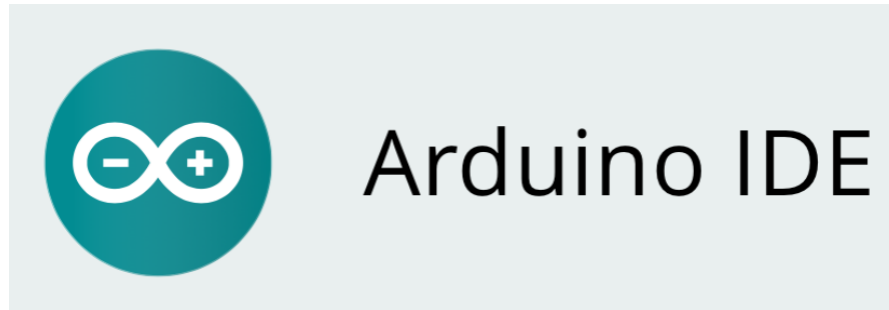


Fig 3.4 : Arduino IDE

The Arduino Integrated Development Environment (IDE) is a software platform designed to simplify the process of programming Arduino microcontroller boards. It provides a user-friendly interface for writing, compiling, and uploading code to Arduino-compatible devices. The IDE is open-source and available for Windows, macOS, and Linux operating systems, ensuring accessibility across different platforms. It features a text editor with syntax highlighting and auto-completion, making it easier for users to write and edit code. The IDE includes a built-in library manager, allowing users to easily add and manage libraries for additional functionalities. Furthermore, it supports a wide range of Arduino boards and variants, providing compatibility with various hardware configurations. The IDE integrates seamlessly with the Arduino hardware, facilitating the uploading of compiled code to the target device via USB or other communication interfaces. Additionally, it includes a serial monitor tool for debugging and monitoring serial communication between the Arduino board and the computer. The Arduino IDE fosters a vibrant community of makers, hobbyists, educators, and professionals, with extensive online resources, tutorials, and forums available for support and collaboration. Its simplicity, versatility, and robustness have contributed to its widespread adoption as the go-to development environment for Arduino projects. Moreover, the IDE's open-source nature encourages continuous improvement and customization, with users contributing to its development and enhancement. Overall,

the Arduino IDE serves as a powerful tool for prototyping, experimenting, and creating innovative projects in the world of electronics and embedded systems.

3.2.2 Matlab:



Fig 3.5 : Matlab

MATLAB is a high-level programming language and interactive environment primarily used for numerical computation, data analysis, and visualization. Developed by MathWorks, it offers a wide range of built-in functions and toolboxes for various domains such as signal processing, image processing, control systems, and machine learning. MATLAB provides an intuitive syntax and graphical user interface (GUI), making it accessible to engineers, scientists, and researchers for solving complex mathematical problems and performing simulations. Its powerful matrix manipulation capabilities enable efficient handling of large datasets and mathematical operations. MATLAB's extensive plotting and visualization tools allow users to create interactive plots, graphs, and animations to visualize and analyze data effectively. Additionally, MATLAB supports integration with external languages such as C, C++, and Java, facilitating interoperability with other software and hardware platforms. Its versatility and robustness make it suitable for a wide range of applications, including academic research, industrial development, and prototyping of algorithms and systems. MATLAB's widespread adoption is further supported by its active community, comprehensive documentation, and online resources, providing users with ample support and learning opportunities.

CHAPTER 4

SYSTEM ANALYSIS

4.1 EXISTING SYSTEM

The existing system introduces an innovative approach for detecting anomalies and changes in remote sensing (RS) images through an unsupervised method. This method leverages a multivariate Gaussianization technique to accurately estimate multivariate densities, addressing a longstanding challenge in statistics and machine learning.

The essence of this methodology lies in transforming highly complex multivariate data into a multivariate Gaussian distribution. This transformation is pivotal as it enables the system to work with data in a form that is more amenable to analysis. By achieving this Gaussianization, the system gains the ability to model the underlying probability distribution of the data more effectively.

A key advantage of this approach is its differentiability. By ensuring that the transformation process is differentiable, the system can apply the change of variables formula. This formula enables the estimation of probabilities at any point within the original domain of the data. This means that the system can assess the likelihood of observing a particular data point within the RS image, allowing for robust anomaly detection and change identification.

By utilizing multivariate Gaussianization, the system can handle diverse and complex patterns present in RS images. This capability is particularly valuable in the context of remote sensing, where images often contain intricate spatial and spectral information. The method's ability to accurately estimate multivariate densities equips the system to discern subtle deviations or anomalies within the data, which may indicate changes or irregularities in the observed scene.

Overall, the proposed unsupervised method represents a significant advancement in RS image analysis. By combining Gaussianization with sophisticated statistical and machine learning techniques, the system offers a powerful tool for detecting anomalies and changes, thereby enhancing the effectiveness of remote sensing applications in various domains such as environmental monitoring, urban planning, and agriculture.

4.2 DIS EXISTING SYSTEM

The envisioned system presents a promising solution aimed at alleviating the burden on human operators by automating surveillance tasks and effectively managing multiple video streams from a centralized storage location. By streamlining the monitoring process, this system enhances operational efficiency and enables more comprehensive surveillance coverage across various locations.

Moreover, the system's capability to accept multiple video streams from a centralized storage location signifies its scalability and adaptability to diverse surveillance environments. This centralized approach simplifies data management and facilitates seamless access to video feeds, enhancing the system's overall usability and effectiveness.

However, it's essential to acknowledge that the system may encounter challenges in achieving consistent and accurate data collection, particularly concerning crowd density and movement analysis. Factors such as fluctuating lighting conditions, complex crowd behaviors, and occlusions can introduce variability and inaccuracies in the collected data.

Despite these challenges, the system represents a significant step forward in automating surveillance operations and enhancing situational awareness. By leveraging advanced algorithms and sensor technologies, future iterations of the system can strive to improve accuracy and consistency in crowd density and

movement detection, thereby enhancing its reliability and usefulness in various applications.

Continued research and development efforts are necessary to refine the system's capabilities and address existing limitations. By incorporating feedback from real-world deployments and leveraging advancements in machine learning, computer vision, and data processing techniques, the system can evolve into a valuable tool for enhancing security, public safety, and operational efficiency across a wide range of surveillance scenarios.

4.3 PROPOSED SYSTEM

The proposed system offers an innovative approach to automating the detection of abnormal events from CCTV camera feeds. By leveraging cutting-edge technologies such as Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks, anomalies can be detected both in supervised and unsupervised manners, enhancing the system's adaptability to various surveillance scenarios.

One of the key features of the system is its ability to send alert messages to authorities upon detecting abnormal events. This real-time notification capability enables swift responses to potential threats or incidents, improving overall security and safety measures.

Furthermore, the system optimizes storage efficiency by storing original quality video snippets of abnormal events in high quality while compressing and archiving normal events in low quality. This approach minimizes storage requirements while ensuring that critical footage is retained for future analysis or investigation.

The design of the system encompasses several modules, each fulfilling a specific function to achieve its overall objectives. The Video Compressor module is responsible for compressing original video feeds to low quality and resolution for

archival purposes. This helps conserve storage space while maintaining access to historical footage.

The Anomaly Detector module employs unsupervised learning methods to identify anomalies in real-time CCTV feeds. By training on a diverse range of data, the system can detect both predetermined and undetermined anomalies, enhancing its versatility and effectiveness in anomaly detection.

The Storage Management module plays a crucial role in organizing and storing both low-quality archived videos and original-quality anomalous video snippets. This ensures efficient retrieval and management of stored footage for future reference or analysis.

Finally, the Alert Management module is responsible for managing and disseminating alerts to relevant personnel or authorities upon the identification of an anomaly. This ensures that appropriate actions can be taken promptly to address the detected event.

In summary, the proposed system offers a comprehensive solution for automating anomaly detection in CCTV surveillance systems. By leveraging advanced technologies and modular design principles, the system reduces human workload, enhances storage efficiency, and improves overall surveillance capabilities, contributing to enhanced security and safety measures in various environments.

4.4 FLESIBILITY STUDY:

Conducting a feasibility study for the implementation of advanced CCTV systems integrating machine vision and machine learning technologies requires a comprehensive analysis of various factors. Firstly, considering the prevalent deployment of CCTV cameras in diverse environments, there is likely substantial market demand for enhanced surveillance solutions capable of real-time anomaly detection. The increasing emphasis on security and safety in public spaces further

underscores the necessity for proactive surveillance measures to mitigate potential risks effectively.

Moreover, from a technological standpoint, the feasibility of such systems is promising. With continuous advancements in machine vision and machine learning algorithms, the necessary technology for developing these systems is readily available. These technologies have already demonstrated success across various domains, indicating their feasibility and potential for application in CCTV systems.

However, alongside technological feasibility, it's essential to assess the cost considerations associated with implementing advanced CCTV systems. This includes expenses related to hardware procurement (such as high-resolution cameras), software development, acquisition of training data, and ongoing maintenance. Conducting a thorough cost-benefit analysis is crucial to ascertain the financial viability of the project.

Furthermore, regulatory compliance is paramount in the development and deployment of advanced CCTV systems. Adhering to local regulations and privacy laws governing video surveillance is imperative to avoid potential legal ramifications. Ensuring compliance with data protection regulations regarding data collection, storage, and usage is essential to safeguard individuals' privacy rights.

Scalability and integration capabilities also need to be evaluated during the feasibility study. The system must be scalable to accommodate varying surveillance needs and seamlessly integrate with existing infrastructure. Compatibility with different CCTV camera models, communication protocols, and data storage systems is crucial for smooth integration and interoperability.

Additionally, assessing the system's performance and reliability is critical. Extensive testing and validation procedures should be conducted to ensure accurate and real-time anomaly detection across different environments and scenarios. User

acceptance is also a key consideration, requiring stakeholder engagement to understand end-users' needs and requirements.

Finally, evaluating the potential return on investment (ROI) associated with implementing advanced CCTV systems is essential. This includes assessing potential cost savings from reduced security incidents, improved public safety, and enhanced operational efficiency. By carefully considering these factors, a comprehensive feasibility study can provide valuable insights into the viability and potential benefits of implementing advanced CCTV systems with machine vision and machine learning capabilities.

4.5 CONSTRAINTS OF THE SYSTEM

Constraints play a crucial role in shaping the development and implementation of any project, including the Human Activity Identification System. Here are some potential constraints that had an impact for this project:

- *Data Collection:* Gathering labeled data for training the model is time-consuming and can be expensive. Quality and quantity of the data are crucial for accurate activity recognition.
- *Feature Extraction:* Selecting relevant features that capture the essence of human activities is challenging. Incorrect or insufficient features can lead to poor model performance.
- *Variability:* Human activities can be performed in various ways by different individuals, leading to intra-class variability. The model needs to generalize well to recognize activities performed differently.
- *Real-time Processing:* For real-time applications, the model must be efficient enough to process data and make predictions quickly, which can be a computational challenge.
- *Overfitting:* Complex models may memorize the training data rather than learning patterns, leading to poor generalization to new, unseen data.
- *Sensor Noise and Errors:* Sensor data used for activity recognition can be noisy or contain errors, affecting the model's performance. Preprocessing steps are needed to handle this.

4.6 TRADE OFF

Trade-offs in human activity identification using machine learning and MATLAB often revolve around accuracy, computational complexity, and real-time processing. Here are some of the key trade-offs:

- *Accuracy vs. Complexity*: Higher complexity algorithms can achieve better accuracy by capturing intricate patterns in the data. Simpler models may sacrifice some accuracy but require less computational power and are easier to interpret.
- *Real-time Processing vs. Accuracy*: Achieving real-time processing often requires simpler models or optimized algorithms, which might sacrifice a bit of accuracy. Highly accurate models might not be suitable for real-time applications due to their computational demands.
- *Feature Engineering vs. Automatic Feature Learning*: Manual feature engineering allows for more control and understanding of the features used for identification but can be time-consuming and may miss relevant features. Automatic feature learning with deep learning can capture complex relationships in the data but requires more data and computational resources.
- *Interpretability vs. Performance*: Simple models like decision trees or linear models are more interpretable but might not achieve the same level of performance as complex models like deep neural networks.
- *Data Quantity vs. Data Quality*: More data generally leads to better model performance, but collecting high-quality labeled data can be challenging and expensive. Balancing the quantity and quality of data is crucial for developing robust models.
- *Generalization vs. Overfitting*: Striking a balance between building a model that generalizes well to new data and avoiding overfitting to the training data is essential. Regularization techniques can help prevent overfitting but might slightly reduce the model's performance.

4.7 STANDARDS IN THIS PROJECT

- IEEE 1855 - Standard for Accessible and Usable Multimedia Content: This standard focuses on making multimedia content accessible and usable, which can be relevant for developing user-friendly interfaces for human activity identification systems.
- IEEE 754 - Standard for Floating-Point Arithmetic: Although more general, understanding floating-point arithmetic can be crucial for implementing machine learning algorithms accurately in MATLAB or other programming languages.
- IEEE 1012 - Standard for System and Software Verification and Validation: This standard provides guidelines for verifying and validating software systems, which can be applied to ensure the reliability and accuracy of human activity identification models.
- IEEE 1451 - Standard for a Smart Transducer Interface for Sensors and Actuators: This standard can be relevant when dealing with sensor data acquisition and integration in human activity recognition systems.
- IEEE 1599 - Standard for Adaptive Systems of Augmented Reality: While focused on augmented reality, this standard can provide insights into designing adaptive and interactive user interfaces for human activity identification applications.
- IEEE 1872 - Standard Ontologies for Robotics and Automation: Although aimed at robotics and automation, the concept of standard ontologies can be useful for defining and categorizing human activities in a structured manner.
- IEEE P7000 - Model for the Consumer Internet of Things (IoT): As IoT devices often collect human activity data, understanding the principles and guidelines outlined in this model can be beneficial for integrating IoT sensors with machine learning models for human activity identification.

CHAPTER 5

METHODOLOGY

❖ MODULE LIST

- ✓ Data Collection
- ✓ Pre-Processing
- ✓ Segmentation
- ✓ Classifications

❖ MODULE DESCRIPTION

5.1 Data Collection : For data collection, we have curated two primary datasets: the UCSD anomaly detection dataset and the Avenue dataset. These datasets serve as crucial resources for training and testing our anomaly detection system. The UCSD dataset comprises videos captured by a camera positioned at an elevated height, overlooking a pedestrian walkway. These videos predominantly feature pedestrian activity.

Within the UCSD dataset, anomalies manifest as non-pedestrian entities entering the walkway, such as bikers, skaters, and small carts. Additionally, abnormal events encompass atypical pedestrian motion patterns, such as individuals traversing the walkway or venturing onto the surrounding grass areas. This dataset is subdivided into two parts: Ped1 and Ped2, each containing instances of anomalous behavior for analysis.

In addition to the UCSD dataset, we incorporate the Avenue dataset into our training and testing procedures. The Avenue dataset complements the UCSD data by providing additional examples of abnormal events in urban environments. By leveraging both datasets, our system can learn a diverse range of anomaly patterns and enhance its robustness in detecting deviations from normal behavior.

By utilizing Ped1, Ped2, and the Avenue dataset for training and testing, our system gains exposure to a wide array of anomalous scenarios, enabling it to recognize and classify abnormalities effectively. This comprehensive approach to data collection ensures that our anomaly detection system is well-equipped to handle various real-world surveillance scenarios, contributing to improved security and safety measures in diverse environments.

5.2 Pre-Processing : In the pre-processing stage of our system, we prepare the training dataset by extracting sequences of normal video frames. These sequences serve as the basis for training our model to reconstruct normal behavior patterns. To streamline the training process and optimize memory usage, we adopt a strategy of subsampling the frames.

Specifically, we select every 5th alternate frame from the video sequence to form our training set. By doing so, we effectively reduce the number of frames that need to be processed and stored, thus mitigating computational demands and memory requirements. This subsampling approach allows us to maintain a representative sample of the original video data while significantly reducing the dataset's size.

Subsampling the frames at regular intervals also helps to capture key temporal dynamics and variations in the video sequences. While reducing the frame rate may result in some loss of temporal information, the selected frames still provide sufficient context for training the model to reconstruct normal behavior patterns effectively.

By employing this pre-processing strategy, we strike a balance between computational efficiency and model effectiveness. The subsampled frames facilitate efficient training of our anomaly detection model while retaining the essential information needed to learn and recognize normal behavior patterns. As a result, our system can effectively detect and classify abnormal events in real-world surveillance footage, contributing to enhanced security and safety measures in various environments.

5.3 Segmentation : In the segmentation phase of our system, we employ specific optimization and initialization techniques to train our anomaly detection model effectively. We utilize the Adam Optimizer, a popular optimization algorithm widely used in deep learning tasks, to minimize the reconstruction loss during training. The learning rate is set to 0.0001, ensuring that the optimization process progresses at an appropriate pace to converge towards the optimal solution.

To prevent the learning rate from becoming too small and stagnating the optimization process, we implement a learning rate decay mechanism. When the training loss stops decreasing, indicating a potential convergence or plateauing of the optimization process, the learning rate is reduced. We apply a decay rate of 0.00001 to gradually decrease the learning rate, allowing the optimization process to fine-tune and refine the model parameters further.

Additionally, we set the epsilon value of the Adam Optimizer to 0.000001. Epsilon is a small constant added to the denominator of the update step to prevent division by zero and stabilize the optimization process, particularly when dealing with small gradients or numerical instabilities.

For weight initialization, we utilize the Xavier algorithm, also known as Xavier initialization or Glorot initialization. This algorithm initializes the weights of the neural network layers in a way that prevents the signal from becoming too small or too large as it propagates through each layer during forward and backward propagation. By ensuring that the weights are initialized appropriately, the Xavier algorithm helps stabilize the training process and facilitates efficient learning.

Overall, by incorporating these optimization and initialization techniques into our segmentation phase, we aim to optimize the training process, prevent overfitting, and enhance the convergence and performance of our anomaly detection model. Through careful parameter tuning and initialization strategies, we strive to develop a robust and

effective system capable of accurately detecting and classifying abnormal events in real-world surveillance footage.

5.4 Classifications : In the classification phase of our system, each video is individually tested to detect abnormal events. We utilize three main datasets for testing: UCSD Ped1, UCSD Ped2, and the Avenue dataset. Each dataset presents unique characteristics and challenges, allowing our system to generalize its anomaly detection capabilities across diverse surveillance scenarios.

In the UCSD Ped1 dataset, we have 36 testing videos, each containing 200 frames. However, to streamline processing and conserve memory usage, we subsample the frames by selecting every 5th alternate frame. This results in a total of 40 frames from each video. Similarly, in the UCSD Ped2 dataset, which comprises 12 testing videos of varying lengths, we adopt the same subsampling strategy to reduce processing time and memory usage.

The Avenue dataset consists of 21 testing videos with varying durations. Similar to the UCSD datasets, we select every 5th alternate frame from each video to mitigate computational demands while maintaining a representative sample of the original data.

To identify abnormal events within the video sequences, we employ a sliding window technique. This involves segmenting the video into sequences of four consecutive frames, after subsampling the frames. For each sequence, we calculate a regularity score, $S_r(t)$, where t represents the starting frame index. If the reconstruction error value exceeds a predefined threshold, our system triggers an alert signal to notify relevant personnel or authorities of the detected anomaly.

Furthermore, upon detection of an abnormal event, the system initiates the storage of frames in their original quality and resolution. This allows for future reference and analysis of the anomaly. We store a predefined number of frames before and after the occurrence of the abnormality to provide a comprehensive view of the event.

Typically, we store 10 to 20 frames before the occurrence and 100 to 120 frames after the occurrence to ensure a clear understanding of the abnormality and its context.

By implementing these classification techniques and storage strategies, our system can effectively detect and respond to abnormal events in real-time surveillance footage, contributing to enhanced security and safety measures in various environments.

CHAPTER 6

SYSTEM TESTING

System testing usually consists of a layered process, including the user interface (UI) layer, the business layer, the data access layer and the database itself. The UI layer deals with the interface design of the database, while the business layer includes databases supporting business strategies.

Databases, the collection of interconnected files on a server, storing information, may not deal with the same type of data, i.e. databases may be heterogeneous. As a result, many kinds of implementation and integration errors may occur in large database systems, which negatively affect the system's performance, reliability, consistency and security. Thus, it is important to test in order to obtain a database system which satisfies the ACID properties (Atomicity, Consistency, Isolation, and Durability) of a database management system.

One of the most critical layers is the data access layer, which deals with databases directly during the communication process. Database testing mainly takes place at this layer and involves testing strategies such as quality control and quality assurance of the product databases. Testing at these different layers is frequently used to maintain consistency of database systems, most commonly seen in the following examples:

Data is critical from a business point of view. Companies such as Google or Symantec, who are associated with data storage, need to have a durable and consistent database system. If database operations such as insert, delete, and update are performed without testing the database for consistency first, the company risks a crash of the entire system.

Some companies have different types of databases, and also different goals and missions. In order to achieve a level of functionality to meet said goals, they need to test their database system.

The current approach of testing may not be sufficient in which developers formally test the databases. However, this approach is not sufficiently effective since database developers are likely to slow down the testing process due to communication gaps. A separate database testing team seems advisable.

Database testing mainly deals with finding errors in the databases so as to eliminate them. This will improve the quality of the database or web-based system. Database testing should be distinguished from strategies to deal with other problems such as database crashes, broken insertions, deletions or updates. Here, database refactoring is an evolutionary technique that may apply.

6.1 WHITE BOX TESTING: From fig 6.1 White-box testing, also known as clear box testing, glass box testing, or structural testing, is a software testing method that focuses on examining the internal structures and workings of an application. Unlike black-box testing, which evaluates the functionality of the software without considering its internal implementation, white-box testing involves a deep understanding of the code and its internal logic.

In white-box testing, testers utilize their knowledge of the internal workings of the system and programming skills to design test cases. They select inputs that exercise specific paths through the code, aiming to uncover errors or defects within the software. This approach is akin to testing individual nodes in a circuit, a concept familiar in in-circuit testing (ICT) in electronics.

White-box testing can be applied at various levels of the software testing process, including unit testing, integration testing, and system testing. While it was traditionally associated with unit testing, it is now increasingly used for integration and system testing as well. This broader application allows white-box testing to evaluate

paths within a single unit, paths between units during integration, and interactions between subsystems during system-level testing.

One of the key advantages of white-box testing is its ability to uncover errors or problems within the codebase, including logic errors, boundary cases, and unexpected behaviors. By examining the internal structure of the software, testers can identify areas of weakness and potential vulnerabilities.

However, white-box testing also has its limitations. Despite its thoroughness, it has the potential to miss unimplemented parts of the specification or requirements that are not adequately addressed in the code. Additionally, white-box testing requires a high level of technical expertise and may be more time-consuming compared to black-box testing methods.

In summary, white-box testing is a valuable approach for thoroughly evaluating the internal workings of software applications. By leveraging internal knowledge and programming skills, testers can design comprehensive test cases to uncover errors and ensure the robustness and reliability of the software.

6.2 BLACK BOX TESTING: From fig 6.1 Black-box testing is a widely used method of software testing that focuses on evaluating the functionality of an application without delving into its internal structures or workings. This approach is applicable across various levels of software testing, including unit testing, integration testing, system testing, and acceptance testing. Sometimes referred to as specification-based testing, black-box testing relies on the external specifications and requirements of the software to design test cases.

Test cases in black-box testing are derived from external descriptions of the software, such as specifications, requirements documents, and design parameters. The primary goal is to ensure that the application behaves as expected and fulfills its intended functionality. Test cases are designed to cover both valid and invalid inputs, as well as edge cases and boundary conditions, to assess the robustness and reliability of the software.

While functional tests constitute the majority of test cases in black-box testing, non-functional tests may also be included to evaluate aspects such as performance, usability, and security. The test designer selects inputs based on the specifications and requirements of the software and determines the expected output without any knowledge of the internal structure or implementation details.

To validate the correctness of the output, a test oracle may be used, which compares the actual output of the software with the expected output. Alternatively, previous results that are known to be correct may serve as reference points for validating the output.

Black-box testing offers several advantages, including its independence from the internal implementation of the software, which allows testers to focus solely on the functional behavior. Additionally, it enables testing to be conducted from an end-user perspective, ensuring that the software meets user expectations and requirements.

Overall, black-box testing is a versatile and effective method for assessing the functionality and behavior of software applications, providing valuable insights into their performance and reliability without requiring knowledge of the internal workings.

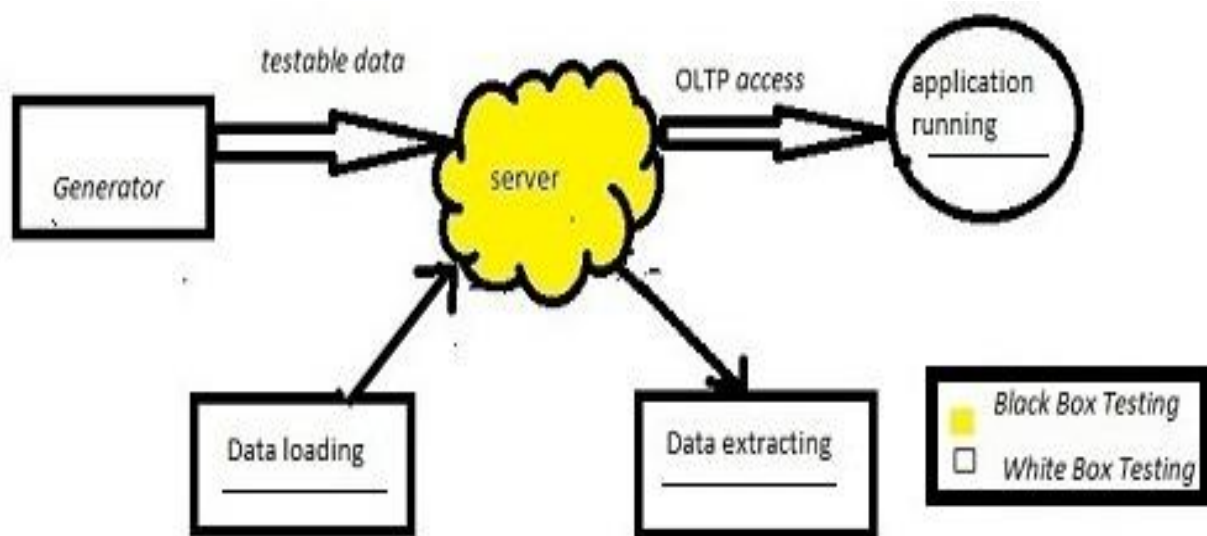


Fig 6.1. System testing diagram

CHAPTER 7

ARCHITECTURE

In the contemporary landscape, the ubiquity of CCTV cameras underscores their pivotal role in surveillance and security measures. While historically utilized for post-event analysis, the evolving security needs of today's world demand a shift towards proactive and real-time monitoring systems. To meet these demands, modern CCTV architectures integrate cutting-edge technologies like machine vision and advanced machine learning algorithms to detect anomalies promptly and alert authorities accordingly.

The integration of machine vision technology enables CCTV systems to interpret visual data in real-time, allowing for the identification of abnormal events and behaviors as they occur. By employing sophisticated machine learning algorithms, these systems can analyze vast amounts of visual data and learn to recognize patterns indicative of potential threats or irregularities.

One area where this technology is particularly impactful is in the analysis of crowd behavior and object detection. In crowded environments, such as public spaces or transportation hubs, the ability to detect and respond to anomalies in real-time is critical for ensuring public safety and security. By leveraging machine vision algorithms, CCTV systems can detect deviations from normal crowd behavior, such as sudden changes in density or unusual movements, which may indicate potential security threats or emergencies.

However, analyzing crowd behavior poses unique challenges due to the dynamic nature of crowds, where individuals may be dispersed across varying locations and move in diverse directions. Extracting meaningful features from this complex data requires sophisticated algorithms capable of capturing the nuances of crowd dynamics.

Despite these challenges, the integration of machine vision and machine learning technologies offers significant benefits in enhancing the effectiveness of CCTV surveillance systems. By providing real-time anomaly detection and proactive alerting capabilities, these systems empower authorities to respond swiftly to potential threats and maintain public safety in an increasingly complex and dynamic world.

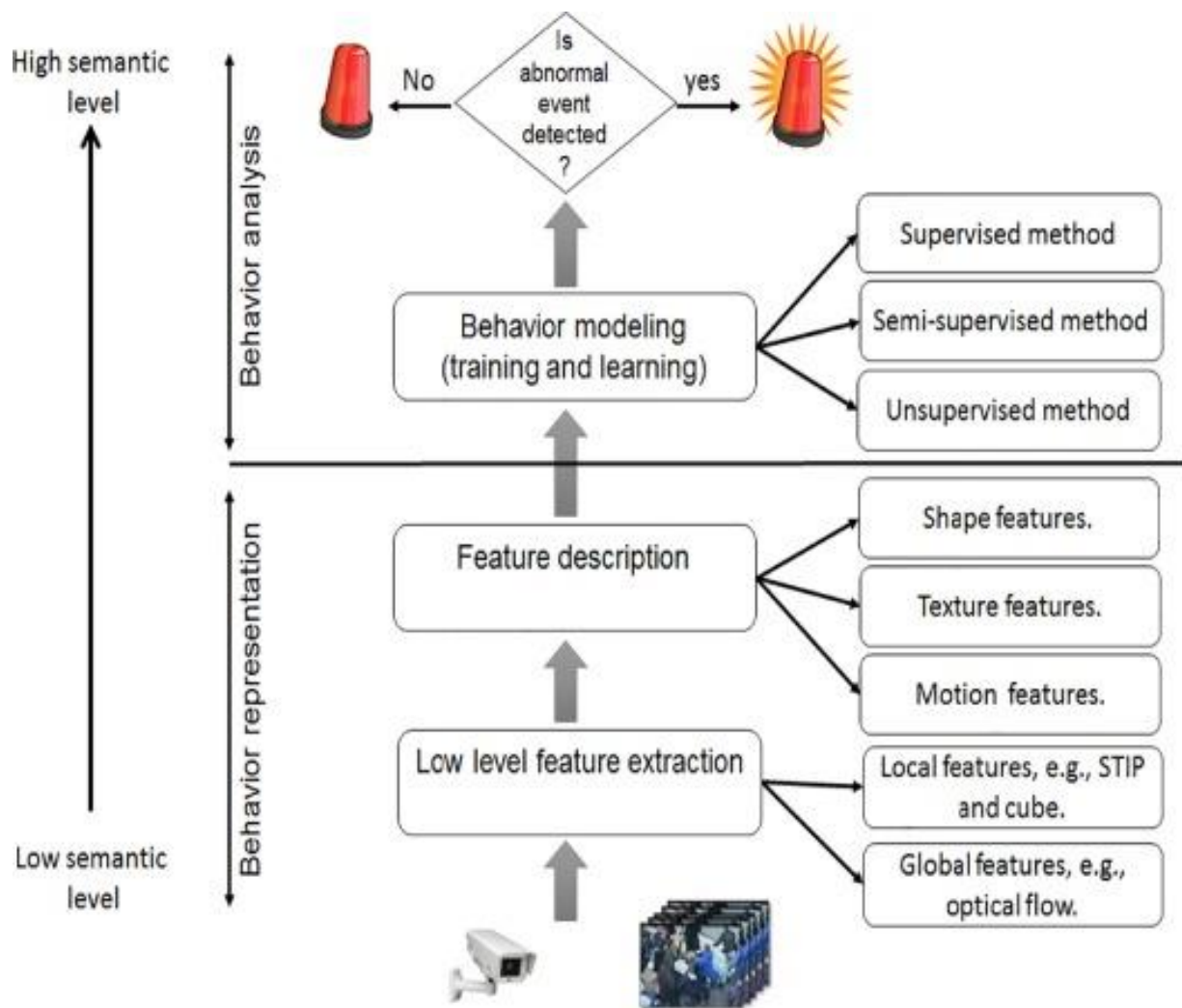


Fig 7.1. Architecture diagram

Within this architectural framework, the seamless integration of machine vision and machine learning technologies serves as the backbone of modern CCTV systems. These systems are designed not just to passively record events, but to actively analyze and interpret visual data in real-time, transforming CCTV surveillance into a proactive tool for security and safety.

The adoption of machine vision technology enables CCTV systems to process and understand visual information, facilitating the detection of anomalies and irregularities as they occur. By leveraging sophisticated algorithms, these systems can identify and classify objects, track movements, and recognize patterns within the visual data stream.

Furthermore, advanced machine learning algorithms play a crucial role in enhancing the intelligence of CCTV systems. Through continuous learning and adaptation, these algorithms can improve the accuracy and reliability of anomaly detection, enabling the system to discern subtle deviations from normal behavior and trigger alerts accordingly.

In the context of crowd behavior analysis, machine vision and machine learning technologies offer unprecedented capabilities for detecting and responding to potential threats or emergencies. By analyzing crowd dynamics in real-time, CCTV systems can identify abnormal behavior patterns, such as overcrowding, aggressive behavior, or suspicious activity, and notify authorities promptly.

However, the effectiveness of these systems relies not only on the sophistication of the underlying technology but also on the seamless integration of various components and modules within the architectural framework. From data acquisition and preprocessing to anomaly detection and alert management, each module must work cohesively to ensure the system's overall performance and reliability.

CHAPTER 8

RESULT AND DISCUSSION

8.1 RESULT

With our system, we can automate the process of detecting abnormal events from CCTV camera feeds. CNN and LSTM technologies are used to detect anomalies in both supervised and unsupervised manner. Alert messages can be sent to authorities on detection of events. Original Quality video snippets of the abnormal event can be stored in high quality and low quality recordings for events which are considered as normal are stored separately. By implementing this system, we reduce the time taken and human workload for detecting anomalies and also the system becomes more storage efficient. The design of this system consists of various modules or parts that have to be integrated together to complete the system. This involves the creation of Video Compressor, Anomaly Detector, Storage Management and Alert Management modules. Video Compressor is the first module in the system which is used to compress the original video in order to store it as low quality, low resolution video for archiving.

Anomaly Detector detects the presence of any kind of anomalies in the live feed from the CCTV system. This module is trained by Unsupervised learning method so that the system can detect all kinds of anomalies (both predetermined and undetermined). Also the video snippet where the anomaly is occurring is stored in the original video quality and resolution. Storage Management module is used to manage the storage of both the low quality archived video and original quality anomalous video snippets. These are stored separately for future reference. Alert Management module is used to manage and send the alert to respective personnel on identification of the anomaly

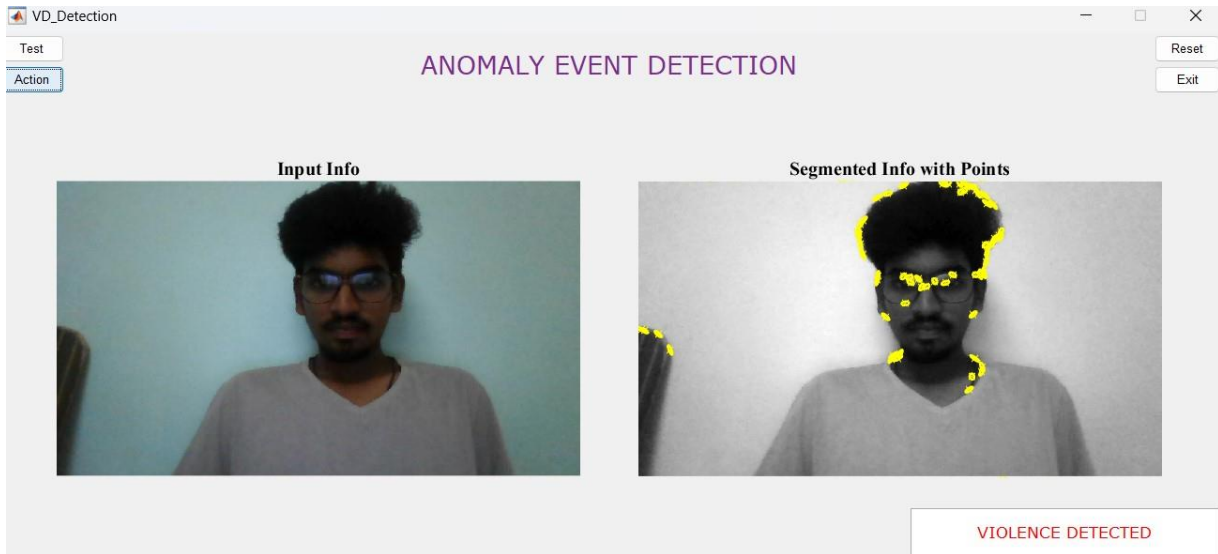


Fig 8.1 Frames

A GSM (Global System for Mobile communication) module is a hardware component that enables devices to communicate via cellular networks. Typically, it integrates a GSM modem along with other necessary components such as a SIM card slot and antenna. These modules support GSM networks worldwide, allowing devices to send and receive SMS messages, make voice calls, and access mobile data services. The module communicates with a microcontroller or computer via serial communication protocols such as UART, enabling seamless integration into various electronic systems. GSM modules are widely used in applications such as remote monitoring, telemetry, security systems, and IoT devices, where access to cellular networks is essential. They offer reliable wireless connectivity, making them suitable for applications in remote or mobile environments where traditional wired communication is impractical. Additionally, GSM modules come in different form factors and variants, including 2G, 3G, and 4G/LTE, providing flexibility and compatibility with different network technologies and frequencies. The integration of GSM modules into electronic projects enables remote control and monitoring capabilities, empowering users to stay connected and informed regardless of their location. Moreover, these modules often come with additional features such as GPS functionality, enabling location tracking and asset management applications. Overall, GSM modules serve as

indispensable components in modern communication systems, facilitating reliable and efficient wireless communication over cellular networks.

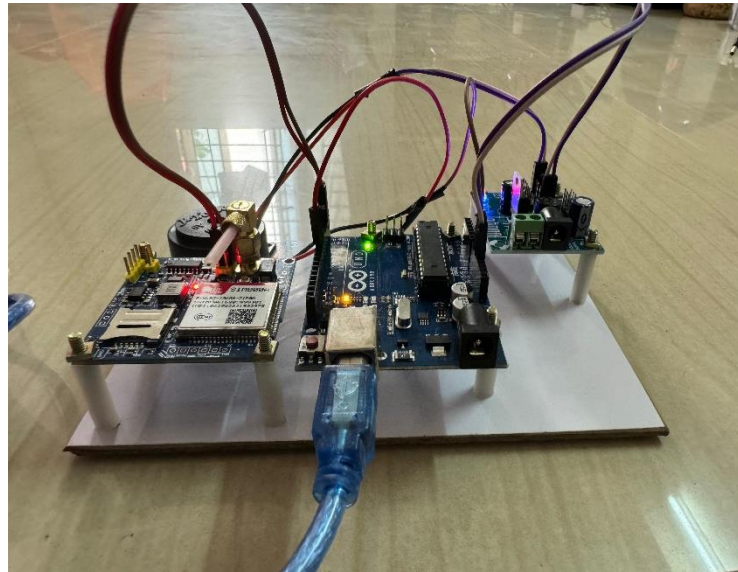


Fig 8.2 Working Model

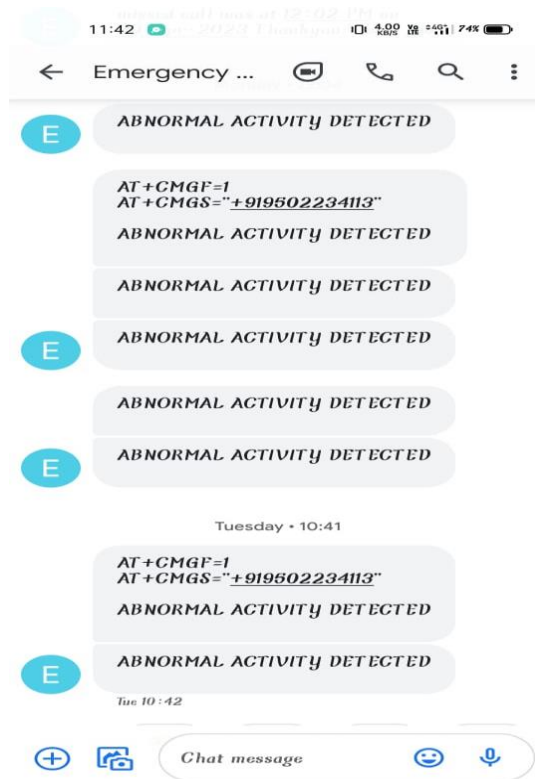


Fig 8.3 Message Alert

```
sketch_apr15b
}

void loop()
{
  if (Serial.available() > 0)
  {
    rx=Serial.readStringUntil('\n');
    if (rx=="AA")
    {
      digitalWrite(7,1);
      Serial.println("AT+CMGF=1");
      delay(1000);
      Serial.println("AT+CMGS=\"919390402430\"");
      delay(1000);
      Serial.println("ABNORMAL ACTIVITY DETECTED");
      delay(1000);
      Serial.println((char)26);
      delay(4000);

      Serial.println("AT+CMGF=1");
      delay(1000);
      Serial.println("AT+CMGS=\"9193902234113\"");
      delay(1000);
      Serial.println("ABNORMAL ACTIVITY DETECTED");
      delay(1000);
      Serial.println((char)26);
      delay(4000);

      digitalWrite(7,0);
    }
    digitalWrite(7,0);
  }
  digitalWrite(7,0);
}

Sketch uses 4126 bytes (12%) of program storage space. Maximum is 32768 bytes.
Global variables use 294 bytes (1%) of dynamic memory, leaving 1724 bytes for local variables. Maximum is 2048 bytes.
```

Fig 8.4 Arduino Code

8.2 Discussion

Identifying human abnormal activities through machine learning, coupled with MATLAB, constitutes a multifaceted approach that encompasses several stages of data processing, model development, and deployment. At the outset, data collection serves as the foundation, where various sources such as video feeds from surveillance cameras, sensor data from IoT devices, or a combination thereof, are aggregated. This data is often heterogeneous, requiring preprocessing to rectify inconsistencies, handle missing values, and ensure uniformity in format. MATLAB's suite of preprocessing functions facilitates tasks like noise removal, filtering, normalization, and interpolation, thus readying the data for subsequent analysis.

Once preprocessed, the data undergoes feature extraction, a pivotal step in distilling meaningful information from raw inputs. Features encapsulate relevant characteristics of human activities, ranging from spatial-temporal cues in video frames to movement patterns gleaned from accelerometer data. MATLAB's extensive library of signal processing, image processing, and computer vision algorithms simplifies

feature extraction, empowering developers to capture intricate nuances in activity representations.

With features in hand, the focus shifts to model selection, where machine learning algorithms are enlisted to discern patterns and classify activities as normal or abnormal. MATLAB offers a plethora of algorithms, including support vector machines (SVM), decision trees, random forests, and deep learning architectures like convolutional neural networks (CNNs). This versatility enables developers to tailor models to the specific requirements of their application domain, leveraging MATLAB's built-in functionalities for model training, evaluation, and optimization.

Training models involves feeding them labeled data, where normal and abnormal activities are annotated, and iteratively refining their parameters to enhance performance. MATLAB's cross-validation utilities aid in robust model evaluation, allowing developers to gauge metrics such as accuracy, precision, recall, and F1-score. Once a model meets satisfactory benchmarks, it can be deployed to real-world scenarios, be it in security systems, healthcare monitoring setups, or industrial safety protocols.

Real-time monitoring is often imperative in such applications, necessitating continuous data acquisition, processing, and visualization. MATLAB provides real-time capabilities, facilitating seamless integration with live data streams and enabling the development of responsive monitoring systems. Moreover, MATLAB's deployment options, including standalone executables, MATLAB Compiler, and MATLAB Coder, cater to diverse deployment needs, ensuring that models can be seamlessly integrated into existing infrastructure.

The journey does not end with deployment; rather, it evolves into a cycle of continuous improvement, wherein the system is fine-tuned based on real-world feedback, retrained with updated data, and iteratively refined to adapt to evolving contexts. This iterative approach underscores the dynamic nature of human behavior analysis, wherein vigilance, adaptability, and innovation are paramount. By harnessing the synergy between machine learning and MATLAB, developers can forge robust systems capable of identifying human abnormal activities with precision and efficacy, thereby enhancing safety, security, and well-being across various domains.

CHAPTER 9

CONCLUSION

Abnormal event detection stands as a critical component in the development of smart CCTV systems, enabling automatic detection of anomalies and timely alert generation. While supervised learning models have been traditionally employed in existing systems, they pose limitations in their ability to detect the diverse range of anomalies encountered in real-world scenarios. As a solution, the integration of unsupervised learning methods offers a promising avenue to effectively train the system without the need for exhaustive labeling of anomaly types.

Moreover, the implementation of such systems not only enhances security measures but also ensures storage efficiency by selectively storing abnormal frames in high quality, while conserving resources by storing recordings in lower quality. This approach optimizes storage space without compromising on the quality of critical footage.

Looking ahead, the future of CCTV systems lies in the integration of both supervised and unsupervised learning methods, leveraging the strengths of each to improve anomaly detection capabilities. Additionally, the incorporation of advanced anomaly identification methods and object detection algorithms holds promise for further enhancing the system's ability to accurately identify and classify anomalies in diverse environments.

By continuously refining and innovating upon existing methodologies, smart CCTV systems have the potential to revolutionize surveillance and security measures, ensuring the safety and well-being of individuals in various settings. Through collaboration and advancements in technology, we can build smarter, more efficient CCTV systems capable of adapting to evolving security challenges and enhancing overall public safety.

In addition to the advancements in learning methods, future iterations of smart CCTV systems may benefit from the integration of advanced anomaly identification techniques. These techniques could enable the system to not only detect anomalies but also classify them into different categories, providing more nuanced insights into potential security threats or abnormal behaviors.

Furthermore, object detection algorithms can play a crucial role in enhancing the capabilities of smart CCTV systems. By accurately detecting and tracking objects of interest, such as vehicles or individuals, these algorithms can provide valuable contextual information that helps in understanding the nature and severity of detected anomalies.

In conclusion, the future of smart CCTV systems is promising, with opportunities for advancements in learning methods, anomaly identification techniques, object detection algorithms, real-time analysis capabilities, and sensor technologies. By leveraging these advancements, we can develop more sophisticated and effective surveillance solutions that enhance security, improve public safety, and address the evolving challenges of the modern world.

REFERENCES

- [1]. Wang, T., Chen, J., & Snoussi, H. Online detection of abnormal events in video streams. *Journal of Electrical and Computer Engineering*, 2013.
- [2]. Li, C., Gan, Z., Cheng, Y., & Liu, J. (2020). MirrorGAN: Learning Text-to-image Generation by Redescription. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(4), 949962.
- [3]. Xu, T., Zhang, P., Huang, Q., Zhang, H., Gan, Z., Huang, X., & Metaxas, D. N. (2018). Attngan: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1316-1324.
- [4]. Zhang, H., Xu, T., Li, H., Zhang, S., Huang, X., Wang, X., & Metaxas, D. N. (2017). StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks. *IEEE International Conference on Computer Vision (ICCV)*, 5907-5915.
- [5]. Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. (2016). Generative Adversarial Text-to-Image Synthesis. *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, 1060-1069.
- [6]. Li, Y., Zhang, H., Zhang, K., Wang, Y., Xie, Y., & Sun, H. (2020). Controllable Text-to-Image Generation. *IEEE Transactions on Multimedia*, 21(6), 1490-1502.
- [7]. Zhang, H., Xu, T., Li, H., Zhang, S., Wang, X., Huang, X., & Metaxas, D. N. (2020). StackGAN++: Realistic Image Synthesis with Stacked Generative Adversarial Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 41(8), 1947-1962.
- [8]. Huang, Q., Zhang, H., Li, W., Gan, Z., Cheng, Y., & Wang, J. (2020). Text-to-Image Synthesis via Fine-Grained Image-Text Dual-Embedding. *Proceedings of the 28th ACM International Conference on Multimedia (MM)*, 1463-1471.
- [9]. Zhu, Y., Zhang, O., Song, Y., Shen, Z., Cao, Z., & Zhang, Z. (2019). DM-GAN: Dynamic Memory Generative Adversarial Networks for Text-to-Image Synthesis. *Proceedings of the 28th ACM International Conference on Multimedia (MM)*, 1572-1580.

- [10]. Li, C., Gan, Z., Cheng, Y., & Liu, J. (2020). MirrorGAN: Learning Text-to-image Generation by Redescription. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 42(4), 949962.
- [11]. Xu, T., Zhang, P., Huang, Q., Zhang, H., Gan, Z., Huang, X., & Metaxas, D. N. (2018). Attngan: Fine-Grained Text to Image Generation with Attentional Generative Adversarial Networks. *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 1316-1324.
- [12]. Zhang, H., Xu, T., Li, H., Zhang, S., Huang, X., Wang, X., & Metaxas, D. N. (2017). StackGAN: Text to Photo-realistic Image Synthesis with Stacked Generative Adversarial Networks. *IEEE International Conference on Computer Vision (ICCV)*, 5907-5915.
- [13]. Reed, S., Akata, Z., Yan, X., Logeswaran, L., Schiele, B., & Lee, H. (2016). Generative Adversarial Text-to-Image Synthesis. *Proceedings of the 33rd International Conference on Machine Learning (ICML)*, 1060-1069.