APPENDIX I

PROJECT REPORT

ON

HUMAN INTERACTION DETECTION USING VIDEO PROCESSING AND MACHINE LEARNING

Submitted by

APOORV SHAH	B150388656
HARSHWARDHAN JADHAV	B150388562
ADITI BHANGALE	B150388503
SHIVANI DANDIR	B150388670

in partial fulfillment for the award of the degree

of

Bachelor of Engineering

of

Savitribai Phule Pune University

IN

INFORMATION TECHNOLOGY

MAHARASHTRA INSTITUTE OF TECHNOLOGY
COLLEGE OF ENGINEERING

2018-19

APPENDIX II

PROJECT REPORT
ON

HUMAN INTERACTION DETECTION USING VIDEO PROCESSING AND MACHINE LEARNING

Submitted By

APOORV SHAH	**B150388656**
HARSHWARDHAN JADHAV	**B150388562**
ADITI BHANGALE	**B150388503**
SHIVANI DANDIR	***B150388670***

Guided by
PROF. ADITI JAHAGIRDAR

INFORMATION TECHNOLOGY
MAHARASHTRA INSTITUTE OF TECHNOLOGY
COLLEGE OF ENGINEERING
PUNE

SAVITRIBAI PHULE PUNE UNIVERSITY

2018- 2019

APPENDIX III

INFORMATION TECHNOLOGY

Certificate

This is to certify that,

APOORV SHAH	**B150388656**
HARSHWARDHAN JADHAV	**B150388562**
ADITI BHANGALE	**B150388503**
SHIVANI DANDIR	***B150388670***

have successfully completed this project report entitled

HUMAN INTERACTION DETECTION USING VIDEO PROCESSING AND MACHINE LEARNING

under my guidance in partial fulfillment of the requirements for the degree of Bachelor of Engineering in Department of Information Technology of Savitibai Phule Pune University, Pune during the academic year 2018-19.

Date: -
Place: -

Prof. Aditi Jahagirdar
Prof. Dr. Krishna Warhade
Internal Guide

Prof.Dr.Krishna Warhade
Head of Department

APPENDIX IV

Acknowledgement

We take this opportunity to thank our project guide Prof. Aditi Jahagirdar and Head of the Department Prof. Dr. Krishna Warhade for their valuable guidance and for providing all the necessary facilities, which were indispensable in the completion of this project report. We are also thankful to all the staff members of the Department of Information Technology of Maharashtra Institute of technology College Of Engineering, Pune for their valuable time, support, comments, suggestions and persuasion. We would also like to thank the institute for providing the required facilities, Internet access and important books.

Apoorv Shah
Harshwardhan Jadhav
Aditi Bhangale
Shivani Dandir

APPENDIX V

CONTENTS

ABSTRACT

This study addresses the problem of recognition of human interactions between two people using Video Processing and Machine Learning. The basic overview of this problem statement initiates with providing video input that involves human interaction as a ground reality. Furthermore, the video is converted into frames which then undergo through the preprocessing activity. Accordingly, the required features are extracted. Feature Extraction procedure is followed by training of the classifiers which will be used for the classification of the type of Human Interaction. The scope of the type of Human Interaction to be detected is limited to the following actions: Push, Kick, Bow, Hug, Handshake, and Boxing. This statement predominantly finds its usage in the domain of security as well as psychology as a measure of understanding the behavioral pattern of a person through continuous monitoring. Furthermore, combination of classifiers will be used to check for better accuracy as compared to the previously achieved accuracy levels.

Chapter 1

INTRODUCTION

Recognizing human activity is an important step in real time monitoring, and has found various applications such as video surveillance, artificial intelligence, etc. But there are several difficulties in analysis of interaction. Mostly, body parts of a person are occluded by the other interacting person, full body detection from videos is a challenging problem.

Past researches proposed algorithms to classify short videos of simple periodic actions performed by a single person (e.g. 'walking' and 'waiving'). In real-world applications, actions and activities are seldom periodic and are often performed by multiple persons (e.g. 'pushing' and 'hand shaking')[3]. Recognition of complex non-periodic activities, especially interactions between multiple persons, will be necessary for several applications e.g., to detect violent activities in smart surveillance systems.

To overcome this problem, we propose a hybrid model using SVM (Support Vector Machine) /CNN (Convolutional Neural Network) for detecting human interactions such as pushing, boxing, hugging, hand-shake, kicking and bowing. We also used dataset that was created using RGBD sensor (e.g. Microsoft Kinect) [3] that has been used previously by researchers. The achieved sensitivity, specificity and accuracy values make it suitable for usage in real time environment.

1.1 Project idea

This application helps in detection of human interaction and classification of the type of interaction.

1.2 Area of Project

Video Processing- it is a form of signal processing, in particular image processing, which often employs video filters and where input is video files or video streams and output of video processing either characteristics or set of parameters related to video.

Machine Learning- it is an interdisciplinary field that uses statistical techniques to give computer systems the ability to learn from data, without being explicitly programmed.

1.3 Statement of scope

The scope of this project is to detect interaction between two people in a pre –processed video. The actions are compared with the already existing set of action in the database.

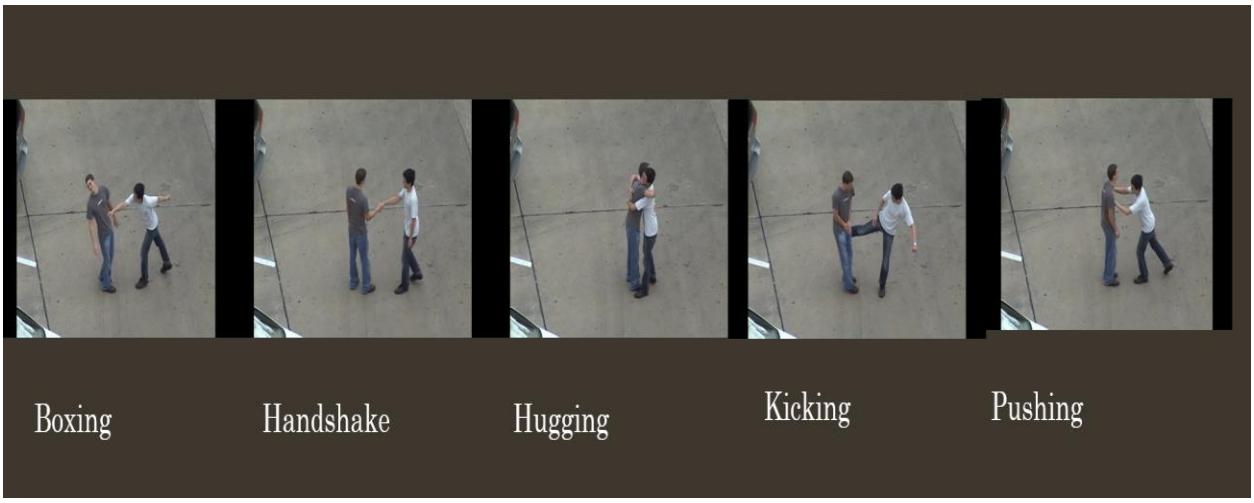
1.4 Project Statement-

To detect basic interactions like boxing, kicking, handshaking, hugging, pushing, etc between two human beings using video processing and machine learning for previously recorded videos.

1.5 Datasets

We have used datasets which were previously used in various applications such as,
1.https://www.kaggle.com/uciml/human-activity-recognition-with-smartphones
2.http://cvrc.ece.utexas.edu/SDHA2010/Human_Interaction.html

Snapshot of Dataset: -



CHAPTER 2

BACKGROUND

2.1 Saeid Motiian[1] proposed

- Representation of interactions by forming temporal trajectories, coupling together the body motion of each individual and their proximity relationships with others.
- Such trajectories are modeled with kernel state-space (KSS) models. Their advantage is being suitable for the interaction detection, recognition, while enabling a fast implementation based on recursive updates.
- For recognition, in order to compare interaction trajectories in the space of KSS models, this paper[1] design so-called pair wise kernels with a special symmetry. For detection, they exploit the geometry of linear operators in Hilbert space, and extend to KSS models the concept of parity space, originally defined for linear models.

2.2 Yu Kong[2] proposed

- Recognition of human interactions between two people, the main difficulties lie in the partial occlusion of body parts and the motion ambiguity in interactions.
- This paper [2] describes an efficient method, which model the action of each person by a large-scale global feature and local body part features, to capture such interdependencies for recognizing interaction of two people.
- A variant of multi-class Adaboost method is proposed to automatically discover class-specific discriminative three-dimensional body parts.

2.3 Kiwon Yun [3] proposed

- Human activity recognition has potential to impact a wide range of applications from surveillance to human computer interfaces to content based video retrieval.
- Recently, the rapid development of inexpensive depth sensors (e.g. Microsoft Kinect) provides adequate accuracy for real-time full-body human tracking for activity recognition applications.
- In this paper[3], they created a complex human activity dataset depicting two person interactions, including synchronized video, depth and motion capture data.
- Experimentally, we find that the geometric relational features based on distance between all pairs of joints out perform other feature choices.

2.4 WANG Jun[4] proposed

- A new method for interaction recognition based on sparse representation of feature covariance matrices was presented.
- Firstly, the dense trajectories (DT) extracted from the video were clustered into different groups to eliminate the irrelevant trajectories, which could greatly reduce the noise influence on feature extraction.
- Then, the trajectory tunnels were characterized by means of feature covariance matrices. In this way, the discriminative descriptors could be extracted, which was also an effective solution to the problem that the description of the feature second-order statistics is insufficient.
- Classification was achieved using multiple instance learning (MIL), which was more suitable for complex environments.

2.5 Rami Alazrai [5] proposed

- Motion-pose geometric descriptor (MPGD) as a human– human interaction representation to capture the semantic meaning of body-parts between two interacting humans.
- The proposed MPGD representation is based on utilizing the concept of anatomical planes to construct a motion profile and a pose profile for each human.
- For the human–human interaction classification problem, they propose a hierarchical classification framework consisting of a representation layer and three classification layers.
- The human–human interaction prediction problem aims to predict the class of ongoing human–human interaction at its early stages. To do so, they propose a prediction framework that utilizes the proposed MPGD to construct an accumulated histograms-based representation for an ongoing interaction.

CHAPTER 3

SPECIFICATIONS

Main Constraints

- The main constraint is the quality of camera which directly impact on the result of application. Good quality camera is needed for recording.
- The resolution of video is another issue which impact on specification and also on implementation of application.

3.1 Hardware Specifications

Processor	- Intel i3 core
Speed	- 1.1 GHz
RAM	- 256 MB (min)
Hard Disk	- 20 GB
Key Board	- Standard Windows Keyboard
Mouse	- Two or Three Button Mouse\

3.2 Software Specifications

Operating System	: Windows7
Coding language	: JAVA (JDK 7.0)
Database	: MySQL 5.0
IDE	: Eclipse Luna

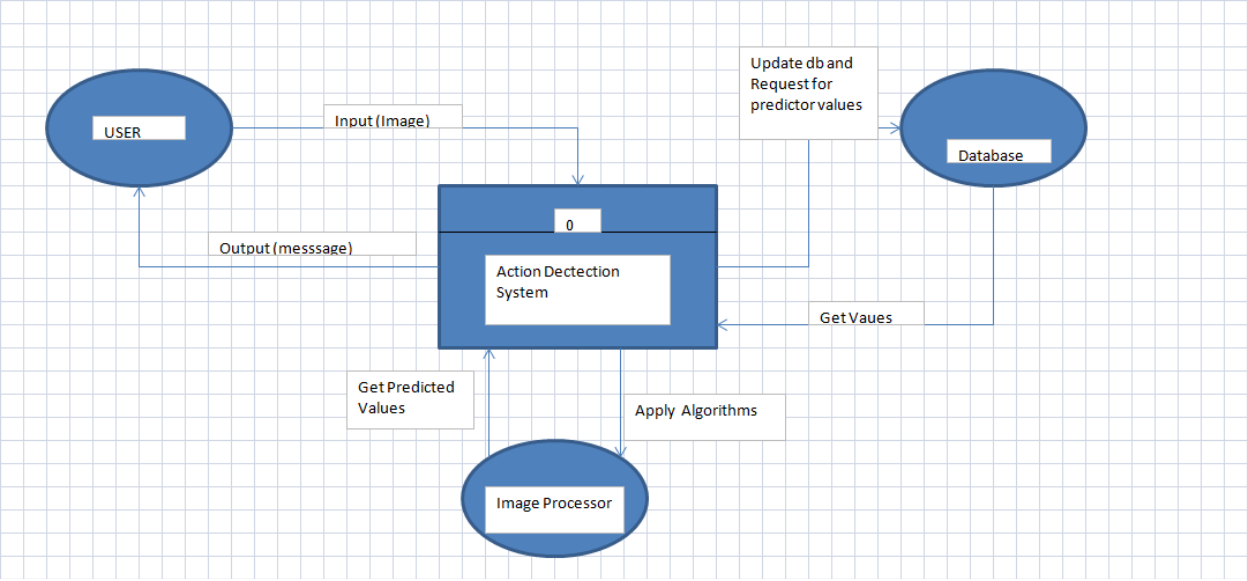
Chapter 4

DESIGN

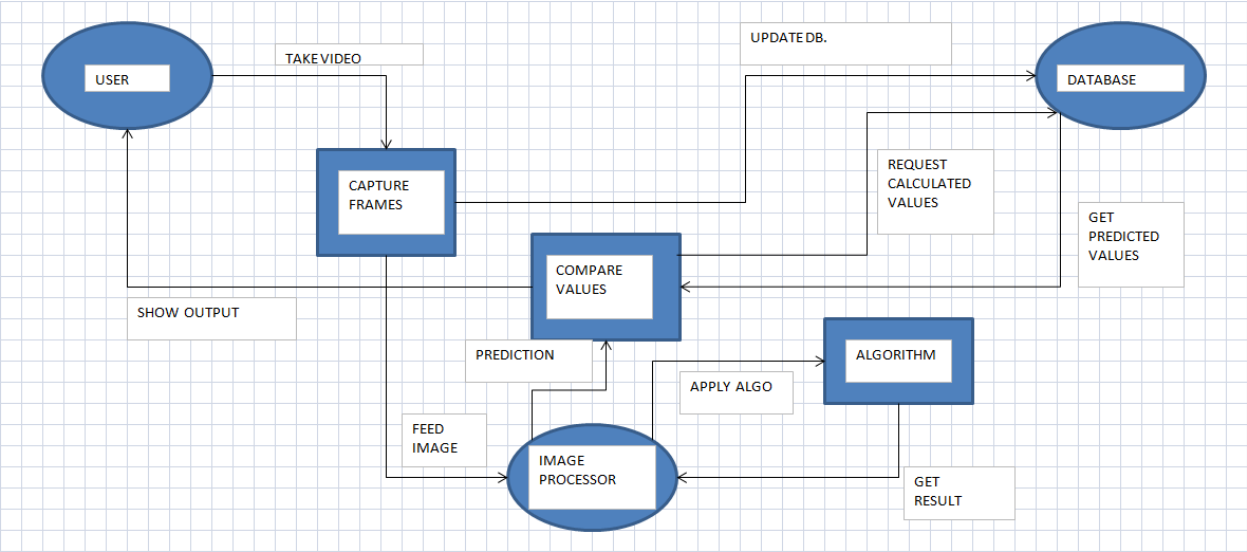
4.1 Data Flow Diagram

A Data Flow Diagram is a graphical representation of the flow of data through an information system, modeling its process aspects. A DFD is often used as a preliminary step to create an overview of the system, which can later be elaborated. A DFD shows what kind of information will be input to and output from the system, where the data will come from and go to, and where the data will be stored.

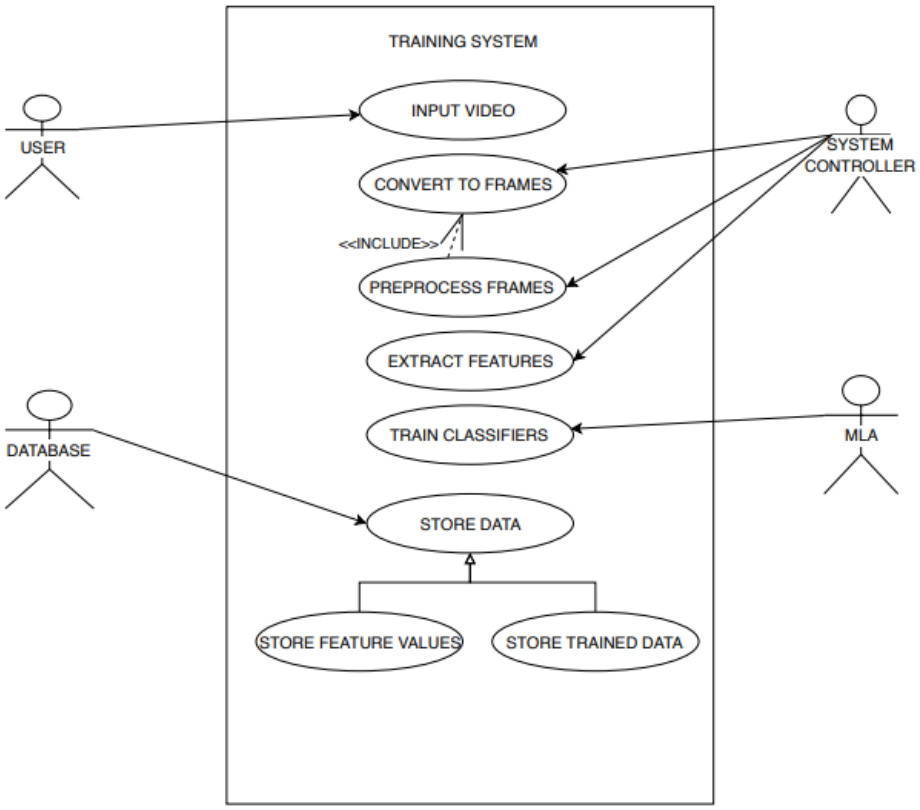
LEVEL 0 DFD

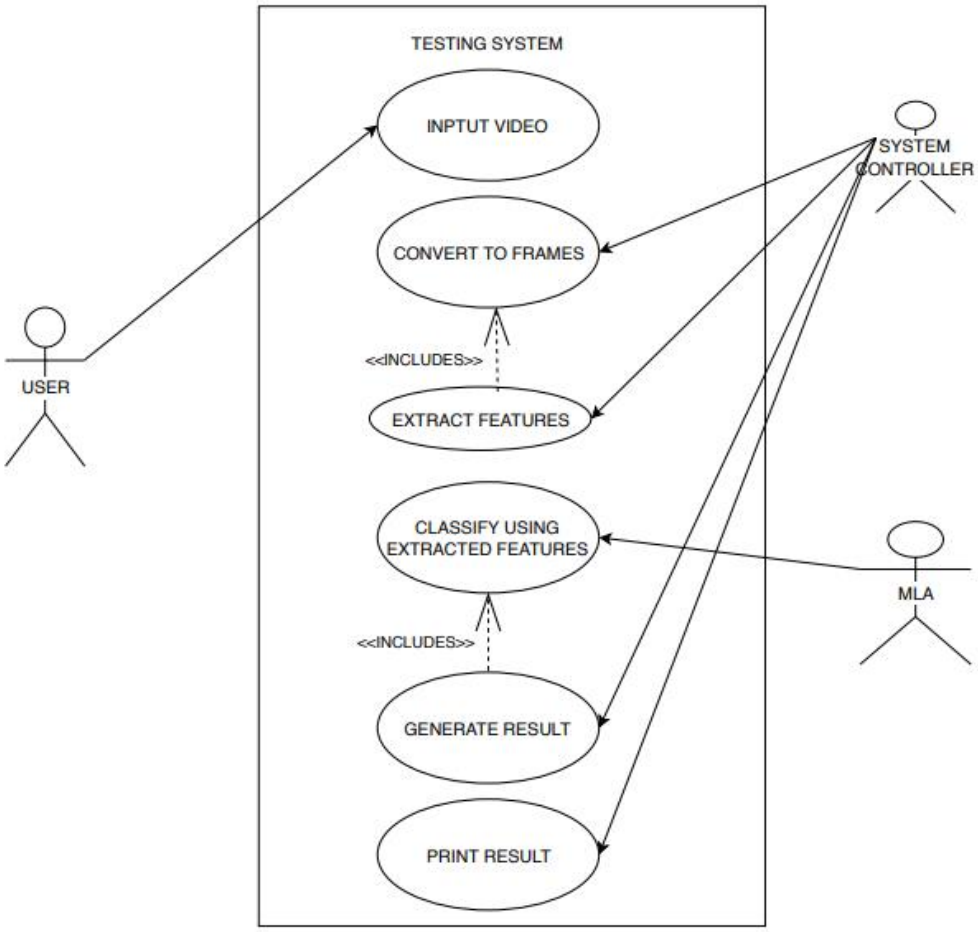


4.2 LEVEL 1 DFD

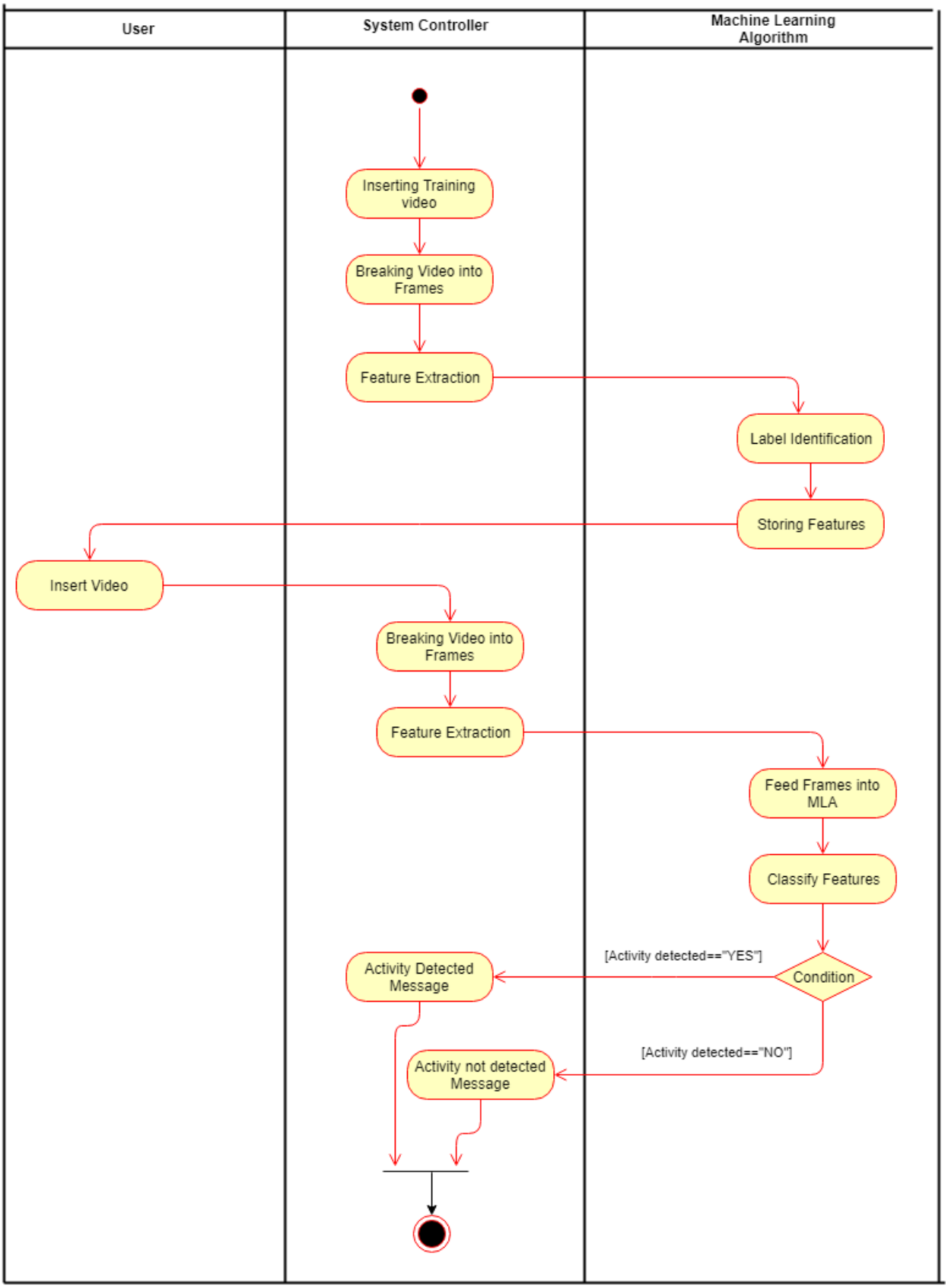


4.3 Use Case Diagram

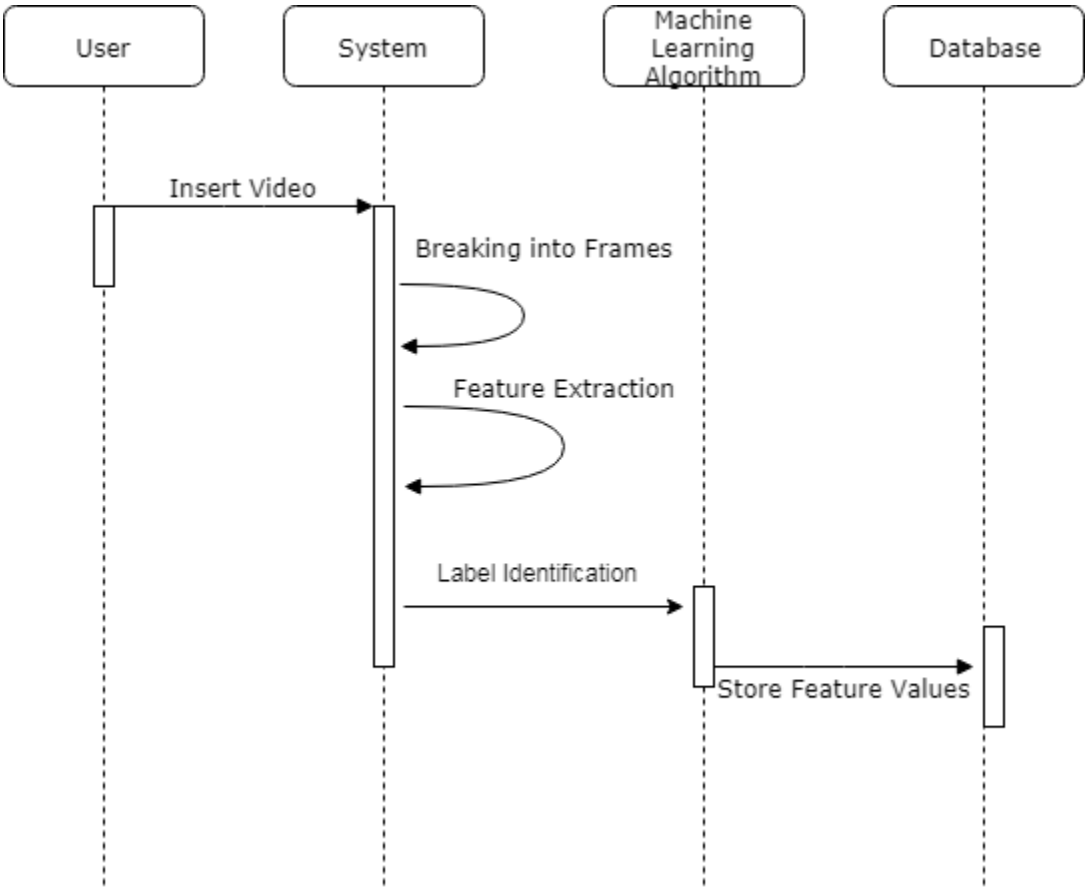


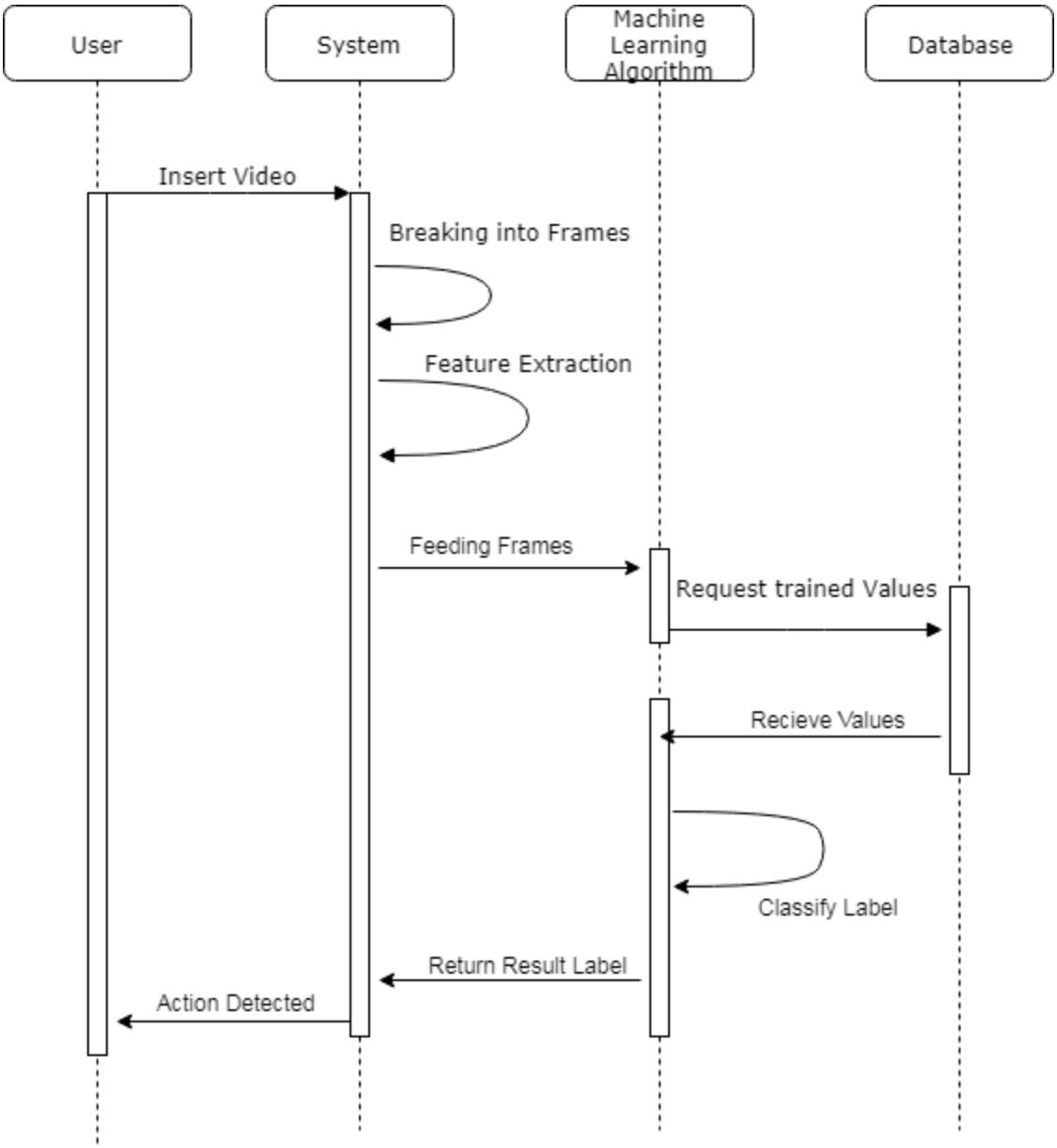

TESTING SYSTEM
INPTUT VIDEO
CONVERT TO FRAMES
<<INCLUDES>>
EXTRACT FEATURES
CLASSIFY USING EXTRACTED FEATURES
<<INCLUDES>>
GENERATE RESULT
PRINT RESULT
USER
SYSTEM CONTROLLER
MLA

4.4 Activity Diagram

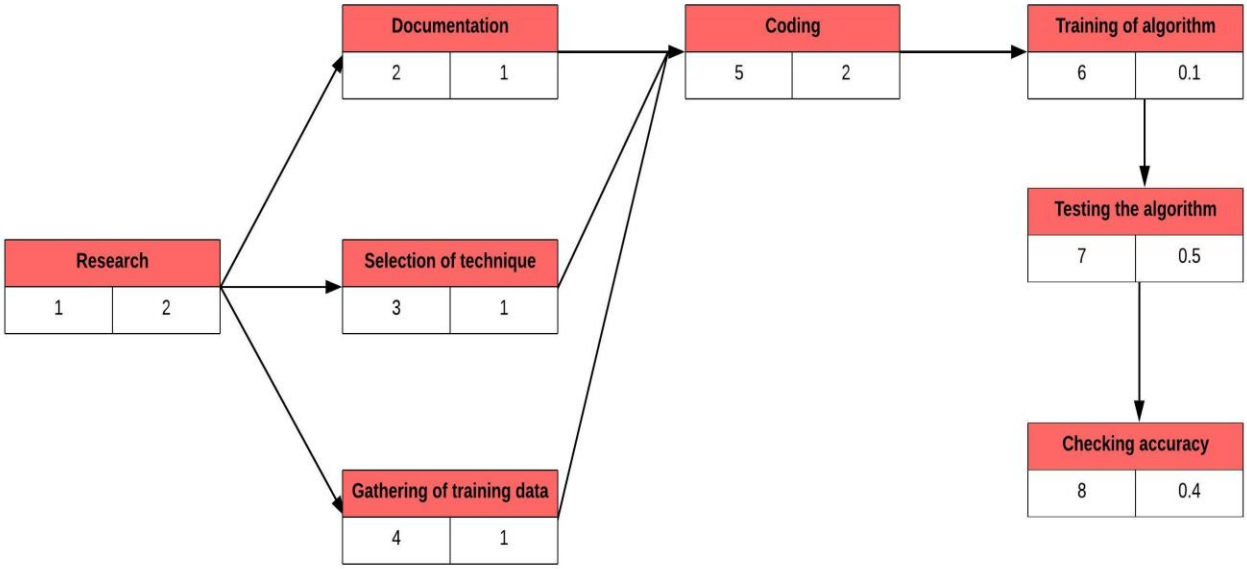


4.5 Sequence Diagram




User
System
Machine Learning Algorithm
Database
Insert Video
Breaking into Frames
Feature Extraction
Feeding Frames
Request trained Values
Recieve Values
Classify Label
Return Result Label
Action Detected

4.6 Pert Chart



Chapter 5

IMPLEMENTATION

5.1 Implemented Modules:

Train model:
Program is trained using categorizes set of videos.

Video to Frames:
The input video is converted to frames and recorded into the database.

ML Algorithm:
After studying various algorithm, we have finally opted for SVM as it is …//advantages

Testing Results:
The program is tested for various actions to identify its precision.

Results:
Results are calculated on every test case and the result is shown on the output window which is accessible to the user.

5.1 Overview of Project Modules

- **User**
  - User first register account in application.
  - After activate account, user should login.
  - User upload video and extract the frames
  - After processing user gets final result i.e. identified human actions.

5.2 Algorithms

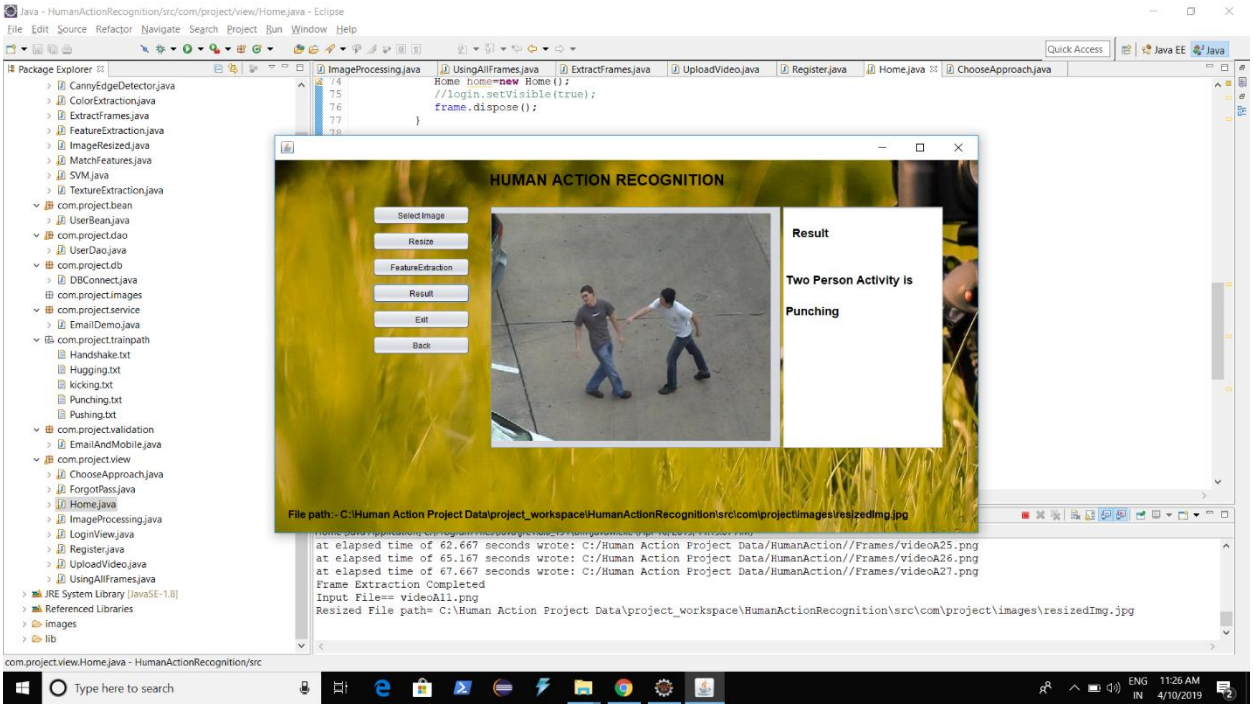
1. Feature Extraction for Face Identification:

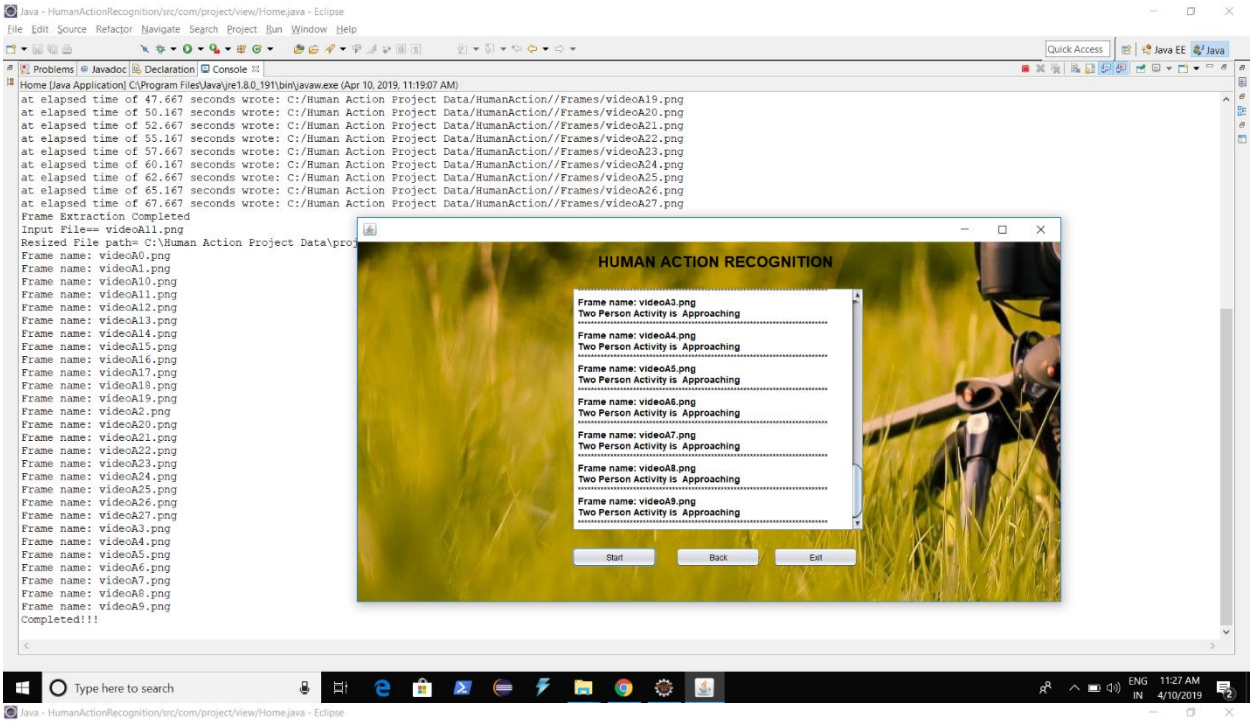
Color feature: As most of the color distribution information can be captured by the low-order moments, using only the first three moments: mean, variance and skewness, it is found that these moments give a good approximation and have been proven to be efficient and effective in representing the color distribution of images.

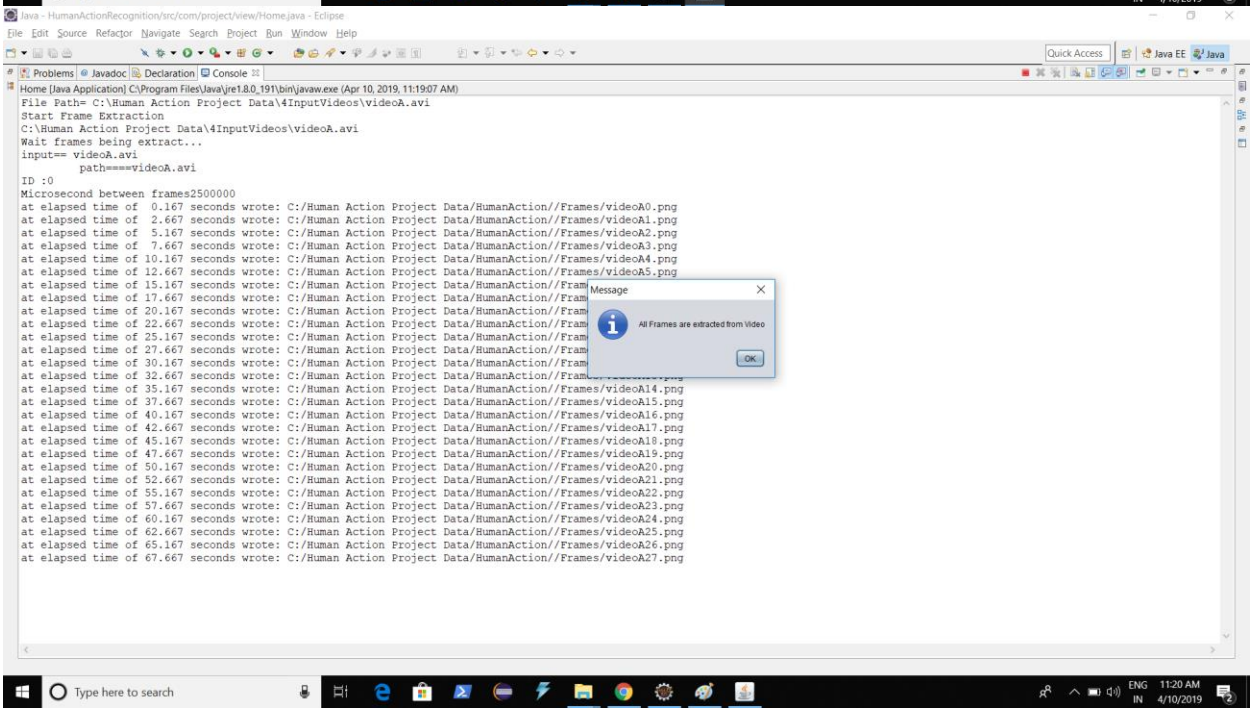
Edge Detection: Most of the shape information of an image is enclosed in edges. So first we detect these edges in an image and by using these filters and then by enhancing those areas of image which contains edges, sharpness of the image will increase and image will become clearer. **Texture feature:** Describes the structure arrangement of surfaces and their relationship to the environment, such as fruit skin, clouds, trees, and fabric. The texture feature in our method is described by hierarchical wavelet packet descriptor (HWVP). A 170- D HWVP descriptor is utilized by setting the decomposition level to be 3 and the wavelet packet basis to be DB2.

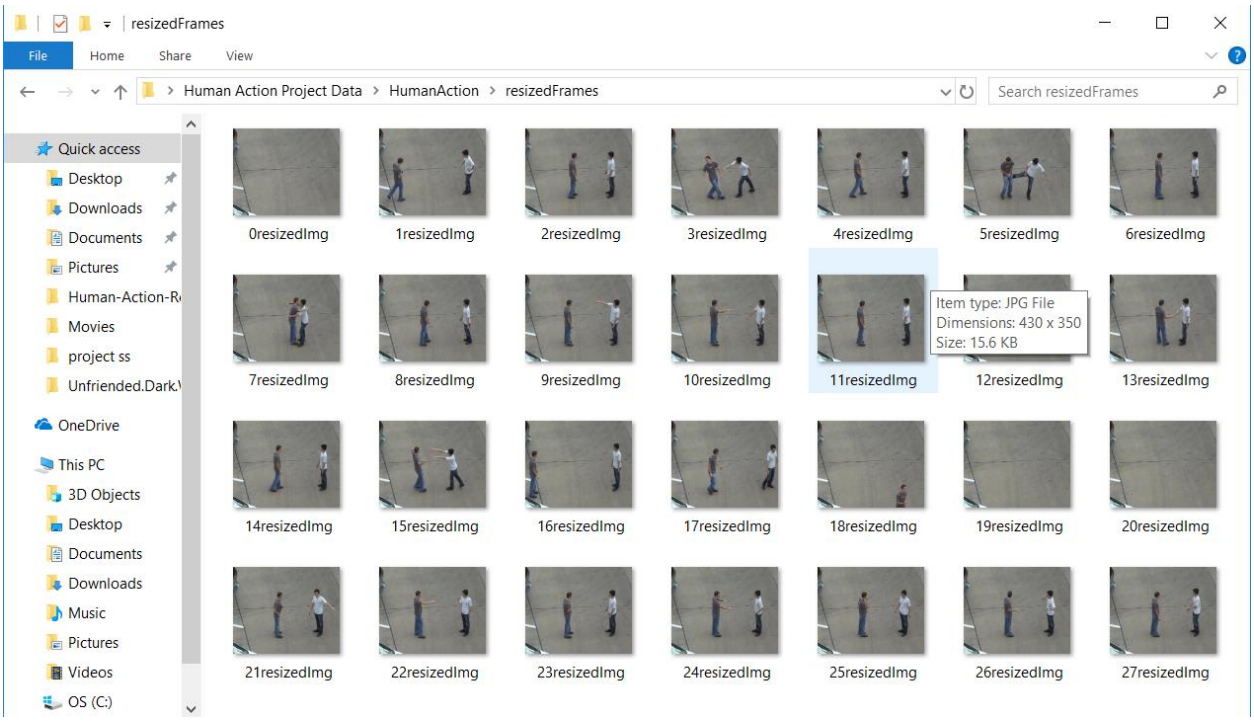
2. Support Vector Machine

- Support Vector Machine (SVM) is used to classify the fruit quality. SVM Support vector machines are mainly two class classifiers, linear or non-linear class boundaries.
- The idea behind SVM is to form a hyper plane in between the data sets to express which class it belongs to.
- The task is to train the machine with known data and then SVM find the optimal hyper plane which gives maximum distance to the nearest training data points of any class




HUMAN ACTION RECOGNITION
Frame name: videoA3.png
Two Person Activity is Approaching
Frame name: videoA4.png
Two Person Activity is Approaching
Frame name: videoA5.png
Two Person Activity is Approaching
Frame name: videoA6.png
Two Person Activity is Approaching
Frame name: videoA7.png
Two Person Activity is Approaching
Frame name: videoA8.png
Two Person Activity is Approaching
Frame name: videoA9.png
Two Person Activity is Approaching
Start
Back
Exit


Message
All Frames are extracted from Video
OK


resizedFrames
File
Home
Share
View
Human Action Project Data
HumanAction
resizedFrames
Search resizedFrames
Quick access
Desktop
Downloads
Documents
Pictures
Movies
project ss
OneDrive
This PC
3D Objects
Desktop
Documents
Downloads
Music
Pictures
Videos
OS (C:)
0resizedImg
1resizedImg
2resizedImg
3resizedImg
4resizedImg
5resizedImg
6resizedImg
7resizedImg
8resizedImg
9resizedImg
10resizedImg
11resizedImg
12resizedImg
13resizedImg
14resizedImg
15resizedImg
16resizedImg
17resizedImg
18resizedImg
19resizedImg
20resizedImg
21resizedImg
22resizedImg
23resizedImg
24resizedImg
25resizedImg
26resizedImg
27resizedImg
Item type: JPG File
Dimensions: 430 x 350
Size: 15.6 KB

Chapter 6

RESULTS AND EVALUATION

Test case For Registration and Login Page:

Project Name: **Human Action Recognition**

.

Total no of test Cases:-04

Total no of test Cases Passed:-04

Total no of test Cases failed:-00

Total no of test Cases executed:-04

Total no of test Cases pending:-00

Test Case ID	Test Case Procedure	Input Data	Expected Output	Actual Output	Test Status
POI - LO-01	Checking the functionality of Login Button	1.Enter valid User name 2 .Enter valid Password 3. Click on Login Button	Login page should be displayed	Login page displayed	Pass
POI - LO-02	Checking the functionality of Login Button	1.Enter valid User name 2 .Enter Invalid Password 2. Click on Login Button	Login page should not be displayed	Login page not displayed	Pass
POI - LO-03	Checking the functionality of Login Button	1.Enter Invalid User name 2 .Enter valid Password 2. Click on Login Button	Login page should not be displayed	Login page not displayed	Pass
POI - LO-04	Checking the functionality of Login Button	1.Enter Invalid User name 2 .Enter Invalid Password 2. Click on Login Button	Login page should not be displayed	Login page not displayed	Pass

Module Name: Register.

Total no of test Cases:-05

Total no of test Cases Passed:-05

Total no of test Cases failed:-00

Total no of test Cases executed:-05

Total no of test Cases pending:-00

Test Case ID	Test Case Procedure	Input Data	Expected Output	Actual Output	Test Status
POI - LO-01	Checking the functionality of Register Button	1.Enter User name 2 .Enter valid email address 3. Enter valid mobile No. 4. Enter valid password i.e. in specific format. 5.Click on Register Button	Register page should be displayed	After Registration, Login page displayed	Pass
POI - LO-02	Checking the functionality of Register Button	1.Enter User name 2 .Enter invalid email address 3. Enter valid mobile No. 4. Enter valid password i.e. in specific format. 5.Click on Register Button	Registration Fail.	Registration Fail.	Pass
POI - LO-03	Checking the functionality of Register Button	1.Enter User name 2 .Enter valid email address	Registration Fail.	Registration Fail.	Pass

		3. Enter invalid mobile No. 4. Enter valid password i.e. in specific format. 5.Click on Register Button			
POI - LO-04	Checking the functionality of Register Button	1.Enter User name 2 .Enter valid email address 3. Enter valid mobile No. 4. Enter invalid password i.e. not in specific format. 5.Click on Register Button	Registration Fail.	Registration Fail.	Pass
POI - LO-05	Checking the functionality of Register Button	1.Enter User name 2 .Enter invalid email address 3. Enter invalid mobile No. 4. Enter invalid password i.e. not in specific format. 5.Click on Register Button	Registration Fail.	Registration Fail.	Pass

Test Cases for Video Upload:

Total no of test Cases:-02

Total no of test Cases Passed:-02

Total no of test Cases failed:-00

Total no of test Cases executed:-02

Total no of test Cases pending:-00

Test Case ID	Test Case Procedure	Input Data	Expected Output	Actual Output	Test Status
POI - LO-01	Checking the functionality of browse Button	1. Upload valid video 2. Click on Upload Button	Uploaded video path should be displayed	Uploaded video path displayed	Pass
POI - LO-02	Checking the functionality of browse Button	1. Upload invalid video 2. Click on Upload Button	Uploaded video path should not be displayed	Uploaded video path not displayed	Pass

Test Cases for Image Upload:

Total no of test Cases:-02

Total no of test Cases Passed:-02

Total no of test Cases failed:-00

Total no of test Cases executed:-02

Total no of test Cases pending:-00

7.

Test Case ID	Test Case Procedure	Input Data	Expected Output	Actual Output	Test Status
POI - LO-01	Checking the functionality of browse Button	1. Upload valid image 2. Click on Upload Button	Uploaded image should be displayed	Uploaded image displayed	Pass
POI - LO-02	Checking the functionality of browse Button	1. Upload invalid image 2. Click on Upload Button	Uploaded image should not be displayed	Uploaded image not displayed	Pass

Software Testing
System Testing

For system testing we test to perform validations and verifications on users every input and then processed such as user personal details, uploaded video format, uploaded image type and more. After performing the integration testing, the next step is output testing of the proposed system. No system could be useful if it does not produce the required output in a specified format. The outputs generated are displayed by the user.

Project aspect:

The entire project was tested and found successful.

6.4 Test Plan

Test Plan Identifier: cost-sensitive

It is use to identify test plan uniquely.

1. Purpose of the Test Plan Document

The main purpose of this document is to fit a particular projects needs. It documents and tracks the necessary information required to effectively define the approach to be used in the testing of the projects product. The Test Plan document is created during the Planning Phase of the project. Its intended audience is the project manager, project team, and testing team.

2. Objective of Test Panning

To find as many defects as possible and get them fix.

3. Items to be tested OR not to be tested

Describe the items/features/functions to be tested that are within the scope of this test plan.
Include a description of how they will be tested, when, by whom, and to what quality standards.
Also include a description of those items agreed not to be tested.

4. Items to be tested

Overall functionality of the application

User Interface of the application.

5. Test Approach

Describe the overall testing approach to be used to test the projects product. Provide an outline of any planned tests. There are many approaches like: Black Box Testing White Box Testing
Here we used Black Box Testing approach. In Black Box Testing we just give input to the system and check its output without checking how system processes it.

6. Test Pass OR Test Fail Criteria

When actual and expected results are same then test will be passed. When actual and expected results are different then test will be failed.

7. Test Entry OR Exit Criteria

Describe the entry and exit criteria used to start testing and determine when to stop testing.

Entry criteria: As soon as we have requirement we can start testing.

Exit criteria: When bug rate fall below certain level we can stop testing.

8. Test Suspension OR Resumption Criteria

Describe the suspension criteria that may be used to suspend all or portions of testing.

Also describe the resumption criteria that may be used to resume testing.

Suspension criteria: If there is large change in application like change in requirements we can suspend work for some time.

Resumption criteria: After resolving the respective problem we can resume work. Item

Testing Type

It describes which testing types we are going to follow in our testing lifecycle. Here we are using:

UI Testing

Integration Testing

9. Test Cases

Test cases are built around specifications and requirements, i.e., what the application is supposed to do. Test cases are generally derived from external descriptions of the software, including specifications, requirements and design parameters. Although the tests used are primarily functional in nature, non-functional tests may also be used. The test designer selects both valid and invalid inputs and determines the correct output without any knowledge of the test object's internal structure. Test Design Techniques

Typical black-box test design techniques include:

- Decision table testing
- All-pairs testing
- State transition Analysis
- Equivalence partitioning
- Boundary value analysis
- Cause effect graph
- Error guessing

Advantages

- Efficient when used on large systems.
- Since the tester and developer are independent of each other, testing is balanced and unprejudiced.
- Tester can be non-technical.
- There is no need for the tester to have detailed functional knowledge of system.
- Tests will be done from an end user's point of view, because the end user should accept the system. (This testing technique is sometimes also called Acceptance testing.)
- Testing helps to identify vagueness and contradictions in functional specifications.

- Test cases can be designed as soon as the functional specifications are complete.

Disadvantages

- Test cases are challenging to design without having clear functional specifications.
- It is difficult to identify tricky inputs if the test cases are not developed based on specifications.
- It is difficult to identify all possible inputs in limited testing time. As a result, writing test cases may be slow and difficult.
- There are chances of having unidentified paths during the testing process.
- There is a high probability of repeating tests already performed by the programmer.

The model implemented will detect various actions like kicking, boxing, hugging, handshake, push, etc. This will be helpful in surveillance domain.

Chapter 7

CONCLUSION AND FUTURE WORK

Below graph shows the action recognition accuracy of our proposed work. Our proposed work successfully classifies the human actions based on various feature extraction techniques. The final accuracy result as shown in following graph and table respectively.

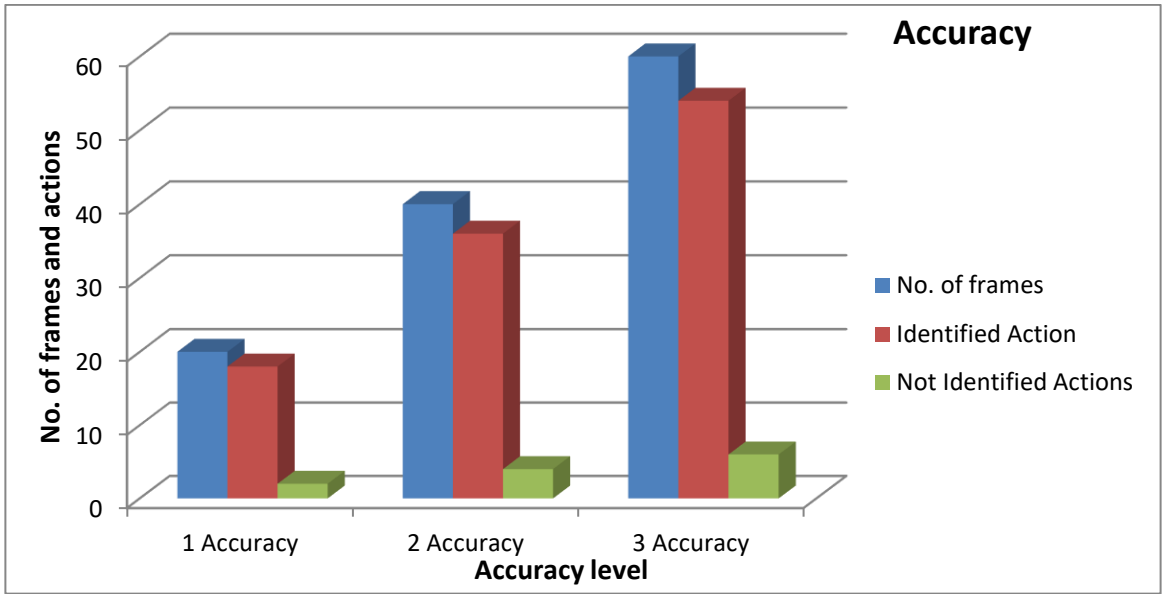


Fig2 Accuracy Graph

Table1: Accuracy Table

	No. of frames	Identified Actions	Not Identified Actions
1 Accuracy	20	18	2
2 Accuracy	40	36	4
3 Accuracy	60	54	6

7.1 Outcomes

- Video Frames Extraction
- Frame Feature Extraction
- SVM Classification
- Human Action Recognition Result

8. Conclusions

8.1 Conclusions

In this project, a feature representation and activity recognition system for human activity recognition for video retrieval system is proposed. The key frames selected to represent a sequence of activity, significantly reduced the computational complexity. The feature extraction

algorithm is implemented using java to identify the activity of person. This results in increased efficiency for our human activity recognition system. The recognition process is performed by using classifier support vector machine (SVM). SVM classifier shows highest recognition results for human activity recognition. Videos similar to input video are retrieved from the database.

8.2 Future Work

In future we will enhance this system to implement an android application also real time video capturing process.

8.3 Applications

- Searching and browsing large video archives
- Indexing and archiving videos
- Easy access to material
- Video content filtering

REFERENCES

[1]. Motiian, Saeid, Farzad Siyahjani, Ranya Almohsen, and Gianfranco Doretto."Online human interaction detection and recognition with multiple cameras.", IEEE Transactions on Circuits and Systems for Video Technology 27, no. 3 (2017): 649-663.

[2]. Kong, Yu, Wei Liang, Zhen Dong, and Yunde Jia. "Recognising human interaction from videos by a discriminative model." IET Computer Vision 8,no. 4 (2014): 277-286.

[3]. Yun, Kiwon, Jean Honorio, Debaleena Chattopadhyay, Tamara L. Berg, and Dimitris Samaras. "Two-person interaction detection using body-pose features and multiple instance learning." In Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on, pp. 28-35. IEEE, 2012.

[4]. Wang, Jun, Si-chao Zhou, and Li-min Xia. "Human interaction recognition based on sparse representation of feature covariance matrices." Journal of Central South University 25, no. 2 (2018): 304-314.

[5]. Alazrai, Rami, Yaser Mowafi, and CS George Lee. "Anatomical-plane-based representation for human–human interactions analysis." Pattern Recognition 48, no. 8 (2015): 2346-2363.