A PROJECT PHASE – I REPORT ON

**“Human Activity Recognition Using Artificial Neural Network”**

SUBMITTED TO THE SAVITRIBAI PHULE PUNE UNIVERSITY, PUNE

IN THE PARTIAL FULFILLMENT OF THE REQUIREMENTS

FOR THE AWARD OF THE

DEGREE

OF

**BACHELOR OF ENGINEERING (INFORMATION TECHNOLOGY)**

**SUBMITTED BY**

Siddhi Shelke BIT47

Shivani Joldapke BIT49

Vikrant Jumle BIT58

Sakshi Chikhale BIT50



**DEPARTMENT OF INFORMATION TECHNOLOGY**

DR. D. Y. PATIL INSTITUTE OF TECHNOLOGY

PIMPRI, PUNE 411018

SAVITRIBAI PHULE PUNE UNIVERSITY

**2022 -2023**



##### 

##### CERTIFICATE

This is to certify that the project report entitles

**“Human Activity Recognition Using Artificial Neural Network”**

Submitted by

Siddhi Shelke BIT47

Shivani Joldapke BIT49

Sakshi Chikhale BIT50

Vickrant Jumle BIT58

is a bonafide student of this institute and the work has been carried out by her under the supervision of **Prof. SONALI PATIL** and it is approved for the partial fulfillment of the requirement of Savitribai Phule Pune University, for the award of the degree of **Bachelor of Engineering** (Information Technology).

**(Prof. Sonali. C. Patil) (Prof. S. A. Nalawade) (Dr. L. K. Wadhwa)**

Project Guide Head Of Department Principal

and

Project coordinator

Seal/ Stamp of the College Examiner 1 :

Place: Pune

Date:

##### ACKNOWLEDGEMENT

It gives us immense pleasure in presenting the preliminary project report on  **“Human Activity Recognition Using Artificial Neural Network”.**  The success and the outcome of this report required a lot of guidance. We are very grateful to our **Prof. Sonali C. Patil** who has provided expertise and encouragement. We thank ma’am who provided vision and knowledge that was very helpful throughout the research. All that we have done is only due to the great guidance. We also express our gratitude to **Prof. S.A. Nalawade**  Head of Information Technology Department, Dr. D. Y. Patil Institute of Technology Pimpri ,Pune, for the valuable support. We would also like to extend our sincere thanks to Principal **Dr. L. K. Wadhwa**, for his dynamic and valuable guidance throughout the project and providing the necessary facilities that helped us to complete our dissertation work.

Siddhi Shelke (BIT47)

Shivani Joldapke (BIT49)

Sakshi Chikhale (BIT50)

Vikrant Jumle (BIT58)

**ABSTRACT**

Human Activity Recognition(HAR) is an active field of research and scientific development in which various models have been proposed using different methods for identification and categorization of activities using Machine Learning. HAR has reached a remarkable milestone in the area of computer vision. Except for applications in human-computer interactions, surveillance systems and robotics, lately it has extended its applicability in the fields of healthcare, multimedia retrieval, social networking, and education as well. It aims to bring latest technologies together to develop complex assistive system with adaptive capability and learning behavior. HAR interprets human motion using computer and machine vision technologies to identify and detect simple and complex actions in real-world. This paper presents research made for surveillance of restricted military areas. Our scope is to develop a live monitoring system for tracking the illegal activities done in the restricted area for border security, which is an issue of concern since decades. In this we have introduced a deep learning model that learns to classify human actions without having prior knowledge. The features of image or video set are extracted and detected for detecting whether the activity is illegal or not. Due to such application of this model many harmful activities can be avoided or at least negative consequences of such activities can be minimized. Finally, these results lead to achieve a promising performance in the activity recognition rate.

**INDEX**

|  |  |  |
| --- | --- | --- |
| **01** | **INTRODUCTION**   * 1. Motivation |  |
| 02 | Literature Survey |  |
| 03 | Problem definition |  |
| 04 | Project Plan  4.1 Time Line of Project  4.2 Feasibility Study  4.3 Risk Analysis / Risk Management  4.4 Lines of Code ( LOC )  4.5 Effort  4.6 Development Time  4.7 Number of people |  |
| 05 | Project Requirement specification  5.1 Software Requirements  5.2 Hardware Requirements  5.3 Functional Requirements  5.4 Non Functional Requirements |  |
| 06 | System Architecture |  |
| 07 | UML / ER Diagrams. |  |
| 08 | High Level design of the project (DFD) |  |
| 09 | System Implementation-  9.1 Algorithm / Methodologies . |  |
| 10 | Working Modules |  |
| 11 | Other Specification  11.1 Advantages  11.2 Limitations  11.3 Applications |  |
| 12 | Conclusions |  |
| 13 | References |  |

**LIST OF ABBREVATIONS**

|  |  |
| --- | --- |
| **ABBREVIATION** | **ILLUSTRATION** |
| LSTM | Long Short – Term Memory |
| CNN | Convolutional Neural network |

**LIST OF FIGURES**

|  |  |  |
| --- | --- | --- |
| Figure No. | Illustration | Page No. |
| 1 | System Architecture of Human Activity Recognition | 18 |
| 2 | Activity Diagram | 21 |
| 3 | Sequence Diagram | 22 |
| 4 | Class Diagram | 22 |
| 5 | Use Class Diagram | 23 |
| 6 | Data flow diagram (DFD0) | 24 |
| 7 | Data flow diagram (DFD1) | 24 |
| 8 | Data Flow diagram (DFD2) | 25 |

**LIST OF TABLES**

**Chapter 1**

**Introduction**

Human activity recognition (HAR) aims to provide information on physical activity and to detect simple or complex actions in real world. Human activity recognition is an active research area in last decades due to its applicability in various fields and increasing need for home automation and convenient service for elderly ones. Human activity recognition has been revolutionized in various domains such as healthcare, sports, entertainment etc. It involves behavior and environment monitoring, activity modelling, data processing and pattern recognition. Artificial intelligence aims towards building machines that are capable to think like humans. An Artificial Neural Network (ANN) in the field of Artificial intelligence attempts to mimic way nerve cells work in the human brain. It makes understanding of computers easy to make decisions similar to a human-brain. The (ANN) is trained to behave like interconnected brain cells by programming computers. It utilizes computer software programs that analyze the audio and images from videos in order to recognize humans, vehicles, objects, attributes, and events. It has been proven a very useful technology to detect illegal occurrences to avoid threats and maintain security. Recognition of human activity is an ability to look over the movements/gestures of the human body and to determine human action/activity. If the majority of routine human tasks can be identified by activity recognition systems, they can either be made simpler or mechanized. The main objective of HAR framework is to observe and analyse human behaviors efficiently, understand the human actions and then to retrieve and process the relative data. Abnormal activity recognition is another area of research for public safety. In public sectors CCTV cameras are used to monitor the crowd activities and track the motion of suspicious activities. HAR has attracted a great amount of attraction in past decades due to its great value in wide range of real-world applications.

In this study we have proposed a HAR system that collects data and uses ANN for classification in military restricted areas. For compound security, a military system must be able to identify specific human behaviors. To secure military installations or other critical infrastructure, early identification of human activity that suggests a potential danger is required. The variety of human actions makes their automated recognition a real issue. There are activities that are performed by person such as running, fighting, picking up an item, exchanging items, digging, entering restricted area etc. The area of human action recognition is related to other ways of research that analyze human actions from images and video. We focus on actions and do not explicitly consider context such as the environment, interactions between persons or objects. Moreover, we consider only full-body movements. The system must be able to represent each of these components in order to be able to recognize a wide range of human actions. To identify the focus of attention, the person is distinguished from the rest of the scene. Based on the position of key points system can generate alerts in real time when potential assault is in progress.

* 1. **Motivation**

The goal of human activity recognition is to examine activities from video sequences or still images. Motivated by this fact, human activity recognition systems aim to correctly classify input data into its underlying activity. As a result, this system should be implemented in the Restricted area for border security purposes,

* 1. **Objective**
* Design a simple, light weight, and accurate system that can learn human activity with minimum user interaction.
* Provides information about human actions which helps in analyzing the behavior of a person in a real environment.
* Reduce the labeling time and labor works using active learning.
* The aim of the system is to continuously track human actions and detect illegal activities for security purposes.

**Chapter 2**

**LITERATURE SURVEY**

In this paper they have used two approaches for human activity recognition Conditional Random Field (CRF) and Long Short term Memory (LSTM) layers. Accelerometers sensors placed on multiple body locations for example ankle, chest and hand for capturing the body motion and later this information use for predicting the output. Another way they used was use machine learning to capture the data and then perform comparison. [1]

In this paper they have used various hardware devices for human activity recognition devices such as fitness trackers, watches, wearable devices and for feature extraction they have used two techniques called using ensemble empirical mode decomposition (EEMD). Then time and frequency domain feature extracted by applying Hilbert-Huang Transform. [2]

In this, wearable devices and smartphone are widely used for detecting human activity recognition hence. In this paper they have used CNN-LSTM network which tackle all the activities which are capture by digital devices they have used. [3]

This paper focuses on both human and robotic activities so that they can collaborate with each other. Human robot collaboration (HRC) is achieved by combining high-level robot control techniques and human action recognition (HAR) into a single control system. It uses artificial neural network classifier for activity recognition. [4]

In the modern world, millions of people pass away every year because they lack knowledge about their health. If the healthcare system focused more on disease prevention through routine health status assessments and the treatment of diseases in their early stages, rising expenditures may be lowered. Due to this circumstance, numerous researchers are working to create technologies that will use IoT to enhance human life quality. According to the research, IoT technologies can be used to monitor a patient's condition in real time or to collect sensitive data that will then be analyzed for a medical diagnosis. These technologies can also help gather information to assess health and provide information to properly trained personnel in diagnosing patients. [5]

Activity detection uses sensor data to recognize and detect both straightforward and intricate actions in natural environments. It is a difficult endeavor since sometimes the sensor data that is produced is unclear regarding the activity that is occurring. Due of this, actions are difficult to interpret.

Sometimes the data that is obtained can also be noisy. Human error or network technology flaw that prevents the network from providing accurate sensor readings can both lead to noise in the data. Such real-world situations demand ways to learn from data, to extract knowledge, and to aid in decision-making because they are full with uncertainty. In addition, using inverse probability, one can forecast and infer unknowns. [6]

In this paper, accuracy and computational cost are typically used to evaluate Human Activity Recognition (HAR). Previous works have attempted to extract hand-crafted features from the accelerometer and gyroscope signals, and have utilized various traditional machine learning algorithms. [7]

In every aspect of modern life, including education, health, security, agriculture, business, and notably health assessments, technology plays a vital role. Technology-related fields are currently the most researchable topics. This study also examines activity detection, which can be used to track human activity in a variety of contexts, including the health sector to track patient activity and smart cities to leverage the Internet of Things to track domestic activity. In addition to being helpful for crowd anomaly detection and object tracking, activity detection can also be used for security considerations. [8]

Due to the complexity and wide variety of human activities, the topic of human activity recognition (HAR) has recently attracted a lot of attention. Due in great part to developments in sensing technologies and wireless sensor networks, sensor-based methods to HAR have become particularly popular in pervasive computing. HAR is a crucial element in a wide number of application fields, including as connected health, pervasive computing, security systems, human computer interface (HCI), and ambient assisted living (AAL) in smart home environments. Other prominent interest domains include human/object detection and recognition based on object analysis and processing, for example, tracking and detection, computer engineering, physical sciences, health-related difficulties, natural sciences, and industrial academic disciplines. [9]

In this Paper, we designed an IMU based activity recognition system that accurately and robustly identifies among 6 different activities. We achieved the state-of-the-art accuracy of 97% by generating a compact 15-D discriminative space, via exploring feature extraction and reduction and discriminative analysis of the resulting space. [10]

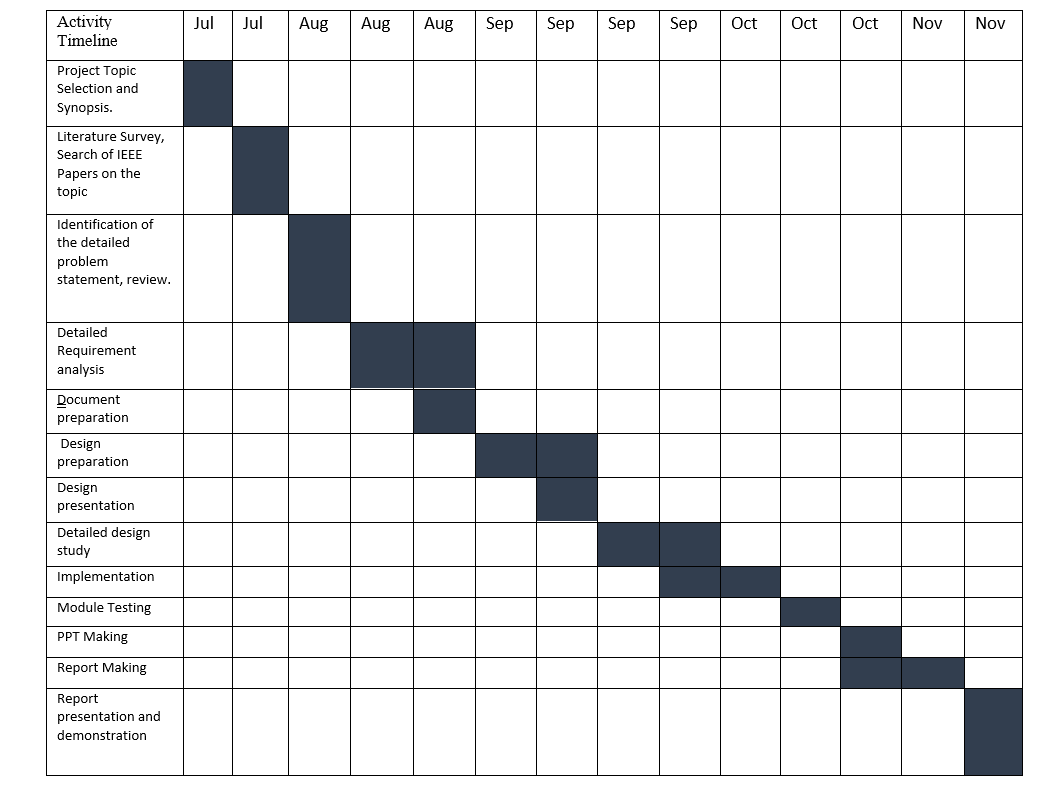
**Chapter 3**

**3.1 PROBLEM STATEMENT**

Human Activity Recognition in the real world has number of applications including intelligent video surveillance, shopping behavior analysis. Video surveillance is an integral part of security it has vast application areas especially for indoor and outdoor places. Today security camera became a part of life for the safety and security purposes. Currently the tracking has been performed by human. Since we are dealing with the huge amount of video data. The aim of this paper is to identify a physical activity carried out by a person depending on a movement recorded in a military restricted area.

Chapter 4

**PROJECT PLAN**

**4.1. PROJECT PLAN**

**4.2. FEASIBILITY STUDY**

The project uses the latest technology stack and the budget of the project will depend on the server and hosting platform. The estimated time the project will take is 8 months.

**4.3. RISK ANALYSIS AND RISK MANAGEMENT**

The main risk in our project is to be true to the timeline. The best way to track that our project functions on time is to set up daily or weekly meetings to track the project functionalities are in sync with the dedicated timeline.

If in case there is a situation where we have missed a timeline, we could increase our working hours to compensate it in the next phase.

**4.4. LINES OF CODE (LOC)**

Estimated Modules - 3

Estimated Lines of Code (per module) – 400 - 500

Chapter 5

**PROJECT REQUIREMENT SPECIFICATION**

**5.1 SOFTWARE REQUIREMENTS**

* Operating System: Windows 11
* Platform: Jupyter Notebook
* Programming Language: Python 3.2

**5.2 HARDWARE REQUIREMENTS**

* RAM: 8 GB
* Hard Disk: 40 GB
* Processor : Intel i5 Processor
* Operating System: Windows 11

**5.3 FUNCTIONAL REQUIREMENTS**

* + 1. **Functional Specification**
* System should be able to process the frame from the video input for processing.
* System should be able to read video sequence as input.
* System should be able to pre-process the frames extracted from the input and resize or crop it to the required threshold size.
* System should be able to compare the frames with the trained weights.

**5.4 Non – FUNCTIONAL ATTRIBUES**

**5.4.1 Performance Requirements**

* The performance of the functions and every module must be well. The overall performance of the software will enable the users to work evidently.
* Performance of encryption of data should be fast.
* Performance of the providing virtual environment should be fast Safety Requirement the application is designed in modules where errors can be detected and easily. This makes it easier to install and update new functionality if required.

**5.4.2 Software Quality Attributes**

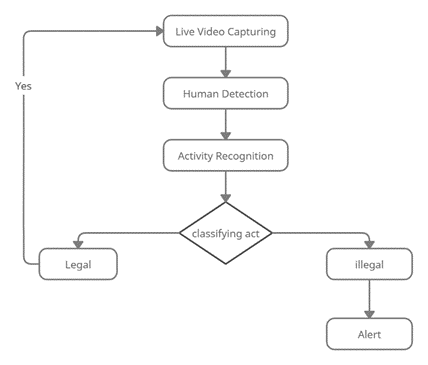
Our software has many quality attribute that are given below: -

1. Adaptability: This software is adaptable by all users.
2. Availability: This software is freely available to all users. The availability of the software is easy for everyone.
3. Maintainability: After the deployment of the project if any error occurs then it can be easily maintained by the software developer.
4. Reliability: The performance of the software is better which will increase the reliability of the Software.
5. User Friendliness: Since, the software is a GUI application; the output generated is much user friendly in its behavior.
6. Integrity: Integrity refers to the extent to which access to software or data by unauthorized persons can be controlled.
7. Security: Users are authenticated using many security phases so reliable security is provided.
8. Test ability: The software will be tested considering all the aspects.

**Chapter 6**

**SYSTEM ARCHITECTURE**

Human activity recognition system helps to detect the human movements in the surroundings. The main aim of this model is to identify and detect the movements happening in the restricted military areas to ensure security.



**Figure1.** System Architecture of Human Activity Recognition

Human activity recognition system helps to detect the human movements in the surroundings. The main aim of this model is to identify and detect the movements happening in the restricted military areas to ensure security.

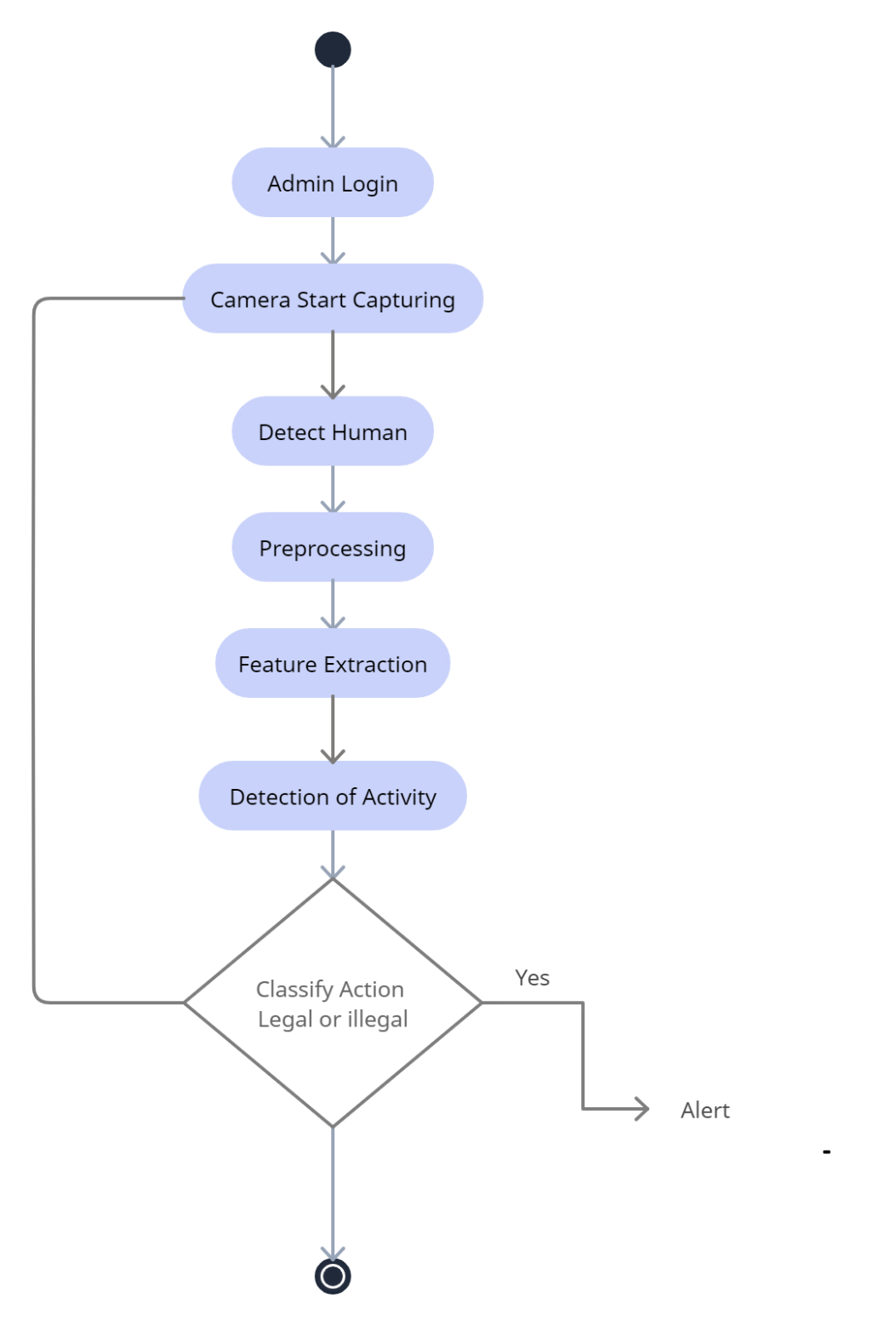
We are using a very useful deep learning technology for detection and identification of human movements and to track their activities. Long Short-term Memory also known as LSTM is used for video capturing, prediction and classification of images. Due to their ability to understand long-term connections between data time steps, LSTMs are frequently used to learn, analyze, and classify sequential data.

We will be using a webcam which will help to continuously monitor the area and track any activities carried out by humans. After capturing the signal from surrounding the system will process with denoising techniques. This denoising techniques will help to remove the noise from a noisy image and restore the original image. Data filtering will help in the process of image detection, face recognition and whatever task involved in computer vision and improve the edge of an image. Normalization techniques will help to change the range of pixels and reducing the scale from 0-255 to 0-1 range and will convert the image into a collection of pixels which is represented by labelled image by segmentation process. Various image pre-processing techniques like normalization, resizing, segmentation is applied on images. Later, feature extraction techniques will be useful in classifying and recognition of images. It will classify into time and frequency domain, and obtain most relevant information from the original image and represent information in low dimensionality space. The extracted features will be later used for detection of the activities carried out by the humans. If the tracked motion or activity is found illegal it will send an alert or alarm to make aware of the illegal activity carried out in a restricted area. This will ensure the security of the military bases.

Chapter 7

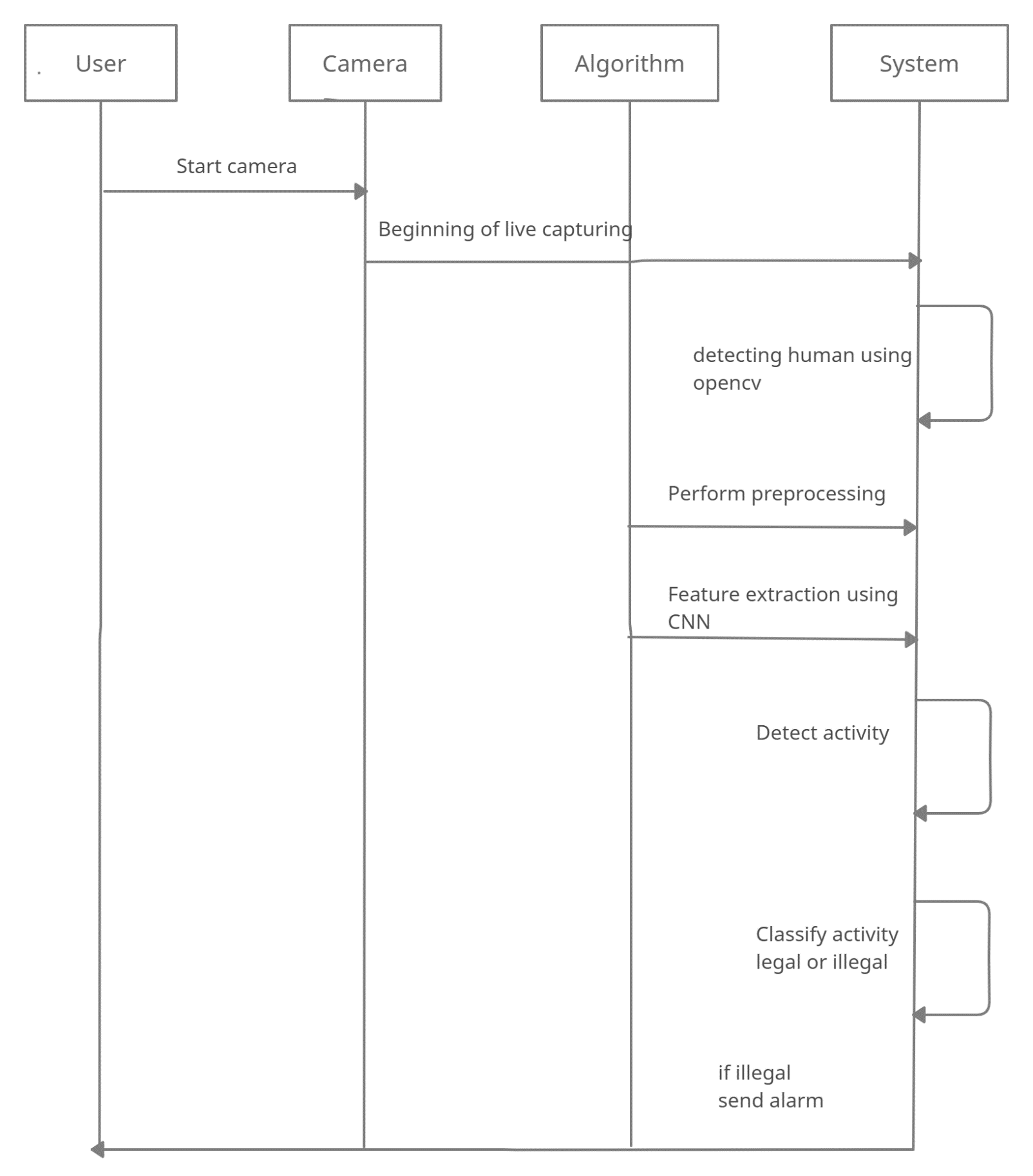
**UML / ER DIAGRAM**

Unified Modeling Language is a standard language for writing software blueprints. The UML may be used to visualize, specify, construct and document the artifacts of a soft- ware intensive system. UML is process independent, although optimally it should be used in process that is use case driven, architecture- centric, iterative, and incremental. The Number of UML Diagram is available.

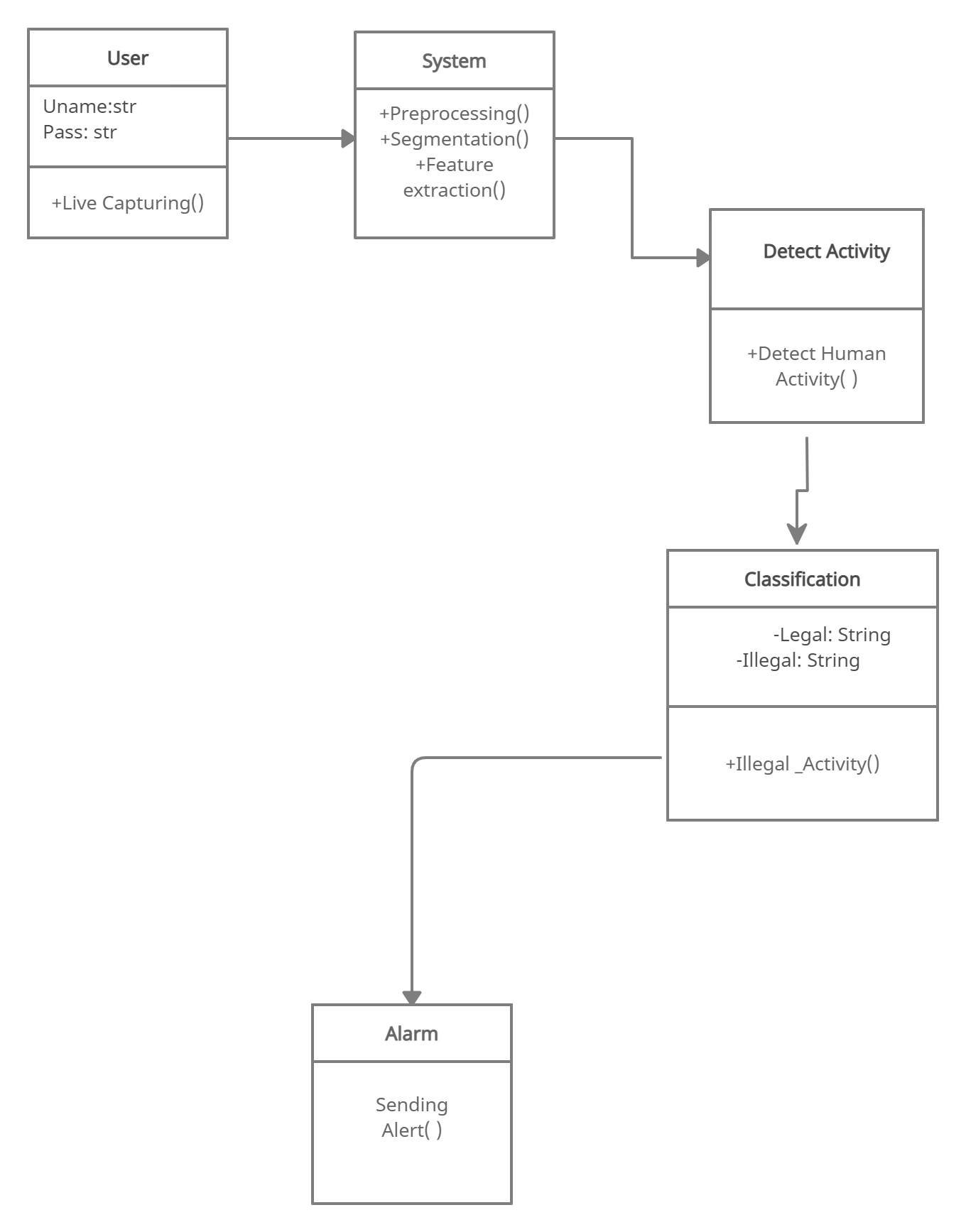
* Activity Diagram

**Figure 2.** Activity Diagram

* Sequence Diagram

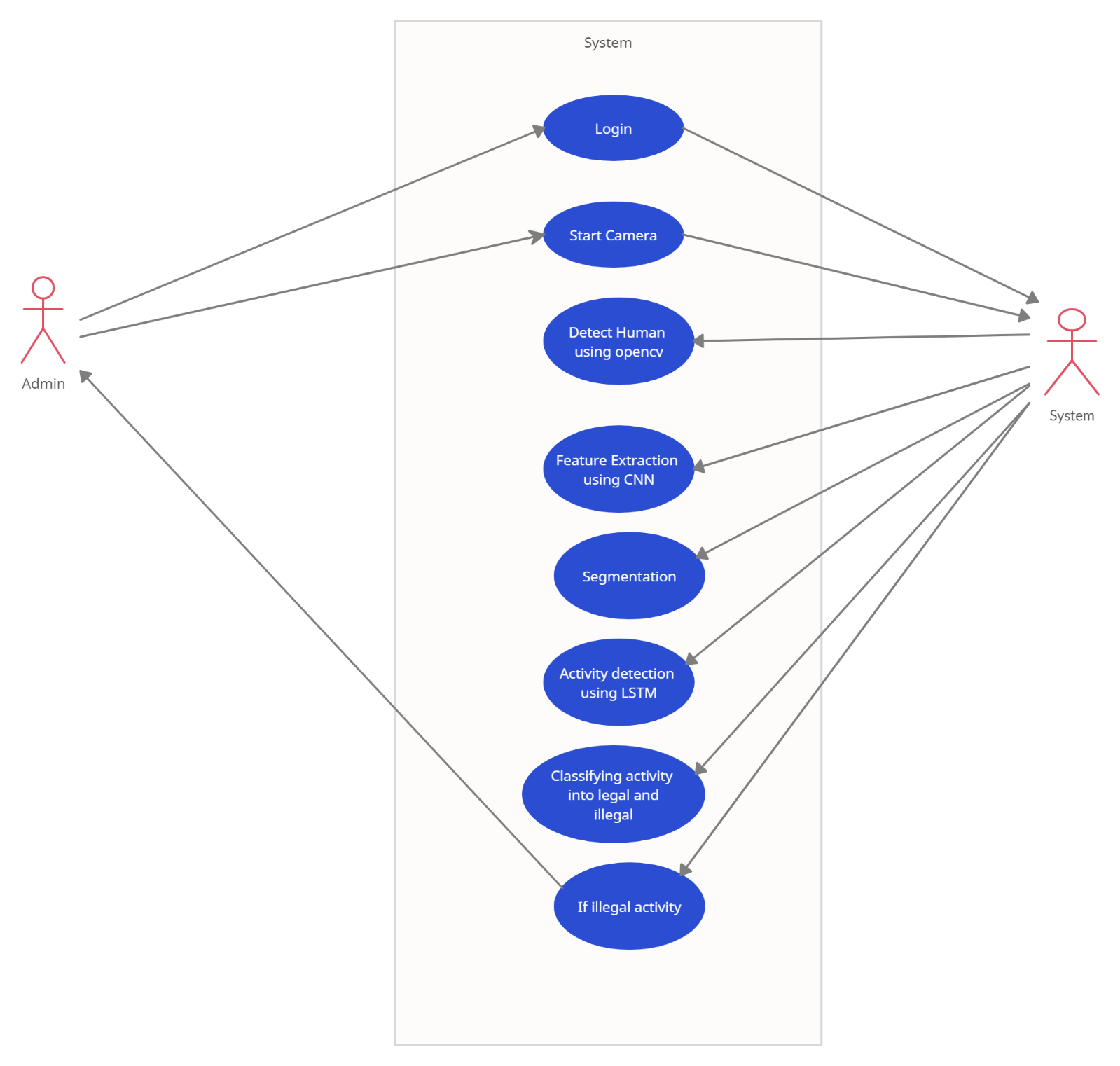


**Figure 3.** Sequence Diagram

* Class Diagram

**Figure 4.** Class Diagram

* Use Case Diagram

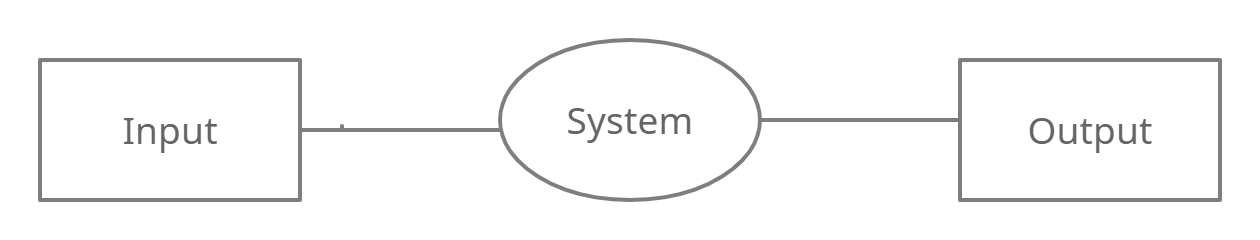


**Figure 5.** Use Class Diagram

**Chapter 8**

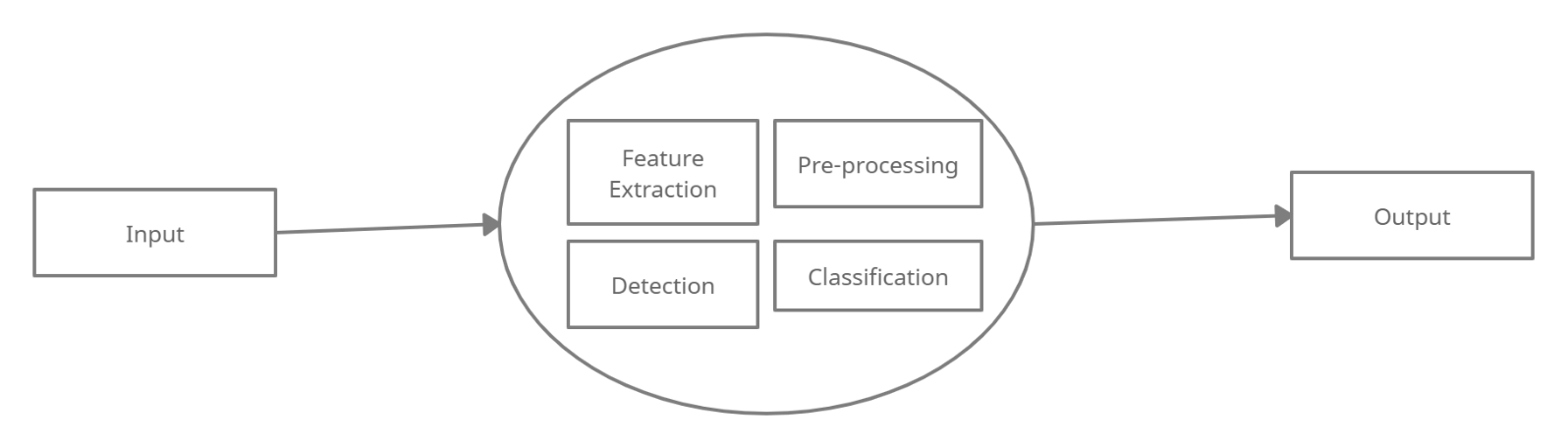
**High Level design of the project (DFD)**

* In Data Flow Diagram, we Show that flow of data in our system in DFD0 in which rectangle present input as well as output and circle show our system.
* It represents the entire system as a single bubble with input and output data indicated by incoming/outgoing arrows.

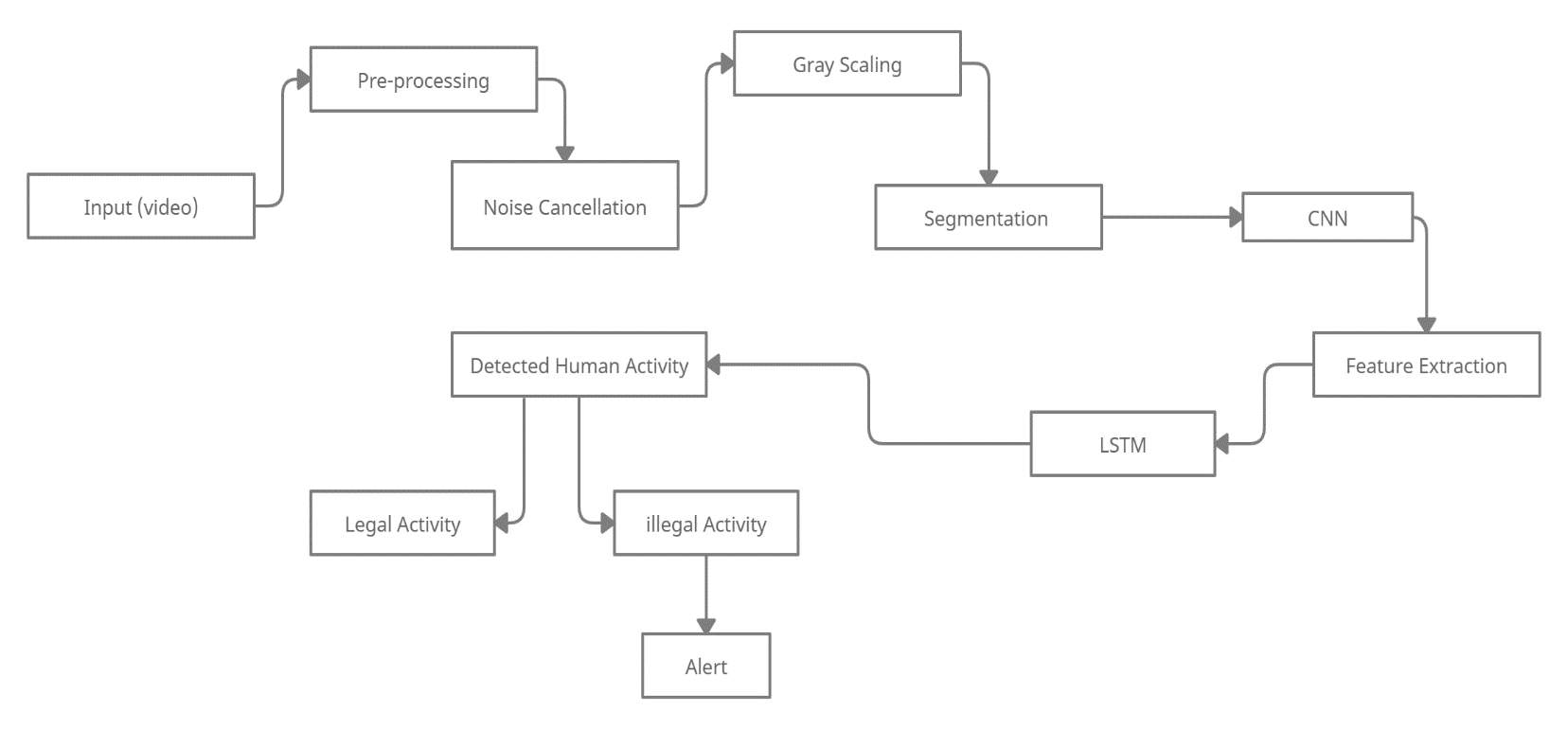


**Figure 6.** Data flow diagram (DFD0)

* **I**n DFD 1 input is given and inside the system the processes like Feature Extraction, Pre-processing, Detection and Classification are carried out and given as the output.



**Figure 7.** Data flow diagram (DFD1)



**Figure 8.** Data Flow diagram (DFD2)

In DFD2, as shown, live capturing takes videos as input for processing. It is divided into chunks for easy processing. Noise cancellation and gray scaling is carried out on it. Later, Segmentation and feature extraction is carried out on images using CNN algorithm. Using LSTM algorithm human activity detection is done and then classified into legal and illegal activities. If activity is found illegal an alarm is sent indicating a threat.

**Chapter 9**

**SYSTEM IMPLEMENTATION**

**9.1.**  **Algorithm**

We are using CNN and LSTM algorithm for implementing this system.

**9.1.1. Convolutional Neural Network (CNN)**

Convolutional Neural Network is a type of artificial neural network, which is widely used for image/object recognition and classification. Deep Learning recognizes objects in an image by using a CNN. Now when we think of a neural network, we think about matrix multiplications but that is not the case with ConvNet. It uses a special technique called Convolution. Now in mathematics convolution is a mathematical operation on two functions that produces a third function that expresses how the shape of one is modified by the other. Convolutional neural networks are composed of multiple layers of artificial neurons. Artificial neurons, a rough imitation of their biological counterparts, are mathematical functions that calculate the weighted sum of multiple inputs and outputs an activation value. When you input an image in a ConvNet, each layer generates several activation functions that are passed on to the next layer.

The first layer usually extracts basic features such as horizontal or diagonal edges. This output is passed on to the next layer which detects more complex features such as corners or combinational edges. As we move deeper into the network it can identify even more complex features such as objects, faces, etc.

Similar to the Convolutional Layer, the Pooling layer is responsible for reducing the spatial size of the Convolved Feature. This is to decrease the computational power required to process the data by reducing the dimensions. There are two types of pooling average pooling and max pooling.

**9.1.2 Long Short- Term Memory**

LSTM networks were designed specifically to overcome the long-term dependency problem faced by recurrent neural networks RNNs (due to the vanishing gradient problem). LSTMs have feedback connections which make them different to more traditional feedforward neural networks.

LSTMs are particularly good at processing sequences of data such as text, speech and general time-series. LSTMs use a series of ‘gates’ which control how the information in a sequence of data comes into, is stored in and leaves the network. There are three gates in a typical LSTM; forget gate, input gate and output gate. How do LSTM Networks Work?

Firstly, at a basic level, the output of an LSTM at a particular point in time is dependent on three things:

▹ The current long-term memory of the network — known as the cell state

▹ The output at the previous point in time — known as the previous hidden state

▹ The input data at the current time step

**9.1.3. Hmm (Hidden Markov Model)**

This method is used for classifying normal and abnormal activity.

It compares the capture video frames with the existing normal frames and identify the similarity between these frames. The unusual event in the video footage could be detect by tracking of peoples.

**9.1.4. DDBN (Discriminative deep belief Network)**

Human beings are detected from the video using background subtraction method and the features extracted using CNN and later these features fed into DDBN (Discriminative deep belief network).

DDBN- labeled videos of some suspicious events are also fed to the DDBN and their features are also extracted. Then the comparison of features is extracted using CNN and features extracted from labeled sample videos of classified suspicious action is done using DDBN and suspicious activities are detected from the given video.

**Chapter 10**

**WORKING MODULES**

1. Human detection:

In this module human detection is done by background subtraction method using OpenCV library.

All other environmental things are excluded in this method and only human is taken into consideration.

1. Prepossessing:

* In this module system will be processing on given input.
* It includes noise removal, resizing data, dividing input into chunks, converting it to grayscale image.

1. Feature Extraction

* It is mainly for extracting important features from raw data.
* In this module large size dataset is converted into small size dataset by dividing into manageable groups.
* It reduces dimensions of the input maintaining accuracy.

1. Activity detection

In this module activity is tracked and analyzed for detection using LSTM algorithm.

1. Activity classification

* In this module after detecting activity, we perform classification of the activities.
* We classify activity into legal and illegal based on the model trained by dataset and send an alarm if found illegal.

**Chapter 12**

**CONCLUSION**

The description of methodologies and methods used in each process stage, including data collection and filtering, data segmentation, feature extraction, dimension reduction, and classification, served to introduce the idea of HAR processes. This paper provides a real time solution for surveillance of restricted military areas. We propose to implement the given solution with the help of deep learning techniques and a comparative study of different methodologies. Deep neural network substantially increases the output by automatically learning features from raw data, making motion tracking a promising application. In addition, we demonstrate that our method is accurate and versatile. It can recognize the human actions and detect the illegal or abnormal activities accordingly assuring security of the areas. Although the proposed method can achieve better results than other methods for target detection, the accuracy of activity recognition still needs to be improved. Other possible problems, such the identification of complex actions, could not be taken into account in this research. Solving the identification and accurate detection of these complex actions will become the next main step for research.

**REFERENCES**

1. Jeremiah Okai, Stylianos Paraschiakos, Marian Beekman, Arno Knobbe and Cl´audio Rebelo de S´a “Building robust models for Human Activity Recognition from raw accelerometers data using Gated Recurrent Units and Long Short-Term Memory Neural Networks.”
2. Enda Wista Sinuraya, Aminuddin Rizal, Yosua Alvin Adi Soetrisno, Denis “Performance Improvement of Human Activity Recognition based on Ensemble Empirical Mode Decomposition (EEMD).”
3. Sakorn Mekruksavanich1 and Anuchit Jitpattanakul “Smartwatch-based Human Activity Recognition Using Hybrid LSTM Network.”
4. Tarik Uzunovic, Edin Golubovic, Zlatan Tucakovic, Yasin Acikmese, Asif Sabanovic. “Task-Based Control and Human Activity Recognition for Human-Robot Collaboration.”
5. Stefan Oniga, József SütĘ **“**Human activity recognition using neural networks.”
6. Deepika Singh1(B), Erinc Merdivan1(B), Ismini Psychoula2, Johannes Kropf1, Sten Hanke1, Matthieu Geist3, and Andreas Holzinger. “Human Activity Recognition Using Recurrent Neural Networks”
7. Piyush Mishra∗, Sourankana Dey†, Suvro Shankar Ghosh‡, Dibyendu Bikash Seal§, Saptarsi Goswami, “Human Activity Recognition using Deep Neural Network”
8. Shahab Geravesh,Vaibhav Rupapara. “Artificial neural networks for human activity recognition using sensor based dataset”
9. Naomi Irvine , Chris Nugent, Shuai Zhang , Hui Wang and Wing W. Y. NG ,“Neural Network Ensembles for Sensor-Based Human Activity Recognition Within Smart Environments.”
10. Omid Dehzangi, Vaishali Sahu . “IMU-Based Robust Human Activity Recognition using Feature Analysis, Extraction, and Reduction.”