ENTIRE HOME SURVEILLANCE APPROACH USING PI AND INTELLIGENT CAMERAS WITH MOBILE APP NOTIFICATION

A PROJECT REPORT

Submitted by

SANTHOSH.M (312417106074)

SRI DHARSHAN.V.S (312417106094)

in partial fulfilment for the award of the degree

of

BACHELOR OF ENGINEERING

in

ELECTRONICS AND COMMUNICATION ENGINEERING

St. JOSEPH'S INSTITUTE OF TECHNOLOGY
CHENNAI 600 119

ANNA UNIVERSITY::CHENNAI 600 025

JULY 2021

ANNA UNIVERSITY:CHENNAI 600 025

BONAFIDE CERTIFICATE

Certified that this project report "**HOME SURVEILLANCE APPROACH USING PI AND INTELLIGENT CAMERAS WITH MOBILE APP NOTIFICATION"** is the bonafide work of **SANTHOSH.M (312417106074)** and **SRI DHARSHAN.V.S (312417106094)** who carried out the project under my supervision.

SIGNATURE

Dr. C.GNANA KOUSALYA, M.E., Ph.D.,

PROFESSOR & HEAD

DEPARTMENT OF
ELECTRONICS
AND COMMUNICATION
ENGINEERING

St. JOSEPH'S INSTITUTE OF
TECHNOLOGY,

OMR,CHENNAI- 600119.

SIGNATURE

Mrs.P. PRINCY MAGDALINE, M.Tech.,

SUPERVISOR & ASSISTANT
PROFESSOR

DEPARTMENT OF
ELECTRONICS
AND COMMUNICATION
ENGINEERING

St. JOSEPH'S INSTITUTE OF
TECHNOLOGY,

OMR,CHENNAI- 600119.

CERTIFICATE OF EVALUATION

College : **St. Joseph's Institute of Technology, Chennai-600 119**

Branch : **Electronics and Communication Engineering**

Semester : **VIII**

NAME OF THE STUDENT	TITLE OF THE PROJECT	NAME OF THE SUPERVISOR WITH DESIGNATION
SANTHOSH.M (312417106074) SRI DHARSHAN.V.S (312417106094)	HOME SURVEILLANCE APPROACH USING PI AND INTELLIGENT CAMERAS WITH MOBILE APP NOTIFICATION	Mrs. P.PRINCY MAGDALINE, M.Tech., ASSISTANT PROFESSOR

The report of project work submitted by the above students in partial fulfilment for the award of Bachelor of Engineering degree in Electronics and Communication Engineering branch of Anna University were evaluated and confirmed to be reports of the work done by the above student and then evaluated.

INTERNAL EXAMINER **EXTERNAL EXAMINER**

ACKNOWLEDGEMENT

At the outset we would like to express our sincere gratitude to our beloved **Chairman Dr. B. BABU MANOHARAN, M.A., M.B.A., Ph.D.,** for his constant guidance and support.

We would like to express our thanks to our respected **Managing Director, Mrs. B.JESSIE PRIYA, M.Com.,** and our **Director Mr. B.SHASHI SEKAR, M.Sc.,** for their kind encouragement and blessings.

We express our sincere gratitude and whole hearted thanks to our **Principal Dr. P.RAVICHANDRAN, M.Tech., Ph.D.,** for his encouragement to make this project a successful one.

We wish to express our sincere thanks and gratitude to our **Head of the Department Dr. C.GNANA KOUSALYA**, **M.E., Ph.D.,** of Electronics and Communication Engineering for leading us towards the completion of this project.

We also wish to express our sincere thanks to our **Project Guide Mrs. P.PRINCY MAGDALINE, M.Tech.,** Assistant Professor**,** Department of Electronics and Communication Engineering for her guidance and assistance in solving the various intricacies involved in the project.

Finally, we thank our parents and friends who helped us in the successful completion of this project.

ABSTRACT

Home surveillance aims to allow the homeowners to observe their home at any time virtually from any location. Home automation has increased in popularity in recent years as a result of its low cost and ease of use, through smartphone and tablet connectivity. Homeowners who want to save money often opt for a home security provider's simple kit. For instance, criminals may find an unprotected entry point into the home if there is only a front and backdoor sensor. By taking a tour around the perimeter of the home to spot any security weaknesses and order the required equipment pieces to protect all areas. In this project the home is protected with multiple Raspberry Pi that communicate among themselves to accomplish a common goal. The beginning may be by building a case for security with an emphasis on the lack of flexibility of commercially available systems, that is where Raspberry Pi are brought to the rescue. Also, this interoperable system will involve Raspberry Pi, cameras and a mobile application for alert by predicting the presence of unknown person. The algorithms used are Single shot multi-box detector algorithm for face detection, Mobile facenet algorithm for face recognition and Image ZMQ algorithm for image transportation. A mobile application is developed using react native to enable live streaming and options are given for the user to alert the police, lock the door or can even ignore when an unknown trespasser enters the property. As a result, the solution is flexible, affordable, and interoperable with other IoT devices and services that are worth paying for.

TABLE OF CONTENTS

CHAPTER	TITLE	PAGE NO
	ABSTRACT	**v**
	LIST OF FIGURES	**ix**
	LIST OF ABBREVIATIONS	**x**
1	**INTRODUCTION**	**1**
	1.1 GENERAL	1
	1.2 TECHNOLOGIES USED	4
	1.2.1 Importance of Deep Learning	5
	1.2.2 Deep Learning Process	6
	1.2.3 Classification of Neural Networks	7
	1.2.4 Feed Forward Neural Networks	8
	1.2.5 Recurrent Neural Networks	8
	1.2.6 Common uses of RNN	9
	1.2.7 Convolutional Neural Networks	9
	1.3 EXISTING SYSTEM	11
	1.4 DISADVANTAGES OF EXISTING SYSTEM	12

2	**LITERATURE REVIEW**	**13**
	2.1 INTRODUCTION	13
	2.2 LITERATURE SURVEY	13
3	**SYSTEM ANALYSIS**	**31**
	3.1 PROPOSED SYSTEM	31
	3.2 SYSTEM ARCHITECTURE	32
	3.3 WORKING	33
	3.4 MODULE DESCRIPTION	34
	3.4.1 Object Detection Module	34
	3.4.2 Face Detection Module	37
	3.4.3 Face Recognition Module	38
	3.4.4 Image ZMQ for Pi to Pi Communication	40
	3.4.5 Mobile App Module	42
	3.5 HARDWARE USED	44
	3.5.1 Raspberry Pi Zero	44
	3.5.2 Servo Motor	45

	3.5.3 Power Supply	46
	3.5.4 Buzzer	47
	3.5.5 Raspberry Pi Camera Module	48
	3.6 SOFTWARE LIST	50
	3.6.1 Microsoft Visual Studio	50
	3.6.2 Python	52
	3.7 ADVANTAGES	53
	3.8 APPLICATIONS	53
4	**RESULTS AND DISCUSSIONS**	**54**
	4.1 FINAL OUTPUTS	54
5	**CONCLUSION AND FUTURE ENHANCEMENT**	**60**
	5.1 CONCLUSION	60
	5.2 FUTURE ENHANCEMENT	60
	REFERENCE	**61**

LIST OF FIGURES

FIGURE NO	NAME OF THE FIGURE	PAGE NO
1.1	DEEP LEARNING LAYERS	5
1.2	DEEP LEARNING PROCESS	6
1.3	TYPES OF DEEP LEARNING NETWORKS	8
1.4	CNN	10
1.5	TYPES OF CNN	12
3.1	SYSTEM ARCHITECTURE	32
3.2	SSD STRUCTURE	36
3.3	MOBILEFACENET ARCHITECTURE	39
3.4	OVERVIEW ARCHITECTURE OF ZMQ	42
3.5	MOBILE APP DEVELOPMENT	43
3.6	RASPBERRY PI	44
3.7	SERVO MOTOR	45
3.8	POWER SUPPLY	46
4.1	HARDWARE SETUP OF THE PROJECT	54
4.2	LOGIN PAGE	55
4.3	HOME PAGE	56
4.4	KNOWN FACE DETECTION	57
4.5	UNKNOWN FACE DETECTION	58
4.6	ABNORMAL INTIMATION	59

LIST OF ABBREVIATIONS

CNN	Convolutional Neural Network
ILSVRC	ImageNet Large Scale Visual Recognition Challenge (),
RNN	Recurrent Neural Networks
SSD	Single Shot Multi-Box Detector
CLI	Command-Line Interface
ROI	Region of Interest

CHAPTER -1

INTRODUCTION

1.1 GENERAL:

Home security is both the security hardware in place on a property as well as personal security practices. Security hardware includes doors, locks, alarm systems, lighting, motion detectors, security camera systems, etc. that are installed on a property; personal security involves practices such as ensuring doors are locked, alarms activated, windows closed, extra keys not hidden outside, etc. According to an FBI report, 58.3 percent of burglaries in the United States involved forcible entry. Per the most recent statistics, the average burglary in the United States lasts for about 90 seconds to 12 minutes and, on average, a burglar will break into a home within 60 seconds. Most thefts target cash first followed by jewels, drugs, and electronics. Common security methods include never hiding extra keys outside, never turning off all the lights, applying small CCTV stickers on doors, and keeping good tabs with neighbours.

convergence of technologies in machine learning and omnipresent computing as well as the development of robust sensors and actuators has brought interest in the development of smart environments to emerge and support valuable functions in Daily Living Activities (ADLs). The need for such technologies to be developed is underlined by population aging, the cost of formal health care, and the importance individuals place on remaining independent in their own homes. Individuals need to be able to complete daily living activities such as eating, dressing, cooking, drinking, reading, taking medicine, sleeping, to function independently at home. Automating activity recognition is a crucial step towards monitoring a smart home resident's functional health and helping them perform these activities effectively. Before smart home technologies can be deployed for these older people, several challenges should be resolved, including data collection, algorithms for activity recognition, etc. This technology can be used wildly in the future if the accuracy is sufficiently higher. There is a research project called the Advanced Studies Center in Adaptive Systems (CASAS) where only passive, non-intrusive sensors are deployed at Washington State University to create an intelligent home environment.

Pros

Energy Savings

Home automation systems have definitely proven themselves in the arena of energy efficiency. Automated thermostats allow you to pre-program temperatures based on the time of day and the day of the week. And some even adjust to your behaviours, learning and adapting to your temperature preferences without your ever inputting a pre-selected schedule. Traditional or behaviour-based automation can also be applied to virtually every gadget that can be remotely controlled – from sprinkler systems to coffee makers.

Actual energy savings ultimately depend on the type of device you select and its automation capabilities. But on average, product manufacturers estimate the systems can help consumers save anywhere from 10 to 15 percent off of heating and cooling bills.

Convenience

In today’s fast-paced society, the less you have to worry about, the better. Right? Convenience is another primary selling point of home automation devices, which virtually eliminate small hassles such as turning the lights off before you go to bed or adjusting the thermostat when you wake up in the morning.

Many systems come with remote dashboard capabilities, so forgetting to turn off that coffee pot before you leave no longer requires a trip back to the house. Simply pull up the dashboard on a smart device or computer, and turn the coffee pot off in a matter of seconds.

Security

Remote monitoring can put your mind at ease while you're away from the house. With remote dashboards, lights and lamps can be turned on and off, and automated blinds can be raised and lowered. These capabilities – combined with automated security systems – can help you mitigate the risks of intrusions: you will be alerted immediately if something uncharacteristic happens.

Cons

Installation

Depending on the complexity of the system, installing a home automation device can be a significant burden on the homeowner. It can either cost you money if you hire an outside contractor or cost you time if you venture to do it yourself.

Complex Technology

Automating everything in life may sound extremely appealing, but sometimes a good old-fashioned flip of the switch is a lot easier than reaching for your smart phone to turn lights on and off. Before you decide which system is right for you, think about how far you really want to take home automation in your household.

System Compatibility

Controlling all aspects of home automation from one centralized platform is important, but not all systems are compatible with one another. Your security system, for example, may require you to log in to one location to manage settings, while your smart thermostat may require you to log in to another platform to turn the air conditioner on and off. To truly leverage the convenience of home automation, you may need to invest in centralized platform technology to control all systems and devices from one location.

Cost

Even though the price of home automation systems has become much more affordable in recent years, the cost to purchase and install a device can still add up. Consumer Reports offers a wide range of information and insights – including costs – on the best home automation systems on the market.

1.2 TECHNOLOGIES USED:

Deep learning is a computer software that mimics the network of neurons in a brain. It is a subset of machine learning and is called deep learning because it makes use of deep neural networks.

Deep learning algorithms are constructed with connected layers.

- The first layer is called the Input Layer
- The last layer is called the Output Layer
- All layers in between are called Hidden Layers. The word deep means the network join neurons in more than two layers.

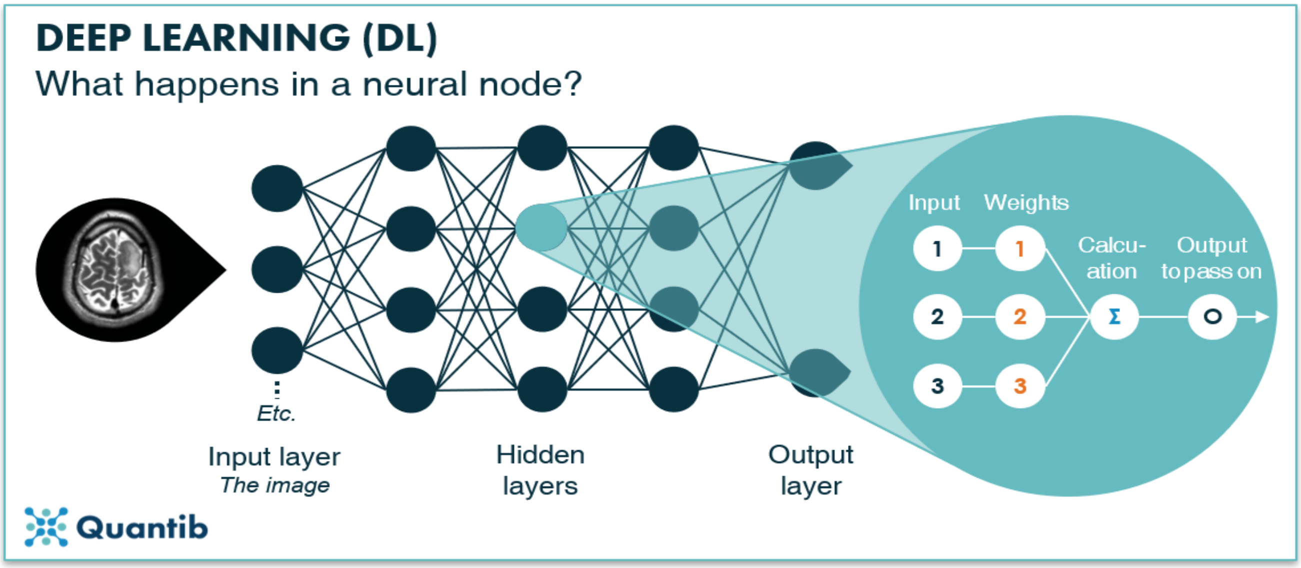

Figure 1.1 Deep Learning Layers

Each Hidden layer is composed of neurons. The neurons are connected to each other. The neuron will process and then propagate the input signal it receives the layer above it. The strength of the signal given the neuron in the next layer depends on the weight, bias and activation function.

The network consumes large amounts of input data and operates them through multiple layers; the network can learn increasingly complex features of the data at each layer.

1.2.1 Importance of Deep Learning:

Deep learning is a powerful tool to make prediction an actionable result. Deep learning excels in pattern discovery (unsupervised learning) and knowledge-based prediction. Big data is the fuel for deep learning. When both are combined, an organization can reap unprecedented results in term of productivity, sales, management, and innovation.

Deep learning can outperform traditional method. For instance, deep learning algorithms are 41% more accurate than machine learning algorithm in image classification, 27 % more accurate in facial recognition and 25% in voice recognition.

1.2.2 Deep Learning Process:

A deep neural network provides state-of-the-art accuracy in many tasks, from object detection to speech recognition. They can learn automatically, without predefined knowledge explicitly coded by the programmers.

Figure 1.2 Deep Learning Process

To grasp the idea of deep learning, imagine a family, with an infant and parents. The toddler points objects with his little finger and always says the word 'cat.' As its parents are concerned about his education, they keep telling him 'Yes, that is a cat' or 'No, that is not a cat.' The infant persists in pointing objects but becomes more accurate with 'cats.' The little kid, deep down, does not know why he can say it is a cat or not. He has just learned how to hierarchies complex features coming up with a cat by looking at the pet overall and continue to focus on details such as the tails or the nose before to make up his mind.

A neural network works quite the same. Each layer represents a deeper level of knowledge, i.e., the hierarchy of knowledge. A neural network with four layers will learn more complex feature than with that with two layers.

The learning occurs in two phases.

- The first phase consists of applying a nonlinear transformation of the input and create a statistical model as output.
- The second phase aims at improving the model with a mathematical method known as derivative.

The neural network repeats these two phases hundreds to thousands of time until it has reached a tolerable level of accuracy. The repeat of this two-phase is called an iteration.

1.2.3 Classification Of Neural Networks:

Shallow neural network: The Shallow neural network has only one hidden layer between the input and output.

Deep neural network: Deep neural networks have more than one layer. For instance, Google LeNet model for image recognition counts 22 layers.

Nowadays, deep learning is used in many ways like a driverless car, mobile phone, Google Search Engine, Fraud detection, TV, and so on.

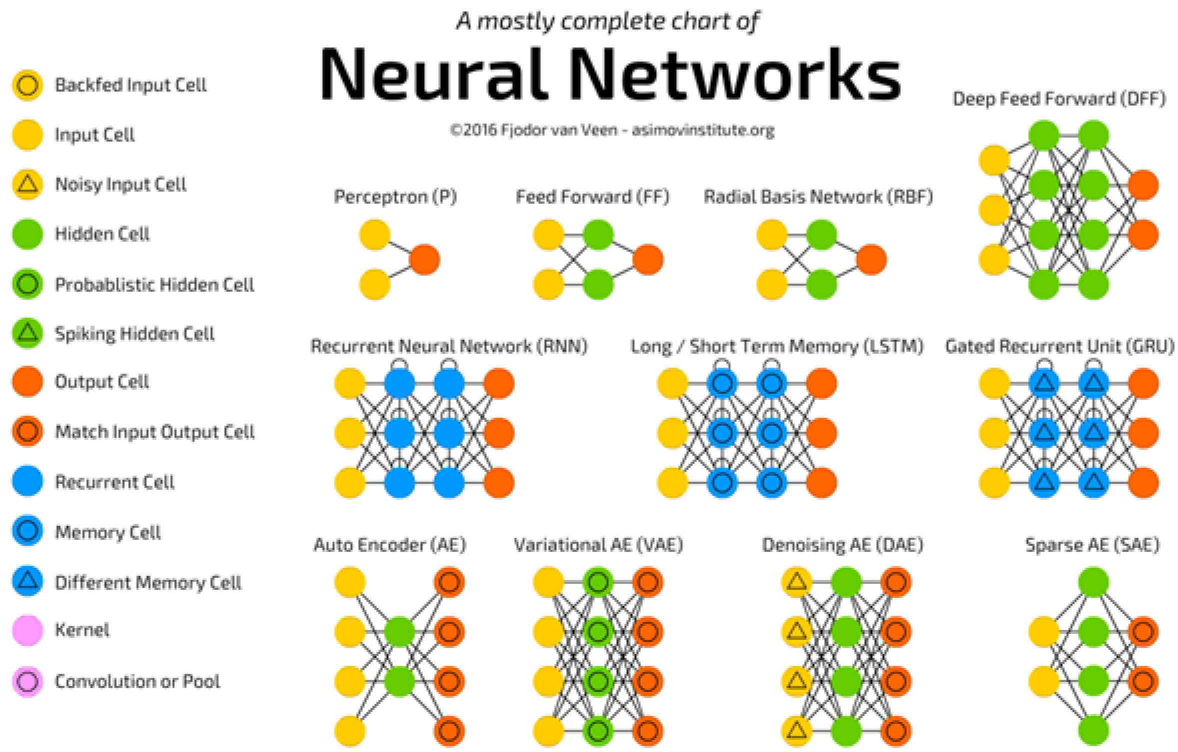

Figure 1.3 Types of Deep Learning Networks

1.2.4 Feed-Forward Neural Networks:

The simplest type of artificial neural network. With this type of architecture, information flows in only one direction, forward. It means, the information's flows starts at the input layer, goes to the "hidden" layers, and end at the output layer. The network does not have a loop. Information stops at the output layers.

1.2.5 Recurrent Neural Networks (RNNs):

RNN is a multi-layered neural network that can store information in context nodes, allowing it to learn data sequences and output a number or another sequence. In simple words it an Artificial neural networks whose connections between neurons include loops. RNNs are well suited for processing sequences of inputs.

- The RNN neurons will receive a signal that point to the start of the sentence.
- The network receives the word "Do" as an input and produces a vector of the number. This vector is fed back to the neuron to provide a memory to the network. This stage helps the network to remember it received "Do" and it received it in the first position.
- The network will similarly proceed to the next words. It takes the word "you" and "want." The state of the neurons is updated upon receiving each word.
- The final stage occurs after receiving the word "a." The neural network will provide a probability for each English word that can be used to complete the sentence. A well-trained RNN probably assigns a high probability to "café," "drink," "burger," etc.

1.2.6 Common uses of RNN:

- Help securities traders to generate analytic reports
- Detect abnormalities in the contract of financial statement
- Detect fraudulent credit-card transaction
- Provide a caption for images
- Power chatbots
- The standard uses of RNN occur when the practitioners are working with time-series data or sequences (e.g., audio recordings or text).

1.2.7 Convolutional Neural Networks (CNN):

CNN is a multi-layered neural network with a unique architecture designed to extract increasingly complex features of the data at each layer to determine the output. CNN's are well suited for perceptual tasks.

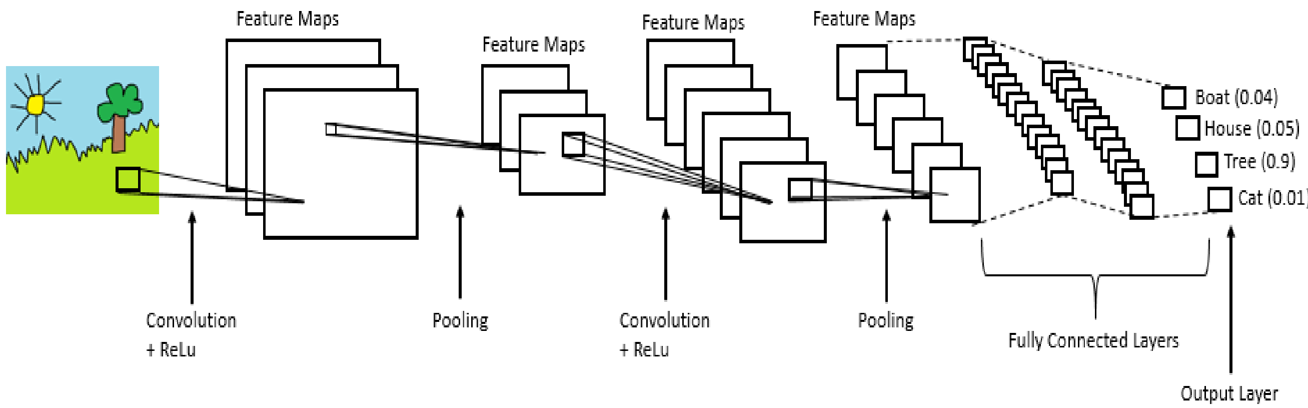

Figure 1.4 CNN

CNN is mostly used when there is an unstructured data set (e.g., images) and the practitioners need to extract information from it

For instance, if the task is to predict an image caption:

- The CNN receives an image of let's say a cat, this image, in computer term, is a collection of the pixel. Generally, one layer for the greyscale picture and three layers for a color picture.
- During the feature learning (i.e., hidden layers), the network will identify unique features, for instance, the tail of the cat, the ear, etc.
- When the network thoroughly learned how to recognize a picture, it can provide a probability for each image it knows. The label with the highest probability will become the prediction of the network.

A Convolutional Neural Network (**CNN**, or **ConvNet**) are a special kind of multi-layer neural networks, designed to recognize visual patterns directly from pixel images with minimal preprocessing.. The **ImageNet** project is a large visual database designed for use in visual object recognition software research. The ImageNet project runs an annual software contest, the **ImageNet Large Scale Visual Recognition Challenge (ILSVRC)**, where software programs compete to correctly classify and detect objects and scenes. Here I will talk about CNN architectures of ILSVRC top competitors.

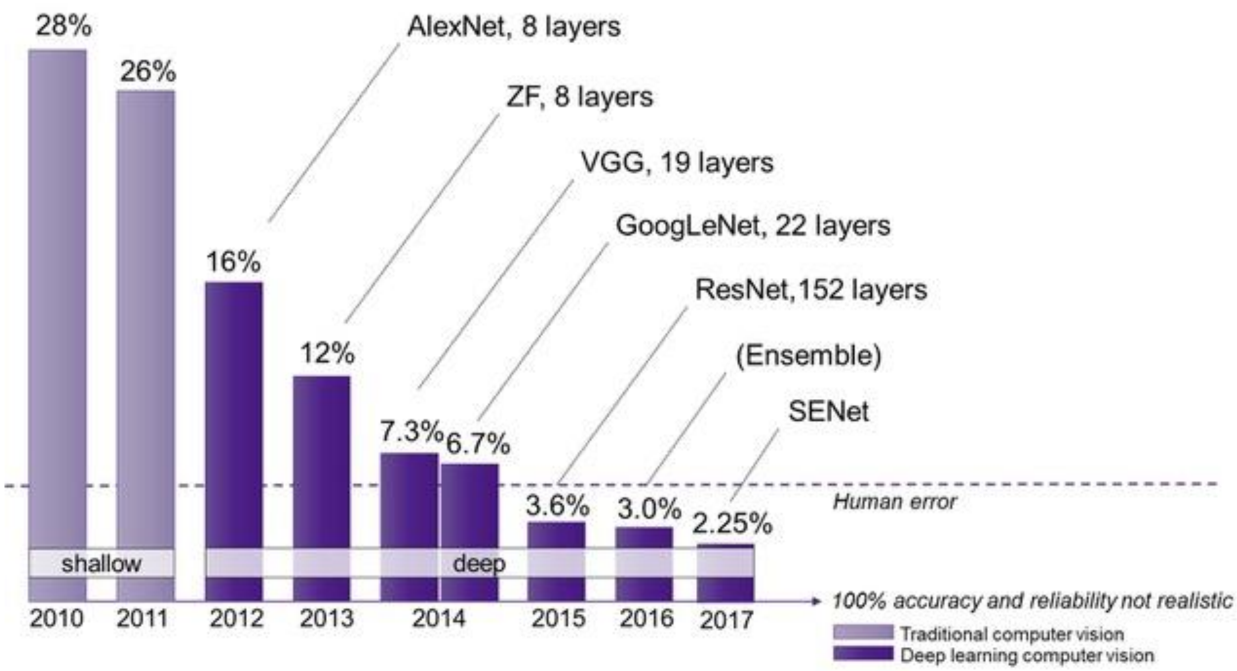

Figure 1.5 Types of CNN

1.3 EXISTING SYSTEM:

Data representation learning is one of the most important problems in machine learning. Unsupervised representation learning becomes meritorious as it has no necessity of label information with observed data. Due to the highly time consuming learning of deep-learning models, there are many machine-learning models directly adapting well-trained deep models that are obtained in a supervised and end-to-end manner as feature abstractors to distinct problems. However, it is obvious that different machine-learning tasks require disparate representation of original input data. Taking human action recognition as an example, it is well known that human actions in a video sequence are 3-D signals containing both visual appearance and motion dynamics of humans and objects. Therefore, the data representation approaches with the capabilities to capture both spatial and temporal correlations in videos are meaningful. Most of the existing human motion recognition models build classifiers based on deep-learning structures such as deep convolutional networks. These models require a large quantity of training videos with annotations. Meanwhile, these supervised models cannot recognize samples from the distinct dataset without retraining. In this article, we propose a new 3-D deconvolutional network (3DDN) for representation learning of high-dimensional video data, in which the high-level features are obtained through the optimization approach. The proposed 3DDN decomposes the video frames into spatiotemporal features under a sparse constraint in an unsupervised way. In addition, it also can be regarded as a building block to develop deep architectures by stacking. The high-level representation of input sequential data can be used in multiple downstream machine-learning tasks, we evaluate the proposed 3DDN and its deep models in human action recognition. The experimental results from three datasets: 1) KTH data; 2) HMDB-51; and 3) UCF-101, demonstrate that the proposed 3DDN is an alternative approach to feedforward convolutional neural networks (CNNs), that attains comparable results.

1.4 DISADVANTAGES OF EXISTING SYSTEM:

- But the disadvantage is that the 3DDN models currently provide low- and middle-level features of video clips, when they only have fewer hidden layers.
- The hyperparameters of the 3DDN models are determined by trial and error, including the number of hidden layers and feature maps will lead to improper results.

CHAPTER - 2

LITERATURE REVIEW

2.1 INTRODUCTION:

The following shows survey did for home surveillance. The most popular of the existing techniques is been discussed as follows.

2.2 LITERATURE SURVEY:

2.2.1 Chun-Yang Zhang , Yong-Yi Xiao, Jin-Cheng Lin, C. L. Philip Chen , Wenxi Liu , and Yu-Hong Tong, "3-D Deconvolutional Networks for the Unsupervised Representation Learning of Human Motions ", 2020.

Data representation learning is one of the most important problems in machine learning. Unsupervised representation learning becomes meritorious as it has no necessity of label information with observed data. Due to the highly time consuming learning of deep-learning models, there are many machine-learning models directly adapting well-trained deep models that are obtained in a supervised and end-to-end manner as feature abstractors to distinct problems. However, it is obvious that different machine-learning tasks require disparate representation of original input data. Taking human action recognition as an example, it is well known that human actions in a video sequence are 3-D signals containing both visual appearance and motion dynamics of humans and objects. Therefore, the data representation approaches with the capabilities to capture both spatial and temporal correlations in videos are meaningful. Most of the existing human motion recognition models build classifiers based on deep-learning structures such as deep convolutional networks. These models require a large quantity of training videos with annotations. Meanwhile, these supervised models cannot recognize samples from the distinct dataset without retraining. In this article, a new 3-D deconvolutional network (3DDN) is proposed for representation learning of high-dimensional video data, in which the high-level features are obtained through the optimization approach. The proposed 3DDN decomposes the video frames into spatiotemporal features under a sparse constraint in an unsupervised way. In addition, it also can be regarded as a building block to develop deep architectures by stacking. The high-level representation of input sequential data can be used in multiple downstream machine-learning tasks, we evaluate the proposed 3DDN and its deep models in human action recognition. The experimental results from three datasets: 1) KTH data; 2) HMDB-51; and 3) UCF-101, demonstrate that the proposed 3DDN is an alternative approach to feedforward convolutional neural networks (CNNs), that attains comparable results.

Pros:

The experimental results from three datasets: 1) KTH data; 2) HMDB-51; and 3) UCF-101, demonstrate that the proposed 3DDN is an alternative approach to feedforward convolutional neural networks (CNNs).

Cons:

The hyperparameters of the 3DDN models are determined by trial and error, including the number of hidden layers and feature maps will lead to improper results.

2.2.2 Geong Sen Poh, Prosanta Gope, and Jianting Ning, "PrivHome: Privacy-Preserving Authenticated Communication in Smart Home Environment", vol. 99, no. 99, may 2019.

A privacy-preserving scheme, PrivHome. It supports authentication, secure data storage and query for smart home systems. PrivHome provides data confidentiality as well as entity and data authentication to prevent an outsider from learning or modifying the data communicated between the devices, service provider, gateway, and the user. It further provides privacy-preserving queries in such a way that the service provider, and the gateway does not learn content of the data. To the best of our knowledge, privacy-preserving queries for smart home systems has not been considered before. Under this scheme is a new, lightweight entity and key-exchange protocol, and an efficient searchable encryption protocol. Our scheme is practical as both protocols are based solely on symmetric cryptographic techniques.

Pros:

The light field face artefact database have revealed the outstanding performance of the proposed PAD scheme when benchmarked with various well established state-of-the-art schemes.

Cons:

There exists no superior PAD technique due to evolution of sophisticated presentation attacks.

2.2.3 Chuanwei Ding ,Hong Hong , Yu Zou, Hui Chu , Xiaohua Zhu, Francesco Fioranelli ,Julien Le Kernec and Changzhi Li , "Continuous Human Motion Recognition With a Dynamic Range-Doppler Trajectory Method Based on FMCW Radar", vol. 57, no. 9, September 2019.

Radar-based human motion recognition is crucial for many applications, such as surveillance, search and res- cue operations, smart homes, and assisted living. Continuous human motion recognition in real-living environment is necessary for practical deployment, i.e., classification of a sequence of activities transitioning one into another, rather than individual activities. A novel dynamic range-Doppler trajectory (DRDT) method based on the frequency-modulated continuous-wave (FMCW) radar system is proposed to recognize continuous human motions with various conditions emulating real-living environment. The DRDT is extracted from these frames to monitor human motions in time, range, and Doppler domains in real time.

Pros:

A smart home enables users to access devices such as lighting, HVAC, temperature sensors, and surveillance camera. It provides a more convenient and safe living environment for users

Cons:

The system produces results very poor in accuracy.

2.2.4 Xueru Bai , Ye Hui, Li Wang, and Feng Zhou ,"Radar-Based Human Gait Recognition Using Dual-Channel Deep Convolutional Neural Network", Volume: 57, Issue: 12, Dec. 2019.

The problem of radar-based human gait recognition based on the dual-channel deep convolutional neural network (dc-dcnn). To enrich the limited radar data set of human gaits and provide a benchmark for classifier training, evaluation, and comparison, it proposes an effective method for radar echo generation from the infrared, publicly accessible motion capture (mocap) data set. According to the different nonstationary characteristics of micro-doppler (m-d) for the torso and limbs, it enhances their distinguishable joint time–frequency (jtf) features by applying the short-time fourier transforms (sftfs) with varying sliding window length and then designs the dc-dcnn structure to achieve refined human gait recognition by separate feature extraction and fusion. An effective radar echo generation method of human gaits from the infrared mocap data and designed the dc-dcnn for refined human gait recognition. In particular, the proposed network could extract enhanced jtf features of both torso and limbs separately and achieve effective feature fusion.

Pros:

The DRDT is extracted from these frames to monitor human motions in time, range, and Doppler domains in real time.

Cons:

The system produces results very poor in accuracy and cannot be used for real time implementation.

2.2.5 Yaxu Xue, Zhaojie Ju, Kui Xiang, Jing Chen, and Honghai Liu, "Multimodal Human Hand Motion Sensing and Analysis - A Review", Volume: 11, Issue: 2, June 2019.

Human hand motion analysis is an essential research topic in recent applications, especially for dexterous robot hand manipulation learning from human hand skills. It provides important information about the gestures, tactile, speed and contact force, captured via multiple sensing technologies. This paper introduces a comprehensive survey of current hand motion sensing technologies and analysis approaches in recent emerging applications. Firstly, the nature of human hand motions is discussed in terms of simple motions, such as grasps and gestures, and complex motions, e.g. in-hand manipulations and re-grasps; secondly, different techniques for hand motion sensing, including contact-based and non-contact-based approaches, are discussed with comparisons with their pros and cons; then, the state-of-the- art analysis methods are introduced, with a particular focus on the multimodal hand motion sensing and analysis; finally, cutting- edge applications of hand motion analysis are reviewed, with further discussion on facing challenges and new future directions. HHM analysis is attracting broad interest in robotics. Two types of the natural HHMs are proposed through the analysis of the current hand motion strategies. In order to realize the human-robot manipulation skill transfer, various sensing techniques in the last decade were used to acquire the information of HHMs, such as the dynamic movement trajectory of finger joints, and the dynamic distribution of finger force.

Pros:

Seam carved images from the same camera, source attribution can still be possible even if the size of uncarved blocks in the image is less than the recommended size of 50×50 pixels.

Cons:

The system is too slow for processing and is not recommended for real time implementation.

2.2.6 Sid Ahmed Walid Talha1, Mounir Hammouche1,2, Enjie Ghorbel1,3, Anthony Fleury*1, Sébastien Ambellouis1,4, "Features and classification schemes for view-invariant and real-time human action recognition", 2018.

Human Action recognition (HAR) is largely used in the field of Ambient Assisted Living (AAL) to create an interaction between humans and computers. In these applica- tions, it cannot be asked to people to act non-naturally. The algorithm has to adapt and the interaction has to be as quick as possible to make this interaction fluent. To improve the existing algorithms with regards to that points, we propose a novel method based on skeleton information provided by RGB-D cameras. This approach is able to carry out early action recognition and is more robust to viewpoint variability. To reach this goal, a new descriptor called Body Directional Velocity is proposed and a real-time classification is performed. We have proposed a novel approach to perform human action recognition using RGB-D sensors. The work is focused on two major challenges for robotic applications, the robustness to viewpoint changes and the early recognition property.

Pros:

Seam carved images from the same camera, source attribution can still be possible even if the size of uncarved blocks in the image is less than the recommended size of 50×50 pixels.

Cons:

The system is too slow for processing and is not recommended for real time implementation.

2.2.7 Samet Taspinar, Manoranjan Mohanty, and Nasir Memon, "PRNU-Based Camera Attribution from Multiple Seam-Carved Images" 2017, VOL. 20, NO. 5.

Photo Response Non-Uniformity (PRNU) noise based source attribution is a well- known technique to verify the camera of an image or video. Researchers have proposed various countermeasures to prevent PRNU-based source camera attribution. Forced seam-carving is one such recently proposed counter forensics technique. This technique can disable PRNUbased source camera attribution by forcefully removing seams such that the size of most uncarved image blocks is less than 50×50 pixels. In this paper, given multiple seam carved images are shown from the same camera, source attribution can still be possible even if the size of uncarved blocks in the image is less than the recommended size of 50×50 pixels. Theoretical analysis and experiments with multiple cameras demonstrate that the effectiveness of our scheme depends on the number of seams carved from an image and the randomness of the seam positions.

Pros:

Theoretical analysis and experiments with multiple cameras demonstrate that the effectiveness of our scheme depends on the number of seams carved from an image and the randomness of the seam positions.

Cons:

The size of the images must be less than 50x50 pixels.

2.2.8 Sujin Jang, Niklas Elmqvist, and Karthik Ramani, "MotionFlow: Visual Abstraction and Aggregation of Sequential Patterns in Human Motion Tracking Data", vol. 22, no. 1, january 2016.

Pattern analysis of human motions, which is useful in many research areas, requires understanding and comparison of different styles of motion patterns. However, working with human motion tracking data to support such analysis poses great challenges. In this paper, MotionFlow, a visual analytics system that provides an effective overview of various motion patterns based on an interactive flow visualization. This visualization formulates a motion sequence as transitions between static poses, and aggregates these sequences into a tree diagram to construct a set of motion patterns. The system also allows the users to directly reflect the context of data and their perception of pose similarities in generating representative pose states. Local and global controls are provided over the partition-based clustering process. To support the users in organizing unstructured motion data into pattern groups, we designed a set of interactions that enables searching for similar motion sequences from the data, detailed exploration of data subsets, and creating and modifying the group of motion patterns. To evaluate the usability of MotionFlow, we conducted a user study with six researchers with expertise in gesture-based interaction design. They used MotionFlow to explore and organize unstructured motion tracking data.

Pros:

The system used Motion Flow to explore and organize unstructured motion tracking data.

Cons:

The system is not very accurate and is not reliable.

2.2.9 Dinesh K Vishwakarma, Kuldeep Singh, "Human Activity Recognition based on Spatial Distribution of Gradients at Sub-levels of Average Energy Silhouette Images", Volume: 9, Issue: 4, Dec. 2017.

A unified framework for human action and activity recognition by analysing the effect of computation of spatial distribution of gradients (sdgs) on average energy silhouette images (aesis). Based on the analysis of sdgs computation at various decomposition levels, an effective approach to compute the sdgs is developed. The aesi is constructed for the representation of the shape of action and activity and these are the reflection of 3d pose into 2d pose. To describe the aesis, the sdgs at various sub-levels and sum of the directional pixels (sdps) variations is computed. The temporal content of the activity is computed through r–transform (rt). Finally, the shape computed through sdgs and sdps, and temporal evidences through rt of the human body is fused together at the recognition stage, which results in a new powerful unified feature map model. The performance of the proposed framework is evaluated on three different publicly available datasets i.e. Weizmann, kth, and ballet and the recognition accuracy is computed using hybrid classifier. The highest recognition accuracy achieved on these data sets is compared with the similar state-of-the- art techniques and demonstrate the superior performance. A human action recognition approach using shape and motion features of the human silhouette in the video sequence is presented, which addresses the problem of less recognition rate under challenging environmental conditions and complex motion pattern. The shape information is computed through modified sdgs and the sdps of aesi.

Pros:

In the video sequence is presented, which addresses the problem of less recognition rate under challenging environmental conditions and complex motion pattern.

Cons:

The sdgs computation is too slow and takes too much time.

2.2.10 Ying-Tsung Lee, Wei-Hsuan Hsiao, Chin-Meng Huang and Seng-Cho T. Chou, "An Integrated Cloud-Based Smart Home Management System with Community Hierarchy", 2016, Vol No: 2162-237X.

A smart home management system in which a community broker role is used for integrating community services, thereby reducing the workload of community management staff, providing electronic information services, and deepening the community's integration with the surrounding environment. At the home end, a home intranet was created by integrating a fixed touch panel with a home controller system and various sensors and devices to deliver, for example, energy, scenario information, and security functions. The community end comprises a community server and community personal computers, and connects to devices (e.g., video cameras and building automation devices) in other community systems and to the home networks. Furthermore, to achieve multiple in home displays, standard interface devices can be employed to separate the logic and user interfaces. This study also determined that the message queuing telemetry transport protocol can provide optimal home control services in smart home systems, whereas hypertext transfer protocol is optimal for delivering location-based information integration services.

Pros:

This study also determined that the message queuing telemetry transport protocol can provide optimal home control services in smart home systems, whereas hypertext transfer protocol is optimal for delivering location-based information integration services.

Cons:

The system is too expensive to implement in real time.

2.2.11 Jinzhu Chen, Rui Tan, Guoliang Xing, , Xiaorui Wang, and Xing Fu, ”Fidelity-Aware Utilization Control for Cyber-Physical Surveillance Systems”, 2012.

Recent years have seen the growing deployments of Cyber-Physical Systems (CPSs) in many mission-critical applications such as security, civil infrastructure, and transportation. These applications often impose stringent requirements on system sensing fidelity and timeliness. However, existing approaches treat these two concerns in isolation and hence are not suitable for CPSs where system fidelity and timeliness are dependent on each other because of the tight integration of computational and physical resources. In this paper, we propose a holistic approach called Fidelity-Aware Utilization Controller (FAUC) for Wireless Cyber-physical Surveillance (WCS) systems that combine low-end sensors with cameras for large-scale ad hoc surveillance in unplanned environments. By integrating data fusion with feedback control, FAUC can enforce a CPU utilization upper bound to ensure the system's real-time schedulability although CPU workloads vary significantly at runtime because of stochastic detection results. At the same time, FAUC optimizes system fidelity and adjusts the control objective of CPU utilization adaptively in the presence of variations of target/noise characteristics. We have implemented FAUC on a small-scale WCS testbed consisting of TelosB/Iris motes and cameras. Moreover, extensive simulations are conducted based on real acoustic data traces collected in a vehicle surveillance experiment. The testbed experiments and the trace-driven simulations show that FAUC can achieve robust fidelity and real-time guarantees in dynamic environments.

Pros:

The testbed experiments and the trace-driven simulations show that FAUC can achieve robust fidelity and real-time guarantees in dynamic environments.

Cons:

These applications often impose stringent requirements on system sensing fidelity and timeliness

2.2.12 Regio A. Michelin, Nadeem Ahmed, Salil S. Kanhere, Aruna Seneviratne and Sanjay Jha,” Leveraging lightweight blockchain to establish data integrity for surveillance cameras”, 2019.

The video footage produced by surveillance cameras is important evidence to support criminal investigations. Video evidence can be sourced from public (trusted) as well as private (untrusted) surveillance systems. This raises the issue of establishing integrity for information provided by the untrusted video sources. In this paper, we present a framework to ensure the data integrity of the stored videos, allowing authorities to validate whether video footage has not been tampered with. The proposal uses a lightweight blockchain technology to store the video metadata as blockchain transactions to support the validation of video integrity. The evaluations show that the overhead introduced by employing the blockchain to create the transactions introduces a minor latency of a few milliseconds.

Pros:

The evaluations show that the overhead introduced by employing the blockchain to create the transactions introduces a minor latency of a few milliseconds.

Cons:

The system uses video for training raising the issue of establishing integrity for information provided by the untrusted video sources.

2.2.13 Amir H. Meghdadi, and Pourang Irani,” Interactive Exploration of Surveillance Video through Action Shot Summarization and Trajectory Visualization”, Volume: 19, Issue: 12, Dec. 2013.

A novel video visual analytics system is proposed for interactive exploration of surveillance video data. The approach consists of providing analysts with various views of information related to moving objects in a video. To do this we first extract each object’s movement path. Each movement is visualized by (a) creating a single action shot image (a still image that coalesces multiple frames), (b) plotting its trajectory in a space-time cube and (c) displaying an overall timeline view of all the movements. The action shots provide a still view of the moving object while the path view presents movement properties such as speed and location. Tools for spatial and temporal filtering are also provided based on regions of interest. This allows analysts to filter out large amounts of movement activities while the action shot representation summarizes the content of each movement. We incorporated this multi-part visual representation of moving objects in sViSIT, a tool to facilitate browsing through the video content by interactive querying and retrieval of data. Based on the interaction with security personnel who routinely interact with surveillance video data, we identified some of the most common tasks performed. This resulted in designing a user study to measure time-to-completion of the various tasks. These generally required searching for specific events of interest (targets) in videos. Fourteen different tasks were designed and a total of 120 min of surveillance video were recorded (indoor and outdoor locations recording movements of people and vehicles). The time-to-completion of these tasks were compared against a manual fast forward video browsing guided with movement detection. The demonstration is how our system can facilitate lengthy video exploration and significantly reduce browsing time to find events of interest. Reports from expert users identify positive aspects of the approach which is summarized in the recommendations for future video visual analytics systems.

Pros:

The evaluations show that the overhead introduced by employing the blockchain to create the transactions introduces a minor latency of a few milliseconds.

Cons:

The system uses video for training raising the issue of establishing integrity for information provided by the untrusted video sources.

2.2.14 Bin Tian, Brendan Tran Morris, Ming Tang, Yuqiang Liu, Yanjie Yao, Chao Gou, Dayong Shen, and Shaohu Tang," Hierarchical and Networked Vehicle Surveillance in ITS: A Survey", Volume: 16, Issue: 2, April 2015.

Traffic surveillance has become an important topic in intelligent transportation systems (ITSs), which is aimed at monitoring and managing traffic flow. With the progress in computer vision, video-based surveillance systems have made great advances on traffic surveillance in ITSs. However, the performance of most existing surveillance systems is susceptible to challenging complex traffic scenes (e.g., object occlusion, pose variation, and cluttered background). Moreover, existing related research is mainly on a single video sensor node, which is incapable of addressing the surveillance of traffic road networks. Accordingly, a review of the literature on the video-based vehicle surveillance systems in ITSs is presented. The existing challenges in video-based surveillance systems are analyzed for the vehicle and present a general architecture for video surveillance systems, i.e., the hierarchical and networked vehicle surveillance, to survey the different existing and potential techniques. Then, different methods are reviewed and discussed with respect to each module. Applications and future developments are discussed to provide future needs of ITS services.

Pros:

The Applications and future developments are discussed to provide future needs of ITS services.

Cons:

The performance of existing surveillance system is susceptible to challenging complex traffic scenes

2.2.15 Frédéric Dufaux, and Touradj Ebrahimi," Scrambling for Privacy Protection in Video Surveillance Systems", Volume: 18, Issue: 8, Aug. 2008.

The problem of privacy protection in video surveillance is analyzed. Two efficient approaches are introduced to conceal regions of interest (ROIs) based on transform-domain or code stream-domain scrambling. In the first technique, the sign of selected transform coefficients is pseudo randomly flipped during encoding. In the second method, some bits of the code stream are pseudo randomly inverted. The cases of MPEG-4 are addressed as it is today the prevailing standard in video surveillance equipment. Simulations show that both techniques successfully hide private data in ROIs while the scene remains comprehensible. Additionally, the amount of noise introduced by the scrambling process can be adjusted. Finally, the impact on coding efficiency performance is small, and the required computational complexity is negligible.

Pros:

The impact on coding efficiency performance is small, and the required computational complexity is negligible.

Cons:

The system is too inaccurate and is not reliable.

CHAPTER - 3

SYSTEM ANALYSIS

3.1 PROPOSED SYSTEM:

The popularity of home automation has been increasing greatly in recent years due to considerable affordability and simplicity through smartphone and tablet connectivity. To save money, homeowners sometimes decide to go with the basic package offered by a home security provider. Although you may pay less upfront costs, not covering all areas of your home can be costly in the long run. For instance, criminals may find an unprotected entry point into your home if you are only buying a front and backdoor sensor. By taking a tour around the perimeter of your home to spot any security weaknesses and order the required equipment pieces to protect all areas. Although installing more cameras may cost slightly more, you'll have the peace of mind to know your property is well protected. In this project we will arm our home with multiple Raspberry Pi that communicate among themselves to accomplish a common goal. We'll begin by building a case for security with an emphasis on the lack of flexibility of commercially available systems — that's where we bring in Raspberry Pis to the rescue. Our interoperable system will involve Raspberry Pi, cameras, a mobile application for alert by predicting the presence of unknown person. A mobile application is developed using react native to enable live streaming and options are given for the user to alert the police, lock the door or can even ignore when an unknown trespasser enters the property. Our solution is flexible, affordable, and interoperable with other IoT devices and services that are worth paying for.

DETAILED DESIGN OF THE PROJECT:

This chapter describes the overall and the detailed architectural design. It also describes each module that is to be implemented along with Data Flow diagram.

3.2 SYSTEM ARCHITECTURE:

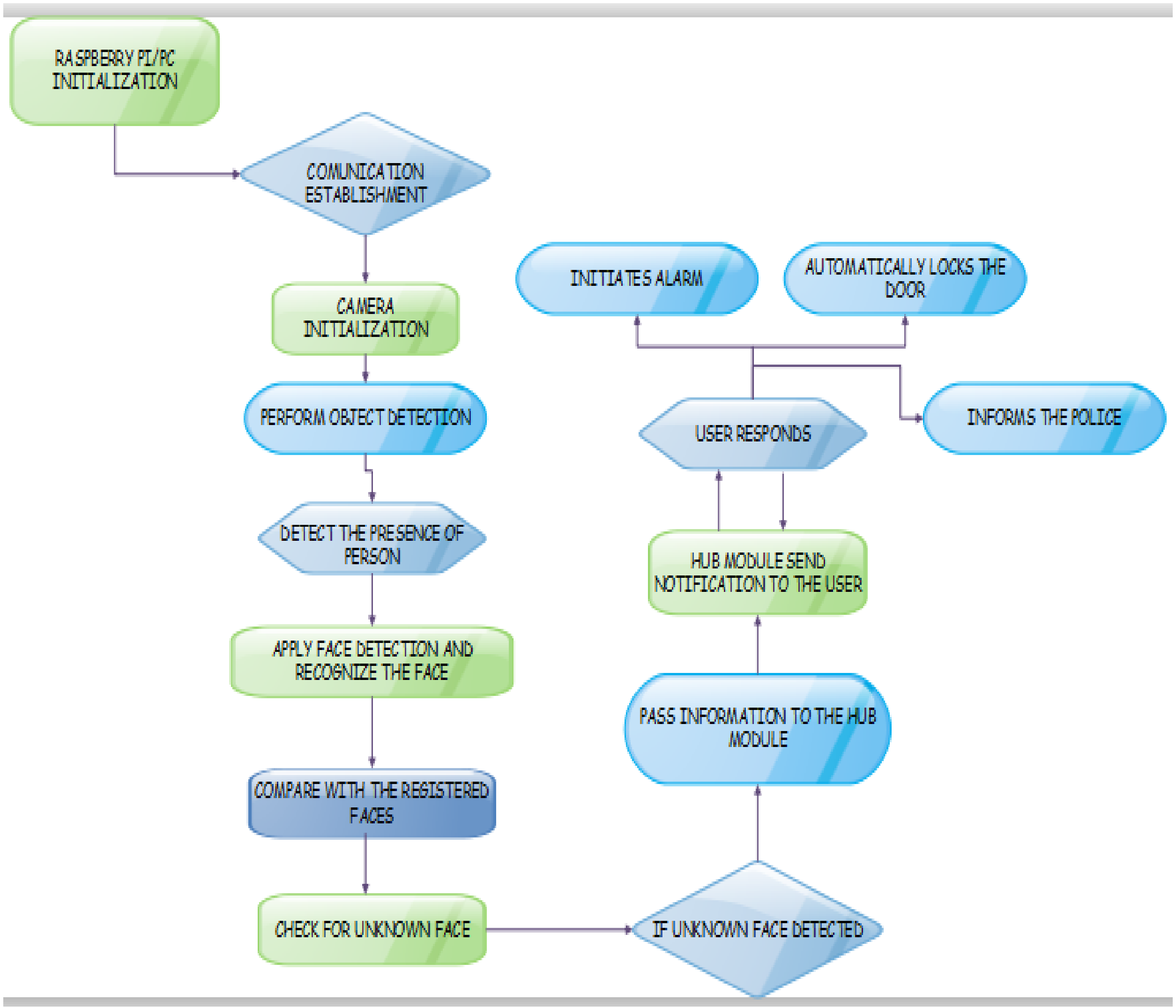

Figure 3.1: System Architecture

3.3 WORKING:

In this project we are going to build a pi to pi communication system for providing home surveillance. So here basically we have raspberry pi in all the room places where monitoring is required and the person who lives in the home can enables the system while they are leaving the home can enable the system and all pis will get activated and start communicating with each other. It automatically connect to the hub module that is main raspberry pi module by which they can share the information regarding situation at the room. If any intruded comes to their home the particular raspberry pi module will be monitoring it by using camera module and as it detects the presence of intruder it automatically communicates with the main hub module which will initiate actions to safeguard the home. As soon as the slave pi module sends an information to the main raspberry pi module it will sent an SMS alert to the home owner, he can automatically trigger the alarm through mobile application which is given to be home user where he will receive notification about the intruded in their home. In this project we will implement the concept using a raspberry pi as a slave device and a PC as a hub module. Using the mobile app they can be able to control the situation in the home by raising alarm, locking the door automatically and informing to police by which they can be able to safeguard their home. This system not only gives security to their home but also helps in identifying and catching the thief as soon as possible. The present system has only cctv cameras by which, it can only monitor are the presence of people but it cannot make an intimation the user but our system effectively makes intimation to user as well as take step against to the intruded and safeguard before the maximum incident occurs. That's we provide end to end security to the homes and it can be applicable for all places wherever security is necessary for example jewelry shops, banks, etc.

3.4 MODULE DESCRIPTION:

- Object Detection Module
- Face Detection Module
- Face Recognition Module
- Image ZMQ for Pi to Pi Communication
- Mobile App Module

3.4.1 Object Detection Module:

SSD (Single Shot Multi-Box Detector) algorithm is used in Object detection model and it is a popular algorithm in object detection. It's generally faster than Faster RCNN. A typical CNN network gradually shrinks the feature map size and increase the depth as it goes to the deeper layers. The deep layers cover larger receptive fields and construct more abstract representation, while the shallow layers cover smaller receptive fields objects and deeper layers to predict big objects, as small objects don't need bigger receptive fields and bigger receptive fields can be confusing for small objects. The following chart shows the architecture of SSD using VGG net as the base net. The middle column shows the feature map sets the net generates from different layers. For example, the first feature map set is generated from VGG net layer 23, and have a size of 38x38 and depth of 512. Every point in the 38x38 feature map covers a part of the image, and the 512 channels can be the features for every point. By using the features in the 512 channels, we can do image classification to predict the label and regression to predict the bounding box for small objects on very point. The second feature map set has a size of 19x19, which can be used for slightly larger objects, as the points of the features cover bigger receptive fields. Down to the last layer, there is only one point in the feature map set, which is ideal for big objects.

For Pascal VOC dataset, there are 21 classes (20 objects + 1 background). You have noticed there are 4x21 outputs for every feature point in the classification results.

Actually, the number 4 comes from the fact we predict 4 objects with different bounding boxes for every point. It's a common trick used in Yolo and Faster RCNN. In SSD, multiple boxes for every feature point are called priors, while in Faster RCNN they are called anchors. For every prior, we predict one bounding box for all the classes, so there are 4 values for very feature point. Beware it's different from Faster RCNN. It may lead to worse bounding box prediction due to the confusion among different classes.

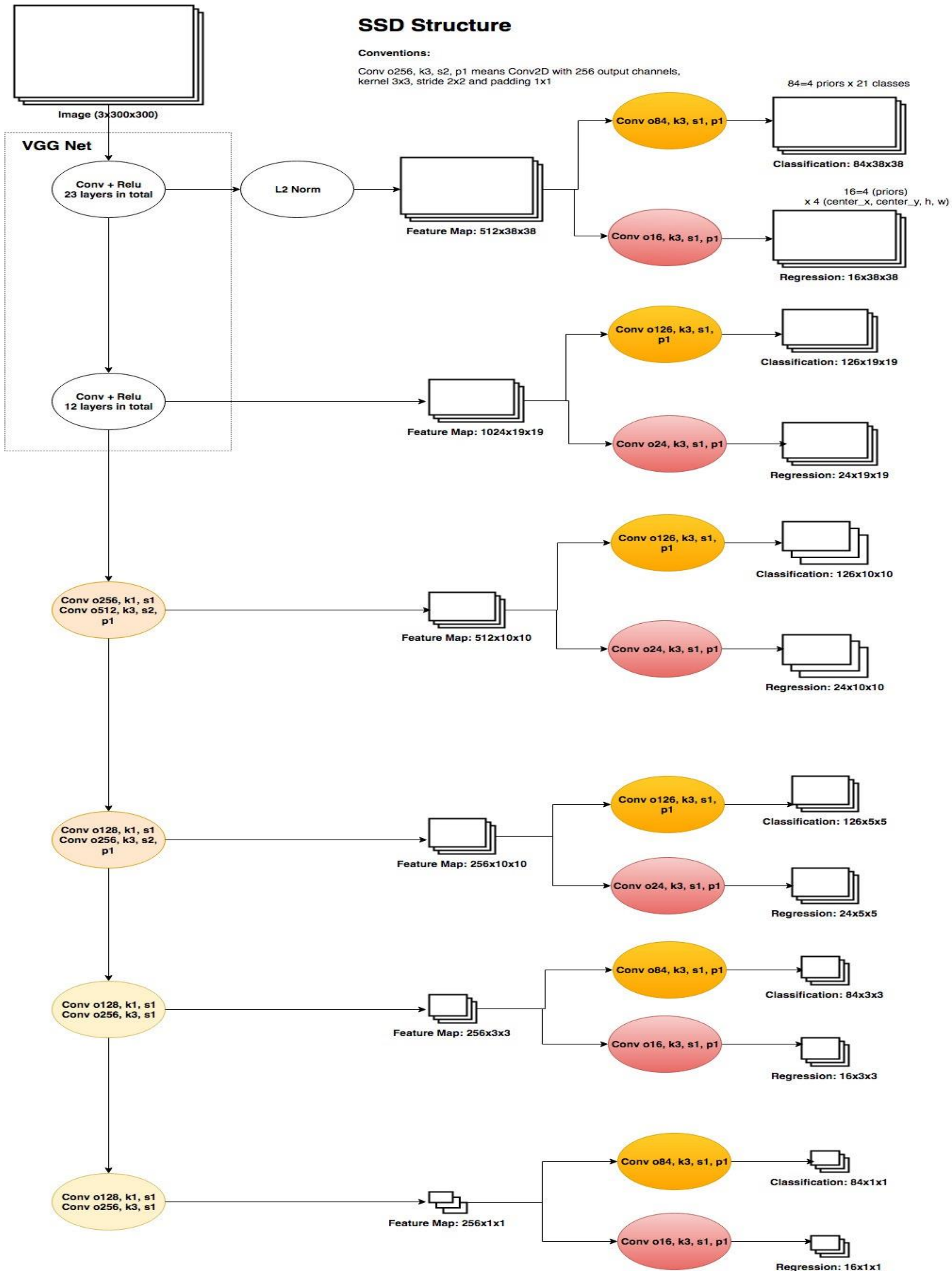

Figure 3.2: SSD Structure

VGG based SSD Architecture. (Notations: Conv o256, k3, s2, p1 means Conv2D with 256 output channels, kernel 3x3, stride 2x2 and padding 1x1. Orange represents classification box, pink represents regression head.

3.4.2 Face Detection Module:

The “Ultra-Light-Fast-Generic-Face-Detector” is designed for general-purpose face detection applications in low-power computing devices and is applicable to both Android and iOS phones as well as PCs (CPU and GPU). The model is a real-time ultra-lightweight universal face detection model designed for edge computing devices or low-power devices. It can be used in low-power computing devices such as ARM for real-time common scene faces.

Facial recognition technology is widely applied in security monitoring, surveillance, human-computer interaction, entertainment, etc. Detecting human faces in digital images is the first step in facial recognition, and an ideal face detection model can be evaluated by how quickly and accurately it performs.

The Face-Detector-1MB stands out in terms of speed — the model’s default FP32 precision (.pth) file size is 1.1MB, and the inference frame int8 is quantized to a size of 300KB. In terms of model calculation, the input resolution of 320×240 is only about 90 to 109 MFlops. The Face-Detector-1MB training process used a VOC dataset generated by the WIDER FACE dataset, a face detection benchmark. WIDER FACE was released in 2015 and consists of 32,203 images and 393,703 face bounding boxes with a high degree of variability in scale, pose, expression, occlusion and illumination.

The 1MB lightweight model comes in a version-slim with slightly faster simplification, and a version-RFB with a modified RFB module for higher precision.

The model was tested on Ubuntu16.04, Windows 10, Python3.6, Pytorch1.2, CUDA10.0, etc.

Features:

- In terms of model size, the default FP32 precision (pth) file size is 1.04~1.1MB, and the inference frame int8 is about 300KB.
- In the calculation of the model, the input resolution of 320x240 is about 90~109 M-Flops.
- There are two versions of the model, version-slim (slightly faster simplification), version-RFB (with the modified RFB module, higher precision).
- Provides pre-training models using wider face training at 320x240 and 640x480 different input resolutions to better work in different application scenarios.
- Support for onxx export, easy to transplant.

3.4.3 Face Recognition Module:

For recognizing the face we use mobile facenet which is a more accurate in classifying face. MobileFaceNet is a neural network and obtains accuracy upto 99.28 percent on labelled faces in the wild (LFW) dataset, and a 93.05 percent accuracy on recognising faces in the AgeDB dataset. The network used around a million parameters taking only 24 milliseconds to run and produce results on a Qualcomm Snapdragon processor. We can compare this performance to accuracies of 98.70 percent and 89.27 percent for ShuffleNet, which has many more parameters and takes a little longer to execute on the CPU.

The researchers have made it easy to replace the global average pooling layer in the CNN with a depthwise convolution layer, which improves performance on facial recognition. This development is really important as the artificial intelligence world searches for efficient models that run on small compute powers which are available on today's mobile phones.

Another approach for obtaining lightweight facial verification models is by compressing pretrained networks by knowledge distillation. Such approaches have achieved 97.32 percent facial verification accuracy on LFW with 4.0 MB model size. The remarkable achievement is that MobileFaceNets achieves comparable accuracy with very small budget.

MobileFaceNet Architectures:

MobileFaceNet architecture is partly inspired by the MobileNetV2 architecture. The residual bottlenecks proposed in MobileNetV2 are used as our main building blocks. The researchers use :

Input	Operator	t	c	n	s
$112^2 \times 3$	conv3x3	-	64	1	2
$56^2 \times 64$	depthwise conv3x3	-	64	1	1
$56^2 \times 64$	bottleneck	2	64	5	2
$28^2 \times 64$	bottleneck	4	128	1	2
$14^2 \times 128$	bottleneck	2	128	6	1
$14^2 \times 128$	bottleneck	4	128	1	2
$7^2 \times 128$	bottleneck	2	128	2	1
$7^2 \times 128$	conv1x1	-	512	1	1
$7^2 \times 512$	linear GDConv7x7	-	512	1	1
$1^2 \times 512$	linear conv1x1	-	128	1	1

Figure 3.3: MobileFaceNet architecture

The primary MobileFaceNet network uses 0.99 million parameters. To reduce computational cost, the researchers decided to change input resolution from 112×112 to 112×96 or 96×96. The the linear 1×1 convolution layer after the linear GDConv layer was also removed from MobileFaceNet. This gives a resulting network called MobileFaceNet-M.

The researchers have used MobileNetV1, ShuffleNet, and MobileNetV2 as the baseline models. All MobileFaceNet models and baseline models are trained on CASIA-Webface dataset from scratch by ArcFace loss, for a fair performance comparison among them. The training is finished at 60K iterations.

To pursue further excellent performance, MobileFaceNet, MobileFaceNet (112×96), and MobileFaceNet (96×96) are also trained on the cleaned training set of MSCeleb-1M database with 3.8 million images from 85,000 subjects. The accuracy of our primary MobileFaceNet is boosted to 99.55 percent and 96.07 percent on LFW and AgeDB-30, respectively.

3.4.4 Image Zmq For Pi To Pi Communication:

Imagezmq is an easy to use image transport mechanism for a distributed image processing network. For example, a network of a dozen Raspberry Pis with cameras can send images to a more powerful central computer. The Raspberry P is perform image capture and simple image processing like flipping, blurring and motion detection. Then the images are passed via **imagezmq** to the central computer for more complex image processing like image tagging, text extraction, feature recognition, etc.

Features

- Sends OpenCV images from one computer to another using ZMQ.
- Can send jpeg compressed OpenCV images, to lighten network loads.
- Uses the powerful ZMQ messaging library through PyZMQ bindings.
- Allows a choice of 2 different ZMQ messaging patterns (REQ/REP or PUB/SUB).
- Enables the image hub to receive and process images from multiple image senders simultaneously.

ZMQ allows many different messaging patterns. Two are implemented in imagezmq:

- REQ/REP: Each RPi sends an image and waits for a REPLY from the central image hub. The RPi sends a new image only when the REPLY is received. In the REQ/REP messaging pattern, each image sender must await a REPLY before continuing. It is a "blocking" pattern for the sender.
- PUB/SUB: Each RPi sends an image, but does not expect a REPLY from the central image hub. It can continue sending images without awaiting any acknowledgement from the image hub. The image hub provides no REPLY. It is a "non-blocking" pattern for the sender

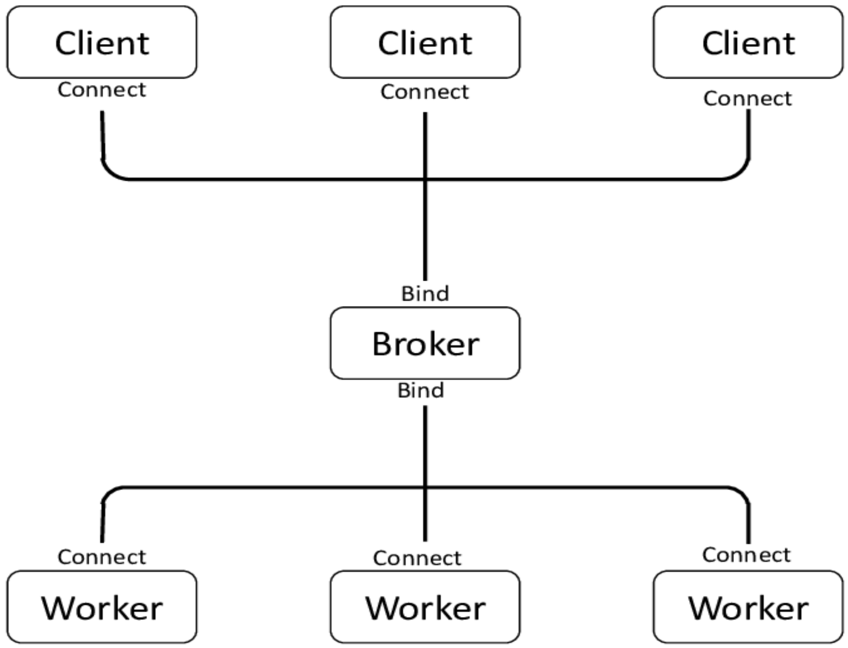

Figure 3.4: Overview architecture of ZMQ

3.4.5 Mobile App Module

In this project react native is used for mobile app development.

React Native is a framework that builds a hierarchy of UI components to build the JavaScript code. It has a set of components for both iOS and Android platforms to build a mobile application with a native look and feel. Mobile development has witnessed unprecedented growth. According to statistics, mobile applications will generate an estimated 188 billion U.S. dollars in revenue via app stores, advertising and in-app purchases by the year 2020. Single and business users require high-standard apps with flawless performance, multiple screens, easy navigation and good design. On the other side, high-performing, good quality native apps are very time-consuming to develop compared to cross-platform apps that provide faster development but compromise on performance and support. React Native seems to be a viable solution for building high-quality apps in a short time with the same performance and user-experience standards that native apps provide.

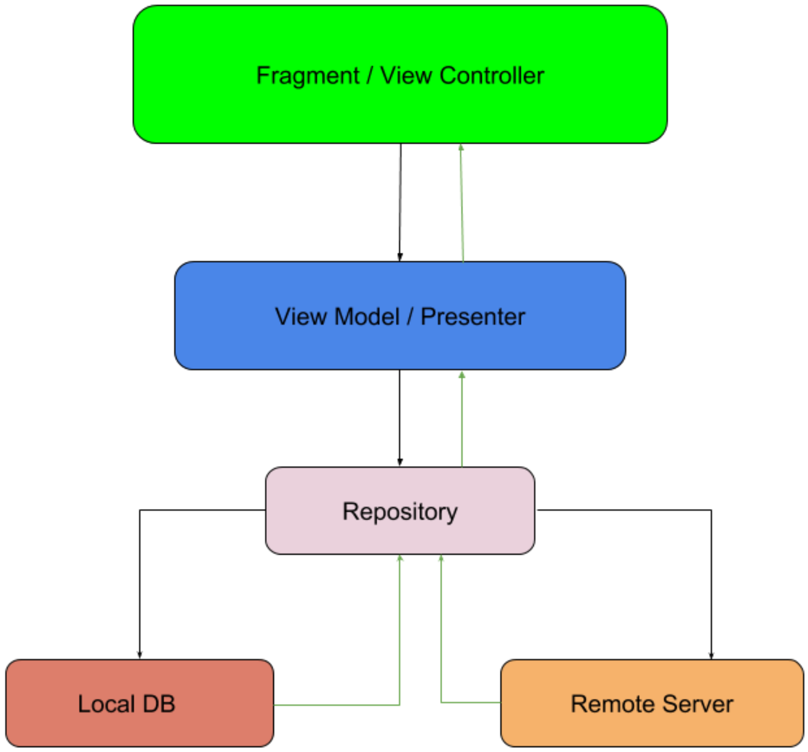

Figure 3.5 Mobile App Development

The architecture of React Native helps us in structuring a project for a multi-platform mobile application, keeping the logic of the business in a reusable and maintained sub-module. In some projects, we can see that the containers could be important to be in the core module but that depends on the applications. React Native uses different mechanisms to create an efficient, consistent and reusable visual identity for the applications.

3.5 HARDWARE USED:

3.5.1 Raspberry Pi Zero:

The Raspberry Pi is a popular Single Board Computer (SBC) in that it is a full computer packed into a single board. Many may already familiar with the Raspberry Pi 3 and its predecessors, which comes in a form factor that has become as highly recognizable. The Raspberry Pi comes in an even smaller form factor. The introduction of the Raspberry Pi Zero allowed one to embed an entire computer in even smaller projects. This guide will cover the latest version of the Zero product line, the Raspberry Pi Zero - Wireless, which has an onboard WiFi module. While these directions should work for most any version and form factor of the Raspberry Pi, it will revolve around the Pi Zero W.

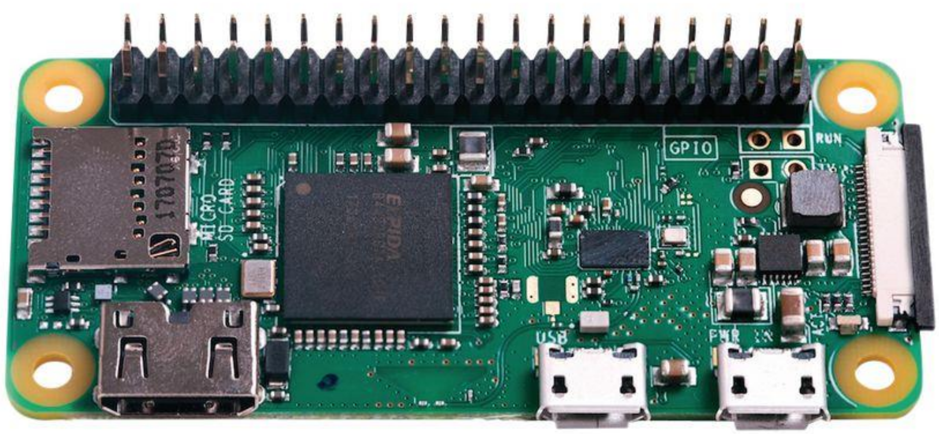

Figure 3.6 Raspberry pi zero

3.5.2 Servo Motor:

A servo motor is an electrical device which can push or rotate an object with great precision. If you want to rotate and object at some specific angles or distance, then you use servo motor. It is just made up of simple motor which run through servo mechanism. If motor is used is DC powered then it is called DC servo motor, and if it is AC powered motor then it is called AC servo motor. We can get a very high torque servo motor in a small and light weight packages. Doe to these features they are being used in many applications like toy car, RC helicopters and planes, Robotics, Machine etc.

Servo motors are rated in kg/cm (kilogram per centimeter) most hobby servo motors are rated at 3kg/cm or 6kg/cm or 12kg/cm. This kg/cm tells you how much weight your servo motor can lift at a particular distance. For example: A 6kg/cm Servo motor should be able to lift 6kg if the load is suspended 1cm away from the motors shaft, the greater the distance the lesser the weight carrying capacity. The position of a servo motor is decided by electrical pulse and its circuitry is placed beside the motor.

Figure 3.7 Servo motor

3.5.3 Power Supply:

The power supply circuit consists of step-down transformer which is 230v step down to 12v.In this circuit 4diodes are used to form bridge rectifier which delivers pulsating dc voltage & then fed to capacitor filter the output voltage from rectifier is fed to filter to eliminate any A.C components present even after rectification.

The filtered DC voltage is given to regulator to produce 12v constant DC voltage. 230V AC power is converted into 12V AC (12V RMS value wherein the peak value is around 17V), but the required power is 5V DC; for this purpose, 17V AC power must be primarily converted into DC power then it can be stepped down to the 5V DC. AC power can be converted into DC using one of the power electronic converters called as Rectifier. There are different types of rectifiers, such as half-wave rectifier, full-wave rectifier and bridge rectifier. Due to the advantages of the bridge rectifier over the half and full wave rectifier, the bridge rectifier is frequently used for converting AC to DC.

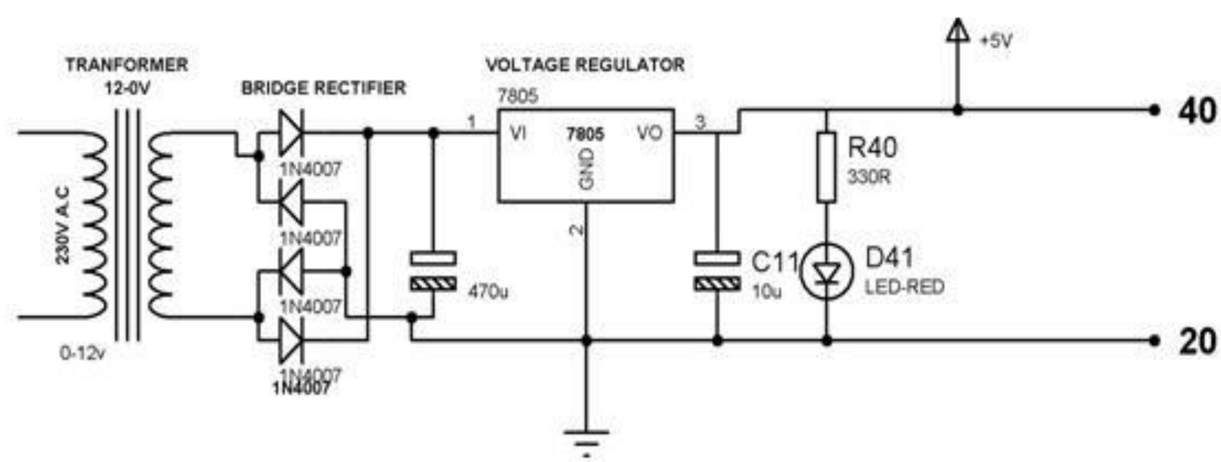

Figure: 3.8 Power supply

3.5.4 Buzzer:

A buzzer or beeper is an audio signaling device, which may be mechanical, electromechanical, or piezoelectric. Typical uses of buzzers and beepers include alarm devices, timers, and confirmation of user input such as a mouse click or keystroke. The fig shows the connection of the buzzer in the circuit.

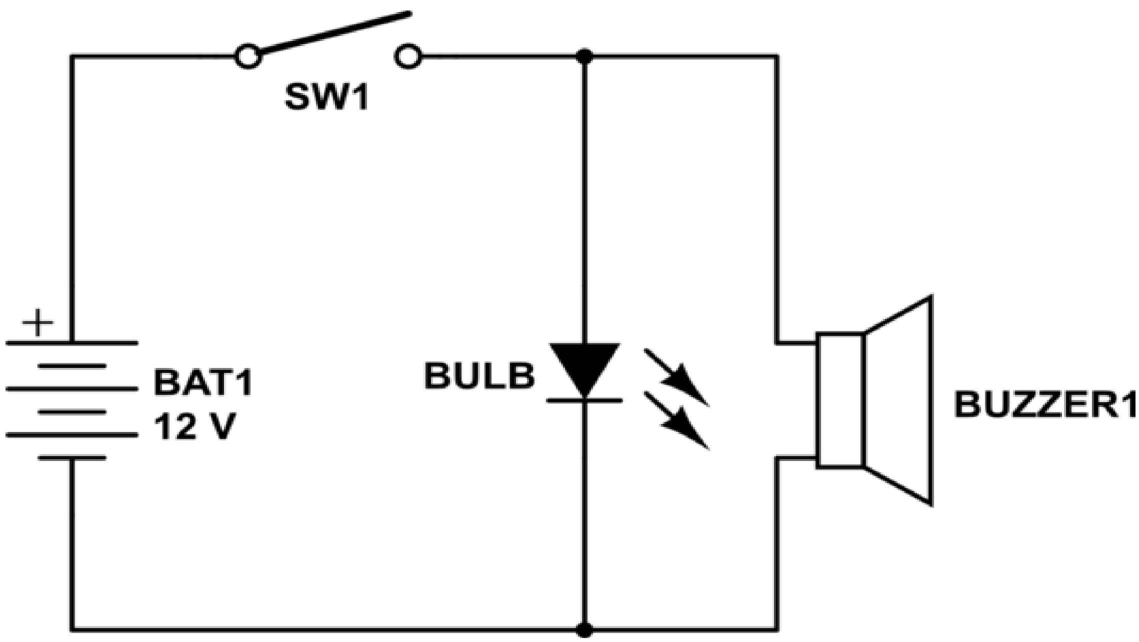

FIG BUZZER

Piezo buzzers are used for making beeps, tones and alerts. This one is petite but loud! Drive it with 3-30V peak-to-peak square wave. To use, connect one pin to ground (either one) and the other pin to a square wave out from a timer or microcontroller. For the loudest tones, stay around 4 KHz, but works quite well from 2 KHz to 10 KHz. For extra loudness, you can connect both pins to a microcontroller and swap which pin is high or low ('differential drive') for double the volume. The PS series are high-performance buzzers that employ unimorph piezoelectric elements and are designed for easy incorporation into various circuits.

Features:

- They feature extremely low power consumption in comparison to electromagnetic units.
- Because these buzzers are designed for external excitation, the same part can serve as both a musical tone oscillator and a buzzer.
- They can be used with automated inserters. Moisture-resistant models are also available.
- The lead wire type (PS1550L40N) with both-sided adhesive tape installed easily is prepared
- A buzzer or beeper is an audio signaling device, which may be mechanical, electromechanical, or piezoelectric. Typical uses of buzzers and beepers include alarm devices, timers and confirmation of user input such as a mouse click or keystroke. Buzzer is an integrated structure of electronic transducers, DC power supply, widely used in computers, printers, copiers, alarms, electronic toys, automotive electronic equipment, telephones, timers and other electronic products for sound devices. Active buzzer 5V Rated power can be directly connected to a continuous sound, this section dedicated sensor expansion module and the board in combination, can complete a simple circuit design.

3.5.5 Raspberry Pi Camera Module:

The Raspberry Pi has seen a number of fantastic improvements since its original release in April 2012, including a new PCB layout, new mounting holes, and a RAM upgrade to 512MB. Following these improvements was the introduction of the low-price model A Raspberry Pi, which was intended as a cheaper model for education and applications that require lower power. Then, in May 2013 (just over a year after the initial release of the model B Raspberry Pi) and after some slight delays, the

Raspberry Pi Foundation officially released its first add-on board: the Raspberry Pi camera module

Figure: 3.9 Raspberry Pi's powerful camera module.

Before the camera module, it was of course possible to access a camera feed on the Raspberry Pi using a suitable webcam. Camera functionality had already worked its way into a variety of projects, from live weather monitoring to robotics. However, the clever folks at the Raspberry Pi Foundation were aware of the fact that, if the highly enthusiastic Raspberry Pi community would repurpose their existing webcams, an add-on camera module stood a strong chance of success.

The camera module, designed specifically for the Raspberry Pi boards, brought with it a fair number of features that make it superior to many sophisticated webcams on the market.

The Raspberry Pi Camera Board plugs directly into the CSI connector on the Raspberry Pi. It's able to deliver a crystal clear 5MP resolution image, or 1080p HD video recording at 30fps! Latest Version 1.3! Custom designed and manufactured by the Raspberry Pi Foundation in the UK, the Raspberry Pi Camera Board features a 5MP (2592?1944 pixels) Omnivision 5647 sensor in a fixed focus module. The module attaches to Raspberry Pi, by way of a 15 Pin Ribbon Cable, to the dedicated 15-pin MIPI Camera Serial Interface (CSI), which was designed especially for interfacing to cameras. The CSI bus is capable of extremely high data rates, and it exclusively carries pixel data to the BCM2835 processor. The board itself is tiny, at around 25mm x 20mm x 9mm, and weighs just over 3g, making it perfect for mobile or other applications where size and weight are important. The sensor itself has a native resolution of 5 megapixel, and has a fixed focus lens onboard. In terms of still images, the camera is capable of 2592 x 1944 pixel static images, and also supports 1080p @ 30fps, 720p @ 60fps and 640x480p 60/90 video recording. The camera is supported in the latest version of Raspbian, the Raspberry Pi's preferred operating system.

The Raspberry Pi Camera Board Features:

- Fully Compatible with Both the Model A and Model B Raspberry Pi
- 5MP Omni vision 5647 Camera Module
- Still Picture Resolution: 2592 x 1944
- Video: Supports 1080p @ 30fps, 720p @ 60fps and 640x480p 60/90 Recording
- 15-pin MIPI Camera Serial Interface - Plugs Directly into the Raspberry Pi Board
- Size: 20 x 25 x 9mm
- Weight 3g
- Fully Compatible with many Raspberry Pi cases

3.6 SOFTWARE LIST:

3.6.1 Microsoft Visual Studio:

In this project we make use of the Microsoft visual studio as the IDE.

Microsoft Visual Studio is an integrated development environment (IDE) from Microsoft. It is used to develop computer programs, as well as websites, web apps, web services and mobile apps. Visual Studio uses Microsoft software development platforms such as Windows API, Windows Forms, Windows

Presentation Foundation, Windows Store and Microsoft Silverlight. It can produce both native code and managed code.

Visual Studio includes a code editor supporting IntelliSense (the code completion component) as well as code refactoring. The integrated debugger works both as a source-level debugger and a machine-level debugger. Other built-in tools include a code profiler, designer for building GUI applications, web designer, class designer, and database schema designer. It accepts plug-ins that enhance the functionality at almost every level—including adding support for source control systems (like Subversion and Git) and adding new toolsets like editors and visual designers for domain-specific languages or toolsets for other aspects of the software development lifecycle (like the Azure DevOps client: Team Explorer).

Visual Studio supports 36 different programming languages and allows the code editor and debugger to support (to varying degrees) nearly any programming language, provided a language-specific service exists. Built-in languages include C, C++, C++/CLI, Visual Basic .NET, C#, F#, JavaScript, TypeScript, XML, XSLT, HTML, and CSS. Support for other languages such as Python, Ruby, Node.js, and M among others is available via plug-ins. Java (and J#) were supported in the past.

The most basic edition of Visual Studio, the Community edition, is available free of charge. The slogan for Visual Studio Community edition is "Free, fully-featured IDE for students, open-source and individual developers".

The currently supported Visual Studio version is 2019

Features:

- Code editor. Like any other IDE, it includes a code editor that supports syntax highlighting and code completion using IntelliSense for variables, functions, methods, loops, and LINQ queries.

- Debugger.
- Designer.
- Other tools.
- Extensibility.
- Previous products.
- Community.
- Professional.

3.6.2 Python:

In this project we make use of the python language as the development language.

In technical terms, Python is an object-oriented, high-level programming language with integrated dynamic semantics primarily for web and app development. It is extremely attractive in the field of Rapid Application Development because it offers dynamic typing and dynamic binding options.

Python is relatively simple, so it's easy to learn since it requires a unique syntax that focuses on readability. Developers can read and translate Python code much easier than other languages. In turn, this reduces the cost of program maintenance and development because it allows teams to work collaboratively without significant language and experience barriers.

Additionally, Python supports the use of modules and packages, which means that programs can be designed in a modular style and code can be reused across a variety of projects. Once you've developed a module or package you need, it can be scaled for use in other projects, and it's easy to import or export these modules.

One of the most promising benefits of Python is that both the standard library and the interpreter are available free of charge, in both binary and source form. There is no exclusivity either, as Python and all the necessary tools are available on all major platforms. Therefore, it is an enticing option for developers who don't want to worry about paying high development costs.

3.7 ADVANTAGES:

- Cheap and effective solution for home surveillance
- Provides alert to the home owner in case of intruder detection
- A proof for police department for finding the thief
- Automatic SMS option for emergency situation
- Mobile application is developed
- App notification to the user
- Live streaming through the mobile application

3.8 APPLICATIONS:

- This project helps in saving people property and money as well as automates the records which helps to punish the thief.
- Implementation of this project also gives rise to many applications such as provide in bank lockers, etc.
- Mainly its used in home for home security purpose.

CHAPTER - 4
RESULTS & DISCUSSIONS

4.1 FINAL OUTPUTS:

The below image shows the complete hardware setup of the project:

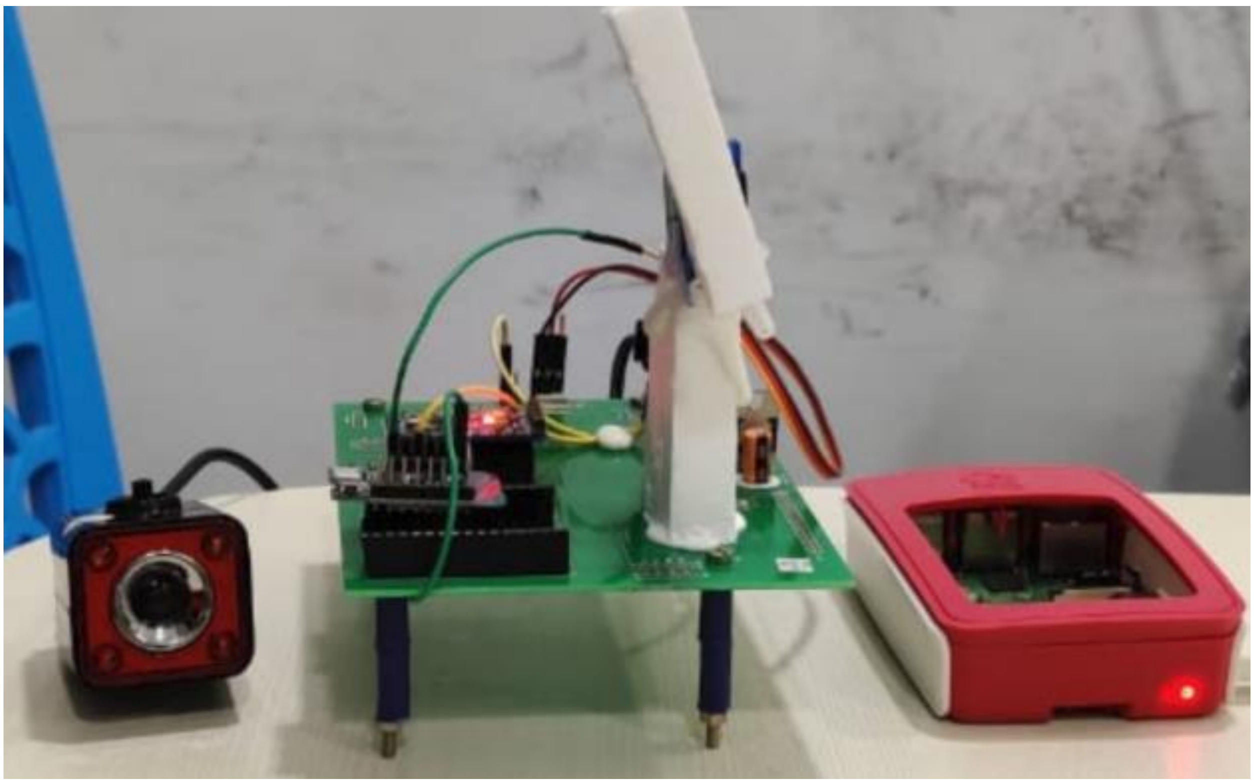

Figure 4.1: Complete hardware setup of the project

The hardware components are namely Raspberry Pi, CSI Camera, Servo motor which acts as the door locker and buzzer which acts as an alarm.

A mobile application using react native is developed for effectively notifying the owner on an unknown person entering into a house. The below image shows the login page of the mobile application:

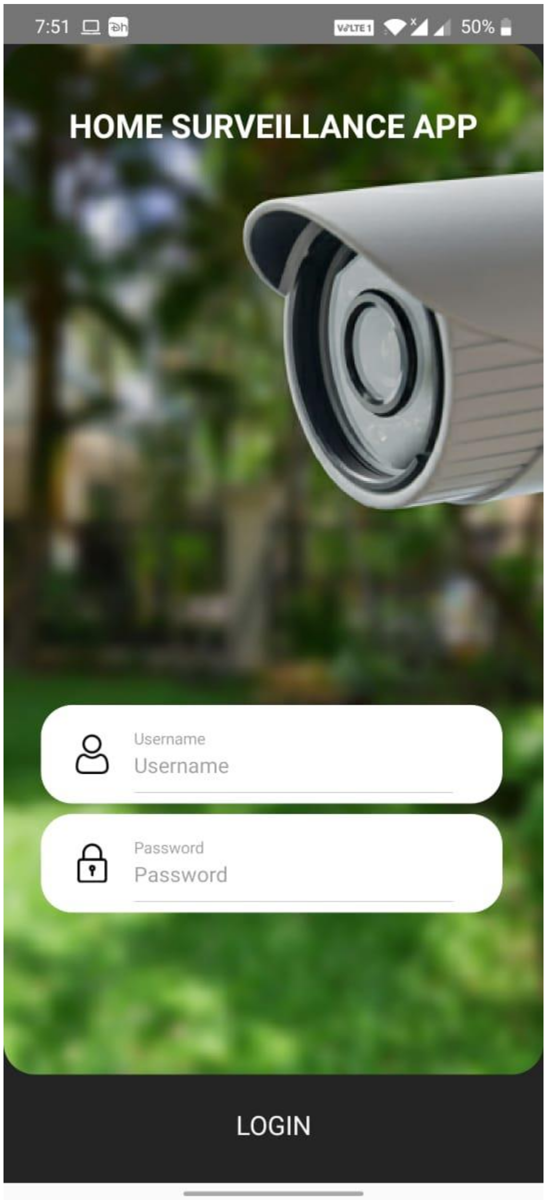

Figure 4.2: Login page of mobile application

Particular user's username and password is entered and logged in to know about the current activity status of the home.

On successful login to the mobile application the home page of the mobile application is displayed:

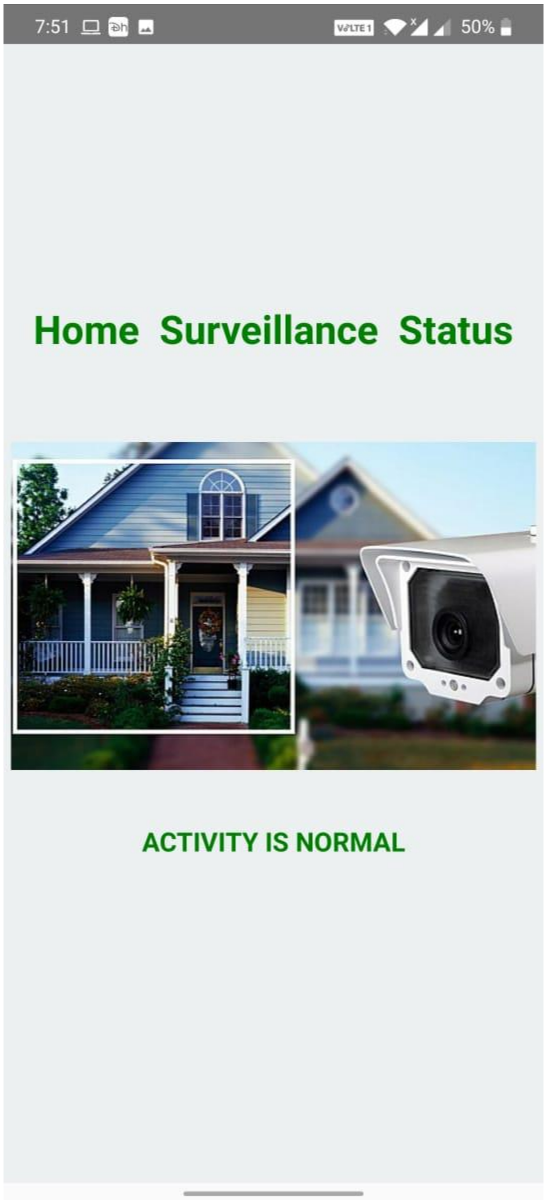

Figure 4.3: Home page of mobile application

This shows that the home is in the normal condition even if the known faces are detected through the camera.

When a known person is detected as seen in the image below the mobile application does no intimation:

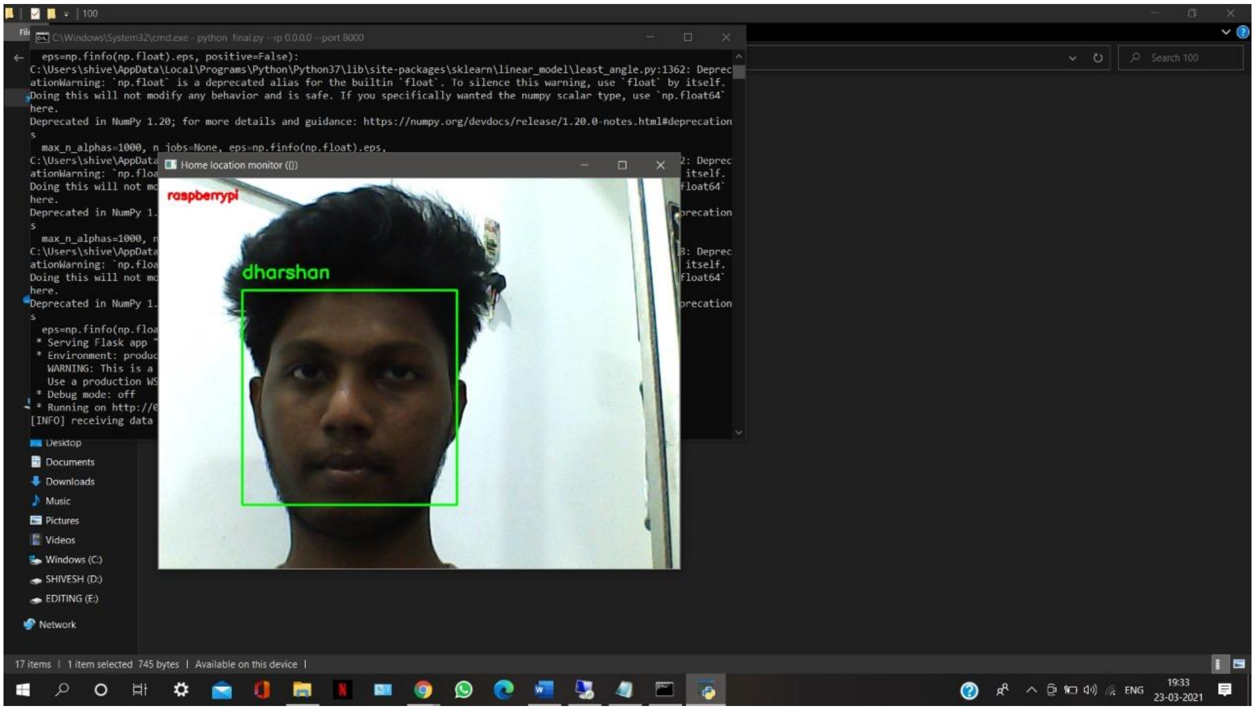

Figure 4.4: Known person detection

This will be shown up in the PC also the hub, by recognizing the known or registered or owner’s face.

The below image shows the unknown person recognition:

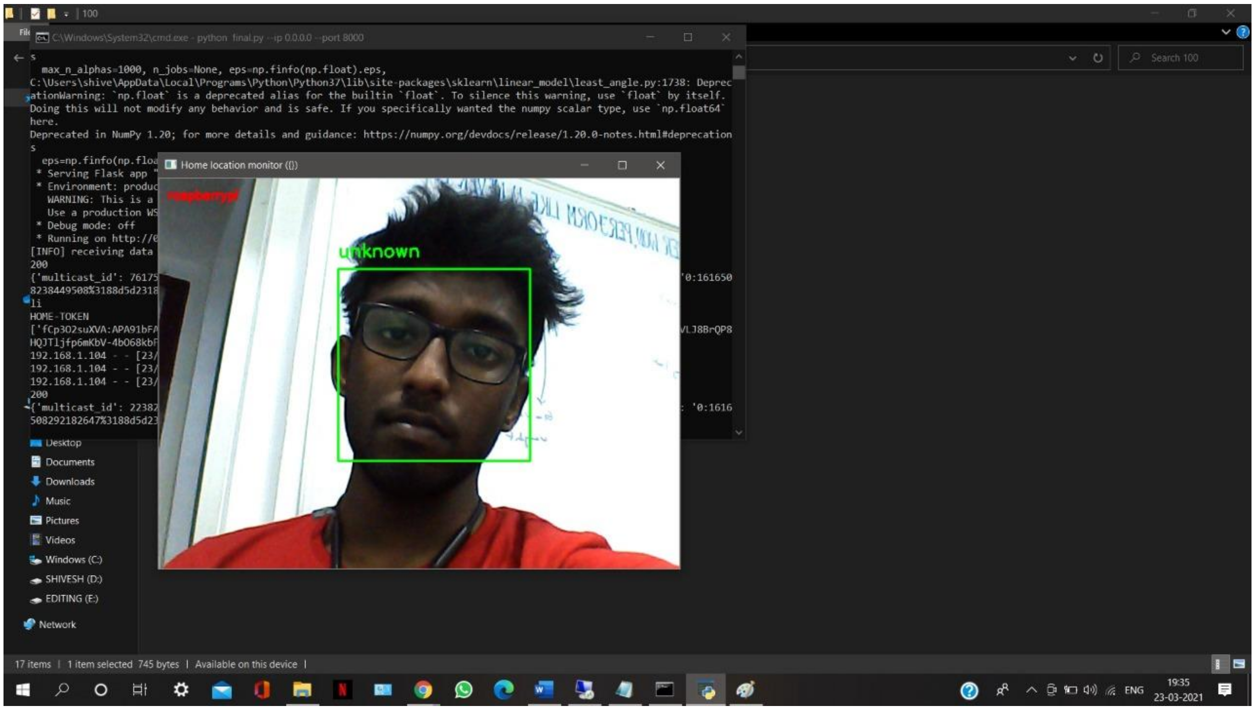

Figure 4.5: Unknown person detection

On unknown person detection the mobile application notifies the owner with a notification, live streaming and actions to perform. This is actually shown in the personal computer or the hub.

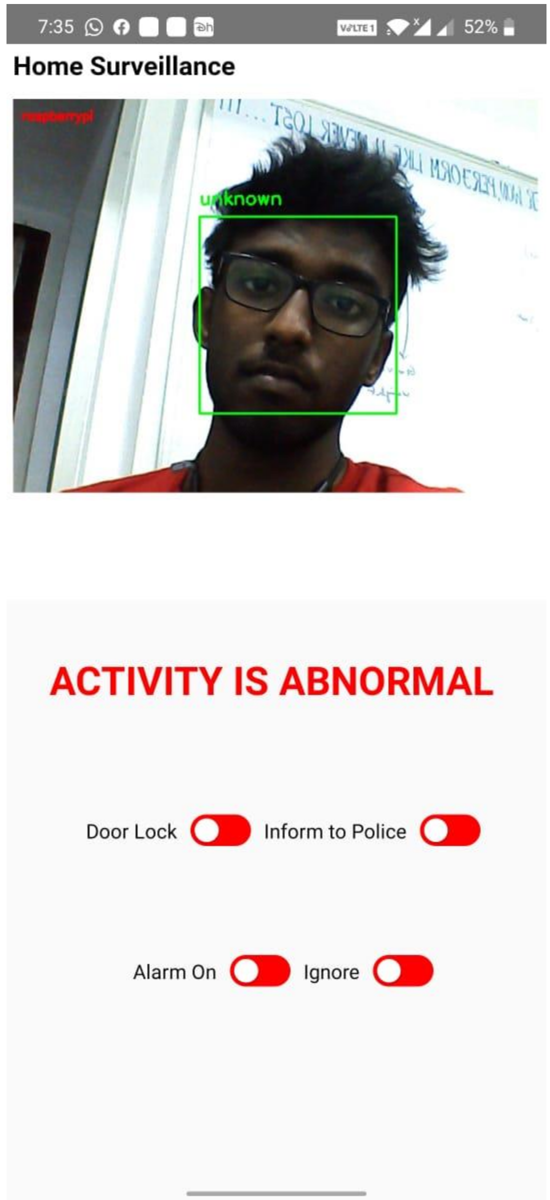

Figure 4.6: Abnormal intimation

This is how live streaming is done where the owner gets to know through the mobile application and can perform any activity with the given options. Thus, from the above results and discussions it is clear that the project has been successfully implemented.

CHAPTER – 5

CONCLUSION AND FUTURE ENHANCEMENT

5.1 CONCLUSION

In this project the home is armed with multiple Raspberry Pi that communicate among themselves to accomplish a common goal. We successfully implemented a full end security system using raspberry pi and intelligent camera with a low cost efficiently. Thus, this project helps the reduce loss of money by thief and provides us a well protected home effectively. Criminals may find an unprotected entry point into the home if just a front and backdoor sensor or a camera is used. By taking a tour around the perimeter of the home to spot any security weaknesses and with the help of cameras, raspberry pi, buzzer, mobile application, surveillance is done. So, in this way of home surveillance by intruder detection, recognition and alerting protection can be ensured.

5.2 FUTURE ENHANCEMENT

In the coming future, the application of the home surveillance system technology is reviewed. It can promote for shop security system with more accuracy and in all kinds of security system. So, there are more chance to develop or convert this project in many ways. And also provide security to home, money, and property. This can not only be implemented in the home areas alone but also in the office areas, areas where high security requirements are needed like security vault, bank lockers, lockers which are transmitted from place to place and many more to ensure properties with more safety.

REFERENCE

[1] Chun-Yang Zhang , Yong-Yi Xiao, Jin-Cheng Lin, C. L. Philip Chen , Wenxi Liu , and Yu-Hong Tong, “3-D Deconvolutional Networks for the Unsupervised Representation Learning of Human Motions “, 2020.

[2] Geong Sen Poh, Prosanta Gope , and Jianting Ning , “PrivHome: Privacy-Preserving Authenticated Communication in Smart Home Environment”, 2019, VOL. 99, NO. 99.

[3] Chuanwei Ding , Yu Zou, Hui Chu , Xiaohua Zhu, Francesco Fioranelli, Julien Le Kernec , and Changzhi Li, “Continuous Human Motion Recognition With a Dynamic Range-Doppler Trajectory Method Based on FMCW Radar”, 2019, VOL. 57, NO. 9.

[4] BSamet Taspinar, Manoranjan Mohanty, and Nasir Memon, “PRNU-Based Camera Attribution from Multiple Seam-Carved Images”, 2017, VOL. 20, NO. 5.

[5] Ying-Tsung Lee, Wei-Hsuan Hsiao, Chin-Meng Huang and Seng-Cho T. Chou, “An Integrated Cloud-Based Smart Home Management System with Community Hierarchy”, 2016, Vol No: 2162-237X.

[6] R. Raghavendra Kiran B. Raja Christoph Busch, ” Presentation Attack Detection for Face Recognition using Light Field Camera”, Volume: 24, Issue: 3, March 2015.

[7] Xueru Bai , Ye Hui, Li Wang, and Feng Zhou ,“Radar-Based Human Gait Recognition Using Dual-Channel Deep Convolutional Neural Network”, Volume: 57, Issue: 12, Dec. 2019.

[8] Yaxu Xue, Zhaojie Ju, Kui Xiang, Jing Chen, and Honghai Liu, “Multimodal Human Hand Motion Sensing and Analysis - A Review”, Volume: 11, Issue: 2, June 2019.

[9] Sid Ahmed Walid Talha1, Mounir Hammouche1,2, Enjie Ghorbel1,3, Anthony Fleury*1, Sébastien Ambellouis1,4, “Features and classification schemes for view-invariant and real-time human action recognition”, 2018.

[10] Sujin Jang, Niklas Elmqvist, and Karthik Ramani, “Motion Flow: Visual Abstraction and Aggregation of Sequential Patterns in Human Motion Tracking Data” , vol. 22, no. 1, January 2016.

[11] Dinesh K Vishwakarma, Kuldeep Singh, “Human Activity Recognition based on Spatial Distribution of Gradients at Sub-levels of Average Energy Silhouette Images”, Volume: 9, Issue: 4, Dec. 2017.

[12] Jinzhu Chen, Rui Tan, Guoliang Xing, Xiaorui Wang, and Xing Fu, “Fidelity-Aware Utilization Control for Cyber-Physical Surveillance Systems” , 2012.

[13] Regio A. Michelin, Nadeem Ahmed, Salil S. Kanhere, Aruna Seneviratne and Sanjay Jha, “Leveraging lightweight blockchain to establish data integrity for surveillance cameras”, 2019.

[14] Amir H. Meghdad, and Pourang Irani, “Interactive Exploration of Surveillance Video through Action Shot Summarization and Trajectory Visualization”, Volume: 19, Issue: 12, Dec. 2013.

[15] Yuqiang Liu, Yanjie Yao, Chao Gou, Dayong Shen, and Shaohu Tang, “Hierarchical and Networked Vehicle Surveillance in ITS: A Survey, Bin Tian, Brendan Tran Morris, Ming Tang, Member”, Volume: 16, Issue: 2, April 2015.

11th INTERNATIONAL CONFERENCE ON SCIENCE AND INNOVATIVE ENGINEERING 2021

ORGANIZED BY

ORGANIZATION OF SCIENCE & INNOVATIVE ENGINEERING AND TECHNOLOGY, CHENNAI

IN ASSOCIATION WITH

JAWAHAR ENGINEERING COLLEGE, CHENNAI

Certificate of Presentation

This is to certify that Dr./Mr./Ms. **V.S.Sri Dharshan** from **St.Joseph's Institute of Technology, Chennai** has presented a paper titled **ENTIRE HOME SURVEILLANCE APPROACH USING PI AND INTELLIGENT CAMERAS WITH MOBILE APP NOTIFICATION** in the "11th International Conference on Science and Innovative Engineering" held on 02nd May 2021.

Secretary

Director

11th INTERNATIONAL CONFERENCE ON SCIENCE AND INNOVATIVE ENGINEERING 2021

ORGANIZED BY

ORGANIZATION OF SCIENCE & INNOVATIVE ENGINEERING AND TECHNOLOGY, CHENNAI

IN ASSOCIATION WITH

JAWAHAR ENGINEERING COLLEGE, CHENNAI

Certificate of Presentation

This is to certify that Dr./Mr./Ms. **M.Santhosh** from **St.Joseph's Institute of Technology, Chennai** has presented a paper titled **ENTIRE HOME SURVEILLANCE APPROACH USING PI AND INTELLIGENT CAMERAS WITH MOBILE APP NOTIFICATION** in the "11th International Conference on Science and Innovative Engineering" held on 02nd May 2021.

Secretary

Director