Non-intrusive Human Sensing - Techniques and Applications

Thesis submitted in partial fulfilment of the requirements for the award of the degree of

Doctor of Philosophy

in

Computer Science and Engineering

by

Sonia

Under the supervision of

Dr. Rashmi Dutta Baruah
Prof. Shivashankar B. Nair

Department of Computer Science and Engineering

Indian Institute of Technology Guwahati

Guwahati - 781039 Assam India

March, 2019

Dedicated to

My Mom, Dad

and

My brothers

Acknowledgements

Firstly, I would like to express my heartfelt gratitude to my supervisors *Dr. Rashmi Dutta Baruah* and *Prof. Shivashankar B. Nair* for their consistent support, indefatigable patience and guidance during my doctoral research. I am incredibly grateful to them for continually motivating me and being a source of inspiration. I thank them for their positive and caring attitude which made me more attentive and a better human being. Their encouragement to always complete that extra mile, to have a strong stand on ethics, and to be impeccable in every respect will reverberate throughout my life.

I am highly obliged to *Prof. Pradip K. Das*, the chairman of my doctoral committee, for his patience, innovative ideas and constant support. I would also like to thank my other thesis doctoral committee members - *Dr. V. Vijaya Saradhi* and *Prof. Harshal B. Nemade* for their insightful comments and timely suggestions that made me improve the quality and clarity of my work.

I want to thank the former and present head of the Department of Computer Science and Engineering during my Ph.D. at IITG - *Prof. Diganta Goswami* and *Prof. S. V. Rao* for allowing me to use the facilities and the available resources including the travel support for the conferences. I thank the Technical staff of the Department of Computer Science and Engineering - Mr. Nanu Alan Kachari, Mr. Hemanta Kumar Nath, Mr. Bhriguraj Borah, Mr. Pranjitt Talukdar, Mr. Raktajit Pathak and Mr. Nava Kumar Boro for handling technical issues. I am deeply thankful to - Mr. Monojit Bhattacharjee, Ms. Gauri Khuttiya Deori and Mr. Prabin Bharali for being a helping hand to control the administrative work. I am obliged to all the faculty members, the staff and security personnel for their constant help and support. I would like to show my gratitude to the Subansiri staff, IITG hospital staff for their great help.

I am thankful to TCS research and innovation lab., Kolkata and IBM research lab., New Delhi for providing me internship opportunities with them. I am grateful to Mr. Avik Ghose and Dr. Tapas Chakarbortay for mentoring me at TCS. I want to thank Dr. Kushal Mukherjee, Rakesh Pimplikar, Dr. Sudhanshu Shekhar Singh, Dr. Gyana R Parija, Udit Sharma, Aayush Shah, Dr. Ritwik Chaudhury and Sarthak Ahuja for all the learnings, good memories and constant support.

I am thankful to Dr. Priya Nair for her positivity and constant support.

I want to thank my senior Mohit Kumar, especially for his motivations. I am indeed thankful to my fellow lab mates at the Robotics Lab. - Mohit, Tushar, Abhay, Manoj, Nikhil, Rahul, Pallabi, Achyut, Deepankar, Bhagat, Suraj Pandey and Divya for their love and support. The regular discussions and tea breaks are indeed special moments to me.

Special mention to my friends - Ashish, Rakesh, Vishal, Sunil, Prateek, Abhishek, Dipika, Nidhi and Sanhita Das who made my stay at IITG easy going and memorable. A special thank to Tushar Semwal for always being there for me and a caring and trustworthy human being. I am thankful to all my colleagues and friends during my journey as a Ph.D. scholar.

I want to thank my childhood friends and college friends - Aasma, Jyoti, Piyush, Himanshu Pandey, Shubham Saxena, Nitesh Singh, Meenakshi Malik, Ankita Aggarwal, Aakriti, Rajiv Pandey, Anushruti, Dhanya for bearing with me and for listening to me with patience.

Finally, yet importantly, I would like to thank Almighty God and my family - Mom, Dad, my elder brothers Rakesh and Rajesh and my younger brother Punit for their selfless love, support, care, motivation, strength, and encouragement throughout all these years. I am truly indebted to them and cannot express my love for them in words.

March 25, 2019 Sonia

Declaration on Plagiarism

I, Sonia (Roll No. 146101008), understand that plagiarism is defined as any one or combination : Uncredited verbatim copying of individual sentences, paragraphs or illustrations (such as graphs, diagrams, etc.) from any source, published or unpublished, including the Internet.

Uncredited improper paraphrasing of pages or paragraphs (changing a few words or phrases, or rearranging the original sentence order). Credited verbatim copying of a major portion of a paper (or thesis chapter) without clear delineation of who did or wrote what. (Source: IEEE, The Institute, Dec. 2004)

I have to the best of my abilities ensured that all the concepts, ideas, text, expressions, graphs, diagrams, etc., which are not the outcome of my research have been credited to those who own the same. Elaborate sentences used verbatim from published work have been clearly identified and quoted. I also affirm that no part of this thesis can be considered as plagiarism to the best of my knowledge and understanding and take complete responsibility if any complain arises.

I am fully aware that my thesis supervisor(s) are not be in a position to check for any possible instance of plagiarism within this submitted work.

Guwahati 29-03-2019 Sonia

March 25, 2019 Sonia

Declaration

I certify that

- The work contained in this thesis is original and has been done by myself and under the general supervision of my supervisors.

- The work reported herein has not been submitted to any other Institute for any degree or diploma.

- Whenever I have used materials (concepts, ideas, text, expressions, data, graphs, diagrams, theoretical analysis, results, etc.) from other sources, I have given due credit by citing them in the text of the thesis and giving their details in the references. Elaborate sentences used verbatim from published work have been clearly identified and quoted.

- I also affirm that no part of this thesis can be considered plagiarism to the best of my knowledge and understanding and take complete responsibility if any complaint arises.

- I am fully aware that my thesis supervisors are not in a position to check for any possible instance of plagiarism within this submitted work.

March 25, 2019 Sonia

Department of Computer Science and Engineering
Indian Institute of Technology Guwahati
Guwahati - 781039 Assam India

Dr. Rashmi Dutta Baruah
Assistant Professor
Email : r.duttabaruah@iitg.ac.in
Phone : +91-361-258-3253

Dr. Shivashankar B. Nair
Professor
Email : sbnair@iitg.ac.in
Phone : +91-361-258-2356

Certificate

This is to certify that this thesis entitled "**Non-intrusive Human Sensing - Techniques and Applications**" submitted by **Sonia**, in partial fulfilment of the requirements for the award of the degree of Doctor of Philosophy, to the Indian Institute of Technology Guwahati, Assam, India, is a record of the bonafide research work carried out by him under my guidance and supervision at the Department of Computer Science and Engineering, Indian Institute of Technology Guwahati, Assam, India. To the best of my knowledge, no part of the work reported in this thesis has been presented for the award of any degree at any other institution.

Date: March 25, 2019
Place: Guwahati

Prof. Shivashankar B. Nair

Dr. Rashmi Dutta Baruah

Abstract

Human sensing refers to the process of differentiating human beings from other living beings or non-living objects through the use of sensory technology. The task of human detection is vital for many applications such as human-robot interaction, surveillance and monitoring, search and rescue, smart homes, etc. However, detecting humans is a challenge due to the facts that humans are non-rigid in structure with various sizes and shapes, dress up in different types of clothing, and are generally non-stationary. Though many methods have been propounded by researchers which can be used in a singular or stand-alone manner [3], none provide a foolproof and guaranteed solution to this problem. In an Internet of Things (IoT), where the devices along with their associated sensors are scattered yet connected as a network, it is imperative to reinvent a new, distributed, co-operative and adaptive paradigm for human sensing. Such a paradigm will aid in promoting a harmony between human beings, societies and smart things. A considerable amount of research has been done in the area of vision-based human detection. However, for many applications vision-based human detection is not desirable. For example, the unattended ground sensors (UGS) systems, used for military purpose, are mainly based on non-imaging sensors like acoustic, seismic, ultrasonic, Pyro-infrared (PIR) sensor; the reasons being that of power efficiency since imaging sensors consume more power. Further, the image or video processing algorithms are slower and resource intensive compared to acoustic or other forms of data. Non-imaging sensors based human detection is also desirable in applications where a subject's privacy is of concern. Other issues with imaging sensors include their sensitivity to the position, orientation, ambient light, and unable to detect human beings behind opaque structural obstructions. Since each sensor has its own limitations, they cannot fulfill the task of human sensing in an efficient manner. For instance, a PIR sensor cannot detect a static human being. One of the feasible solutions is to use the concept of sensor fusion. Techniques such as Kalman filtering, fuzzy logic, probabilistic approaches, etc. are used to fuse the

data from multiple sensors at the various levels such as data, abstract or decision level. This thesis discusses the challenges in human sensing and endeavors to solve some of the major issues. The thesis commences with a description of various sensors that could be used for human sensing and the manner in which the features obtained from the raw signal from an ultrasonic sensor can be used for non-intrusive sensing. Combining the information from multiple heterogeneous sensors to improve the accuracy of human sensing has also been described. The above combination of sensors along with a camera was mounted on-board a robot and the efficacy of the multi-modal human sensing system was studied in a real environment. This along with investigations into dynamic environments forms the next part of the thesis. In order to reduce redundancy in the set of features extracted thereby reducing computational costs, an immuno-inspired mechanism has been proposed, next. This mechanism was implemented using a Pyro Infrared (PIR) sensor and an Analog Ultrasonic Sensor (AUS). One of the prominent areas where human sensing is required is in the domain of health and wellness. For this it is essential to extract the behavior of a human being given his/her profile. The subsequent part of the thesis thus proposes an approach for detection early symptoms of cognitive impairment using spatio-temporal data which has been obtained using human sensing. Since it is not feasible to capture spatio-temporal data for every possible scenario, a method to use this learned information in other scenarios, where there is no/insufficient data, is highly essential. The last part of the thesis thus describes the manner in which transfer learning can be used to achieve the same.

The first contribution of the thesis deals with identifying sensors for non-intrusive human sensing. An Ultrasonic (US) and a PIR sensor were separately used to differentiate human beings from non-human things. As mentioned earlier, every sensor has its own inherent limitation(s). Therefore, to improve the accuracy of human sensing, an algorithm to combine the information from both these sensors has been proposed in the thesis. The proposed approach was tested by using a combination of US and PIR sensors in an indoor environment. The results obtained using this combination achieve better accuracy for human sensing than methods incorporating only single sensors.

The previous contribution confirmed the fact that a sensor is susceptible to fail under certain scenarios. The credit goes to the technological limitations of a sensor. One viable solution to compensate for this constraint could be to augment with other complementary sensors. The next contribution, thus presents a multi-modal human sensing approach that facilitates autonomous sensor selection based on the changes in the environment. Experiments were carried out using a camera, a PIR and US sensors in different environments. The results obtained after using this approach proved its efficacy.

An underlying model for human detection is expected to adapt and perform well in terms of accuracy and detection time. The number of features that can be extracted from the raw signals forms one of the parameters that define the computational complexity and time complexity of a given system. Therefore, the selection of the minimal number of features useful in correctly detecting a human being, is a non-trivial issue. An approach to find such features from a give set of features constitutes the second contribution. This contribution applies the proposed feature-selection approach to the data received via US and PIR sensors for human sensing. A comparison of results obtained by using a classifier with and without using feature selection, has been presented. One of the conclusions arrived at is that a change in the environment causes the set of selected features to change. The approach has been substantiated by experiments carried out in the real world.

Behaviour profiling of a human depends on the success of human sensing. Different sensors can be embedded in the living space of human beings so as to track their activities. From the data obtained, concerning to an application, behavior of a person is analyzed. The next contribution takes human sensing to a further level and is towards the early symptoms detection of cognitive impairment in humans. In the fourth contribution, the use of unobtrusive, non-contact activities of daily living sensors for early detection of Mild Cognitive Impairment in geriatric population has been explored. The feasibility of using deep learning techniques to make such inferences has been shown. Further, this information is used to design a classifier to predict the future cases of illness.

Delivering accurate and helpful information on the following action to be performed by a person is an essential factor in pervasive computing. However, the available algorithms are dependent on the availability of training data. Collecting training data is not feasible for every scenario and also it is a time consuming task. To cater to this issue the concept of *Transfer Learning* in the domain of smart homes is explored in the last contribution of this thesis.

Citation to Published Work

Chapter 3 is based on the following papers:

1. **Sonia**, Manish Singh, Rashmi Dutta Baruah and Shivashankar B Nair, ***A Voting-Based Sensor Fusion Approach for Human Presence Detection***, 8th International Conference on Human Computer Interaction, 12-13th Dec. 2016, Pilani, India.

2. **Sonia**, Achyut Mani Tripathi, Rashmi D. Baruah,Shivashankar B. Nair, ***Ultrasonic Sensor-based Human Detector using One-class Classifiers***, IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS 2015), Douai, France.

Chapter 4 is based on the following papers:

1. **Sonia**, Rashmi Dutta Baruah, Shivashankar B. Nair***Multi-Modal Human Sensing***, 2018 ACM Transactions on Intelligent Systems and Technology(TIST) (Under review)

2. **Sonia**, Prateek Vij, Rashmi Dutta Baruah, ***An Experience of Multi-Sensor Robot for Adaptive Human Sensing***, 2018 Symposium Series On Computational Intelligence (SSCI), 18-21 November, 2018, Bengaluru, India.

Chapter 5 is based on the following papers:

1. **Sonia**,Shivashankar B. Nair,Rashmi D. Baruah ***An Immuno-inspired Online Feature Selection Mechanism***, International Conference on Systems, Man and Cybernetics (SMC 2017), Banff, Canada, October 05-08, 2017.

2. **Sonia**, Rashmi D. Baruah,Shiva shankar B. Nair, ***A Reward and Penalty Based Approach for Online Feature Selection***, IEEE International Conference on Cybernetics (Cybconf 2017), Exeter, England, UK, 21-23 June 2017.

Chapter 6 is based on the following patents and papers:

1. Avik Ghose, **Sonia**, Arijit C., Systems and Method For Early Detection of Mild Cognitive Impairment in Subjects. Application No: 201821007810,Date of Filing: 01/03/2018

2. Avik Ghose, **Sonia**, System and Method for Authenticating Humans Based on Bahavioral Pattern. Application No: 201821039448, Date of Filing: 17/10/2018

3. **Sonia**, Avik Ghose, ***Unobtrusive and Pervasive Monitoring of Geriatric Subjects for Early Screening of Mild Cognitive Impairment***, PerHealth, IEEE International Conference on Pervasive Computing and Communications (PERCOM) 2018, Athens, Greece.

4. **Sonia**, Avik Ghose ***Analysis of Analog signal from PIR-sensors for human analytics***, Bird, IEEE International Conference on Pervasive Computing and Communications (PERCOM) 2019, Kyoto, Japan.

5. **Sonia** ***Recognising Human Beings from Their Behavior Pattern***, IEEE SIU-IoT 2018.

Chapter 7 is based on the following patents and papers:

1. **Sonia**, Rashmi Dutta Baruah ***Transfer Learning in the Domain of Smart Homes***, 2019 IEEE Internet of Things (Under review)

Contents

List of Figures

List of Algorithms

List of Tables

List of Symbols

Symbols	Description
Ab	Antibody
Ag	Antigen
θ_{aff}	Affinity
β	Affinity towards network
(σ_{anti})	Antigenic Capture
α	Activation
P	Class
G	Centroid
α_c	Constant for learning rate
π	Concentration
c	Cluster center
r	Cluster radius

$Hull$	Convex Hull
$\mathcal{DF}$	Data Frame
d	Datastream
DS	Density of a convex hull
α_3	Deviation in routines
$\mathbb{F}$	Decision Making feature vector
E	Environment
$\mathcal{F}$	Eliminating feature vector
Ep	Epitope
D_x	Euclidean distance of the antigen
f	Feature value
F	Feature Vector
$\tau(x)$	Firing strength of a fuzzy rule
x^*	Focal point of a fuzzy rule
x	Input variable
α_4	Intra House Transfer Learning

Id	Idiotope
L	Least dense point
$\mathbb{M}$	Machine Learning Model
ϕ	Mapping function
T_{max}	Maximum lifetime
M	Number of features
N_{fr}	Number of fuzzy rules
N_S	Number of sensors
N_{clust}	Number of clusters
N_f	Number of frames or signals
N_v	Number of vectors
$\lvert\chi_t\rvert - \lvert\chi_{s_i}\rvert$	Number of paired sensors in target domain
α_2	Number of residents and Data Sampling Rate
y	Output
$\mathcal{P}$	Probability
μ	Performance

Pt	Paratope
Tg_{prm}	Parameters in Target
Sr_{prm}	Parameters in Source
ψ	Ratio
RD	Relative potential
α_r	Relation
fq	Sampling frequency
S_k	Sensor
S	Set of sensors
Q	Set of centers and radii of clusters
S^U	Set of useful sensors
S^B	Set of blinked sensors
B	Set of possible subsets
α_1	Sensor Modality and Physical Space
F'	Subset of better suited features
Sti	Stimulation received

rho	Scaling factor
δ	Suppression
Sr	Source
T	Task
θ	Threshold
z	Test data feature representation
Z	Test data vector
t	Time period
Tg	Target
N	Total length of a data frame
U	Utility value
V	Vote
$\mathcal{V}$	Vector
W_{old}	Weight matrix of the model trained for the old resident

List of Abbreviations

Terms	Abbreviations
FRB	Fuzzy Rule-Based
SVM	Support Vector Machine
IoT	Internet-of-Things
PIR	Pyro Infrared
US	Ultrasonic
UGS	Unattended Ground Sensors
TPR	True Positive Rate
TNR	True Negative Rate
FPR	False Positive Rate
FNR	False Negative Rate
MS/s	Million samples per second
bps	bits per second

DWT	Discrete Wavelet Transform
ZCR	Zero Crossing Rate
TDF	Time Domain Features
FDF	Frequency Domain Features
E	Energy
SD	Standard Deviation
RMS	Root Mean Square
Max	Maximum
AUS	Analog Ultrasonic Sensor
BIS	Biological Immune System
AIS	Aritificial Immune System
TL	Transfer Learning
iFS	Immuno-inspired Online Feature Selection
RFID	Radio-Frequency IDentification
HVAC	Heating, Ventilating, and Air Conditioning
FSP	Feature Selection Problem

MCI	Mild cognitive impairment
GA	Genetic Algorithm
PCA	Principal Component Analysis
FFT	Principal Component Analysis
AUS	Analog Ultrasonic Sensor
kbps	kilo bits per second
ML	Machine Learning
CNN	Convolution Neural Network
AAL	Ambient assisted living
ADL	Activities of Daily Living
HMM	Hidden Markov Model
RNN	Recurrent Neural Network
LSTM	Long Short Term Memory

Glossary of Terms

Terms	**Description**
Human Sensing	It is a process of differentiating human beings from non-humans.
Non-intrusive Sensing	A sensing where sensors do not invade in the privacy of human beings lives.
Antigens	Foreign substances which invades a living body and induces an immune response.
Antibodies	Y-shaped proteins used by immune system to neutralize foreign antigens.
Internet of Things	Devices including sensors connected to the Internet.
Epitope	An antibody binding part of an antigen.
Paratope	A portion of antibody which binds with the matching epitope of an antigen.

"You have to dream before your dreams can come true."

A.P.J. Abdul Kalam (1931 - 2015)

Indian scientist and leader

1

Introduction

The advent of cheaper and smaller embedded systems has paved a way for technology to creep into domestic habitats. Now it is possible for us to sit back at home and enjoy the fruits of such automation. It is thus now possible for your refrigerator to indicate to you of items that need to be replenished, your car can inform you that it is time for a service while the fans, lights, the music player, the air-conditioner, the desktop computer which you left on, are turned off the moment you leave your home to save energy. Even the fact that your elderly mother has fallen off the bed could be noted and conveyed to you as also the doctor via a message.

These are only a few manifestations of what is known today as the Internet of Things (IoT). In all these scenarios some direct or indirect human action triggers a sensor which in turn initiates a communication of the data over a network to a monitoring device that aggregates and analyzes it to flag the relevant situation.

In an IoT, devices can generate and share data amongst each other or send the same directly to a Cloud Server [119]. The *things* could include a plethora of devices such as sensor nodes, Personal Computers, embedded devices, robots, etc. The concept of an IoT by itself opens up research areas such as device networking, target identification and target detection, privacy control and so on. Smart houses, smart hospitals, energy management, homeland security, etc., are few of the initial applications thrown open by this concept. IoT thus plays a vital role in adding automation and smartness to an environment. Since people play a major role in the use of an IoT, devices constituting it need to be empowered with human sensing capabilities. To achieve this, researchers have investigated the use of multiple sensors coupled with different means of communication including Bluetooth, Wi-Fi, etc. [28]. Most of these techniques however cater to a specific application or environment [64]. The need of the day is an interoperable and standardized generic system that can greatly reduce the false alarm rate. One method is to exploit the distributed nature of sensors and associated data in an IoTized environment and then realize mechanisms that can use such data to reckon the presence or absence of a human being.

Human sensing deals with sensing the presence of at least one human being in the designated area. According to Teixeira et al. [137],"Human Sensing is a process of extracting an information regarding the people in any environment."

The fundamental questions that arise with human sensing are:

Why is human sensing required?

Human sensing is required to answer questions such as - *How many persons are present in a room? Where is the person X? Where was he earlier? Which path has he used to reach the destination? Who is X among a group of persons?*

How can human sensing be realized?

An object can be identified or differentiated by its traits. Human beings also have some traits that can be sensed via various sensors to realize human sensing. These traits are:

1. **Appearance:** Humans have a specific appearance which differentiates them from other living beings. This trait can be captured with the help of a camera. Researchers have performed the task of human recognition, human tracking and human localization with the help of this vision sensor. One of the limitations of such sensors is that the presence of light is mandatory.

2. **Heartbeat:** Human beings have a specific heart count which can be measured with the help of skin penetrating radio and ultrasonic signals. There are many health-care instruments which sense the heartbeat of a human being and provide information about the health issues or the requirements of the

person.

3. **Heat dissipation:** A human body emanates heat waves of a certain frequency which falls in the infrared region. This property has been exploited in many applications involving human sensing through sensors such as infrared cameras, PIR sensors, etc.

4. **Specific absorption coefficient:** A human body has a specific absorption coefficient for different type of waves. Ultrasonic based human sensing exploits this property of a human body to sense human presence.

5. **Vibrations:** Vibrations are pressure waves that are produced by a human being voluntarily or involuntarily. They can be measured using accelerometers and microphones, acoustic sensors, etc.

Some other traits that recently have been used by researchers are scent [23], gait [70], etc.

The above mentioned human characteristics can be sensed using different sensors such as Ultrasonic sensors, Pyro Infrared (PIR) sensors, vibration sensors and pressure sensors.

1.0.1 Non-Intrusive Human Sensing

Detecting the presence of a human being is vital to many real-world applications. To sense the human presence non-intrusively, sensor being motion sensors (PIR) and proximity sensors (scalar infrared range-finders) are used for this purpose [151] [49] [137]. In office scenarios, where people could be made to wear devices such as RFID (Radio-Frequency IDentification), GPS, accelerometer, gyroscope, etc. can provide for human information and hence their subsequent detection. However, in the case of elders, the ailing and those in staying in assisted living or smart houses, enforcing such wearable devices may not be viable because of their health concerns, forgetful behaviors, etc. In such scenarios, human sensing needs to be facilitated in a non-intrusive manner. The non-intrusive human sensing approaches used in various applications may be categorized as:

1. Single modality approaches

   Here different types of sensors such as Pyro InfraRed (PIR) sensor, UltraSonic (US) sensor, vibration sensor, etc. are used by researchers to detect the presence of a human being [137],[94],[51]. Unfortunately, all these sensors have their limitations and fail to detect human beings under some circumstance or the other. It is intuitive that given all kinds of scenarios, a single sensor proves to be inadequate to sense the presence of a human being.

2. Multi-modality approaches

   Due to uncertainty, missing observation, and incompleteness of data reported by a single sensor (as cited above), there is a growing need to integrate and fuse multi-sensory information to achieve robustness in sensing [84][116][11]. Multi-modal human sensing is the method of using information from multiple sensors to detect human presence.

Human sensing involves the detection of the presence of at least one human being in a given environment. It also pertains to the process of extracting information about the people in the given environment [54].

This sensing involves a variety of technologies most of which do not mandate any intentional participation on part of the human, being detected. Though many methods have been propounded by researchers which can be used in a singular or stand-alone manner [54] none provide a foolproof and guaranteed solution to this problem. In an IoT, where the devices along with their associated sensors are scattered yet connected as a network, it is imperative to reinvent a new, co-operative and adaptive paradigm for human sensing. Such a paradigm will aid in promoting a harmony between human beings, societies and smart things [37].

1.1 Research Challenges

Some of the research challenges in non-intrusive human sensing are described below:

- ***Sensor Selection under the constraints of low-cost, portability and non-intrusiveness***

  Teixeira et al. [137] claim that false-positive and false-negative rates of the best approaches typically lie near the 10 percent mark for human sensing in uncontrolled environments. Tracking and identification are still a big research problem, as the base problem (presence of human being) is still not completely solved. To build a system which can sense human presence with high accuracy will thus be an achievement for the researchers. Human beings exhibit a dynamic motion pattern which makes it very difficult to predict their activities. This dynamism causes error and failure in a human sensing system. Human nature varies from person to person because of which their sensing is an open-ended problem till date. Ease to carry and deploy is an essential quality of a system. Besides, it should have low-cost and offer non-intrusive sensing. These features put an additional constraint and pose a challenge to the designer.

- ***Can human beings be identified using non-intrusive sensing?***
  Finding similarity between humans or the identification problem by itself is a research area. Finding the identity of a specific human being is a non-trivial problem [144]. Tracking or identifying a person involves differentiating the person from others around using similarity measures. The technical limitations of the sensors themselves often lead to a further loss of personally-identifying information in the acquired signal [137].

- ***How can a human sensing cope up with the change in environment, dynamic behavior and active deception of human being?***
  Unexpected or sudden changes in environmental conditions can cause errors in sensing. For instance, radar signals, can be dampened by rain or fog while those from PIR sensors are often triggered by heat currents flowing from HVAC (Heating, Ventilating, and Air Conditioning) systems. Kooij et al. [141] claim to solve the problems induced by dynamic environments using adaptive sensor fusion for human stance control. It is thus felt that there needs to be further investigation into such fusion to improve results in the domain of human sensing. Also, in adversarial scenarios, it is important to consider possible attack vectors which could fool or debilitate a human-sensing system [137]. Jamming signals, for instance, are often used in military scenarios to disable the enemy's radars and communication systems. Other deceptive techniques may be

as simple as turning off the lights in an area covered by cameras or movements of strolling animals to fool motion sensors. A truly robust human-sensing system is still by and large an unrealized goal. Humans are very dynamic. They can easily disappear from sensing area of sensors, and tracks are often terminated or, even worse, incorrectly extended in the face of ambiguities. Research on human sensing based applications (e.g. localization, tracking, and identification) is still in progress due to limitations or constraints in the proposed systems [102][82].

- ***Likewise, from each sensory data multiple features can be extracted to feed algorithm. Consequently, computational complexity goes high with increase in number of sensors. How to select the important features to reduce computation, complexity, minimize redundancy and to improve the accuracy of human detection is still an open problem.***

  Given a set of features for human sensing, the process of selecting the best-suited features from the current environment remains a challenge. The problem deteriorates when one has to deal with dynamic environments. Also, with the gained attention, multi-modal approach presents a challenge to design an algorithm which can process different sensory data with the minimized delay to cater to real-time applications. Designing an algorithm using a multi-modal approach to process data from different sensors in minimum time to cater to real-time applications such as fall detection, human detection during hazards, etc. presents a challenge.

- ***Can data for human beings be collected in every possible scenario under realistic conditions? If sufficient training is not available, how can a model be trained?***

  A significant amount of data is required to train a model. Being a dynamic creature, it is very difficult to capture sufficient amount of data for every possible condition including posture and motion of a human being. Thus, data collection is one of the hard challenges in human sensing.

  Researchers in the Artificial Intelligence community have struggled for decades trying to build machines capable of matching or exceeding the mental capabilities of humans. One capability that continues to challenge researchers is designing systems which can leverage experience from previous tasks into improved performance in a new task which has not been encountered before. When the new task is drawn from a different population than the old, it is considered to be transfer learning [26]. The benefits of transfer learning are numerous; less time is spent learning new tasks, less information is required of experts (usually human), and more situations can be handled effectively. These potential benefits have lead researchers to apply transfer-learning techniques to many domains [109] with varying degrees of success. In the domain of human sensing, where it is difficult to capture data for every human in every possible scenario, the concept of transfer learning [120] can be a useful tool.

- ***How do we select attributes of a human being that required to be measured?***

  There are various attributes of a human being which can be sensed using different sensors. With the changing environment, all the attributes of human beings cannot be sensed in all possible scenarios. For example, the weight of a human being cannot be measured everywhere. Similarly, the shape of a human being cannot be estimated in dark unlit conditions. Therefore, the selection of human attributes that need to be sensed to detect them is still an open challenge.

1.2 Research Aims and Objectives

This thesis aims to lay a foundation for catering to the challenges enumerated above. The thesis involves real-world experiments and proposes algorithms to solve the problem of human sensing in different scenarios. Such a solution can be implemented for elderly care, digital health care, assisted living, etc. The different objectives that have been tackled in this thesis have been listed below:

1. Hardware Selection for non-intrusive human sensing

2. Methodology to analyse the sensory data

3. Real world implementation (case study)

1.2.1 Hardware Selection for non-intrusive human sensing

Selecting hardware (sensors) to sense human presence under the constraint of low-cost, portability and robustness is itself an open challenge. The objective is choose sensor(s) in such a way that it should be able to sense human presence in different environments such as dark, illuminated, foggy areas both indoors as well as outdoor environments. This requires a feasibility study of all the available sensors along with their advantages as well as failures. What needs to be found this is a sensor or a combination of sensors that can sense human presence with reduced number of false alarms. In addition, the selected sensors should compliment each other in a way that if one fails the other will still cope to ensure success.

1.2.2 Methodology to analyse the sensor data

The meaning of the sensory data is dependent on the methodology used to interpret the data. Llata et al. [80] have used two ultrasonic sensors to achieve object classification. To complete this task the authors have proposed a fuzzy expert system having a dual knowledge base - one statistical and the other a standard rule based one. Using this technique, they extracted and used 25 features including parameters like the beginning time, mass center, etc. V.Matz et al. [2] describe the use of a Support Vector Machine (SVM) to classify objects into two classes. When ultrasonic homogeneous waves strike different types of surfaces inhomogeneous waves are reflected by these different surfaces. Authors have used the characteristic property of these reflected ultrasonic waves to extract features such as the mean and root mean square values, standard deviation and the absolute values to classify the objects. The second objective of the thesis is to provide a methodology that enable human sensing based on the incoming data with better selected features. Work using multi-sensor approaches demand survey of sensor fusion algorithms/models where challenges of sensor fusion were encountered. Adaptive sensor fusion techniques [117] form a feasible approach. The objective of this research is to explore the adaptive human sensing.

1.2.3 Real-world Implementation (Case Study)

The smartness of a living space of human beings has proved its advantage in several domains such as elderly care, traffic management, intruder detection, automation of devices in smart homes,etc. Smart environments could perform better if they are aware of the nature and behaviour of the human beings that inhabit them. This however is highly dependent on the mechanisms used for human sensing. The third objective of the thesis is to implement the human sensing system for real applications.

This thesis proposes solutions for the different challenges of non-intrusive human sensing. A brief description of the contributions of the thesis is provided in the next section.

1.3 Contributions of the Thesis

The significant contributions of the thesis are described in brief below.

1.3.1 Sensor Selection for Human Sensing

There are many sensors such as PIR, US, vibration, camera, etc. that can sense different traits of a human being. However, an ultrasonic sensor is an unexplored sensor for the detection of human presence. An ultrasonic sensor is mainly used to find the distance of an obstacle in front and to find the deformity in the metals. In this contribution, an analog ultrasonic signal has been exploited to differentiate human beings from non-humans in an indoor environment. This work involves analyzing the patterns of the ultrasonic waves reflected from the object located in front of the sensor. The task of human detection is performed using a Fuzzy Rule-Based (FRB) one-class classifier. Experiments performed with different human beings, cushioned chairs, cupboards, wooden ply boards, cardboard, glass panels, and cement and stone walls in the sensing environment have been described and results explained.

To improve the accuracy of human sensing, a voting based sensor fusion approach has been proposed. This approach fuses the sensory data from different sensors so as to aid the detection of a human being. To apply this approach, it was assumed that all the sensors point to the same target at the same instant of time and that data obtained from all the sensors is buffered at the same speed. Experiments were performed using both ultrasonic and PIR sensors. A comparison of results obtained with those of from other classification techniques such as Support Vector Machine (SVM) prove the efficacy of the proposed approach. This approach also outperforms the results obtained using individual sensors.

1.3.2 Multi-modal Human Sensing

A system incorporated with knowledge and intelligence can analyze, interpret and understand its surrounding environment and take appropriate actions or decisions. Such intelligent systems fundamentally rely on their sensory inputs and the underlying machine learning models, for understanding the situation and arriving at the relevant decisions or actions to be taken. In this contribution, a multi-modal sensing algorithm has been proposed that facilitates autonomous useful sensor selection for a given environment. A camera was used to show that this algorithm can work in scenarios where the use of intrusive sensors is not an issue. The analysis of the walking speeds of human beings is walking (in motion) was also studied. This in turn endorses the robustness of the approach.

1.3.3 A Feature Selection Mechanism

The process of Feature Selection has, of late, gained considerable attention because of a tremendous increase in the rate at which data is being generated. It is applied to a wide range of applications wherein the dimensionality and heterogeneity of the data involved are high. Given a set of candidate features for a real application, selection of the best-suited features from the current environment remains a challenge. For robotic applications where the environment can change dynamically, features selected *a priori* may not perform equally well in different environments. In this contribution, two different approaches for feature selection have been proposed. The first one is a *Reward and Penalty based online approach* while the second is an *Immuno-inspired Online Mechanism.*

A Reward and Penalty Based Approach for Online Feature Selection

This involves an approach that allows a system to learn the best-suited features *on-the-fly*. Once the system senses the change in the environment, it starts to learn the best-suited feature for the new or changed environment. This contribution focuses on the selection of the best-suited features out of an initial set of candidate features for a given environment. The proposed approach was tested for a service robot that needs to recognize a human being in various indoor environments. Features are suppressed and enhanced on the basis of rewards and penalties in a supervised scenario. A comparison of classifiers with online feature selection and classifiers without feature selection was performed to assess their performances in terms of processing time and accuracy. Results obtained revealed that for the given set of sensors, the combinational rule-based system outperforms the Support Vector Machine (SVM). Also the accuracy of the system trained in a specific environment falls, when it is tested in a different environment. It was found that the performance of the system is high if the training and the testing environments are the same. This essentially indicates that the system is dependent on the environment and needs to be trained again for every change in the environment. This would mean an addition of a new database to the system every time the environment changes. Such a system would eventually run out of memory. What is thus desired is a system that can adapt to the new environmental conditions without the need for retraining.

An Immuno-inspired Online Feature Selection Mechanism

This contribution describes a novel Immuno-inspired Online Feature Selection (iFS) approach to select the best-suited features for human sensing. The approach uses an Immune Network [59] as a base to realize an autonomous and self-learning solution to online the Feature Selection Problem (FSP). Unlike the previous approach which relied on a fixed set of predefined training data, this approach does not require off-line training of the model and can respond to the incoming data (antigen) and deliver the best combination of features (set of antibodies).

Experimental results validate the efficacy of the proposed method. The experiments were carried out so as to classify human beings from non-human where results are obtained with a reduced false positive rate. A comparison of results with that using an SVM indicate the proposed approach to be far better.

1.3.4 Human Behavior Analysis for Health and Wellness Detection

With the different available sensors, human presence can be sensed in different scenarios. At a high level, human sensing can be used to detect the different activities performed by a human being. When recorded with time, the data of the different activities performed can be analyzed for his/her behavior profiling. The behaviour profiling can further be analyzed for his/her health and wellness detection. This contribution focuses on an application of human sensing in the domain of health and wellness detection. Nutritious living and access to opportune health care are of major concern for the geriatric and solitary population. Health care given at early stages of any disease has proved to be life-saving. Consequently, the sooner the symptoms of a diseases are detected, better is the care that can be provided to the patient. Elders are more prone to diseases for conspicuous reasons and consequently, the geriatric population demands more attention and care. Mild cognitive impairment (MCI) is one of the symptoms commonly found in elders. Patients with MCI and recollection complaint are consistently more likely to progress to dementia, concretely of the Alzheimer's variety. However, differentiating between normal aging and the development of MCI is still a research challenge. If activities of a human can be monitored perpetually, then the deviation in the routines of the elderly can be used to track the progression of MCI. In this contribution, an auto-encoder based approach has been proposed, to reduce the data dimensions for the detection of deviations in the daily routines of the person. Abnormal behavior is indicated if the abnormality in the behavior persists over time. Results were cross checked with the medical data of patients which seemed to agree with the predicted results.

1.3.5 Transfer Learning in Smart Home Scenario

As mentioned earlier, it is not possible to collect data for each human in every possible scenario. This is a cause of insufficient or no training data required for a model to sense human presence in some scenarios. To deal with this issue, transfer learning has been explored. The concept of transfer learning is inspired from human brain which can leverage the knowledge gained from previous task to accomplish or improve the performance of tasks which have never been encountered before. This requires a transfer of useful knowledge from the source domain to the target domain where training data is insufficient which holds true when a new user enters in a resident space. To apply the concept of transfer learning, different parameters such as architecture similarity, number of members, modalities of the sensors, a similarity in the routines, etc. were calculated. These parameters decided the source domain from where knowledge can be transferred. Experiments were carried out using collected data. Experimental results indicate clearly that the use of transfer learning provides for flexibility to a system when there is either no training data or insufficient data.

1.4 Thesis Organization

The chapter wise organization of the thesis is given below:

- **Chapter 1:** This chapter provides an introduction and overview of the thesis.
- **Chapter 2:** This chapter provides a survey and motivation behind the research. For pedagogical reasons, a few terminologies and their origins are discussed.
- **Chapter 3:** This chapter presents the use of an ultrasonic sensor followed by a multi-sensor approach to realize human sensing. The chapter is based on the work published in [130] [129].
- **Chapter 4:** This chapter presents a multi-modal human sensing approach where a system autonomously selects the sensor with changes in the environment. The contents of this chapter are based on the published work reported in [132].
- **Chapter 5:** This chapter presents approaches for feature selection to improve the accuracy of human sensing by shedding the redundant features. The work presented herein proves that in different environments different features are selected. The contents in this chapter have been published in [123] [128].

- **Chapter 6:** This chapter introduces the application of human sensing to detect early symptoms of mild cognitive impairment. The approach is tested in the real-world, where predicted results were found to agree with the medical reports. The work is published in [127] [126].

- **Chapter 7:** A mechanism for Transfer Learning in the domain of smart homes is proposed herein. The approach adds flexibility to the system so as to support a new entry.

- **Chapter 8:** This final chapter highlights the conclusions arrived at. A summary of the contributions made is presented and new avenues for future research have also been described.

2

Literature Survey

With the growing internet usage, billions of devices are getting populated and networked over the internet. Because of ease of data exchange among devices, systems are getting smarter and more efficient. In other words, environments are becoming smart for example smart traffic systems, smart working area, smart living room and smart hospitals are the highlighted problems of the current research. Smart devices are contributing towards smart systems by interacting with environment through sensors. The sensor network and ubiquitous computing communities increasingly focus on creating environments that are seamlessly aware of and responsive to the humans that inhabit them, the need to sense people will become ever more pressing. Human sensing encompasses issues from the lowest level instantaneous sensing challenges all the way to large-scale data mining for behaviour analysis. The simplest applications of human sensing make direct use of such information to, for instance, sense the human presence in dark to turn lights on/off when a room is occupied/empty, or lock a computer machine when the user moves away. This chapter presents a literature survey of the techniques and applications used in this thesis.

Applications like intruder detection, elderly fall detection, health care applications, human tracking, etc. solely depend on the detection of human beings within the given environment. Based on the application at hand, researchers have used a variety of sensors for human presence detection. Sensors to be used for human sensing can be categorized in intrusive and non-intrusive sensors as shown in Figure 2.1. Intrusive sensors are the sensors which needs active participation of users for the output. Non-intrusive sensors are the sensors which do not intrude in a human life but still a part of human daily living.

2.1 Intrusive Sensors

Intrusive sensors can further be classified in two subcategories that are vision based sensors and wearables.

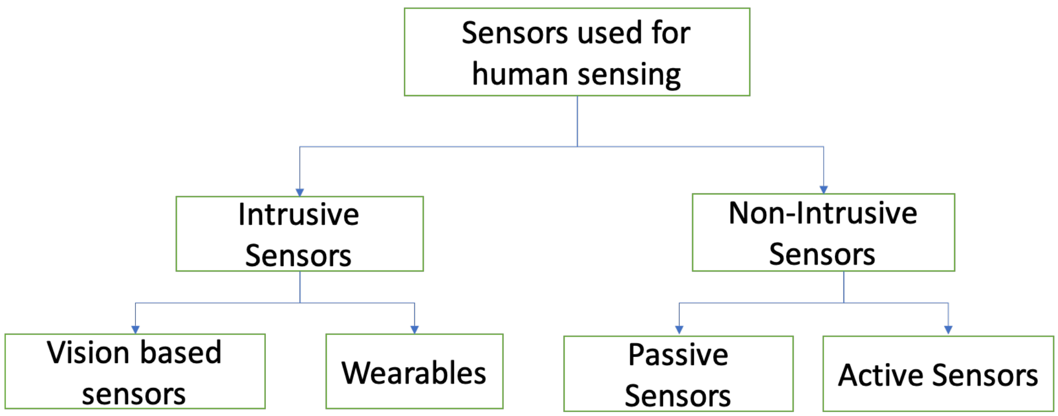


Figure 2.1: Categorization of sensors used for human sensing

2.1.1 Vision Based Sensors

Concerning human sensing, vision based sensors are considered as the most reliable and accurate sensors [137]. Visualization of human movement and human presence has become an important area in the domain of computer vision. The motivation comes from the plethora of applications lie ahead and the technical enhancements which enables the real-time capture, transfer and processing of vision based data. In 1999, Gavrila et al. [44] presented a survey on the usage of vision based sensors for motion analysis to sense human presence which shows the progress in the imaging techniques over a decade. 2D imaging and 3D imaging techniques has been introduced to improvise the vision based human sensing [58][153]. Starting from pedestrian detection, occupancy detection, intruder detection, human identification, human computer interaction through assisted living, border security, robotic navigation and many more, vision based sensors are exploited by researchers and claim to achieve a good accuracy [58][153][135][148][20]. However, in each of these cases the associated computations that include image segmentation, morphological operations and image filtering, are computationally heavy and thus time consuming [137]. Since illumination is mandatory, these techniques cannot work in the dark. These sensors however are unsuitable for applications wherein privacy is a matter of concern. Smart homes, where a fall inside a washroom needs to be detected, fall under this category. Vision sensors also fail to serve their purpose in rainy or dusty environments. Such sensors can also be fooled into detecting idols, mannequins or a robot, as human beings. Though the technique described therein using vision based sensors are effective, the cost of the gadget (such as kinect sensor) and the related computations involved does not make it a practically viable solution.

2.1.2 Wearable Sensors

Humans may themselves be accompanied by smart devices such as mobile phones, use surface-mounted devices (wearable computing) and contain embedded devices (e.g., pacemakers to maintain a healthy heart operation or AR contact lenses). Wearables are the smart devices that are either attached with the clothes or worn by human beings [113]. Wearables have been extensively used in the domain of healthcare by considering multiple practical challenges [121]. As a consequence, there are now numerous applications [121], services [121], and prototypes in the field. Rawassizadeh et al. [113] have addressed the current era as an era of smart watches whose evidences are provided by the Pebble story [16]. With the increasing demand of devices such as Google Glass, Pebble, Fitbit, etc., wearables market is expected to have massive growth in the next few years [105]. With the multiple functionality such as health data, message display, tracking, fitness accountability wearables are an integral part of Internet of things to estimate number of people, number of cars, device automation, etc. Using wearable devices, an user's and his surroundings data can be collected ubiquitously and continuously. This raises questions and present challenges on privacy concerns of a human being [10]. Motti et al. [93] have mentioned about different privacy concerns a human being might face. Despite of ease, wearable devices are not appreciated in the domain of non-intrusive human sensing. For a wearable user, it is mandatory to wear that particular device which is always not possible in case of an elderly person or a sick person [137]. As mentioned earlier, wearable devices can invade the privacy of a human being therefore can not be used in every scenario.

2.2 Non-intrusive Sensors

These sensors gained the attention of researchers working in the area of a digital citizen for health and wellness detection of a human being. PIR, US, vibration, laser, etc. sensors falls under this category. Unlike vision based sensors, these sensors do not invade the privacy concerns of a human being. Therefore, non-intrusive sensors are preferred over intrusive sensors. Concerning human sensing, non-intrusive sensors can be categorized into Passive sensors and Active sensors.

2.2.1 Passive Sensors

Passive sensor is a device that does not emit or illuminate the target but senses the vibration, heat, temperature, etc. in the target's environment. Passive sensors includes Pyro Infrared (PIR) sensor, temperature sensor, gas sensor, light sensor, vibration sensor, water sensor, etc. These sensors are embedded in the living environment of human beings to sense their presence and absence at different locations. A PIR sensor is one of the widely used sensors to detect the presence of a human being in an indoor environment [92]. They have also been used for other purposes such as detection of a the falling of a person [107], to count the number of human beings [142] and human tracking [152]. A PIR works on the principle of sensing the IR waves emitted by living beings. PIR sensors [151] [49] have been used for human detection by researchers working in the domian of human sensing but the system fails for still objects. Further PIRs can detect only one human being per change [137]. PIR works based on heat differential and hence fails to distinguish between objects - living or nonliving - which emit similar heat waves. This can mislead the system into generating a false alarm.

Vibration sensors, which work on the amount of force applied, have been used to learn walking patterns [4], fall detection [3], etc. Similarly, other passive sensors are also used for human presence detection in different scenarios and for different applications such as occupancy detection [53], border security [65], human tracking [124] etc.

Laser sensors have been used by Shao et al. [122] for human detection. The outline of the human being is used to differentiate humans from non-humans. The use of lasers may however pose health issues.

2.2.2 Active Sensors

Unlike passive sensors these sensors respond to some input for example Ultrasonic Sensor (US), radar, laser, etc. As on date, not much work has been carried out to explore the use of a US for human detection. A US uses the difference between the energies of the emitted and reflected ultrasonic waves. Waves of a specified ultrasonic frequency are bombarded on target objects (which are placed at equal distance from source of ultrasonic waves). The reflected waves from these objects are heterogeneous in nature. This heterogeneity arises from the fact that the objects have different absorption coefficients for ultrasonic waves. Such sensors are thus widely used in detecting deformity in metals [160] and differentiating surfaces [68]. Still face detection using ultrasonic sensing [90] have also been carried out for a limited number of faces, using features such as count above a certain threshold, average acoustic area, etc. However, just like the others, these sensors too fail to provide the desired high true positive rate (TPR) and low false positive rate (FPR). Ultrasonic sensors are mainly used for obstacle avoidance in the robotic domain.

A human being can be considered to be very dynamic in the sense that s/he can assume numerous body postures, walk at varying speeds, wear a range of widely varying clothing material and even think of ways to escape detection [137]. The complexity of detecting human presence increases when we take into consideration the aspect of predicting human behavior. Unfortunately, all sensors have their own limitations and fail to detect human beings under some circumstances. One may thus conclude that merely using one sensor for human detection may not be the right solution. Data from numerous heterogeneous sensors could be used to arrive at a more reliable and conclusive evidence of human presence. Fusion of sensory data thus becomes a sine qua non. The need of the day is a fusing sensor values generally would mean the construction of better feature vectors [156] which in turn can lead to more accurate detection. To increase the accuracy and efficiency of the smart devices multiple sensors are required. As humans, either individually or collectively, inherently form a smart environment for devices.

2.3 Multi-modal Human Sensing

Researchers have been known to use a range of other sensors for human detection which include PIR, vibration sensor, acoustic [21], etc. As can be observed, sensors have their respective and distinct advantages and limitations. Because of individual limitations, a single sensor is incapable to sense human presence in every possible scenario. Sensor fusion [137] is a possible alternative, which researchers are now exploring [21] [1]. This paradigm uses a combination of sensors, such as a PIR and a camera [29], PIR and a vibration sensor [159], etc. and interprets the data received from each of them to eventually conclude on a detection. Several applications exist where a combination of PIR and other sensors have been used for human detection. These include surveillance [7], positioning [38], etc. Ahmet Yazar et al. [150] have made use of a combination of PIR and vibration sensors , while Nadee et al. [96] have used infrared and ultrasonic sensors to detect the fall of elderly persons. Home appliances have been automated based on task activity detection of human beings using multiple sensors in [101]. A survey by Avci et al. [4] cite the use of multiple sensors such as PIR, ultrasonic, vibration, temperature sensor, etc., to monitor daily activities of human beings. Similarly, a survey on wearable devices presented by Lara et al. in [76] emphasizes the utility of multiple sensors to monitor the health and wellness of human beings. While adhering to the basic concepts of building a system for human detection through successive layers, the task-achieving behavior of a system can be fragmented into many smaller decision-making units [115]. Each unit has an input that needs to be converted to the output using analytical data techniques. Data received via different resources need to be processed using different techniques [40]. Matthews et al. [87] describe a system where different algorithms are implemented to process the data from different sensors. A voting based approach is explored in [129] to process the data from both ultrasonic and PIR sensors to detect the human presence. One may thus conclude from the literature that data received via a variety of sensors need to be processed in different ways so as to reliably detect human presence. Yang et al. in [149] have used the face, body appearance and silhouette via a Kinect sensor and multiple color cameras to detect human beings in a health-care application. Using a Kalman filter they claim their multi-modal approach to be more effective than those reported by others. However, their sensor combination has been tested only in a controlled environment. Human detection has also been performed in robotics where sensors mounted on a robot try to detect a human presence for assistance. Human detection has been performed by fusing the data from a laser sensor (to detect leg structure) and a camera [9]. Unattended ground sensors have been used by Jin et al. in [63] for human sensing. In the approach proposed herein the primary level features from individual sensor signals are extracted and then combined to form composite patterns based on relational dependencies between them. The advantage of this approach is that the system can function (with reduced accuracy) even in case of a failure of a sensor. Vibration sensors come to the rescue when the PIR sensor fails to differentiate between a human being and animals. Occupancy detection which solely depends on human detection has been demonstrated by Candanedo et al. [15]. They have presented a detailed analysis of all the sensors and pairwise sensor combinations via a correlation matrix. Analysis shows the performance of the different statistical models concerning different combinations of sensors to sense the human presence. The authors focuse on choosing the combination of sensors with the highest accuracy.

Applications targeting smart homes [54] [79], surveillance systems [7], etc., all use a combination of ultrasonic and PIR sensors. Apart from choosing a suitable combination of sensors, the selection of an appropriate classification technique to suit the application at hand is crucial. Maslov et al. [85] present the management of sensory data for real time analysis of the environment, wherein they use Dempster-Shafer theory [95], fuzzy rules [1], Bayesian Belief Network [118], etc. techniques to fuse the data from the sensors. In order to classify the data among different classes they use several algorithms which include Fuzzy Rule based classifiers [131], Support Vector Machines [41], etc.

The use of vibration sensors is however infrastructure dependent [159] and thus can be used only in some specific scenarios. On the other hand, ultrasonic sensors are not dependent on infrastructure. The better option possibly is to combine the use of both PIR and ultrasonic sensors. As can be observed, both PIR and ultrasonic sensors have their respective and distinct advantages and limitations. This means that their combined use can result in the realization of a better system.

As discussed, every sensor has its limitations. Sensor fusion is one of the viable solution to overcome the limitations of an individual sensor. It can be inferred that likewise to sense human presence in every scenario multiple sensors can be populated together. However, higher the number of sensors more will be the complexity of the system. An adaptive system which can autonomously select the sensor(s) corresponding to an environment is a challenge in front of researchers.

2.4 Adaptive Human Sensing

Research performed only with single perceptual sensor has its inherent limit of capabilities. Due to the possible weaknesses of uncertainty, missing observation, and incompleteness of single sensor (as discussed above), there is a growing need to integrate and fuse multisensory information for advanced systems with high robustness and flexibility. Sensor fusion is a method of integrating signals from multiple sources. It allows extracting information from several different sources to integrate them into single signal or information. The proposed method for multi-modal human sensing is required to integrate with Adaptive Technology mechanisms to enable it to dynamically adapt to changes that may occur in real time resulting in better data analysis and processing by fusion techniques and, consequently, in a data series with better quality for decision-making systems. The method focuses on sensor fusion in order to obtain inferences about the environment in a complementary way, trying to infer information from different sensor sources (different properties) in a dynamically changing environment. In order to exploit advantages of both fuzzy-logic as an outstanding intelligent method, and Kalman filter as an efficient fusion method, suggests a hybrid Kalman filter [146] fuzzy logic adaptive multisensory data fusion architectures. In an Internet of Things (IoT), where the devices along with their associated sensors are scattered yet connected as a network, it is imperative to reinvent a new, co-operative and adaptive paradigm for human sensing. Such a paradigm will aid in promoting a harmony between human beings, societies and smart things [42].

Working with sensors refers to feature extraction from the received sensory data. For example Working with ultrasonic sensors refers to feature extraction from the received wave which is reflected by the confronted object. Likewise several features can be extracted from sensory data. However, all the features extracted might not be useful for a built model to sense human presence or absence. This calls for an appropriate method for feature selection for further enhancement of functioning of a model. Feature Selection has recently gained considerable attention because of a tremendous increase in the rate at which data is being generated. It is applied to a wide range of applications wherein the dimensionality and heterogeneity of the data involved is high. The feature selection problem refers to the selection of a subset comprising a fixed number of candidate features that optimizes on the evaluation measures while also ensuring that there is no degradation in performance compared to when the supersets are used [47]. It thus strives to reduce the dimensionality of a given set of features without compromising on performance. Given a set of candidate features for a real application, selection of the best suited features from the current environment remains a challenge. To build a predictive model, classification algorithms such as decision tree, Support Vector Machine (SVM), etc. relies on the feature set to be used for classification. Efficiency of predictive modelling can be improved by shedding redundant features, thereby alleviating the ill effects of dimensionality. This also results in improved performance of the model especially in terms of its learning rate and accuracy. Feature Selection has also proved its relevance in Internet of Things based (IoT) scenarios, where an increasing number of sensors generate high amounts of data. With each node in an IoT having limited resources, both in terms of computing and memory, a reduction in features is bound to create an impact on its performance and utilization. A number of statistical and machine learning algorithms for feature selection have been proposed [66],[30],[17]. However most of them rely on features selected *a priori* thereby depicting the importance of online feature selection.

Feature Selection can be performed either as a preprocessing step [39] or by removing redundant features during the learning phase of the model [86], [125]. Conventional methods of feature selection rely mostly on prior selection of the features wherein it is assumed that some features are not that important and are removed due to their redundancy. Such methods aim to compute the best suited features off-line and hence the associated training data needs to be made available a priori. This off-line methodology fails in scenarios where the data evolves over the time and is thus not available a priori. Such online feature selection mechanisms can be of high utility in real-world applications where one has to manage with high dimensional data which streams in continuously such as in an online intrusion detection system. The conventional batch selection algorithm [66] cannot be applied in such cases. For IoTized scenarios where the environment changes dynamically, online feature selection remains a challenge. The goal is thus to develop an online feature selection algorithm that learns the pattern of incoming data and select the best suited subset of given candidate features.

2.5 Feature Selection

Over the years feature selection has been an effective means to deal with high dimensional data. Conventional batch selection algorithms [30] fall under the category of feature selection methods termed preprocessing/filtering. These take into account distance, dependency and consistency for the feature selection process. Later any classifier [52] can be used for the predictive modelling. The wrapper type methods for feature selection [157] rely on the heuristic search within the candidate feature search space. In this category, the features selected are dependent on the classifier used [30][143]. One of the most effective methods in this category is Genetic Algorithm (GA) based wrapper [57]. This method relies on the initially selected population and improves on it to deduce the best suited subset of features. These methods however entail heavy computations. Zhu et al [157] have proposed a method to combine the filter and wrapper methods to reduce such computational complexity while maintaining high prediction accuracy. This method however does not keep track of the older populations that were rejected. It is thus possible that the entities in the older population reappear later causing a never-converging loop. GA is dominated by Artificial Immune System (AIS) which has been applied widely in many areas. AIS is an optimization algorithm, which maintains the diversity well as compare to GA [57].

Kuo et al [73] describe an artificial immune system based feature selection method using back-propagation neural network [14]. This method considers the affinity between antigen and antibody to calculate the weights in neural network. Similar to earlier proposed methods, this approach too depends on prior training data named as memory cells and based on the affinity between antigen, antibody and new antibodies, memory cell mutation occurs to select the best suited features. Others have attempted to use the Clonal Selection theory [12] to select the best suited features [5]. Clonal selection considers the attraction between antigen and antibody as a driving force to calculate the best antibody which further proliferate and clone itself. Whereas, Jerne [60] proposed an Idiotypic or Immune Network theory which takes into account the interaction among antibodies that eventually form a network unlike clonal selection. The proposed mechanism has been used for a range of applications which include Robotics [62][97], Intrusion detection [32], [138] etc. Lu et al. [83] present an AIS based feature selection but have not specifically pointed as to where a network is formed. They have not defined the stimulation and the suppression of antibodies based on the antigenic attack, and simply check the attraction of antigen toward antibodies thereby based on that selected the highest attracted antibody. The highest selected antibody proliferate and clone itself. However, in an immune network, even in the absence of an antigenic attack candidate antibodies interact with each other to form an idiotypic network and as a result of this interaction among antibodies, antibodies are stimulated and suppressed. This stimulation and suppression can further affect the life of an antibody, which plays a major role to decide the concentration of different antibodies. Removing redundancy in features must not be compared with other methods which are used for reduction of dimensionality such as Principal Component Analysis (PCA) since good features may not be dependent on the rest of the data.

Advances in sensor technology has resulted in the availability of a range of sensors that can sense a wide variety of parameters. Selecting the right sensor(s) for an environment can at times pose a challenge. Features are the extracted values from the raw sensor data such as maximum, median, average etc. Complexity of such a problem further increases when a human being enters the scenario. This is so because a human being is a highly dynamic entity and has virtually complete autonomy. Finding patterns in the manner of movement of a human being is thus a convoluted task.

As discussed last decade has marked an exceptional development of microelectronics and computer systems, enabling sensors and mobile devices with unprecedented characteristics. The advent of small size and low-cost sensors have now made it possible to embed them in the living space of human beings to add comfort to their daily living. This originates the research area of *Ubiquitous Sensing* which involves analyzing the pattern and behavior of human beings using data collected from such pervasive sensors [36]. Within this domain of Ubiquitous Sensing, the recognition and prediction of human activities in smart homes, hospitals, military fields, workspaces, etc., has become a task of high interest. To monitor the in-house activities of a human being, sensors are embedded in his living space. Patterns in the activities of a human being are studied for medical purposes such as fall detection, early symptoms of dementia, etc. Gu *et al.* [46] using the emerging patterns from the sensor values that could be associated with the recognition of activities. In smart homes researchers are doing human behavior analysis to health analysis, appliances automation, activity tracking, energy management, assisted living, etc. Smart homes technology has proved its significance in the domain of health care. Healthy living and access to advantageous health care are of significant concern for the geriatric and unaccompanied population. Health care given at early stages of any disease has proved to be life-saving. Consequently, sooner the symptoms of any diseases are detected, better the care can be provided to the patient. Elders are more prone to diseases for conspicuous reasons, and consequently, the geriatric population demands more attention and care. Mild cognitive impairment (MCI) is one of the symptoms commonly found in elders. MCI with recollection complains, and deficits are consistently to have most likely progression to dementia, concretely of the Alzheimer's variety. However, differentiating between healthy aging and development of MCI there-in is a research challenge. MCI affects the behavior of the human being. Thus this abnormality in the action can be a key for early detection of dementia and symptoms thereof. Also, it has been proved by researchers that there is a markable difference in the performance of daily routine activities carried out by a healthy aging person when compared to that of an MCI patient. To continue with the conception if activities of a human can be monitored continuously, then diversion or eccentric activities of the elderly can be used to track the progression of MCI in an elder. To monitor the activities of the denizens different kind of sensors are being embedded in the infrastructure. In literature, researchers have endeavored several techniques to detect the symptoms of MCI. One of the most prevalent approaches used is to interview the targeted individual. This approach has shown good results, but this approach may discomfort the elder by asking questions again and again. Furthermore, one may feel offended by the specific type of question which vary with the individual. Another approach to deal with the issue is to embed the sensors in the living area of humans to monitor the activities of the human unobtrusively. A routine of the human being can be analyzed via pervasive computing for health and wellness detection. However, as sensors are circuitry devices have chances of failure which might lead to information loss and can increase the number of false alarms.

Also, a numerous number of sensors are available that can be used to monitor activities. For obvious reasons more the number of sensors, more accurately the activities of the human being can be monitored. However, with an increasing number of sensors, the amount of data to be processed is also increasing exponentially. Therefore, machine learning models that are capable to deal with an enormous amount of data are required for the processing. Deep learning based models are preferred over other machine learning techniques to deal with huge data-sets.

2.6 Mild Cognitive Impairment

Gaulthier et al. and Peterson et al. [43] [104] provide the clinical interpretation of MCI along with the related concepts and pathways. However, Mattson et al.[88] provides a study on the prevalence of Alzheimer's biomarkers in patients with MCI. Furthermore, a review of the progression of MCI and how behavioral interventions can help at various stages is explained by researchers in [139]. These prior works help to establish MCI as a severe clinical issue in geriatric population and further indicate that the progression of the disease is slow and occurs over several years. In recent times, the study of ADL based detection of cognitive impairments have gained momentum due to their ability to track patient progress in pervasive environments. Riboni et al. [114] provides a study on detecting mild cognitive impairment using particular activity parameters; however, elucidating that level or detail requires extensive monitoring which may prove cost ineffective. Furthermore, authors in [27] provides insights into the development of a tool to evaluate MCI by using ADL grammar functions. However, the detailed longitudinal study with the practical use of sensors is not discussed at length. Lotfi et al. in [81] provides insights into sensor data modeling for smart home environments to help patients with dementia, including generating alerts on aggravation. A study on smart home based MCI detection is provided in [140]. Kosucu et al. in [71] uses the paradigm of opportunistic sensing for detection of MCI, which may become important when using low power or battery-less wearable that harvest energy, to detect MCI in smart home environments. The paper provide a foundation for how ADL can be used to perform indicative studies on MCI, which leads to the belief that a longitudinal study on ADLs may be useful in early detection of the advent of Mild Cognitive Impairment. Eagle et al. in [35] shows that routine is a fundamental part of human behavior and hence, gives us the principle that variance in routine can be a behavioral marker for MCI. We have based on work partly on the assumption that a higher variance in the routine of an individual is indicative of a cognitive decline. However, the work also shows how every individual is different and hence has to be baselined against self. So, the detection process should involve a gradual but steady decline in the orderliness of the ADL or daily/weekly routine of the subject. Langevin et al. in [75] provides a review of techniques for modeling time-series data using principles of deep learning, we have adopted similar methods in our current work to express human routine as a time-series inference. We also use time-series prediction techniques to fill-in missing sensor values in cases where the sensor data is not available due to fault in the sensor circuit or data acquisition system.

Research in the domain of ubiquitous sensing, to sense the behaviour trends of a human being is at a mature stage. This requires sufficient amount of training data. As every human being is different so are their varying needs. Therefore, a proposed solution should be tailored to an individual. As per conventional approaches, realization of individual specific solution requires substantial amount of data which is a challenging task. Therefore, for such a solution, first all the sensors should be in-place to collect the training data for individual human being. For example, let us assume that an activity recognition module is in place in a smart home and this module is trained on data of current residents. In a scenario where new residents enter the smart home, the activity module might not provide desired results. One inefficient solution is to collect data for the new residents and retrain the module. A better alternative is to design a system that can leverage the previous knowledge for the performance improvement of the current task.

Human beings has the mental capability to use the previously gained knowledge to perform a new task which is never encountered before. Over the last decade researchers are struggling to make intelligent machine to match human brain capabilities which can leverage the past experience into the performance of a newly assigned task. When the experience of old task(s) is utilized to solve a new problem, this is considered to be *Transfer Learning (TL)*.The concept of transfer learning is useful in case of insufficient or no data. To implement this concept knowledge from source domain can be utilized in the target domain. However, it depends upon the similarity between source domain and target domain.

In the particular case of smart homes, the concept of transfer learning can be explored to set up a new house. The experience gained from the already running smart homes can be utilized to predict the sensors reading in the new house. This concept wave off the time required by conventional approaches for collection of training data. Cook et al. [26] presented a survey on transfer learning for the activity recognition where the authors emphasizes the challenge of calculation of similarity between the source domain and the target domain. The level of similarity between source domain and target domain is dependent on number of sensors, amount of the data in the source domain, modality of the sensors, the relative placement of sensors, etc. Specific to the domain of smart homes, transfer learning depends on the amount of labeled data available in the source domain. Unlike this, in other domains researchers are successful to use the concept of TL by leveraging the unlabeled source data to make improvements in the target domain [26][100].

As applicability of transfer learning, two natural questions related to transfer learning arise. First, can a generalized method can be proposed to calculate the difference between the source and target populations? To calculate the differences some domain-specific distances have been proposed in the past. However, those can be applied while working in a particular domain. While defining a generalized method to calculate the differences between a source domain and target domain need to indicate differences in term of feature spaces, temporal spaces, label spaces, etc. This measure can provide comparisons between various TL approaches and can indicate that whether TL can be applied to a given situation. Second, can we detect and prevent the occurrence of negative transfer effects. TL can also degrade performance instead of increasing performance. Above two points are related to each other, because an accurate distance metric may provide an indication of affects of transfer learning.

2.7 Transfer Learning

The area of transfer learning addresses the problems when there is insufficient or no training data for example a scenario where a new person enters a house, there is no or insufficient data corresponding to his activities. In the case of lacking data, it is difficult to predict the behavior of the new entry. To cater to this problem, Transfer Learning (TL) is adopted by the researchers for a system to be flexible to support a new entry. Inspired by human intelligence, transfer learning can be defined as an ability to identify the deep, subtle connection between two contexts/domains/tasks. Transfer Learning term is first used by Thorndike and Woodworth [103]. Researchers[109] focused on development of Transfer Learning algorithms to reduce the labeling efforts. This requires a transfer of useful knowledge from the source domain to the target domain where training data is insufficient which holds true when a new user enters in a multi-resident space. The concept of transfer learning for activity recognition has been successfully applied to set up a new smart environment [112]. The concept of transfer learning is used by researchers [56][55] [26] in the field of activity recognition data collected using vision based sensors. However, processing the data of vision based sensors raises issues related to the privacy concerns of the people in the source domain.

The knowledge is transferred from the houses where ample amount of labeled sensory data for each activity is available. Researchers have successfully applied the concept of TL within the sensors of the same modality. However, problem of cross-modality transfer learning is still a challenging problem to solve. Kurz et al. [74] proposes a student/learner model to address the problem of transfer learning in cross sensor modality domains. Similarly, to cater to the similar problem Hu et al. [56] proposed a transfer learning approach. However, the concept of transfer learning to predict and analyze the activity routine for health and wellness detection using unlabeled data is new. Work focusing on transferring across difference in time [99] [98] [72], human difference [111][50][19][110], and devices difference [154][155] has been published. Transfer learning does not limit the number of resources. Therefore, number of sources can vary from single to multiple. Because of the involvement of physical settings, finding relation between source domain and target domain is harder as comapre to other domains. In this domain, type of sensors used, placement of sensors, number of sensors, way in which a human being performs an activity also plays an important role in the calculation of similarity between a source domain and a target domain [26]. The one of first sensor which is used to learn the activities of a human beings is video camera. Based on the data received, similarity between a source domain and target domain is calculated using spatio-temporal features [147]. However, this sensor invades into the privacy of a human being. Also, for a camera to track a person, its position, angle of orientation also affect the collection of data pertaining to the activities of a human beings [91]. Similarly, wearable sensors and non-intrusive sensors are also used to capture the data for the activities of a human being [34]. Sampling rate, sensor modalities can be considered as spatial features which can be further used to calculate the differences and similarity between the source and the target domains. At the same time, apart from spatial features, temporal features, sensor types, labels and devices cannot be neglected to calculate the similarity between a source and a target domain [26]. As compared to other considered factors, amount of labeled data and transferring the knowledge across different labels has gained attention of researchers [111][34]. However, the question arises whether transfer learning can be performed using unlabeled data, which is being discussed in this paper. The proposed approach relies on the representation of data in a vector form in such a way that labelling is not required. It can possibly give output in terms of time, sensors active and active sensors location. The same representation has been exploited for transfer learning where data is not available.

One of the most accepted work for transfer learning in activity recognition is Teacher/Learning TL [26]. As per proposed approach, earlier trained model work in parallel with new model where old model provides labels to train the new model. The mention approach requires sufficient amount of labeled data to train the Teacher (old) model. Two different types known as *inductive* and *transductive* learning are defined by Pan and Yang [100] for TL techniques. As explained by researchers, in inductive learning, knowledge is transferred regarding model parameters and predictive functions. On the other hand, in transductive learning knowledge of data instance is transferred to the target domain. However, most of the work in this area leverage the amount of labeled data to calculate the distance between a source domain and a target domain. Despite the abundance of work in this domain, research which is built on the unlabeled data in the source domain is very sparse.

2.8 Chapter Summary

This chapter discussed about the non-intrusive human sensing. The chapter mention about the previous work done by researchers using single as well as multiple sensors for human sensing. To improve on the accuracy, the concept of feature selection and its related literature is explained in this chapter. This chapter also discusses about the mild cognitive impairment which is an application human sensing. However, in the latter part of the chapter, concept of Transfer Learning in the domian of smart homes has also been explored.

"Research is to see what everybody else has seen, and to think what nobody else has thought."

Albert Szent-Gyorgyi (1893 - 1986)
Hungarian physiologist

3

Sensor Selection for Human Sensing

With advances in technologies, environments and human habitats are all set to become smarter. Such environments, which include smart homes, hospitals, campuses, etc., are oriented towards human comfort and safety. To realize such a smart environment, one of the fundamental tasks is to sense the presence of human beings in the environment automatically. Detecting the presence of a human being in spaces within such environments forms a major challenge. Though there are sensors that can perform this task, they are not without limitations. Using information derived from either single or multiple sensors separately is not sufficient to distinguish human beings from other objects within the environment. Reliable detection of human presence can be achieved only by fusing information obtained from multiple sensors. This chapter discusses an approach for human detection.

Ultrasonic sensors are mainly used for obstacle avoidance in the robotic domain. They have also found use in recognizing human activity. The main concept behind its effective use lies in the heterogeneity of the reflected wave. They have also been used to detect the defects in machine parts. These sensors have also been helpful in human face detection in a limited way. This chapter discusses an approach to sense whether the object in front is a human (dressed up in different types of clothing) or some other non-living object, using an ultrasonic sensor. Based on the heterogeneity in the absorption coefficient of ultrasonic waves, how a human being can be distinguished from other commonly available things in an indoor scenario has been shown experimentally. One may conclude that given all kinds of scenarios, a single sensor proves to be inadequate to sense the presence of a human being. Sensor fusion is a possible alternative, which researchers are now exploring. This chapter attempts to present the concept of non-intrusive human sensing using a single sensor viz. an ultrasonic sensor. It also provides an analysis of the accuracy of human detection in an indoor environment using this sensor. However, due to the possible weaknesses of uncertainty, missing observations, and incompleteness of a single sensor, there is a growing need to integrate and fuse multi-sensory information for advanced systems with high robustness and flexibility. Sensor fusion is a method of integrating signals from multiple sources. It allows extracting information from several different sources and to integrate them into a single signal or information. In the latter part of this chapter, a voting based approach has been discussed to improve on the accuracy of human sensing which is tested experimentally. This chapter provides the details about experimental setup, data collection in real-world scenario.

3.1 Human Detection using an Ultrasonic Sensor

The aim of this research is to detect human being in presence of other common household things such as furniture etc. In this thesis, human sensing problem is reduced to one-class classification problem. Reason for the same is that for every non-human thing data collection is not feasible. In this chapter, the task of human detection is performed using a fuzzy rule-based (FRB) one-class classifier. The rules of the FRB classifier are of zero-order TSK type [78] as given below:

$$Rule_i \;: \boldsymbol{if} \;\; (x_1 \sim x_{i1}^*) \;\; \boldsymbol{and \ldots and} \;\; (x_n \sim x_{in}^*) \;\; \boldsymbol{then} \; class \; P$$

where $Rule_i$ is the i^{th} fuzzy rule, $i = [1, N_{fr}]$, N_{fr} is the number of fuzzy rules, $x = [\ x_1,\ \ x_2,\ \ \ldots,\ \ x_n\ ]$ is the input variable represented as a vector of features, is the $\left(x_j \sim x_{ij}^*\right)$ j^{th} fuzzy set of the i^{th} fuzzy rule, *j=[1,n]*, x_i^* is the focal point of the i^{th} rule, and $class\ C$ is the class label (positive or target class).

The focal points of the fuzzy rules are identified using subtractive clustering [108]. In subtractive clustering each data point is considered as potential cluster center. The potential of each data point is measured by its potential value which depends on the count of data points present in its neighborhood (which is defined by a circle of constant radius). Higher the count of the data points in neighborhood higher will be the potential value of that data point. Effect of points present outside the neighborhood is negligible. The data point with highest potential value is set as first cluster center. After setting the first cluster center the potential value of remaining data points is reduced by a specific amount which depends on the distance between the data point and the cluster center. The reduction in the potential value of the data point which is close to the cluster center is more which in turn reduces the chances of becoming the next cluster center. Now, the data point with next maximum potential will be the next cluster center. The process will terminate on the basis of two threshold values. One if the ratio of the potential value of the current cluster center (c) to the previous cluster center is greater than an upper threshold value then c is accepted as cluster center else if the ratio is below the lower threshold value than c will be rejected and process will terminate. If the ratio exists in between upper threshold value and lower threshold value then there is a trade-off between sufficient potential value and distance of c from the existing cluster centers (it should be at sufficient distance from other cluster centers). Each of the cluster center forms the basis of fuzzy rule. Radius is a user defined parameter in subtractive clustering (suggested value is in range of 0.2 to 0.5).

The Gaussian membership function was used. The decision ($y\left(\boldsymbol{x}\right)$) is taken using a simple strategy based on the firing strength ($\tau_i\left(\boldsymbol{x}\right)$) of rules. The rule with the maximum firing strength is selected. If the maximum firing strength is more than a threshold value (θ) then the input data sample is considered to be of target class, as given below:

$$y\left(\boldsymbol{x}\right) = \begin{cases} 1 \text{ if} max\left(\tau_i\left(\boldsymbol{x}\right)\right) \ \geq \ \theta \\ 0 \text{ if} max\left(\tau_i\left(\boldsymbol{x}\right)\right) \ < \theta \end{cases} \tag{3.1}$$

where $\tau_i = \prod_{j=1}^{n} \mu_{ij}(x_j), \ i = [1, N_{fr}], \ j = [1, n]$ and

$\mu_{ij} = e^{\frac{\|\boldsymbol{x}-\boldsymbol{x}_i^*\|}{2(r_{ij})^2}}$, r_{ij} is the spread of the membership function and represents the zone of influence of the fuzzy rule.

3.1.1 Data Collection

In order to collect the data we used a PING))) ultrasonic sensor in an indoor set-up. Data was collected by focusing the ultrasonic wave emanating out of the sensor onto objects viz. a human being, a cushioned chair, a wooden door, and a glass door. For human 200 echo signals and for each object 20 echo signals were captured at a sampling rate of 200 MS/s (Million samples per second). Each echo signal consists of 4000 samples. The distance of different objects with respect to the sensor was fixed to one meter. To train the one-class classifiers we used 130 data samples of only human class. The remaining 70 data samples of human and 80 (4×20) data samples of objects constituted the test data. Figure 3.1 shows the typical signals obtained from the four objects. It can be inferred from the Figure that for cushioned chair and human mostly it is absorbed and not much is reflected where as for other two cases it is reflected back.

3.2 Training a Classifier

Initially, the features from the signals in the time and frequency domains were extracted. To extract the former features, the signals were filtered using a single level one-dimensional Discrete Wavelet Transform (DWT). Figure 3.2 shows one filtered signal from the human class. The resulting signal is used to calculate six different features as represented by equation 3.2 through equation 3.7.

$$Mean = \sum_{i=1}^{N} x_i \tag{3.2}$$

$$Standard\ Deviation\,(SD) = \sqrt{\frac{1}{N-1}\sum_{i=1}^{N}(x_i - Mean)^2} \tag{3.3}$$

$$Root\ Mean\ Square\,(RMS) = \sqrt{\frac{1}{N}\sum_{i=1}^{N} x_i^2} \tag{3.4}$$

$$Absolute\ Value\,(A) = \sum_{i=1}^{N} |x_i| \tag{3.5}$$

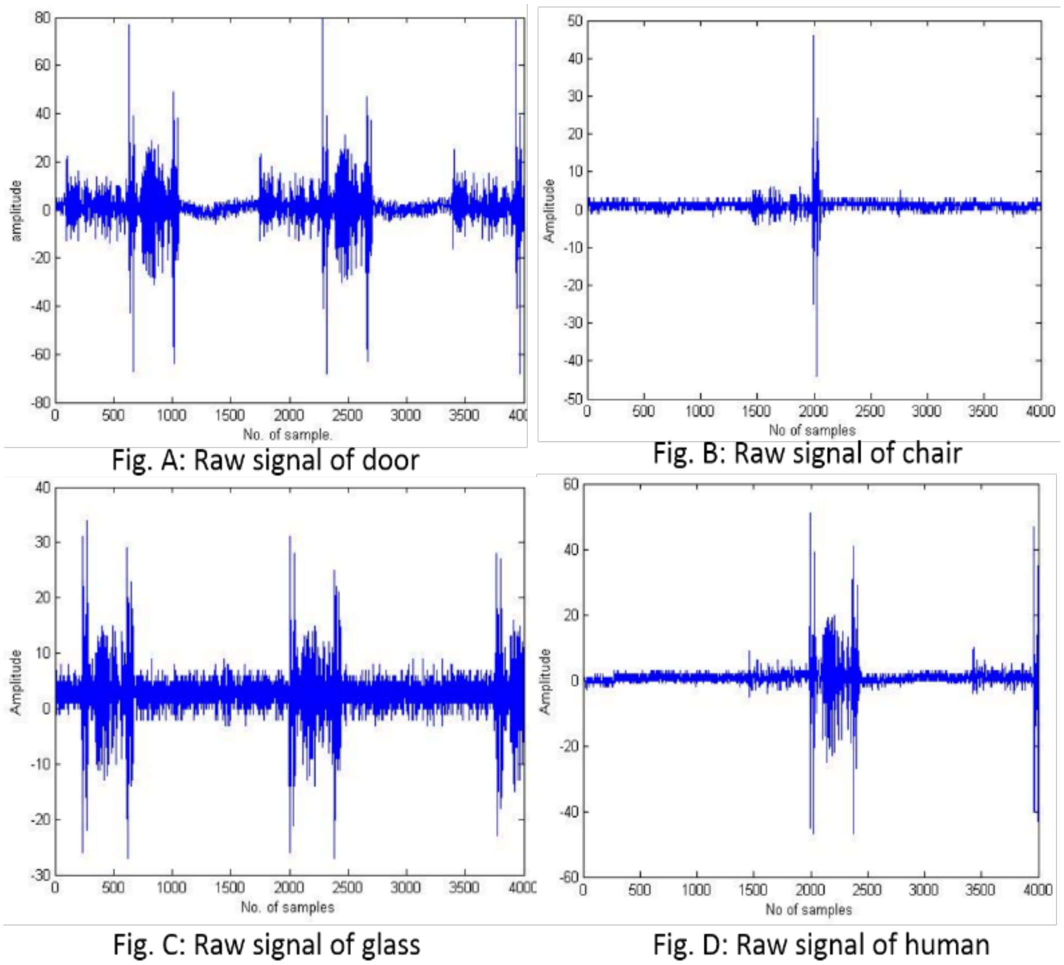


Figure 3.1: Typical signals obtained of the four objects viz. door, cushion chair, glass, human

$$Energy\,(E) = \sum_{i=1}^{N} x_i^2 \tag{3.6}$$

$$Maximum\ \ (Max) = \ \max(x_i)_{i=1}^{N} \tag{3.7}$$

Where, N denotes number of samples (which is 4000 in our case) and x_i denotes i^{th} sample for i = 1, 2, 3. . . .4000

To obtain the frequency domain features, Fast Fourier Transform (FFT) was used and only coefficients in the range 0-40MHz is considered as shown in Figure 3.3. The selected frequency range is divided into 10 bins each comprising a band of size 4MHz. For each of the 10 bins, two features viz. maximum and median, were

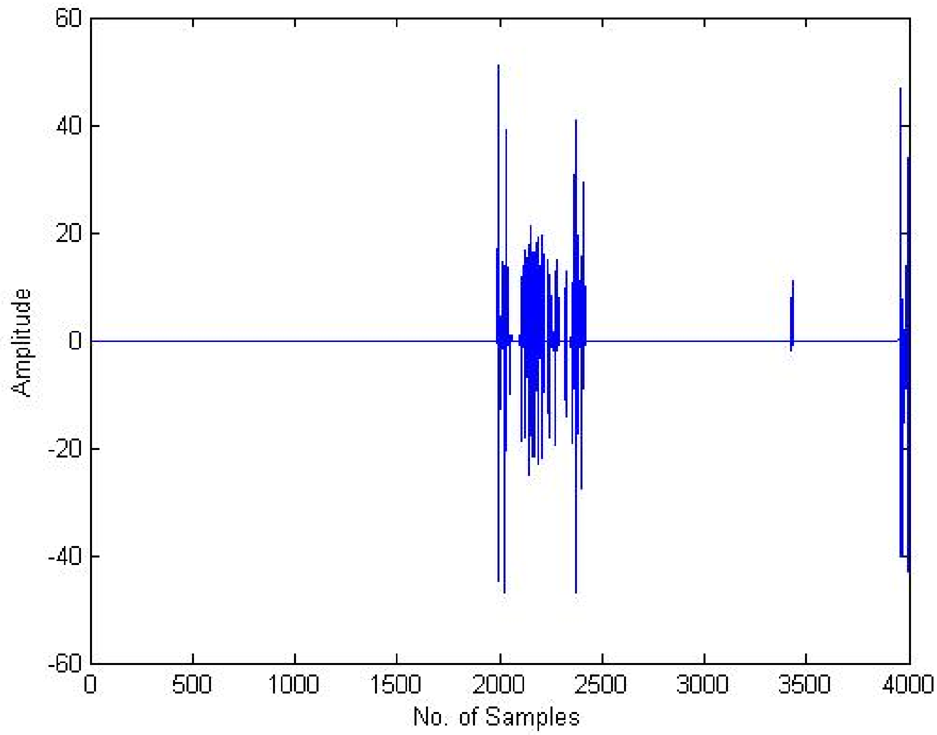


Figure 3.2: Filtered signal for human class (using DWT)

calculated. Each signal was thus represented in terms of 21 features (10 max. + 10 median + 1 ZCR). The ZCR (zero crossing rate) feature is calculated from the time domain of the signal.

3.2.1 Experiments and Results

Several experiments were conducted to evaluate the performance of the one-class FRB classifier for the task of human detection. First the Time Domain Features (TDF) were considered. Based on correlation analysis we found that out of the six features (explained in equation 3.2 to equation 3.7), only four viz. SD, RMS, E, and Max were found to be useful. The input parameter of subtractive clustering was chosen empirically and set to 0.5. The suggested range of values of this parameter is 0.2-0.5 [108]. Clustering resulted in three fuzzy rules.

Table 3.1 shows the results for classification in terms of accuracy, True Positive Rate (TPR) and False Positive Rate (FPR) for various values of θ (threshold of fuzzy rules). From Table 3.1, it is apparent that as θ increases both TPR and FPR

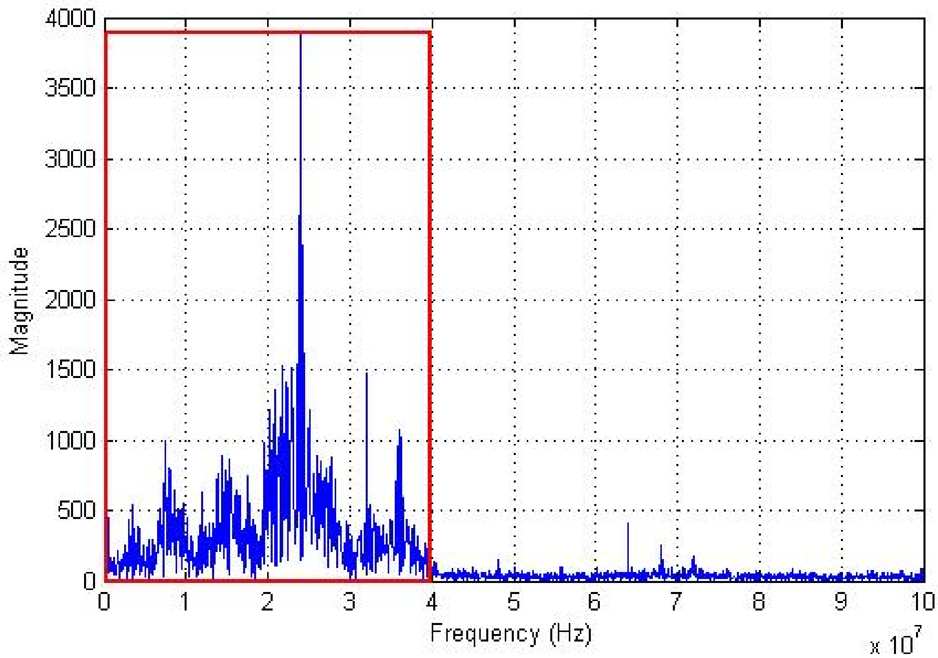


Figure 3.3: Frequency spectrum of a signal from human class.

Threshold	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
Accuracy	0.80	0.84	0.85	0.85	0.82
TPR	0.93	0.93	0.93	0.90	0.84
FPR	0.35	0.27	0.23	0.20	0.20

Table 3.1: Results (Time Domain Features)

decreases. θ thus decides the affinity of a data sample to the target class. If it is set to a high value then the value of the fuzzy membership function for the input data sample should also be high in order for the sample to be qualified as a positive class. For the time domain features the accuracy is high when θ equals to 0.3. Table 3.2 shows the number of human samples that were correctly identified by the classifier for θ equals to 0.3.

From Table 3.2 it can be inferred that out of the 70 samples taken for the human being, 65 were correctly detected as a human while the remaining 5 were misclassified as an object. The Wooden door was classified accurately as an object. Similarly, only one sample of the cushioned chair and glass door was misclassified.

The second set of experiments was performed using frequency domain features (FDF). Table 3.3 shows the results for the frequency domain features in terms of

	Object	**Human**
Wooden Door	20	0
Glass Door	19	1
Cushioned Chair	19	1
Human	5	65

Table 3.2: Classification based on TDF

Threshold	**0.1**	**0.2**	**0.3**	**0.4**	**0.5**
Accuracy	0.90	0.90	0.88	0.88	0.86
TPR	0.89	0.84	0.81	0.80	0.74
FPR	0.08	0.03	0.03	0.17	0.00

Table 3.3: Results (Frequency domain features)

accuracy, TPR and FPR. It is apparent from the table that similar to TDF results, the value of TPR and FPR decreases as threshold is increased.

From table 3.4 it can be inferred that for human out of 70 samples it is able to classify 62 correctly as human and fails for 8 cases. It correctly classify the cushion chair and wooden door but fails to classify the glass door for 5 cases out of 20.

From Tables 3.1 and 3.3 it can be inferred that when θ equals to 0.1 accuracy is more for frequency domain features than time domain features whereas TPR and FDR are high for time domain features.

A comparison between one-class FRB classifier and widely used Support Vector Machine (SVM) one class classification [89] was also performed.

For time domain features TPR obtained is 0.50 and FPR is 0.31 and accuracy is 0.58 when θ is 0.1 as shown in table 3.5. From table 3.6 it is apparent that SVM classify 35 humans correctly out of 70 and fails in 35 cases. For non-human objects it shows correct results for 41 samples and fails for 19 samples θ equals to 0.1.

For frequency domain features accuracy is 0.5, TPR is 0.64 and FPR is 0.67 at θ equal to 0.1 as per Table 3.7. Analysis of Table 3.8 data says that system recognizes 45 human samples and misclassified 25 sample, while for non-human samples it classify 20 objects correctly but 40 non-human samples are misclassified.

From Table 3.2 and Table 3.5 (For TDF), Table 3.4 and table 3.7 (for FDF) it seems that one class classifiers are better to detect the human as well non-human objects for on both time domain features and frequency domain features. From table 3.1, 3.2 and Table 3.5, 3.7 it can be concluded that fuzzy rule based one-class

	object	Human
Wooden door	20	0
Glass door	15	5
Cushioned chair	20	0
Human	8	62

Table 3.4: Classification based on FDF

Models(Threshold = 0.1)	**One-class SVM**	**One-class FRB**
Accuracy	0.58	0.80
TPR	0.50	0.93
FPR	0.31	0.35

Table 3.5: One-Class SVM VS. One-Class FRB(TDF)

classifiers are better than SVM in terms of accuracy, TPR and FPR for time domain features. From table 3.3, 3.4 and 3.7, 3.8 it is apparent that even for frequency domain features fuzzy rule based one-class classifiers are better than SVM in terms of accuracy, TPR and FPR.

Experiments were performed for human sensing and analysed using only a limited number of objects. Despite of having few objects false alarms were detected which are likely to increase with increase in the number of objects. Also, experiments were performed in a restricted manner that is human beings were standing still. One of the reasons for false alarm is that reflected wave might not be received by an US sensor. Also, it is difficult to differentiate between two objects which have nearly same absorption coefficient for ultrasonic waves. One of the feasible solutions to fuse data from multi-sensors. However, sensors should compliment each other in a way so that number of false alarms can be reduced. In the next section, an approach is proposed to fuse data from multiple sensors. This approach is a base line approach for comparison with advance approach where multiple sensors are used.

3.3 Multi-sensors Fusion : An approach and Experiments

This approach intend to overcome the limitations of an individual sensor and attempts to improve the accuracy of human sensing.

	Human	**Non-Human**
Human	35	35
Non-Human	19	41

Table 3.6: Classification based on TDF by SVM

Model (Threshold = 0.1)	**One-class SVM**	**One-class FRB**
Accuracy	0.50	0.90
TPR	0.64	0.89
FPR	0.67	0.08

Table 3.7: One-Class SVM VS. One Class FRB (FDF)

3.3.1 Methodology

Voting-based Sensor Fusion and Detection

This approach is used to fuse the sensory data from different sensors so as to aid the detection of a human being. To apply this approach, it is assumed that all the sensors point to the same target at the same instant of time and that data obtained from all the sensors is buffered at the same speed. Step 1 to step 3 of methodology explains the model development process whereas, step 4 to step 6 depicts the detection process.

Step1: Order the sensors $S_1, S_2, ..., S_{N_S}$ on the basis of their reliability in detecting human presence. Here, S_1 has the highest reliability (in terms of human detection) while S_{N_S} has the least. Data from the sensor with highest reliability is considered first for processing. Define an upper threshold θ_k for each sensor S_k (k= 1 to N_S). This threshold is determined experimentally and indicates the probability of an object being classified as a human being.

Step2: Data captured from a sensor S_k for a period T can be conceived to be a signal. M features are extracted from each of a total of N_f signals. f_{ij}^k is defined as the j^{th} feature extracted from i^{th} signal which is received from k^{th} sensor. Thus, for all $F^k(k = 1, 2, 3, ..., N_S)$ for sensor k is represented as a matrix F^k as given

	Human	**Non-Human**
Human	45	25
Non-Human	40	20
FPR	0.67	0.08

Table 3.8: Classification based on FDF by SVM

below. The matrix F^k is used as training data for developing the model.

$$F^k = \begin{bmatrix} f_{11}^k & f_{12}^k & . & . & . & f_{1M}^k \\ f_{21}^k & f_{22}^k & . & . & . & f_{2M}^k \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ f_{N_f1}^k & f_{N_f2}^k & . & . & . & f_{N_fM}^k \end{bmatrix} \tag{3.8}$$

Step3:Let,Y_j^k be a vector which is represented as :

$Y_j^k = \left\{ f_{1j}^k \quad f_{2j}^k \quad . \quad . \quad . \quad f_{N_fj}^k \right\}$ where, $j = 1$ to M

Apply mean-shift clustering technique [25] for each Y_j^k. The mean-shift is a non-parametric method and can be used to find the local maxima of a density function. In this technique a single parameter called the bandwidth is used which is internally determined by the median of all the pairs of elements of Y_j^k ($j = 1$ to M). The proximities are measured using the Gaussian method. We have used the flat kernel [106] to calculate the kernel density function. This function is used to determine all the local maxima that are cluster centers.

Applying mean shift results in a set of clusters along with their corresponding centers and radii for each Y_j^k. Thus for each sensor S_k, there will be M sets of such clusters. Let Q_j^krepresents the set of centers and radius of the clusters of Y_j^k.

$Q_j^k = \left\{ (c_{1j}^k, r_{1j}^k) \quad (c_{2j}^k, r_{2j}^k) \quad . \quad . \quad . \quad (c_{N_{clust}j}^k, r_{N_{clust}j}^k) \right\}$ where, c_{pj}^k is the p^{th} center of the p^{th} cluster of Y_j^k, r_{pj}^k is the p^{th} radius of the respective cluster of Y_j^k, N_{clust} is the number of clusters obtained.

Step4: Let Z^k represent all the M features extracted from a test signal captured via the sensor S_k when, an obstacle is in front of the sensors.

$Z^k = \left\{ z_1^k \quad z_2^k \quad . \quad . \quad . \quad z_M^k \right\}$

It may be noted that here just one test signal is considered unlike what was done in step 2 where G number of signals were considered. In the detection process, sensor signals are processed following the same order as defined in step1. So, the first sensor S_1 we obtain the following cluster centers and radii corresponding to vector Y_j^1 (according to step 3).

Thus, $Q_j^1 = \left\{(c_1^1, r_1^1) \quad (c_2^1, r_2^1) \quad . \quad . \quad . \quad (c_{N_{clust}}^1, r_{N_{clust}}^1)\right\}$

Now, votes are calculated based on the membership of each z_j^1 to the respective clusters. If z_j^1 belongs to any one of the clusters associated Q_j^1,
then the vote for the feature j, of sensor S_1, $V_j^1 = 1$, else $V_j^1 =$ -1 (In general, the vote for feature j extracted from signal of sensor S_k is represented as Vj^k). Similarly, votes for all the M features are calculated.

Step5: Count the number of positive votes and calculate the probability of the object to be classified as human being using equation 3.9.

$$Probability(\mathcal{P}^k) = \frac{\Sigma_{j=1}^{M}(V_j = 1)}{M} \tag{3.9}$$

Step6: If the calculated probability (P^k) adheres to the threshold (θ_k) condition for the sensor S_k then the confronted object is detected as a human being. If,the calculated probability (P^k) does not satisfy this threshold condition , then the step 4 through step 6 are repeated for the signal received from the next sensor.

The complete process is also described in algorithm 1. Input to the algorithm are sensor priorities and their respective thresholds as defined in the step 1 of the methodology section. $calc\_features(DataFrame)$ is a function which calculate the feature from the given dataframe of a sensor as per step 2. Once feature matrix is ready, clustering is performed as per step 3 which correspond to $Feature\_wise\_clustering(F^k)$ function. Function $Calculate\_radii\_centres(Y_j^k$ calculates the radii and centers of the clusters formed. Once cluster centers and radii are computed then features from the test data frames of the respective sensors are computed using $calc\_features$. Based on the membership of the calculated features to the respective feature clusters probability is calculated using $Calculate\_probablity()$ function and compared with the prior calculated thresholds (input to the algorithm).

Thus human presence and absence is being confirmed as per the latter part (19 to 23) of the algorithm

Algorithm 1 Algorithm for sensor fusion

input : Number of sensors and their respective priority (in terms of human detection) and upper threshold.
output : Human Presence or absence.

1: **for all** Data frames received from N sensors **do**
2:     $calc\_features(DataFrame)$;
3: **end for**
4: **for all** Feature **do**
5:     $Feature\_wise\_clustering(F^k)$
6:     $Calculate\_radii\_centres(Y_j^k)$
7: **end for**
8: **for** high priority Sensor **do**
9:     $calc\_features$ for the test data
10: **end for**
11: **for all** Features of a dataframe **do**
12:     **if** z_i^k lies in any of the clusters of the respective feature **then**
13:         $Vote = 1$
14:     **else if** z_i^k does not lies in any of the clusters **then**
15:         $Vote = 0$
16:     **end if**
17: **end for**
18: $Calculate\_probablity(\mathcal{P})$
19: **if** $\mathcal{P} \geq threshold$ **then**
20:     $Humanispresent$
21: **else**
22:     $Humanisabsent$
23:     $goto8fornextsensor$
24: **end if**

3.4 Experiments and Results

The voting based sensor fusion approach is applied to sense the human being using the data from ultrasonic and PIR sensors to enable human detection. The following hardware was used in the experiments conducted:

1. Ultrasonic Ping Sensor

2. PIR Sensor

3. Maxbotix Analog Ultrasonic Sensor (AUS)

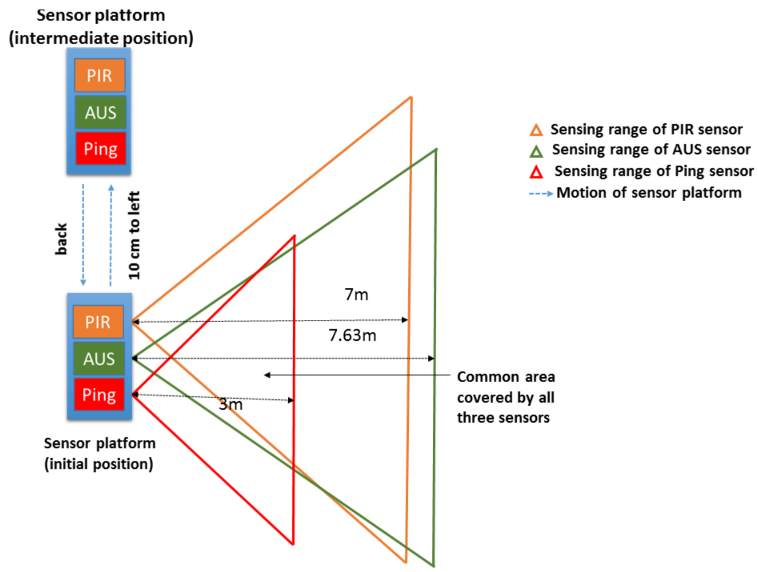
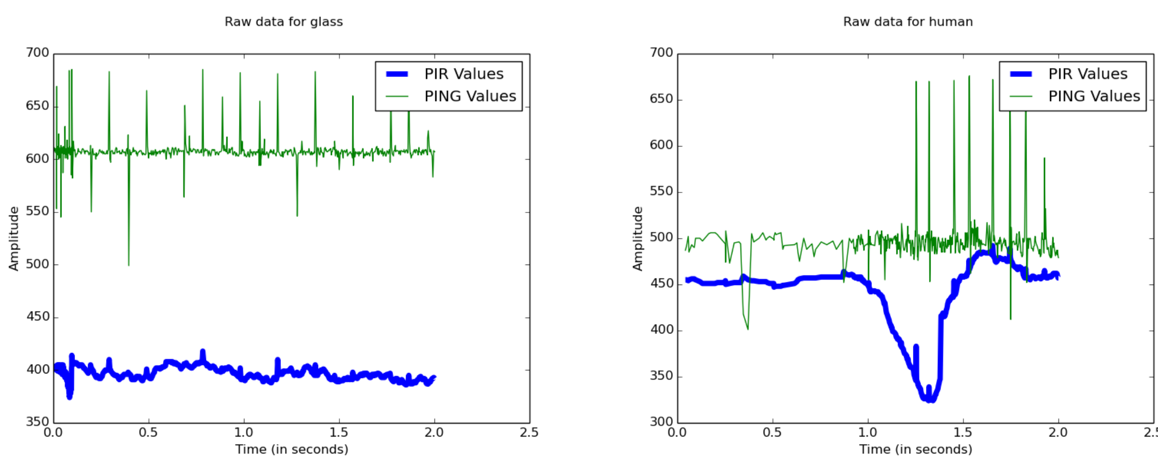


Figure 3.4: Sensing range of the different sensors used

4. Arduino Mega 2086 board

A PIR sensor works on the temperature difference and was used to sense the human motion in front of the sensor while the ultrasonic PING sensor was used to measure the distance of the obstacle confronted. The Maxbotix Analog ultrasonic sensor was used to capture the raw ultrasonic signal reflected from the obstacle in front. The Arduino Mega was used as the interfacing device to read the sensory data through a serial port. The experiments performed were based on the analysis of raw signals received from the PIR sensor as well as the AUS. As mentioned earlier the PIR sensor fails to detect the human being when s/he is stationary. This issue was resolved by analyzing the signal received from the AUS. However, for some objects the difference between two received ultrasonic signals for two different objects was found to be marginal which made it difficult to differentiate the two objects. This difference is also affected by noise. It was observed that merely on the basis of the AUS, one cannot differentiate the human being from other objects. To overcome the limitations of both the PIR sensor and the AUS, the data received from the sensors was combined before the final decision of classification was made. Instead of using these sensors separately, sensors can be used simultaneously to improve the accuracy (defined later by equation 3.14) and to reduce false alarms. A false alarm is said to be generated when a non-human object is classified as a human being. In the experiment, the three sensors were mounted on a moving platform in such a way that the sensing range of the PIR sensor overlapped that of the AUS and the PING sensor as shown in Figure 3.4. The platform was made to move to and fro along a 10 cm straight line.The reason behind doing so was to sense the same target with all sensors at the same instance of time. Since the associated data sheets mention that the AUS could provide erroneous results at distances less than 50 cm from the sensor, data received from the sensors was processed only when the target was detected approx. one meter away. The distance of the obstacle was continuously checked via data received from the PING and the analog ultrasonic sensors. It was observed that when the distance of the target was greater than approx. one meter, the signal strength from the AUS was too small to be read by the Arduino board. In the experiments we have thus tried to detect the human being within the range of approx one meter. This is a fair distance considering the fact that a robot needs to perform the detection. As soon as an obstacle was confronted within this specified range, the data received from the sensors was processed.

Data Acquisition: Data acquired from various sensors was buffered at 112.5 kilo bits per second (kbps). As mentioned earlier, the data was buffered only if the object was in the range of approx. one meter.



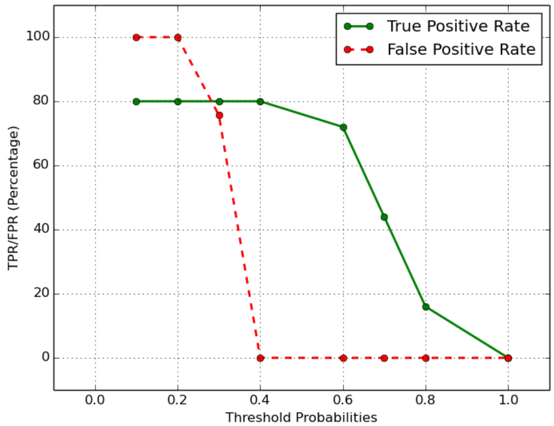
(a) (b)

Figure 3.5: Raw signals of PIR sensor and Ultrasonic sensor for [a] For Glass [b] For Human Being

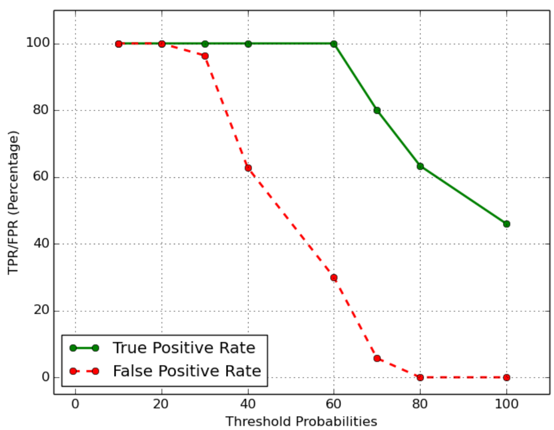
Figure 3.5 shows the nature of the raw signals (amplitude vs. time) received from the AUS and PIR sensors when used to detect various objects including a human being. While there is a marked visual difference between the PIR signals obtained from human and non-human objects the same is not true for those from the AUS. As described in Step 1, the order and thresholds for the PIR sensor data (θ_{PIR}) and AUS (θ_{AUS}) need to be defined. In order to set the thresholds for both the PIR (θ_{PIR}) and AUS (θ_{AUS}) sensors, two experiments viz. Experiment 1 and Experiment 2 were performed. Experiment 1 was performed with a PIR sensor for human detection. The same task was also performed with a AUS in Experiment 2. Assuming different probabilities the values of TPR and FPR were found by conducting experiment 1 and 2.Both experiments were carried out using different objects placed in front of the sensors. Both experiments were repeated several times. The TPR and FPR values were calculated for different probabilities found using equation 3.9 . As can be seen in Figure 3.6a and 3.6b the TPR is maximum and FPR is minimum when the probability is 0.4 for the PIR sensor while the same is true at probability 0.7 for the AUS. Therefore, the thresholds θ_{PIR} and θ_{AUS} were set to 0.4 and 0.7 respectively for the detection experiments. Further, the priority of the PIR sensor was set higher than that of the AUS based on the TPR and FPR values attained.

To train the model 150 signals are captured and each signal is captured for 4 seconds. A signal is buffered for 4 seconds to ensure that buffering time of signal is greater than or equal to the processing time of signal. The data acquired as mentioned in Step 2 of section methodology is only for the human being since the remaining non-human objects are countless and collecting data for this set is beyond the scope of this work. The data obtained from the PIR and AUS sensors was buffered separately and represented in form of matrices as shown in step of methodolgy section. Data was collected from each sensor for 600 seconds using only the human being as a target. Human beings dressed in different clothing formed targets for detection. Features, as mentioned in Step 2 in methodology section, were calculated for both the PIR and ultrasonic signals using the equations 3.2 through 3.4, equations 3.7 and equations 3.10 through 3.13.

$$Median = median\,(x_i)_{i=1}^{N} \tag{3.10}$$



(a)



(b)

Figure 3.6: [Performance of [a] PIR sensor with varying probabilities [b] AUS sensor with varying probabilities

$$Kurtosis = N \frac{\Sigma_{i=1}^{N} (x_i - average)^4}{\left(\Sigma_{i=1}^{N} (x_i - average)^2\right)^2} \tag{3.11}$$

$$CrestFactor = \frac{\frac{1}{2} (Maximum - Minimum)}{RootMeanSquare} \tag{3.12}$$

$$AveragePeaktoPeakDistance(PPD) \tag{3.13}$$

where, N is number of data samples and a_i is $i^{th} sample$.

In order to reduce the inherent noise in the raw signal, Discrete Wavelet Transformation (DWT) was performed on the AUS data before extracting its features. Figure 3.7 shows the nature of the ultrasonic signal reflected from a human being after applying DWT. As mentioned, the features were extracted from such transformed signals.

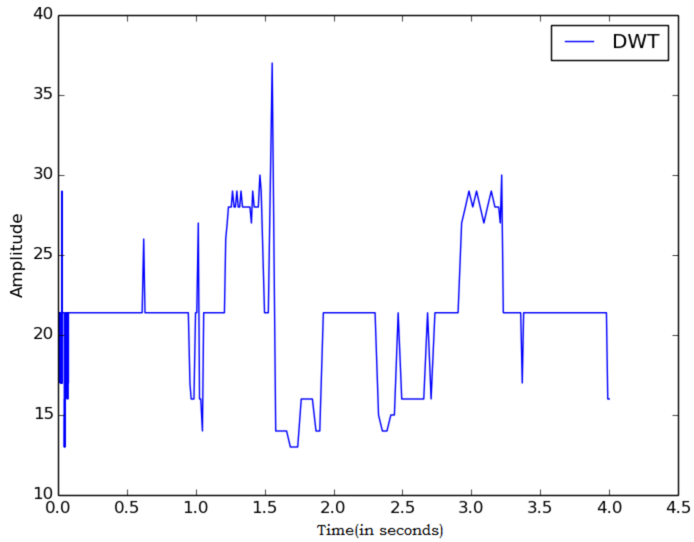


Figure 3.7: Nature of the signal after DWT (Object: Human being)

Once the procedure defined in Step 2 of methodology section was completed, Step 3 was performed for different objects commonly present in an office and cafe environment.

At first features defined by equations 3.2 through 3.4, equations 3.7 and equations 3.10 through 3.13 are extracted from the PIR sensor data. Using procedures detailed in Steps 4 through 6, the voting and probability of the confronted object are calculated on the basis of the PIR sensor signal. If this probability is found to be greater than θ_{PIR} then the object confronted is classified as a Human else as a Non-Human. If the object is classified as a Non-Human, the signal data received via the ultrasonic sensor for the same object is retrieved and its features are extracted after applying DWT. The associated probability is also calculated and voting performed accordingly. If the probability calculated for the AUS signal is less than θ_{AUS}, then the object is categorized to belong to the Non-Human class; else to the Human. To confirm the presence or absence of the human being, the platform on which the sensors were mounted was made to move to their left by a distance of

10 cm and then move back to the original position. During this motion, the new data is acquired via the sensors for the same obstacle present in front of the sensors. From this new data, the PIR sensor signal is once again analyzed in a similar way as defined earlier. If the new probability calculated for the PIR sensor is greater than θ_{PIR}, the presence of the human being is confirmed; else the object is deemed not belong to the Non-human class. Experiments were performed with different human beings, cushioned chairs, cupboards, wooden ply boards, cardboard, glass panels and cement and stone walls. Table 3.9, shows the results which indicates that out of 80 human beings, all were classified correctly with no false alarms(In column 1 of Table 3.9, the values in the brackets denote the number of obstacles confronted by the sensors). All non-human objects were also classified with high accuracy (100 percent in this case) calculated as defined by equation 3.14 .

$$Accuracy = \frac{TPR + TNR}{TPR + TNR + FPR + FNR} \tag{3.14}$$

where, TPR is True Positive Rate, TNR is True Negative Rate, FPR is False Positive Rate and FNR is False Negative Rate.

	Human	**Non-Human**
Human (80)	80	0
Non-Human (120)	0	120

Table 3.9: Classification results of Voting based sensor fusion

The results obtained using the voting based sensor fusion approach was compared with those obtained from Experiments 1 and 2 as well. When experiments are performed with the PIR sensor, it fails to detect a stationary human being (as expected) which caused a reduction in TPR. In case of the experiment with the AUS, only the cushioned chair and the glass panels were classified as human beings. However, these cases were correctly resolved when the voting based sensor fusion

	TPR	**FPR**
Voting-based Sensor Fusion	100	0
AUS(θ_{AUS} = 0.7)	80	5.71
PIR(θ_{PIR} = 0.4)	80	0

Table 3.10: Comparison of various approaches

approach was used. Fusion allows detection of a moving human being using the PIR sensor signal and if stationary, the AUS performs this task. The double checking using AUS and the PIR sensor data obtained on movement of the robot back and forth provides for better and more reliable results. A comparison of the results obtained from the voting based sensor fusion with the results of classification obtained from the PIR as well as the AUS are shown in Table 3.10. The TPR and FPR values for the fusion approach were found to be 100 and 0 respectively. The same for the AUS and PIR at the predefined thresholds, were found to be 80 and 5.71 and 80 and 0 respectively. This indicates the superiority of the combined fusion based approach. Computations in the fusion based approach are performed only when the target is within a prescribed range. Further, if the PIR signal facilitates classification of the target as a human being, then the AUS signal is not processed. These aspects save on computational overheads. Computations in the fusion based approach are performed only when the target is within a prescribed range. Further, if the PIR signal facilitates classification of the target as a human being, then the AUS signal is not processed. These aspects save on computational overheads.

3.5 Chapter Summary

This chapter presented an approach to differentiate human beings from non-humans using an ultrasonic sensor. A human being was differentiated from other objects (when confronted by ultrasonic waves) using one-class FRB classifiers. Experimental results validate the effectiveness of the approach. It may thus be concluded that a human being could be classified as a class distinct from other tested materials which are commonly present in the premises of a coffee shop type of environment. The results obtained are promising and thus pave the way to further investigate the technique using a variety of materials that are commonplace in such environments, so as to facilitate better and more accurate human detection. A combination of ultrasonic and PIR sensor was also used for human sensing and a *A voting based sensor fusion approach* was proposed to fuse information from these sensors. From the experiments performed to detect the presence of a human being, using PIR, AUS and PING sensors, one may conclude that the voting based sensor fusion approach performs far better than cases when individual sensors are used for the same task. The high TPR and low FPR values form the basis of this statement. Detection was not done unilaterally but was re-verified by other means using sensor data and also by physical movement relative to the target. The current approach does not however take into consideration the aspect of speeding objects. For such scenarios, the platform on which the sensors are mounted may need to match speeds in terms of mobility to bring down the relative speed and thus enable the detection.

The work explained in this chapter confirms the fact that a single sensor is susceptible to failure under certain scenarios. One viable solution to compensate for this is to augment the system with other complementary sensors. However, this requires additional processing data received from all the sensors. It can be inferred that more the number of sensors more will be the computational complexity. One of the possible solutions to reduce computational complexity is to have an adaptive system, where it can select the sensor autonomously based on the current environmental conditions. This lays the foundation of the next chapter, where a multi-modal human sensing system is explained. The system autonomously selects the sensor(s) (corresponding to an environment) from which information regarding presence or absence of a human being can be extracted.

"Nature holds the key to our aesthetic, intellectual, cognitive and even spiritual satisfaction."

Edward Osborne Wilson (1929)
American biologist

4

Multi-modal Human Sensing

As discussed in the previous chapter, mechanisms for human sensing are not foolproof because of the dynamic nature and ability of human beings to deliberately mislead the sensors. In addition, sensors have their own inherent limitations, due to either the mechanism of sensing or the environmental conditions. Both these contribute to the failure of the detection. In the last two decades, a significant amount of prototypes and solutions based on Machine Learning (ML) techniques have been proposed by researchers to cater to the mentioned issues.

In this chapter an adaptive multi-modal human sensing mechanism is proposed which can autonomously identify and ignore data which cannot be used for human sensing from a set of sensors thereby reducing computation complexity, reducing false alarm rate and yielding better performance. The effect of sensing when the human being is in motion has also been studied. The results portrayed in this chapter prove the efficacy of the proposed multi-modal system over its single sensor counterpart when used in changing environments.

The contributions of this chapter are:

- Combine data from multiple sensors situated in different environments (indoor and semi-open).
- Analyze the in-built characteristics of every sensor so as to automate the process of finding a blinked sensor(s) thus reducing computation time. Blinked

sensors are sensors from which information regardinh human sensing can not be collected.

- Propound an algorithm for human detection based on multi-modal sensing in both scenarios where sensors are static and mounted on a mobile robot. It also covers both the cases of stationary human as well as moving human.
- Analyze the walking speed (which is a part of dynamic nature of human) versus accuracy of human detection.

In the following section, problem definition is provided following which methodology is explained. Towards the end of this chapter experiments and results obtained were discussed which is followed by chapter summary.

4.1 Problem Definition

- A sensor (S_i) is defined as a *blinked sensor* in a particular environment (E_j) if information obtained from it cannot be used to accomplish task T.
- For a given classification task T, let $S = S_1, S_2, ..., S_{N_S}$ be a set of N_S sensors used to collect the data. In this thesis, the task T is to sense human presence.

$$S_i^U = S - S_i^B \tag{4.1}$$

where, S_i^U is the set of sensors from which data extracted can be made use of accomplishing task T, S_i^B is the set of blinked sensors from which data extracted cannot be used to accomplish task T .

It may be noted that, as environment changes, S_i^B will also differ causing a corresponding changes in S_i^U. Thus, the problem herein is to find the set of blinked sensors ($S_i^B \in S$) based on the raw data received from all sensors in S without any human intervention.

4.2 Description of Method

The following terms have been defined to facilitate the explanation of the proposed method for human sensing.

4.2.1 Data Representation

The incoming data received from the sensors in S are processed in frames. Each frame consists of sensor data buffered for a time of t seconds. Let $\mathcal{DF}_j^k$ is defined as the j^{th} frame of the k^{th} sensor, S_k. Only those frames which contain useful data pertaining to task T are processed while others are discarded. The process of identifying useful frames is described in subsequent sections.

4.2.2 Elimination and Decision-making Features

Every sensor has its limitations due to which it could fail to sense a human being in a particular environment [137]. For example, in dark areas information retrieved from a camera cannot be used to detect a human being. Under this condition, the maximum and minimum pixel values within the concerned frame are equal to zero. Since the image is entirely black, one may conclude that the frame cannot be used to detect human presence. Likewise, if the raw data from a PIR sensor frame cannot be used for human detection when the maximum and minimum values of signal amplitude lie in the range between 150 and 500, since it indicates that either the human being is stationary or not present.

In order to eliminate the processing of information contained in a frame $\mathcal{DF}_j^k$ of the sensor S_k, we need to extract some typical features. Let, $\mathcal{F}_j^k$ represents the feature vectors of eliminating features of j^{th} frame of the k^{th} sensor. *Characteristic features* are the answer to the question : How can a system identify whether a sensor is blinked or not? Therefore, *characteristic features* are the features extracted from the raw data of sensors which decides whether the data received from that sensor should be processed or not for final decision.

$\mathcal{F}_j^k = f_1^{jk}, f_2^{jk}, .., f_M^{jk}$. where, f_x^{jk} is the i^{th} feature of the j^{th} frame of the k^{th} sensor.

Thus, if we consider all the N_S sensors then the complete feature($\mathcal{F}_j$) vector that decides the elimination, may be represented as -

$\mathcal{F}_j = f_1^{j1}, f_2^{j1}, .., f_{M_1}^{j1} f_1^{j2}, f_2^{j2}, .., f_{M_2}^{j2} ... f_1^{jN_S}, f_2^{jN_S}, .., f_{M_{N_S}}^{jN_S}$.

Where M represents the number of features. The data from the remaining non-eliminated frames are used to extract features that will aid in accomplishing the task of human detection. For example, color could be the feature used to make a decision in a camera-based model that differentiates between an apple and an orange. However, some of these features could be redundant and may mislead (for example it can increase the number of false alarms) the overall decision process. Decision features($\mathbb{F}_j^k$) obtained from a non-eliminated frame can be represented as a feature vector $\mathbb{F}_j^k$.

It can be represented as-

$\mathbb{F}_j^k = f_1^{jk}, f_2^{jk}, .., f_M^{jk}$

where, $\mathbb{F}_j^k$ is the decision making feature vector of the j^{th} frame obtained from the k^{th} sensor .

4.2.3 Decision Model Selection

Decision features are used to train an ML model so as to detect human presence or absence. The heterogeneity of the data received from different sensors causes different ML models to perform differently. Selection of the best ML model for data from a sensor needs to be done based on its performance. For instance, to find an appropriate model for the PIR sensor, one could use SVM, Linear regression or Decision tree for training and then choose the best performing one. Thus if models $\mathbb{M}_1^k, \mathbb{M}_2^k, \mathbb{M}_p^k$ are initially considered for sensor S_k and $\mu_1, \mu_2..\mu_p$ are their respective performance measures. Then the model with the highest value of μ is selected for the detection process.

4.2.4 Online Clustering

Clustering techniques group similar objects to form a cluster. Based on the different clusters obtained for different objects clustering can be used to distinguish two different entities/classes/objects. The distribution of data in an n-dimensional space can be represented in the form of clusters. This forms the basis of classification algorithms such as SVM, k-means, decision trees, etc. In the proposed approach, the online clustering technique proposed in [8] has been used and used the elimination characteristics for deciding sensors that should not be taken into consideration while

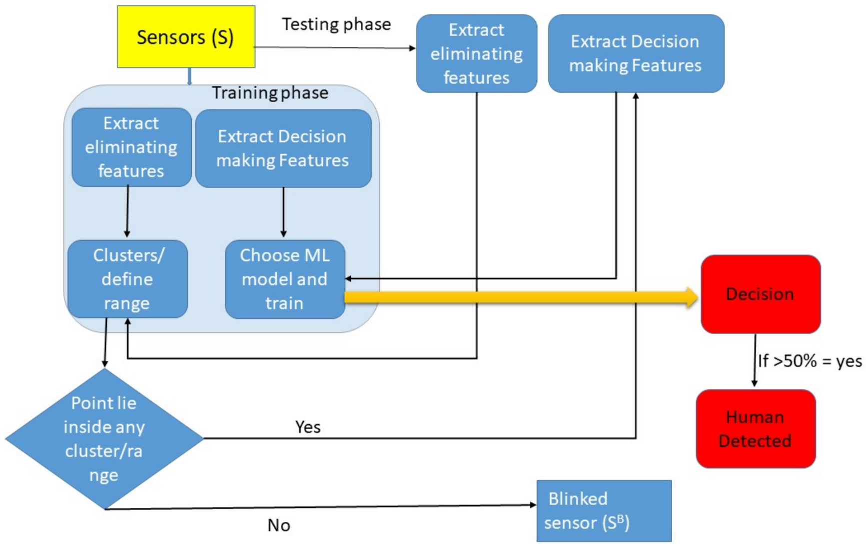


Figure 4.1: Overview of the multi-modal human sensing system

arriving at a decision. The used approach fix the radius of the cluster at initial stage and update the cluster centers in an online manner. The cluster centers are updated based on euclidean distance between data points of a cluster.

4.2.5 Methodology

This proposed methodology is based on one-class classification to detect the presence of a human being in different environments. Figure 4.1 shows an overview of the proposed method. As shown in the figure, during training phase both eliminating and decision-making features are extracted. Output of eliminating features are data clusters obtained via clustering algorithm. However, decision making features extracted from each sensor S_k are used to train their respective machine learning models. It can also be seen from the diagram that during testing phase first eliminating features are extracted. Extracted feature vector is used to find the blinked sensors. From the remaining sensors decision features are extracted to be used for final decision making using their respective ML models. If greater than 50 percent sensors are detecting human than final decision is considered as positive for human detection.

Training Phase

As discussed earlier since a single sensor is not sufficient for human sensing, multi-modal human sensing was considered. For such sensing, the data obtained from the sensors needs to be preprocessed and models generated, individually. For the proposed multi-modal sensing, training phase can be divided in two steps:

- **Selection and training of ML model for each sensor respectively:** Let for $S_k \in S$, d_k be the data stream which is buffered in term of frames, $\mathcal{DF}_i^k$. In the current multi-modal human sensing method, for every sensor $S_k \in S$, an ML model needs to be generated for the classification process. The manner by which the appropriate model (M_k) for sensor (S_k) is chosen has been described earlier in the *Decision model selection* section. While capturing the data required for training the respective models, all the sensors should point to the same human being and sense concurrently.
  Let N_f be the number of frames collected for every sensor. From the buffered frames, the individual feature vectors used for elimination and decision making, viz. $\mathcal{F}_j^k$ and $\mathbb{F}_j^k$ are calculated for all the sensors. $\mathbb{F}_j^k$ are used for training the selected ML model for a sensor.

- **Obtain clusters:** The clustering technique [8] used was provided with all the feature vectors used for elimination, $\mathcal{F} = \mathcal{F}_j^1, \mathcal{F}_j^2, ..., \mathcal{F}_j^{N_S}$, so as to determine the radii and centers of the clusters generated.
  $\mathcal{F}_j^k$ represents the eliminating feature vector of j^{th} frame of the k^{th} sensor.

$$\mathcal{F} = \begin{bmatrix} f_1^{11} & \dots & f_M^{11} & \dots & f_1^{1N_S} & \dots & f_m^{1N_S} \\ f_1^{21} & \dots & f_M^{21} & \dots & f_1^{2N_S} & \dots & f_M^{2N_S} \\ \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots & \dots\dots \\ f_1^{N_f1} & \dots & f_M^{N_f1} & \dots & f_2^{N_fN_S} & \dots & f_M^{N_fN_S} \end{bmatrix} \tag{4.2}$$

$\mathbb{F}$ calculated from the N_f frames from sensor S_k is used for training the associated model M_k. The associated decision feature vectors taking into consideration all the N_F frames can be represented as a matrix $\mathbb{F}_k$ as -

$$\mathbb{F}_k = \begin{bmatrix} f_1^{11} & \dots & f_M^{11} & \dots & f_1^{1N_S} & \dots & f_m^{1N_S} \\ f_1^{21} & \dots & f_M^{21} & \dots & f_1^{2N_S} & \dots & f_M^{2N_S} \\ \multicolumn{7}{c}{\dotfill} \\ f_1^{N_f1} & \dots & f_M^{N_f1} & \dots & f_2^{N_fN_S} & \dots & f_M^{N_fN_S} \end{bmatrix} \tag{4.3}$$

This matrix is used to train the model M_k for the k_{th} sensor.

Algorithm 2 Algorithm for elimination of blinked sensors

input : Prior calculated Cluster centers and radii (which are output of clustering technique).
input : Sensor dependent range of the features.
output : Eliminated sensors S^B.

for all Data frames received from N_S sensors **do**
    $\mathcal{F}_j = calc\_features$;
end for

if $\mathcal{F}_j$ lies in any of the clusters **then**
    No sensor can be eliminated.
    $Update\_Clusters(\mathcal{F}_j)$
else if $\mathcal{F}_j$ does not lies in any of the clusters **then**
    for all S_k **do**
        $Check\_range()$
        if (S_k is out of defined range) **then**
            $Add\_sensor\_S_i^B(S^B)$
        else if No sensor is out of range **then**
            $Update\_Clusters(\mathcal{F}_j)$
        end if
    end for
end if

Distinction Phase

Concluding on the presence or absence of a human being involves:

- Elimination of processing of frame received from the blinked sensors
- Concluding on whether the target is a human being

Algorithms 2 and 3 explain these two processes. Algorithm 2 takes in the cluster centers and radii computed in the training phase along with the sensor specific ranges for deciding whether or not to categorize a sensor as blinked. Also, as mentioned

Algorithm 3 Algorithm for detection of Human Presence

input : Pre-trained models (M_k) for sensor (S_k)
output : Human presence yes or no.

S^B = Algorithm for identification of blinked sensors.
$S^U = S - S^B$
for all Data frames (F_j^k) received from S^U sensors **do**
  $\mathbb{F}_k = calc\_features(F_j^k)$
  $Output = Predict\_output(M^k, \mathbb{F}_k)$
end for
$Count = Calculate\_positive\_count(Output)$
if Count is $>$ 1/2(Count of sensors in S^U) **then**
  Human detected.
  $Update\_Positive\_Models(M^k, \mathbb{F}_k)$
else if Count is $<=$ 1/2(Count of sensors in $\$^U$) **then**
  Human not detected
  $Update\_Negative\_Models(M^k, \mathbb{F}_k)$
end if

earlier, based on the values within the frames the corresponding sensor is either eliminated (blinked) or considered using $Check\_range()$ function. Similar to training phase, here too, data from N_S sensors is buffered in the form of frames. The data within these is used to ascertain the elimination features using $calc\_features()$ which in turn outputs $\mathcal{F}^t$.

If $\mathcal{F}^t$ lies within any of the clusters then frames from all the N_S sensors are taken into consideration for human sensing. $\mathcal{F}^t$ is also used to update the cluster centers as in [8] using $Update\_Clusters()$, to facilitate dynamic evolution of the clusters. In this step center and radii of the clusters are updated by using the incoming data.

On the contrary, if $\mathcal{F}^t$ lies outside these clusters, the values, the algorithm 2 finds the sensor(s) whose data was responsible for making it an outlier. This is done by inspecting the associated features within the relevant frames based on the sensor specific ranges that are already available to the algorithm. If any of the value(s) of any of the feature(s) pertaining to a sensor S_k within a frame is out of the predefined feature-specific range then this sensor is deemed to be blinked and added to the set of blinked sensors, S^B. If $\mathcal{F}^t$ does not lie within any of the clusters and none of the associated sensors have been deemed to be blinked then a new cluster if formed using $Create\_Cluster()$.

The Algorithm 3 depicts the process of detection of human presence. It takes the already available pre-trained models for each sensor as its input and uses the set of blinked sensors, S^B obtained from algorithm 4 to eventually find the set of useful sensors(S^U). The associated frames from the sensors in S^U are used to compute the decision-making features that is $\mathbb{F}_t^k$. These features are used by the sensor-specific ML model to decide whether or not the target is a human being. A human being is said to be detected only if majority of sensor-specific models report this detection to be positive. As its last step, the algorithm 3 update the sensor-specific models based on the final output by re-training, on-the-fly. Function $Update\_Positive\_Model()$ updates the positive model which gives the correct decision in case of human presence. Similarly, $Update\_Negative\_Model()$ updates the negative model which gives the correct decision in case of human absence.

4.3 Experiments and Results

Experiments to validate the efficacy of the proposed human sensing methodology were conducted in two different indoor set ups.

4.3.1 Experiment 1:

The hardware used in the experimentation included -

- **Analog Ultrasonic Sensor (XL-MaxSonar -EZ/AE) (AUS):** The output of this sensor is analog voltage envelope of return acoustic waveform. Data is buffered in the form of frames. If the standard deviation of such a frame is not equal to zero, then data of this frame can be used to sense the presence or absence of human being.

- **Two Pyro infrared Sensor(PARALLAX PIR sensor (Rev B)) (PIR):** The analog output of the sensor is buffered in frames. For a moving object (man or pet), which is in the sensing range of a PIR sensor, the maximum and the minimum values output of the frame are greater than 500 and less than 150, respectively.

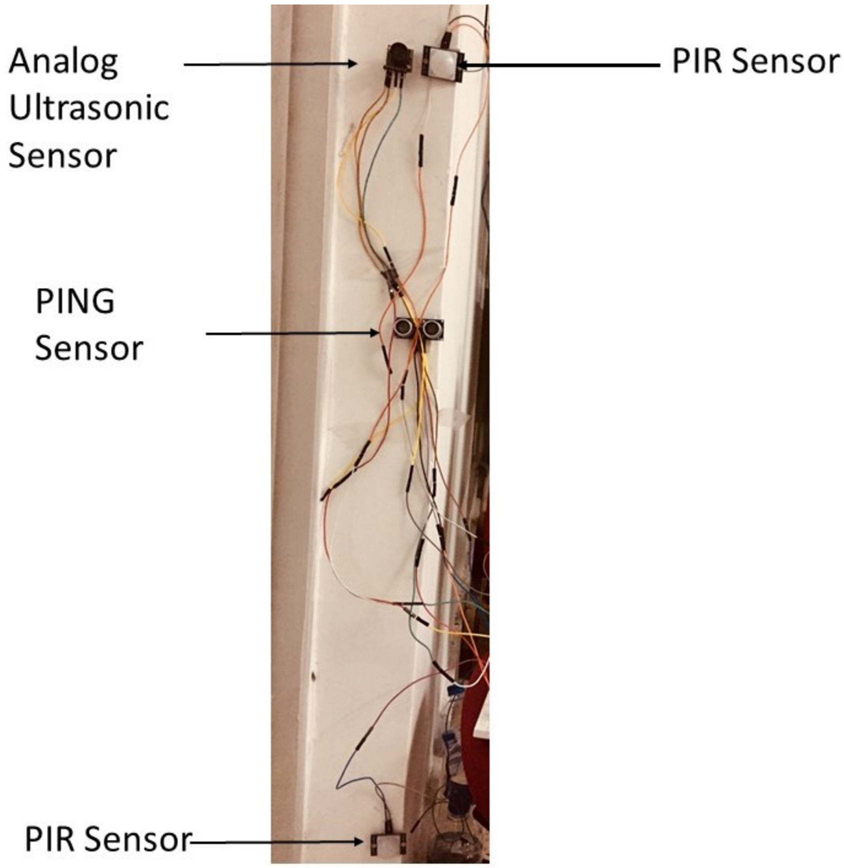


Figure 4.2: Sensors setup on a wall

- **Ping Sensor (PARALLAX PING)))) (US):** This sensor outputs the distance of the obstacle in front. The data from AUS and PIR sensors are processed depending on the output of US sensor.
- **Arduino Mega 2560:** All the sensors are connected with the Arduino board to read the sensory data.

The sensors - Ping, AUS and PIR(PIR_1) - one each, were placed on a wall (near to a door) at a height of 72 cm from the ground as shown in the Figure 4.2. Another PIR(PIR_2) sensor was placed on the same wall at a height of 20 cm from the ground. All sensors were placed in such a way that they all pointed at the same target. In this setup PIR_1 is prioritized over PIR_2 sensor PIR_1 is placed at an height of 72 cm which is above average height of cats and dogs. That is why, for the final decision of human detection PIR_1 is prioritized. Therefore, PIR_1, AUS, PIR_2 is the sequence of sensors in the decreasing order of priority for the finalization of decision of human sensing.

Apart from the distance of the obstacle in front, from analog AUS signal, a comprehensive analysis can be performed to differentiate human from non-humans. Similarly, unlike binary output of PIR sensor analog signal of the PIR sensor can be analyzed for the direction of motion and speed of motion.

The data of PIR sensors and AUS is processed only when an object is detected at a distance of approx. one meter from the sensors.

To train the system, the incoming data was buffered into data frames. One data frame of a sensor consisted of all the data obtained from a sensor in a time period of 4 seconds. When human being was standing at a distance of approx. one meter, the sensory data of PIRs and AUS was buffered in the form of frames. Features were extracted from every frame of AUS and both PIR sensors. For training purpose, both *eliminating* and *decision-making* features were extracted from each frame. Equations 4.5 and 4.4 represent the eliminating features for the PIR sensors while equation 4.5 through 4.12 represent the decision making features. Equation 4.12 represents the eliminating feature for AUS . However, equations 4.5 through equation 4.12 represent the decision making features for the same. Table 4.1 represents the eliminating features and their specified ranges for different sensors.

Sensor	**Eliminating Features**	**Range for human sensing**
PIR_1	Maximum and Minimum	$>$ 500 mV and $<=$ 0 mV
PIR_2	Maximum and Minimum	$>$ 500 mV and $<=$ 0 mV
AUS	Standard Deviation	> 0

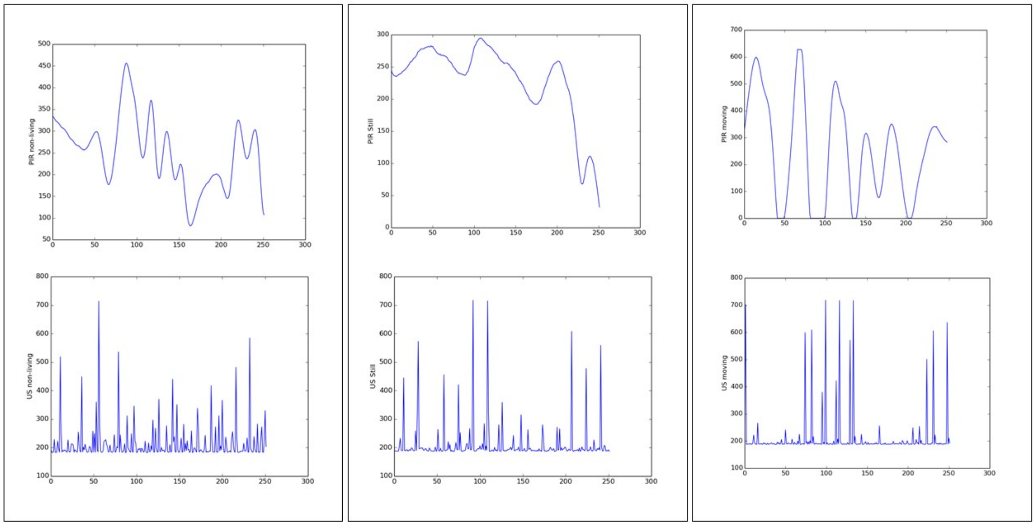
Table 4.1: Eliminating features and their specified ranges for PIR sensors and AUS for Experiment 1

$$Minimum = Min\,(x_i)_{i=1}^{N} \tag{4.4}$$

$$Maximum = Max\,(x_i)_{i=1}^{N} \tag{4.5}$$

$$Median = median\,(x_i)_{i=1}^{N} \tag{4.6}$$

$$Mean = mean(\Sigma_{i=1}^{N} x_i) \tag{4.7}$$



PIR and US signals for non-living, stationary human being and moving human respectively.

Figure 4.3: Input to the system under various conditions

$$Kurtosis = N \frac{\Sigma_{i=1}^{N} (x_i - Mean)^4}{\left(\Sigma_{i=1}^{N} (x_i - Mean)^2\right)^2} \tag{4.8}$$

$$Energy = \Sigma_{i=1}^{N} (x_i * x_i) \tag{4.9}$$

$$CrestFactor = \frac{\frac{1}{2} (Maximum - Minimum)}{RootMeanSquare} \tag{4.10}$$

$$RootMeanSquare = \sqrt{\frac{1}{N} \Sigma_{i=1}^{N} x_i} \tag{4.11}$$

$$StandardDeviation = \sqrt{\frac{1}{i-1} \Sigma_{i=1}^{N} (x_i - Mean)^2} \tag{4.12}$$

where, N is the number of data samples in a data frame and x_i represents the i^{th} sample.

Therefore, $\mathcal{F} = max(PIR_1), min(PIR_1), max(PIR_2), min(PIR_2), Standarddeviation$

For training purposes, the data obtained when 15 different human beings stood/moving at a distance of 60-70 cm. away from the sensors were collected. Human beings wore different types of clothing and stood in different postures. 100 data frames were collected for each human being in order to train the system. Human presence was confirmed in a supervised way. After every frame, human presence

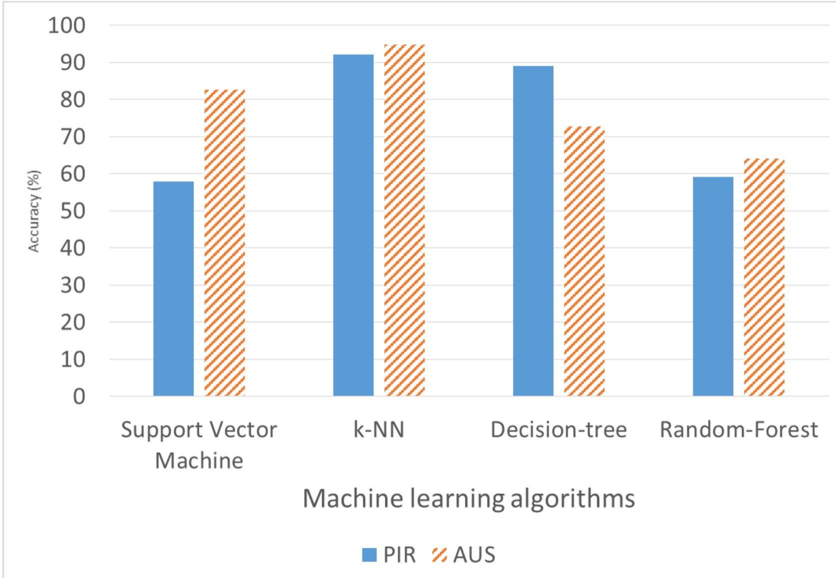


Figure 4.4: Accuracies of different ML algorithm for AUS and PIR sensor

is confirmed by pressing a physical switch. If switch was pressed then that particular frame was retained else discarded. As per methodology section $\mathcal{F}$ and $\mathbb{F}$ were calculated. Initially, based on the literature survey, for PIR and AUS multiple ML algorithms were considered. Models showing highest accuracies from among conventional machine learning algorithms were experimentally selected for each of the sensors. To select such a model for PIRs and AUS, extracted decision making ($\mathbb{F}$) were considered. The collected data was divided in the ratio 7:3 to train and test for a particular model . Also, it should be emphasized that for the testing purpose data frames for cats and dogs were also included. Graph in figure 4.4 shows the initially considered models along with their respective accuracies for PIR sensors and AUS.

It can be seen from the figure 4.4 that kNN performs better for PIR data as compared to SVM, Decision-tree, and Random-Forest. Similarly, it also reveals that k-NN clustering performs better for AUS data. Therefore, k-NN was the selected ML model for PIRs and AUS. In parallel, centers of the clustered were also calculated using extracted feature vectors $\mathcal{F}$. The cluster centers were calculated in an online manner as described in the methodology section. Table 4.2 shows the number of clusters and the average cluster density for different chosen values of r.

For the detection of human sensing in the given set up, the value of r was taken

Radius	Number of clusters	Average cluster density
0.01	12	55.96
0.05	10	45.96
0.1	10	39.06
0.15	9	36
0.2	7	28.34

Table 4.2: Cluster numbers and their average density at varying radius (r)

to be 0.05 mm. Thus, a trained multi-modal system (which consists of two PIRs, one AUS and One Ping sensor) was deployed for testing. The system was tested for human presence detection. When the Ping sensor reported an obstacle at a distance of approx. one meter, the data was buffered in the frames for each sensor. The two features viz. $\mathcal{F}_t$ and $\mathbb{F}_t$ were extracted for all sensors from the respective data frames of the sensors.

If the $\mathcal{F}_t$ (which represents the vector of eliminating features of all the sensors) was inside the clusters built a priori during the training, then the decision making features($\mathbb{F}_t$) were extracted from the data frames of all the sensors. If $\mathcal{F}_t$ was outside all the clusters then eliminating features of all the sensors were checked to find whether they lie within the prior defined sensor specific sensing range or not. If the value of any of the eliminating features of a sensor S_k is found to be out of sensor specific sensing range , then the data received from that sensor was ignored in the decision making process for human detection and sensor is added to S^B. For the remaining sensor (S^U), the decision making features were extracted from their data frames . After the decision-making features were extracted, the output was predicted using the sensor specific models. The outputs of these models are either a 0 or a 1 (0 indicates that the data of the processed data frame does not belong to a human being and 1 suggests that it belongs to the human class).

It is possible that no sensor has failed (i.e. S^B is empty) and $\mathcal{F}_t$ does not lie in any of the already available clusters. In such a situation, a new cluster is formed and the corresponding decision making features are extracted from all the sensors. The output of all the respective models is then considered for human sensing.

For the current setup, if AUS is turned off, PIR sensors fails to sense the presence of a stationary human being. Similarly, if PIR sensors are turned off, AUS fails to sense the human being if he walks out at a high speed in front of the sensor.

However, if all sensors (AUS, PIRs) are turned on, stationary human beings (when PIR cannot sense) are identified by AUS. Similarly, human walking at high speed can be sensed by PIRs. The above statement is supported by results presented in figure 4.5. Results shows that that PIRs failed to detect the presence of stationary

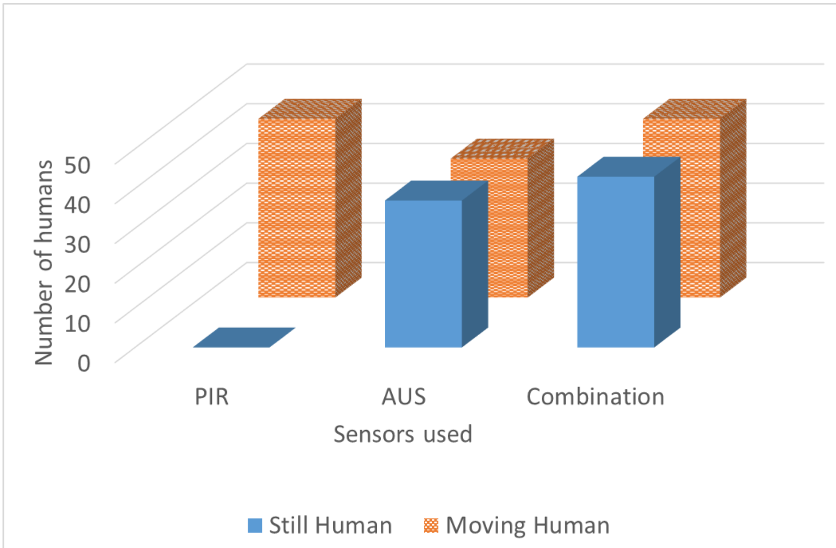


Figure 4.5: Number of human correctly classified for experiment 1 (total number = 45)

human 45 times from a total of 45 and succeed to detect the moving human an all the cases. Also, AUS detected stationary human 37 times from a total of 45 and detected moving human being 35 times from a total of 45. However, using combination of PIR and AUS 43 time stationary human was detected correctly out of 45 times and moving human was detected correctly in all the cases.

4.3.2 Experiment 2:

For the second experiment following hardware was used:

- **Pioneer Amigobot :** This robot has 8 sonar sensor covering an angle of 360 degree, which makes obstacle avoidance possible while moving in an environment.

- **Analog Ultrasonic Sensor (XL-MaxSonar -EZ/AE)(AUS):** This sensor outputs the envelop of the reflected wave which is used to analyze the presence/absence of a human being.

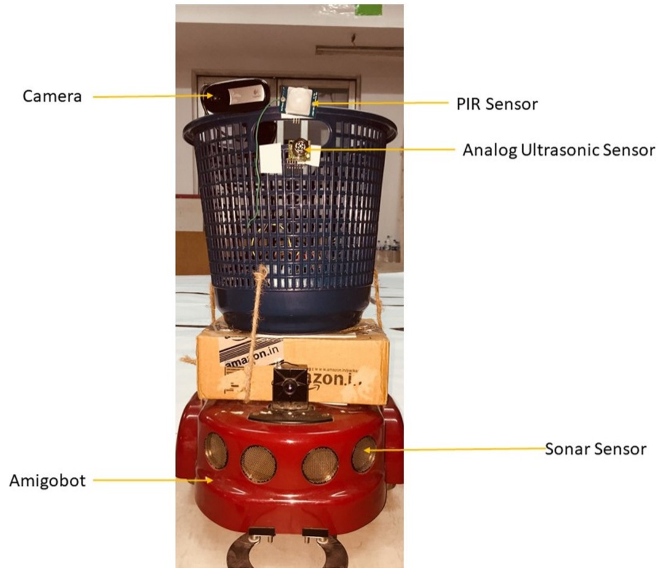


Figure 4.6: Sensors Setup on a Mobile Robot

- **Pyro Infrared Sensor(Parallax PIR sensor (Rev B)):** As prior explained this sensor works on infrared radiations emitted from the body of living beings. This sensor is helpful to find a living being in motion.

- **Camera:** The vision based sensor is used to analyze a given environment. The camera clicks the pictures of the objects in front at a defined frequency to investigate the presence or absence of the human being in front.

In this experiment, the sensors were mounted on a mobile robot. The experiments were performed both in an indoor as well as an outdoor environment. All the sensors (AUS, PIR and camera) were mounted on the robot at a height of 70 cm from the ground as shown in figure 4.6. The mounting was done in a manner that all sensors point to the same obstacle at a given instant of time. The data was processed only when an obstacle was encountered by the robot at a distance of approx. one meter. The robot was programmed in a way that whenever an obstacle was encountered, its speed decreased to 0.05m/sec. The speed of the robot was changed based on the inference derived from the outputs of the sensor data frames. Similar to the previous experiment, a data frame consists of data buffered from the sensors for 4 seconds. All the sensors operated concurrently. Thus, a data frame of all the sensors consisted of information of the same obstacle at a given instant of time. The models chosen for the experiment for PIR, AUS and camera sensors were

SVM, kNN and Convolution Neural Network (CNN) respectively.

The CNN for the camera was trained off-line for the human class using 1000 pictures of human beings in different poses and dresses. The training of the other models for the PIR and AUS sensors was performed on-line where the robot was made to move in an environment which was congenial to the camera (i.e., in daylight) as also the other sensors. As soon as an obstacle was detected at a distance of 60-70 cm, the camera was enabled to capture its image. The data obtained from other sensors(AUS, PIR) was buffered into frames of 4 second while, camera clicks every fourth second when conditions are met.

Features were extracted from the data frames of other sensors and stored to train their respective models, only if the output of the CNN classified the data frames from the camera image to be of the human class. Just as in the previous experiment, the eliminating and decision-making features were extracted from the frames obtained from all sensors.

Sensor	**Eliminating Features**	**Range for human sensing**
PIR	Maximum and Minimum	$>$ 500 mV and $<=$ 0 mV
AUS	Standard Deviation	> 0
$Camera$	Maximum Pixel Intensity and Minimum Pixel Intensity	> 0

Table 4.3: Eliminating features and their specified ranges for PIR sensors, AUS and Camera for Experiment 2

Table 4.3 shows the eliminating features for PIR, AUS and camera with their respective ranges. As per entries in the table 4.3, PIR data should be analyzed only if the calculated maximum value of a data frame is above 500 and minimum value is less than or equal to zero. Similarly, for ultrasonic sensor, the calculated value of the standard deviation of a data frame should be higher than zero. However, for a camera, an image captured cannot be analyzed for human presence if it is completely dark that is both minimum and maximum pixel intensity are equal to zero.

For the PIR sensor, equations 4.4 and 4.5 were used to find the eliminating feature(s) while the equations 4.5 through 4.12 were used likewise for the decision making features. Similarly, for the AUS, the equation 4.12 were used to calculate the eliminating features while equations 4.5 through 4.12 were used for the decision making features. Similar to the experiment 1, multiple machine learning algorithms were tested to select the best one for each of the sensor (to be used for experiment) respectively.

Figure 4.7 shows the calculated accuracies obtained while using different machine learning models for the PIR sensor and AUS respectively. Unlike experiment 1, for experiment 2 sensors were mounted on the robot even for the preprocessing phase of experiment. This includes selection of a machine learning algorithm which outperforms other chosen algorithms for each of the sensor respectively.

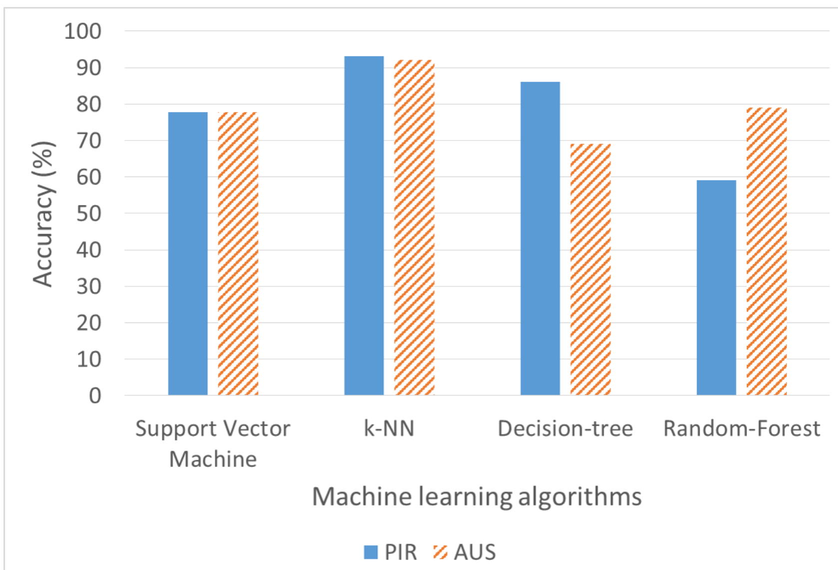


Figure 4.7: Accuracy of different ML algorithms for both AUS and PIR sensor mounted on a mobile robot

It can be seen from the graph in figure 4.7 that K-NN performs better for both PIR as well as AUS as compared to SVM, Decision-tree and Random forest.

For the camera too, the eliminating features were extracted for every image whose output was a human class. The eliminating features of an image captured with the help of a camera are maximum pixel intensity and minimum pixel intensity. Therefore for the complete system eliminating feature vector equation can be defined as below:

$$\mathcal{F} = max, min, SD, maximumpixelintensity, minimumpixelintensity \tag{4.13}$$

.

To train the complete system, the mobile robot was made to move in an environment where rate of human beings to be detected was high. Data was captured for 150 human beings and the eliminating features extracted from the buffered data frames of each sensor were fed to an online clustering methodology so as to form clusters of a defined radius r. With chosen radius equals to 0.1, cluster centers are found. Cluster centers were found using 0.1mm as the radius. Decision-making features were used to train the models for PIR sensor and AUS respectively. A system with three different sensors (viz. PIR, AUS and Camera) and their respective trained models was built tested both indoors as well as outdoors with varying light intensities. For obvious reasons, in dark area the camera fails to detect anything. However, in that scenario, the decision is made by PIR and AUS readings as per proposed approach. For the comparison purpose, the robot is tested in the same environment with PIR sensor alone, AUS alone and camera alone. Accuracies of the three experiments are shown in the figure 4.8. It can be concluded from the accuracy obtained that the accuracy of human sensing is enhanced when a combination of PIR, AUS and camera is used as compare to individual sensors.

4.3.3 Experiment 3

Human being shows dynamism in behavior such as he can walk at various speeds, he can wear different clothings, he can have show multiple posses, etc. Therefore, in this experiment, considering the varying walking speed of human, the robustness of the system was tested. For this experiment, pedometer was used to count the number of steps per minute. To perform the experiment, humans were made to walk at different speeds in front of the sensors in a particular direction as shown in figure 4.9 and 4.10. Figure 4.9 shows the movement of human being w.r.t. sensor set-up of experiment 1 and figure 4.10 shows the direction of movement of both robot and

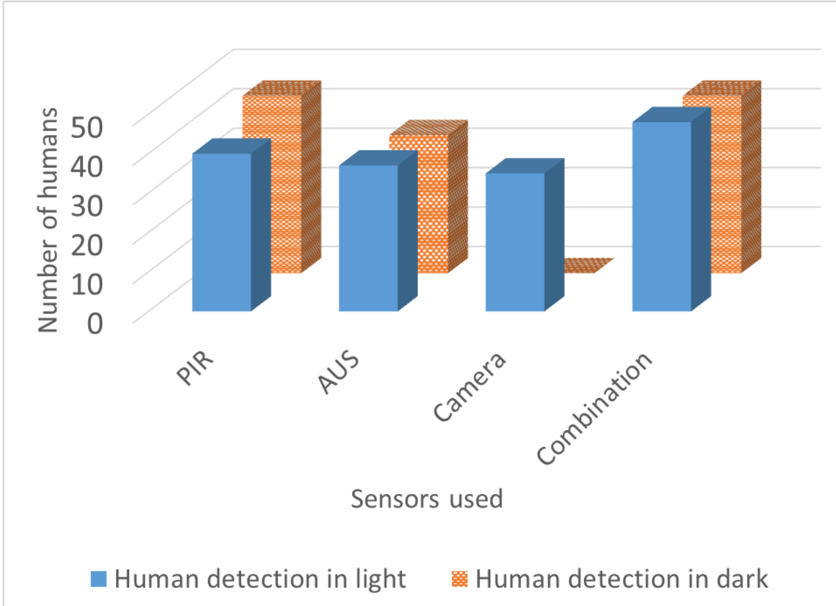


Figure 4.8: Number of human correctly classified for experiment two. (total number = 45)

human being with respect to the sensors set-up of experiment 2. Figure 4.10 also shows the sensing zone i.e. when sensory data is considered for further processing.

Results are compiled in a graph as shown in figure 4.11 and 4.12. It can be concluded from the graph in figure 4.11 that as the walking speed of human increases, True Positive Rate (TPR) decreases. Similar pattern can be observed from the graph in figure 4.12.

The reason for the decreased TPR as observed is that with high walking speed human move away from sensors so quickly that feature values extracted from the data frame are not able to detect human presence. Thus, because of hardware limitations there is an upper bound on the walking speed of human for human detection.

4.4 Chapter Summary

Since human beings, in general, are constantly on the move the use of a dedicated sensor could fail to detect human presence especially when the ambient parameters around the sensor change. In this chapter, a multi-modal human sensing approach has thus been prescribed to overcome this issue. The work focused on automat-

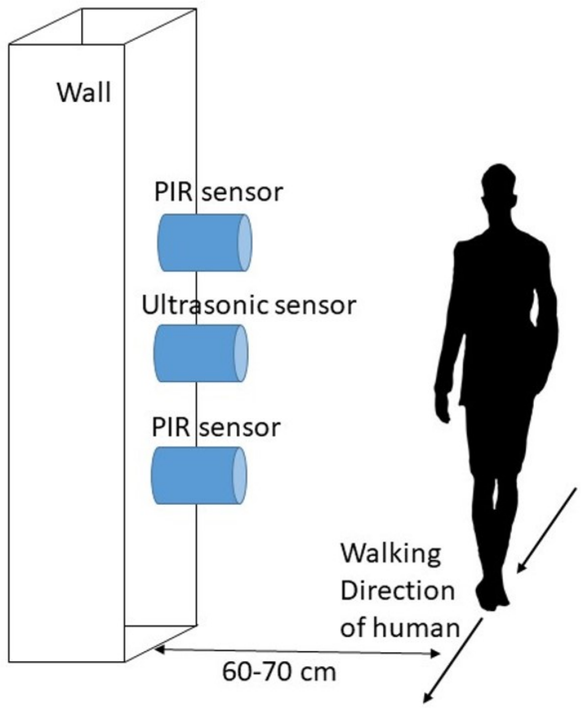


Figure 4.9: Direction of motion of human being during experiment

ing the identification of inappropriate data relayed by some sensors under certain environmental conditions. The proposed approach relied on two different types of features viz. eliminating features and decision making features extracted from the raw sensor signal. For a given environment, eliminating features were utilized to obtain the set of blinked sensors. Ignoring the data received from blinked sensors, decision making features were extracted from the raw signal of remaining sensors. Decision making features were used for the final decision on human presence or absence. Experiments reported include both cases - when the sensors are mounted on a static unit (door frame) and also on a mobile robot. The corresponding results reveal that a combination of sensors outperforms the use of individual dedicated sensors for human detection. Analyses of the walking speed of a human being has also been studied which in turn endorses the robustness of the approach.

An underlying model for human detection is expected to adapt and perform well in terms of accuracy and detection time. The number of features that can be extracted from the raw signals forms one of the parameters that define the computational complexity and time complexity of a given system. Therefore, the selection

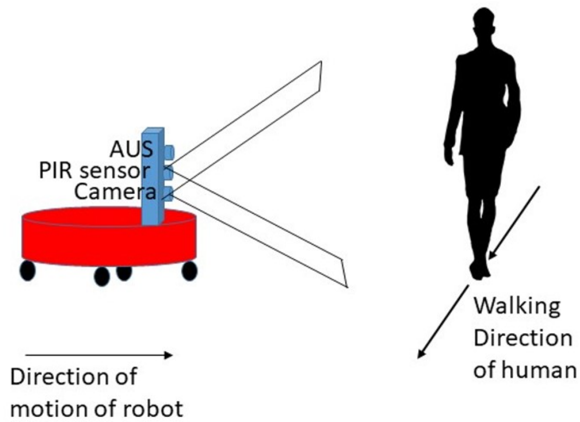


Figure 4.10: Direction of motion of mobile robot and human being during experiment

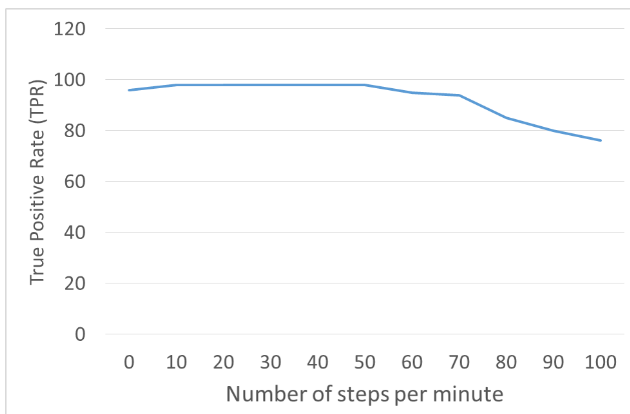


Figure 4.11: Direction of motion of human being during experiment

of the minimal number of features useful in correctly detecting a human being, is a non-trivial issue. An approach to find such features from a give set of features can further improve the accuracy of human detection. The next chapter is based on such a feature selection mechanism. Results obtained from performed experiments highlights the importance of the feature selection.

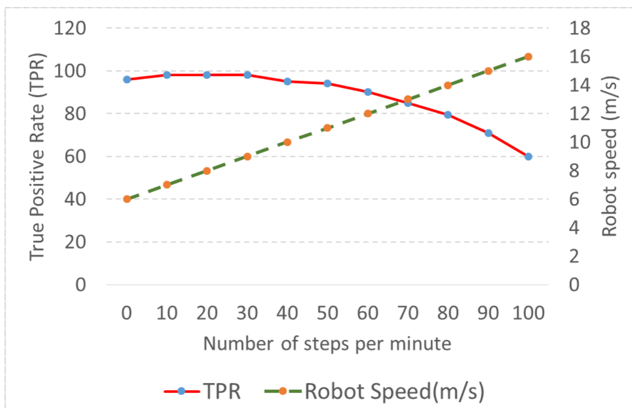


Figure 4.12: Direction of motion of mobile robot and human being during experiment

"Nature holds the key to our aesthetic, intellectual, cognitive and even spiritual satisfaction."

Edward Osborne Wilson (1929)
American biologist

5

Feature Selection Mechanism

As discussed in the previous chapters, the use of multi modal human sensing is one feasible way to overcome the limitations of individual sensors. Multiple features can be extracted from a given sensor data. However, more the number of features higher is the computational complexity. There could also be redundancy in the initially selected features. To remove redundant features and reduce computational complexity, mechanisms for feature selection has recently gained considerable attention. Feature selection has been applied to a wide range of applications wherein the dimensionality and heterogeneity of the data involved is high. In order to build a predictive model for such scenarios, classification algorithms such as Decision trees, Support Vector Machines (SVM), etc. have been used. These algorithms rely on the feature set to be used for classification and prediction.

The efficiency of predictive modeling can be improved by shedding redundant features, thereby alleviating the ill effects of dimensionality. This also results in improved performance of the model especially in terms of its learning rate and accuracy. The feature selection problem refers to the selection of a subset comprising a fixed number of candidate features that optimize on the evaluation measures while also ensuring that there is no degradation in performance compared to that when the super-sets are used [48]. It thus strives to reduce the dimensionality of a given set of features without compromising on performance. Given a set of candidate features for a real application, selection of the best-suited features from the current environment remains a challenge.

This chapter discusses two different approaches for feature selection. The first one is a *Reward and Penalty based online approach* while the second is an *Immuno-inspired Online mechanism.*

5.1 Introduction

Online selection of features along with online learning is a challenging problem.Further, to realize a self-adaptive or evolving system both online learning and online feature selection is crucial. Though, the problem of online learning has received considerable attention from the researchers, the issue of online feature selection needs to be addressed.

A Reward and Penalty Based Approach for Online Feature Selection

This involves an approach that allows a system to learn the best-suited features on-the-fly. A simple but effective approach is presented here that is based on the concept of reward and penalty. Once the system senses the change in the environment, it starts to learn the best-suited feature for the new or changed environment. This contribution focuses on the selection of the best-suited features out of an initial set of candidate features for a given environment. The proposed approach was tested for a service robot that needs to recognize a human being in various indoor environments. Features are suppressed and enhanced on the basis of rewards and penalties in a supervised scenario. A comparison of classifiers with online feature selection and classifiers without feature selection was performed to assess their performances in terms of processing time and accuracy. Results obtained indicated that for the given set of sensors, the rule-based system performs better than Support Vector Machine (SVM). Also the accuracy of the system, which is trained in one specific environment fails when it is tested in a different environment. It was found that the performance of the system is high if the training and the testing environments are the same. This essentially indicates that the system is dependent on the environment and needs to be trained again for every change in the environment. This would mean an addition of a new training data to the system every time the environment changes. Such a system would thus eventually have run out of memory. What is thus desired is a system that can adapt to the new environmental conditions without the need for retraining it.

An Immuno-inspired Online Feature Selection Mechanism

This contribution describes a novel Immuno-inspired Online Feature Selection (iFS) approach to select the best-suited features for human sensing. The approach uses an Immune Network [61] as a base to realize an autonomous and self-learning solution to online Feature Selection Problem (FSP). Unlike the previous approach which relied on a fixed set of predefined training data, this approach does not require off-line training of the model and can respond to the incoming data (antigen) and deliver the best combination of features (set of antibodies). As per Jerne [61], even in the absence of an antigenic attack candidate antibodies within the body of a vertebrate, interact with one another to form an idiotypic network. As a result of this interaction, the corresponding antibodies are stimulated or suppressed. These stimulations and suppressions further affect the life of an antibody and has direct impact on their respective concentrations. Removing redundancy in features must not be compared with other methods, such as Principal Component Analysis (PCA), which are used for reduction of dimensionality of a feature space and such methods are used for feature extraction rather feature selection. This is so because good or useful features need not be dependent on the rest of the data. In the proposed method, given a set of features, the network goes about finding the best of its subsets. The incoming data from the sensors which need to be classified by the subset(s) of the given set of features form the antigens. These subsets constitute the antibodies which eventually evolve to form the immune network. Unlike the existing approaches [158], this immune-inspired feature selection mechanism solution is capable of learning from scratch and evolving while the system provides more data via its associated sensors. Just as in an Immune Network, here too the subsets of features (antibodies) stimulate or suppress one another thereby effecting their respective concentrations accordingly. Those having very low concentrations are eventually removed and replaced by newer ones from a repertoire of subsets based on their affinities to those in the network. This ensures that the redundant ones are either removed or not allowed to enter the network.

Experimental results validate the efficacy of the proposed method. The experiments were carried out so as to classify human beings from non-human where results are obtained with a reduced false positive rate. A comparison of results with that using an SVM indicate the proposed approach to be far better.

5.2 Feature Selection Problem (FSP)

In its simplest form, the Feature Selection Problem (FSP) refers to the selection of the minimal subset of the better suited features F' from a set F of candidate features for a given classification problem T. The set F contains measurable characteristics for each instance of T where an instance constitutes the incoming data from sensors. With respect to a given classification algorithm, the feature set F decides the class (P) of the incoming data it belongs to. Redundancy in F can lead to misclassification causing an increase in the false alarm rate. A minimal feature subset $F_m \subseteq F$, is said to be consequential with respect to T, if it outperforms other such competing feature subsets belonging to F in the classification process. Performance of a subset can be measured on the basis of a certain set of predefined parameters. This could for instance, be the distance of the incoming data from the densest point of C, ratio of distance to radius, etc. As an example imagine three competing feature subsets - $F_1 = \{Max, Mean, StandardDeviation\}$, $F_2 = \{Min, Median\}$ and $F_3 = \{RootMeanSquare, Mean\}$, then the features are extracted from the test data associated with F_1, F_2 and F_3 respectively. Based on the distance of the incoming sensory data from the center of these competing feature subsets, it may be ascertained as to which of F_1, F_2 and F_3 dominate. The dominant one is considered to be the better suited subset of features for this instance of incoming data. Selecting the subset(s) of the better suited features is thus an FSP. Most of the proposed approaches [18] require prior training or test data to build a model on top of which the feature selection is performed by interpreting the instances of the P with respect to training data. Such approaches fair adversely when data is heterogeneous in nature. In such scenarios, a mechanism to solve the FSP without the requirement of prior data thus seems mandatory.

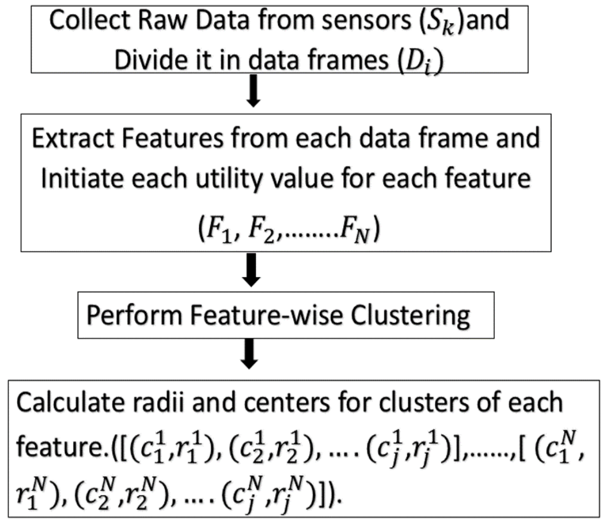
5.3 Proposed Method

5.3.1 A Reward and Penalty Based Approach for Online Feature Selection

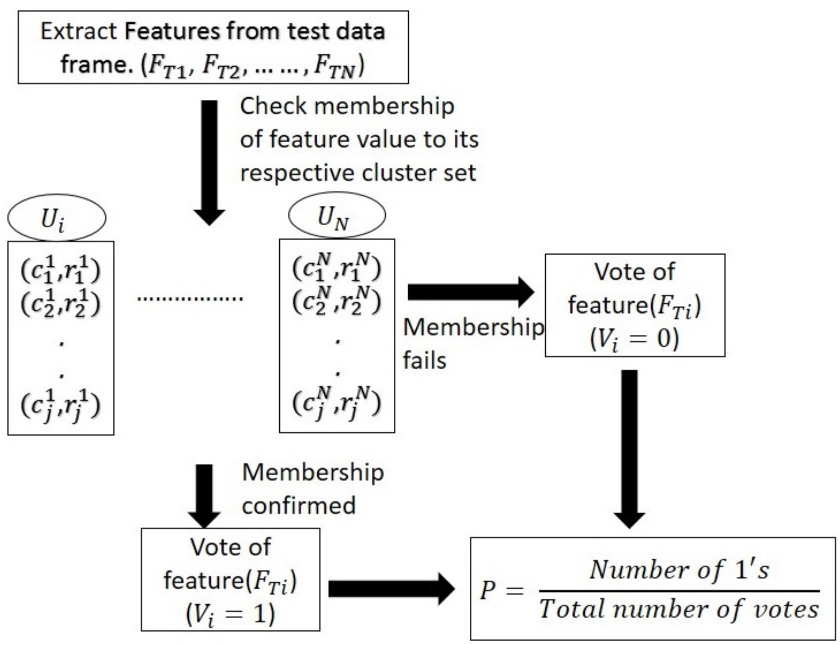
Based on the classification results of Voting-based sensor fusion approach explained in chapter 3, A Reward and Penalty Based Approach for Online Feature Selection is explained in this chapter which is an extended version of the voting-based sensor fusion and detection approach, which is used to fuse the sensory data from different sensors so as to aid the classification process. With the extended approach, a subset of features which are best suited for a given environment can be deduced. It is assumed that all the sensors point to the same target at any instant of time and that the data obtained from all the sensors is buffered at a uniform speed. The initial model is developed in a conventional manner (offline manner) after which the online feature set adaptation is performed. Also, this approach is used if the feature extracted are independent of each other. If two or more features are dependent on each other then combination of these features is assumed as one feature. Some variables used to explain the process which are:

1. $\mathcal{DF}_j^k$ is the j^{th} data frame buffered from k^{th} sensor and j varies from 1 to N_f
2. $i = 1$ to M (where M is maximum number of features extracted)
3. $k = 1$ to N_S (where N_S is the maximum number of sensors used)
4. f_{ji}^k is the i^{th} feature of j^{th} frame from k^{th} sensor
5. N_f = Maximum number of data frames

Collection of data forms a vital step in the development of a model. The corresponding training steps for the system are portrayed in the flow diagram in figure 5.1a. This model is applicable to both scenarios where a single sensor as well multiple sensors could be used. To apply voting based sensor fusion approach, the sensors used need to be prioritized on the basis of their reliability. Reliability can be attained through the literature survey and the classification results of a single sensor. The threshold for each sensor (S_k) can be determined by performing the classification task using the voting based approach [134] at varying probability defined by equation5.2. The probability value at which sensor shows the maximum True Positive Rate (TPR) and minimum False Positive Rate (FPR) can be set as threshold for the sensor S_k. Step 1 through step 4 briefly explains the voting based sensor fusion and classification approach whereas, step 5 through step 7 explains the process of online feature selection process for each sensor.The feature selection process depends on the external feedback received.



(a)



(b)

Figure 5.1: Flow chart for(a) the training steps (b) the voting process

Step 1: The first step is to collect the data for each sensor(s) (to be used). A Data Frame ($\mathcal{DF}_j^k$) can either be defined as data captured from a sensor S_k for a time period of T or framing can be done on the basis of number of instances in a data frame. From each such frame, M number of candidate features can be extracted. The data accumulated for sensor S_k can be represented in the from of a matrix F^k given by.

$$F^k = \begin{bmatrix} f^k_{11} & f^k_{12} & . & . & . & f^k_{1M} \\ f^k_{21} & f^k_{22} & . & . & . & f^k_{2M} \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ f^k_{N_f1} & f^k_{N_f2} & . & . & . & f^k_{N_fM} \end{bmatrix} \tag{5.1}$$

Let Y_i^k be a vector represented as (first column of F^k) :

$Y_i^k = \left\{ f^k_{1i} \quad x^k_{2i} \quad . \quad . \quad . \quad f^k_{N_fi} \right\}$ where, $i = 1$ to M

Step 2: Apply mean-shift clustering technique [33] for each Y_i^k. The Gaussian method has been used to calculate the proximities while the kernel method is flat kernel [22] to calculate the kernel density function. This function is used to determine all the local maxima that form the cluster centers. Clustering technique gives centers and radii of the clusters obtained for each Y_i^k. An initial utility value (U_i) is assigned to each feature(F_i). The upper and lower limit of the utility value need to be defined (that may vary from application to application). The upper and lower limits are used to prune the features later in step 7.

Let, $C_i^k$is the set of cluster centers and radii obtained for Y_i^k which are results of mean-shift clustering. $C_i^k = \left\{ (c^k_{1i}, r^k_{1i}) \quad (c^k_{2i}, r^k_{2i}) \quad . \quad . \quad . \quad (c^k_{N_{clust}i}, r^k_{N_{clust}i}) \right\}$ where,

c^k_{1i} is the center of the 1^{st} cluster of Y_i^k,

r^k_{1i} is the radius of the respective cluster of Y_i^k,

N_{clust} is the number of clusters obtained.

Now, the model built can be tested for classification purpose. The Voting approach for testing process can be explained with the help of flow diagram in figure 5.1b.

Step 3: In order to classify, extract the M number of features from testing data frame of sensor S_1, say $z_1, z_2, , z_M$. For each z_i calculate the membership to its respective clusters. If it is a member of any of the clusters associated with Y_i^1 then vote of that particular feature V_i equals to 1, otherwise V_i equals to 0.

Step 4: For sensor S_k, count the probability of testing data frame to be classified as a member of a particular class by using equation5.2:

$$Probability(\mathcal{P}) = \frac{\Sigma_{i=1}^{N} (V_i = 1)}{N} \tag{5.2}$$

Based on the predefined defined threshold and calculated probability for sensor S_1, the classification result can be decided. From the results obtained and if required, the same procedure is repeated for next sensor in the sequence.

Step 5: If testing data frame is classified correctly, and feedback is also positive then, the utility value (U_i) of the features, who voted 1 is increased to reward that particular feature, whereas utility value (U_i) of the features who voted 0 is decreased to penalize that feature. Otherwise, for the wrongly classified data frame the utility value (U_i) of the features, who voted 0 is increased to reward that particular feature, whereas utility value (U_i) of the features who voted 1 is decreased to penalize that feature . Increment and decrement in utility values can be calculated by measuring the ratio Ψ_i^k of the distance of the feature value z_i where $i = 1$ to M from the cluster center c_i^k to the radius of that cluster r_i^k. If a feature valuez_i does not lie inside any cluster then consider the nearest cluster to find ratio Ψ_i^k .

Step 6: If object confronted belongs to actual class then update the feature utility values U_i^k by adding reward or penalty by using following rules and vice-versa.

1. If $\Psi_i^k <= 0.25$ then $reward = 0.5$.
2. If $0.25 < \Psi_i^k <= 0.75$ then $reward = 0.25$.
3. If $0.75 < \Psi_i^k <= 1$ then $reward = 0.125$.
4. If $1.00 < \Psi_i^k <= 1.25$ then $penalty = -0.125$.
5. If $1.25 < \Psi_i^k <= 1.75$ then $penalty = -0.25$.
6. If $\Psi_i^k > 1.75$ then $penalty = -0.5$.

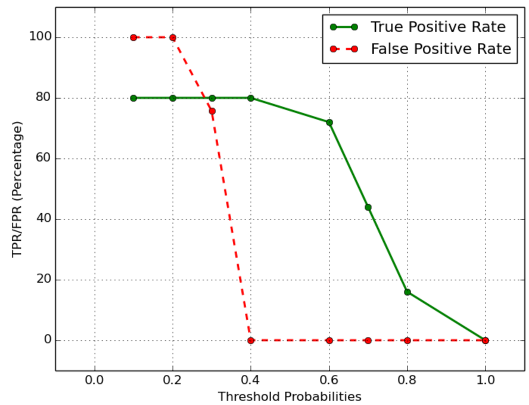
Step 7: Once the feature utility attains its minimum value then that particular feature is exempted from voting for further evaluations. Finally a subset of features with maximum utility can be obtained after multiple test cases and system will attain stability for a particular environment. There exists some features, whose utility value neither increases nor decreases. Such a pattern indicates that that feature sometimes votes correctly and sometimes incorrectly. Obviously such a feature does not fit the best suited subset of features for the specified environment and need to be shed off after sufficient number of iterations.

5.4 Experiment

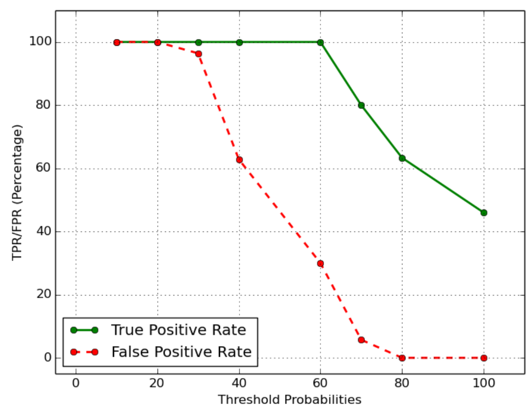
In order to test the proposed approach, human sensing was performed in two different scenarios. The experiment was performed with a robot with sensors mounted on top. The feature selection approach is used to find the best suited features both for PIR and AUS sensors which are used for experimentation. To perform the experiment the following hardware was used :

1. Analog ultrasonic sensor(AUS) Maxbotix MB1300 XL-MaxSonar-AE0
2. Pyro Infrared Sensor (PIR) Parallax 555-28027
3. Pioneer Amigobot

Experiments were performed in a closed door environment. Sensors were placed at a height of 52 cm from the ground and mounted on Amigobot. Signals were acquired from the AUS and PIR sensors. To set the threshold and priority of sensors two experiments were performed using each sensor individually. One experiment was performed by using PIR sensor and the other was performed by using AUS sensor. Classification was performed by using Voting based approach at varying probability. As shown in figure 5.2a and figure 5.2b True Positive Rate (TPR) is high and False Positive Rate (FPR) is low at the probability of 0.4 for PIR. Likewise the same pattern can be observed at a probability of 0.6. Thus, Θ_{PIR} (threshold required for PIR sensor as per methodology section) was set equal to 0.4 and θ_{AUS}(threshold required for AUS sensor as per methodology section) = 0.6. Depending on the number of correctly classified human beings from non-living things results, the PIR sensor seems to be more reliable. Human sensing module was placed on a Pioneer Amigobot Robot. The Robot is set to move in wandering mode and it was made to process the data only when an obstacle is detected in a range of approx. one meter. As per methodology section the system was trained for 300 signals each of 4



(a)



(b)

Figure 5.2: Performance at varying probabilities of (a) PIR sensor (b) AUS sensor

seconds. Therefore, system was trained against data of 1200 seconds. The system was trained only for human class, where different humans in different clothing and in different poses were made to come in front of the sensors at a distance of approx. one meter. To train the system the following features were extracted based on equation 5.21 through 5.28 which are also defined in the previous chapter. The clusters were obtained by mean shift clustering technique. Also, each feature was assigned an initial utility value equals to 50. The minimum possible utility value was set to 0 and the maximum possible utility value was set to 100.

$$Maximum = Max\,(x_i)_{i=1}^{N} \tag{5.3}$$

$$Median = median\left(x_i\right)_{i=1}^{N} \tag{5.4}$$

$$Mean = \Sigma_{i=1}^{N} x_i \tag{5.5}$$

$$Kurtosis = N\frac{\Sigma_{i=1}^{N}\left(x_i - mean\right)^4}{\left(\Sigma_{i=1}^{N}\left(x_i - Mean\right)^2\right)^2} \tag{5.6}$$

$$Energy = \Sigma_{i=1}^{N}\left(x_i * x_i\right) \tag{5.7}$$

$$CrestFactor = \frac{\frac{1}{2}\left(Maximum - Minimum\right)}{RootMeanSquare} \tag{5.8}$$

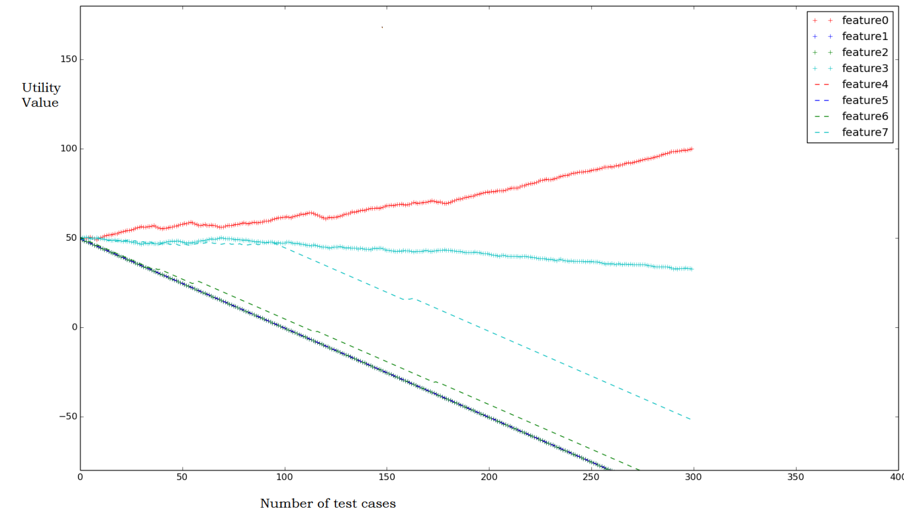
$$RootMeanSquare = \sqrt{\frac{1}{N}\Sigma_{i=1}^{N} x_i} \tag{5.9}$$

$$StandardDeviation = \sqrt{\frac{1}{N-1}\Sigma_{i=1}^{N}\left(x_i - Maximum\right)^2} \tag{5.10}$$

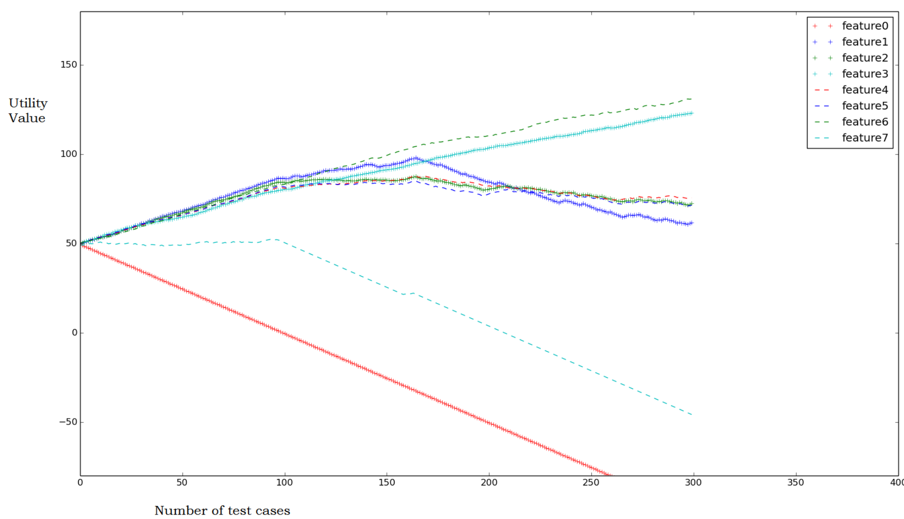
where, N is the number of data samples in the buffered signal (Data Frame) and x_i represents a i^{th} data sample of the buffered file. To perform testing, the robot was made to move in wandering mode. If an obstacle was confronted by the robot at a distance of approx. one meter, robot slows down and start processing the buffered data. Features were extracted from the PIR sensors signals mentioned in equation 5.3 to equation 5.10. Votes for each feature were calculated. Vote was made equal to 1 if the extracted feature value from the test signal lay within anyone of the clusters obtained for that particular feature else the vote for that feature was assigned 0. As explained in methodology section probability of the confronted object to be classified as human being was calculated. If the calculated probability was higher than Θ_{PIR}, then object was classified as human else process was repeated for AUS sensor signal. Once the object was classified as human after consideration of one/two sensors (depending on the probability calculated from each sensor), the system waits for the external feedback. Feedback was received if human said hello to robot. If feedback was positive then the features which voted 1 were rewarded while the others were penalized. Reward and penalty values were decided based on the equations mentioned in the step 6 of methodology section. If system received no external input, which means a negative feedback then, the features which voted 0 were rewarded and others were penalized. After multiple test cases, utility value for each feature changed. Those features, for which utility value reached 100 were selected as the best suited feature for that particular environment. However, if this utility value reached to the value of 0, then that particular feature was suppressed for that particular environment and exempted from the voting. Also, after sufficient number (100) of test cases utility value for some features neither increases nor decreases. It means those features are not best suited for required classification. After multiple test cases subset of features which are best suited for the robot wandering environment were deduced. Also processing time of robot was noted with and without feature selection.

5.5 Results and Discussions

To calculate the results, several experiments were carried out. Robot is kept on wander mode. Experiments were performed in an indoor environment and in a corridor. Initially 8 features were considered as candidate features defined by equation 5.3 to equation 5.10. Figure 5.3a and graph Figure 5.3b shows the variations of utility values for AUS and PIR respectively of various features in closed door environment whereas, Figure 5.4a and Figure 5.4b shows the variation of utility values of various features in corridor. Table 5.1 and 5.2 shows the selected features for PIR and AUS respectively. It is clear from the results obtained that for the same sensor different features behave differently, when environment changes. As shown in table1 best suited feature extracted from PIR sensor for closed door scenario is Maximum whereas for corridor best suited features are maximum, median, mean, kurtosis, energy, crest factor, root mean square and standard deviation. Similarly, as shown in Table 5.2 the best suited features extracted from AUS sensor for closed door experiment are maximum, root mean square, standard deviation whereas, for corridor best suited features are Maximum, mean, kurtosis, energy, root mean square, standard deviation. Also processing time for voting based approach is calculated before and after feature selection. Table 5.4 shows the results of voting based classification technique before and after feature selection. To check the reliability of the proposed approach, procedure of offline feature selection is carried out using combinatorial approach as presented in table 5.1 and 5.2. As per results obtained the subset of best suited features is same in both online and offline scenario. Therefore, it shows the reliability of the proposed approach.

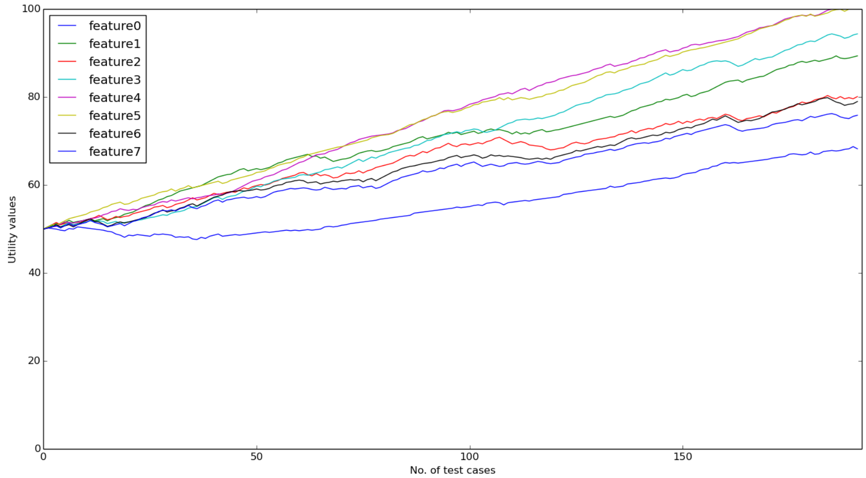


(a)

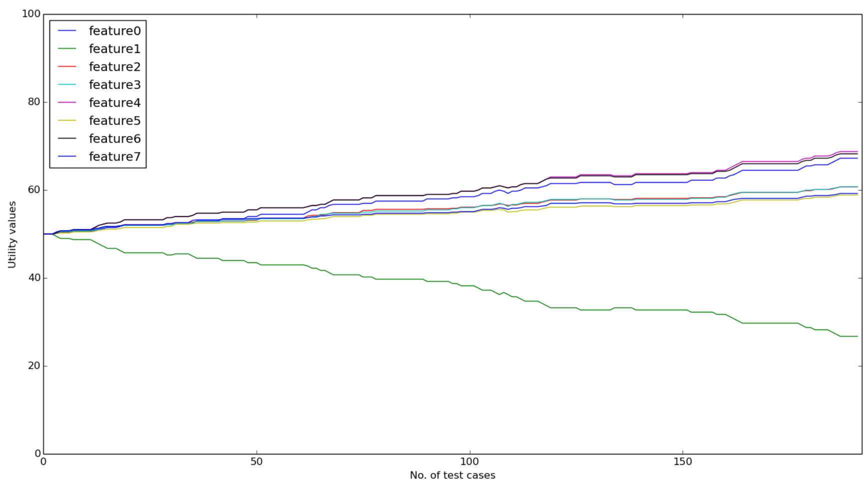


(b)

Figure 5.3: Variation of utility value for features in indoor scenario (a) Variation of utility value for features extracted from PIR sensor data (b) Variation of utility value for features extracted from AUS sensor data



(a)



(b)

Figure 5.4: Variation of utility value for features in corridor (a) Variation of utility value for features extracted from PIR sensor data (b) Variation of utility value for features extracted from AUS sensor data

With reduced number of features, processing time decreases significantly. During experimentation processing time was recorded before and after feature subset selection which are shown in Table 5.3.

As per obtained results, discussed approach is overpowering other approaches. However, this approach requires sufficient amount of training data. Therefore, with change in environment, the discussed approach is effective only if training data

Table 5.1: Results from AUS raw signals in two different scenarios

Candidate Features	**Indoor Scenario**(online)	**Indoor Scenario**(offline)	**Corridor** (online)	**Corridor** (offline)
Maximum	✓	✓	✓	✓
Median	×	×	×	×
Mean	×	×	✓	✓
Kurtosis	×	×	✓	✓
Energy	×	×	✓	✓
Crest Factor	×	×	×	×
Root Mean Square	✓	✓	✓	✓
Standard Deviation	✓	✓	✓	✓

Table 5.2: Results from PIR raw signals in two different scenarios

Candidate Features	**Indoor Scenario**(online)	**Indoor Scenario**(offline)	**Corridor** (online)	**Corridor** (offline)
Maximum	✓	✓	✓	✓
Median	×	×	✓	✓
Mean	×	×	✓	✓
Kurtosis	×	×	✓	✓
Energy	×	×	✓	✓
Crest Factor	×	×	✓	✓
Root Mean Square	×	×	✓	✓
Standard Deviation	×	×	✓	✓

Table 5.3: Processing time of a data frame for classification

	Processing time(before feature selection)	**Processing time(after feature selection)**
Indoor Scenario	10.123ms	5.325ms
Corridor	10.125ms	7.453ms

is available. Collection of training data for every environment beforehand is not feasible. To overcome the mentioned limitation of A Reward and Penalty based approach for Online Feature Selection, needed an approach that can adapt to the changing environment and does not require training data. In the next section, this

Table 5.4: Classification results before and after feature selection using voting based approach

	TPR(before feature selction)	TPR(after feature selection)
Indoor Scenario	92	99
Corridor	83	97

chapter discusses about an Online Immuno-inspired Feature Selection Mechanism which does not require training data.

5.6 Online Immuno-inspired Feature Selection Mechanism

The proposed approach herein, an immuno- inspired approach for online feature selection based on the immune network theory proposed by Jerne [61]. Given a set of features, the network goes about finding the best of its subsets. The incoming data from the sensors which need to be classified by the subset(s) of the given set of features form the antigens. These subsets constitute the antibodies which eventually evolve to form the immune network. Unlike the existing approaches, this immune-inspired feature selection mechanism solution does not necessarily require training. It can learn from scratch and evolve as the system provides more data via its sensors. Just as in an Immune Network, here too the subsets (antibodies) stimulate or suppress one another thereby effecting their respective concentrations accordingly. Those having very low concentrations are removed and replaced by newer ones from a repertoire (of size 2^F) of subsets based on their affinity to those in the network. This ensures that the redundant ones are either removed or not allowed to enter the network.

To explain the Immuno-inspired feature selection mechanism we consider the classification problem C. Let for a given classification problem C, F_x represents the feature vector for a given instance x of C and F is the set of candidate features. Let B be a set comprising all the possible subset of elements e such that $e \epsilon F$. Thus the number of elements of B is 2^F. For the given problem P, the goal of this online iFS mechanism is to provide a set of features F' based on a selection mapping $S(f_x)$

Table 5.5: Immunological metaphors in the Feature Selection

Immune /Biological Entity	**Metaphor in the Feature Selection**
Antigen	Input instances
Epitope	Feature Vector of the problem instance
Antibody	A cluster of all the data instances for one of the possible combination of candidate features
Paratope	Feature Vector associated with an antibody
Idiotope	ID numbers and Stimulation received by an antibody
Affinity	Euclidean distance
Network Dynamics	Equation that governs the stimulations and suppressions of antibodies
Network Metadynamics	Equations that govern the insertion and deletion of antibody in the network
Activation	Sum of stimulations and suppressions in accordance to Farmer's equation

such that:

1. F' does not include any redundancy.

2. The performance (time complexity and accuracy) of the system is enhanced.

In other words, $\phi(S(f_x), B) = F$' where, $F' \subseteq F$ and $\phi(a, b)$ is the performance mapping of a with respect to b.

5.6.1 Immune Metaphors - Definition, Significance and Functionalities

In order to comprehend the proposed mechanism, the analogy between the entities of the biological and the artificial immune network for the FSP is provided in table 6.1. In addition some of these terms are defined below.

1. Antigen (Ag): Each antigen Ag is an instance of the data from a sensor or a set of sensors that needs to be classified.

2. Epitope (Ep): A problem instance x is characterized by its associated feature vector, F_x. A feature vector is an n-dimensional real-valued vector, where n represents the total number of features which epitomizes the problem instance.

The epitope Ep corresponds to one of the subsets of the feature vector, F_x of the given problem instance.

3. Antibody (Ab): The set of data associated to one or more combinations of the elements of the set of candidate features F constitutes the Ab. The structure of an antibody can be represented as given in figure 5.5 also mentioned in [145]. In the figure 5.5, ID represents a unique number assigned to each Ab_i,

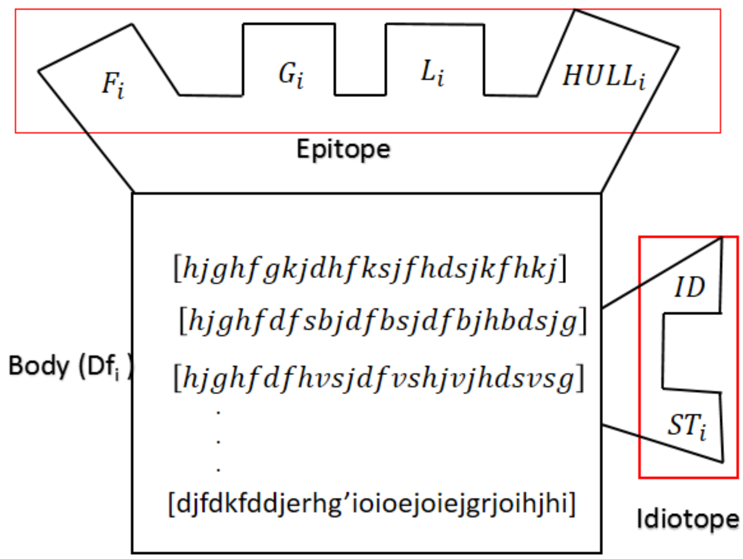


Figure 5.5: Structure of an antibody

DF_i represents the set of data points associated to the subset of the candidate feature set F which represents the antibody, Ab_i and Sti is the stimulation received by Ab_i. F_i represent the subset of F associated with Ab_i. G_i is the centroid of the data cluster formed by the points $\mathcal{DF}_i$ of Ab_i, L_i is the least dense point of this data set and $Hull_i$ represents the convex hull formed by these data points. The overall process starts with $\mathcal{DF}_i = Null$. Data is added incrementally to this set due to which the shape of the convex hull changes dynamically thereby making the antibody evolve with time.

4. Paratope (Pt): This is formed by the, Feature names, centroid, least dense point and convex hull associated with an antibody also highlighted in figure 5.5. The Pt for an Ab provides the information as to which features of the Ag can be tackled by Ab.

5. Idiotope (Id): The Idiotope consists of ID number and the stimulation of an Ab as shown in figure 5.5.

6. Affinity(θ_{aff}): Affinity is a criteria which determines whether or not an antibody Ab has recognized an antigen Ag. If the epitope Ep of an antigen Ag exhibits high affinity towards its paratope Pt, then it essentially means that the Ag is been recognized by Ab. For biological inspired system, extent of complementation between the shape of the Pts of the Ab and shape of the Eps of the Ag decides the affinity between the Ab and the Ag. For a biologically inspired system, the extent of complementarity between the shape of the paratope of an antibody and epitope of an antigen decides the affinity between them. In the computational world this could be proportional to the Euclidean distance between the centroid (G) of an antibody, Ab and the epitope Ep of the antigen, Ag. Affinity $\theta_{aff}(Ab, Ag)$ can be given by,

$$\theta_{aff}(Ab, Ag) = \sqrt{\sum_{i=1}^{n}(G_i - Ep_i)^2} \tag{5.11}$$

where, n is the length of the feature vector F_x, representing the epitope of the concerned antigen. A combination of features forms a candidate antibody, if the antigen, Ag lies within the convex hull boundaries of the antibody cluster, which in turn depends on the calculated affinity between the Ab and the Ag. Lower the value of θ_{aff}, more will be the attraction between the Ag and the Ab and vice versa.

7. Performance Measure(μ) Depending on the affinity between an Ab_i and an antigen Ag_x, a performance measure μ can be calculated for an antibody which is present in the network. For an antibody Ab_i value of its performance measure μ can be calculated by equation 5.12,

$$\mu(Ab_i) = \rho * 1/D_x * 1/DS_i * RD_x \tag{5.12}$$

where,ρ is a scaling factor. D_x is the Euclidean distance of the antigen Ag_x from centroid G_i of the Antibody Ab_i,

DS_i is the density of the convex hull ($Hull_i$) of the antibody Ab_i , $RD_x = d_x/(d_{(G_i)} - d_{(L_i)})$ represents the relative potential of the antigen Ag_x, where, d_x is the potential of the Ag_x with respect to cluster of the Ab_i , $d_{(G_i)}$ and $d_{(L_i)}$ are the respective potential of the centroid and least dense point of the Ab_i. And the potential d of a given point p_i can be calculated as following:

$$d(p_i) = 1/\sum_{i=1}^{n}(EucledianDistance(p_i, a_i)) \tag{5.13}$$

where a_i represents a data point of the data set of the antibody

8. Antigenic Capture(σ_{anti}): It may happen that corresponding to a single antigenic attack multiple antibodies can respond at a time and recognize the same antigen. Such antibodies together constitute the set of candidate antibodies for this antigenic attack. The one that has highest affinity value amongst these antibodies is chosen for the capture. This results in an antigenic stimulation σ_{anti}, of this antibody, proportional to the affinity between the two given by,

$$\sigma_{anti}(Ab_i, Ag_x) = \mu(Ab_i) \tag{5.14}$$

where $\mu(Ab_i)$ is the performance measure of the antibody Ab_i which shows the highest affinity.

The corresponding stimulation received by the antibody is paired with ID and stored in the Idiotope.

9. Activation (α): It is the overall sum of the stimulations (σ) and suppressions (δ) of an antibody in accordance with Farmer's equation. The antibody will purge if its activation is less than a certain lower threshold β.

$$\alpha = \alpha_{prev} + \sigma - \delta \tag{5.15}$$

Where, α_{prev} is the previous activation value of the antibody.

10. Concentration: Rationality of the nature is reflected in terms of the survival of the fittest. Only the fittest one can survive and produce an offspring and thus, increase in population. Here, the term concentration represents the same

concept. The concentration, π of the antibody with the highest value of affinity (θ_{aff}) is increased based on the equation given below

$$\pi = \alpha + y \tag{5.16}$$

where, y is the count of purged antibodies (defined later in this section)

11. Apoptosis: The life time of an antibody is driven by its activation value which also represents its energy level and the number of antigenic attacks it can take. With every attack, the life time T of all non-candidate antibodies, is decreased by a certain factor which is given by -

$$T(I) = \begin{cases} T_{max}, & \text{if } I = 0 \\ T(I-1) - \mu & 1 \leq \text{I} \leq T_{max} \end{cases} \tag{5.17}$$

where,μ represents the performance measure of an antibody and T_{max} is the maximum life time .

Apoptosis occurs when an antibody within the network has either $\alpha < \theta$ or $T(I) <= 0$.

5.6.2 Immune Network Formation Dynamics

Antibodies interact with each other in terms of their performance measure, thus forming an idiotypic or immune network. When an antigenic attack occurs, that antibody, from the subset of antibodies within the network (candidate antibodies) capable of engaging the antigen, which exhibits highest performance and also accepted by the environment, suppresses the other candidate antibodies, while the latter stimulate it. Otherwise, if that antibody is discarded by the environment, then the other candidate antibodies suppress the selected antibody and themselves will get stimulated. Acceptance and discarding of an antibody is done by a third entity in the environment which testifies the success or failure of the antigenic capture. The stimulation, σ_i^j received by an antibody Ab_i from an antibody Ab_j within the network is given by:

$$\sigma_i^j = \theta_{aff}(Ab_i, Ab_j) + \gamma * \alpha_j, 0 < \gamma < 1 \quad (5.18)$$

$\theta(Ab_i, Ab_j)$, represents the affinity between interacting antibodies Ab_i and Ab_j, γ is the scaling factor and α_j is the activation value of Ab_j which is calculated by equation 5.15.

The suppression δ_i^j received by the antibody Ab_j from the antibody Ab_i within the network is given by:

$$\delta_i^j = \sigma_i^j - \mu(Ab_j), 0 < \gamma < 1 \quad (5.19)$$

A biological system follows the principle of survival of fittest. Accordingly those antibodies which fail to compete with the other candidate antibodies, eventually die. Similarly, in the computational model the activation decides the survival of an antibody within the network. If the value of α of an antibody goes below a specified threshold (θ), it is purged from the network. A new antibody from the antibody repertoire thus enters the Immune network. This antibody is selected based on its affinity (β) with the antibodies present in the local network. Insertion of an antibody present in a repertoire is bound to two conditions - 1.) The antibody should either be a superset or subset of at least one of the Ab present in the network. 2.) The antibody present in the repertoire should not have been a part of network earlier or should not have been one that was purged earlier from the network.

The affinity of an antibody Ab_i, which is present in the repertoire and which satisfies the prior mentioned conditions, with any antibody Ab_j which is present in the network can be given by:

$$\beta(Ab_i, Ab_j) = \alpha_j / Z, \quad (5.20)$$

where Z is the number of subsets and supersets of Ab_j in the repertoire.

Figure ?? shows the overview of the immune network which is implemented in this chapter. As can be seen from the figure in the top left corner a set of N candidate features is shown, where its possible subsets constitutes the antibodies, which are stored in the repertoire(in figure only the epitopes are used for the presentation). Whereas, in the right most corner an antigen is shown which is an instance of the problem. Under the problem instance there is a feature vector which represent the epitope of the antigen. In the middle there is an immune network, where a scenario of an antigenic attack is shown. In the network there is a red colour circle which is an antigen whereas, purple ones are the candidate antibodies and the glowing purple is the one candidate antibody (which among all candidate antibodies shows an highest affinity with the antigen) that captures the antigen. If environment accepts the antibody that captures the antigen, then this antibody will suppress the other candidate antibodies which is shown with the help of glowing black lines and also epitope of the antigen is appended to the data set of the antibodies. It is also shown that after an antigenic attack affinity of eligible antibodies present in the repertoire is updated.

5.6.3 Algorithm for online iFS Mechanism

In biological world environment act as a supervisor which provides rewards and penalties that guides the action of a being. Similarly, AIS based system evolves over a period of time in terms of stimulations and suppressions received by an antibody in an Immune Network. These antibodies are stimulated and suppressed based on the feedback received from environment. To calculate the better suited features, an antibody repertoire R needs to be maintained. This stores the paratopes of all antibodies (all the possible combinations of features) together with their respective affinities towards the antibodies which are currently present in the network. The Algorithm 4 is written to describe the Online iFS.

In the given algorithm

- $calc\_affinity(Ab, Ag)$ is a function that calculate the affinity between antibody and antigen using 5.11.
- $Antigenic\_stimulate()$ stimulates the antibody which shows highest affinity towards the antigen using equation 5.14.
- If environment also selects the selected (by antigen) antibody then *Stimulate*

$\_best\_antibody()$ will stimulate the highest performing antibody by using equation 5.18.

- Via. $Suppress\_other\_candidate\_antibody()$ other candidate antibodies will be suppressed by highest performing antibody using equation 5.19.
- If environment discards the selected antibody, then via. $Suppress\_best\_antibody()$ highest performing antibody will get suppressed by using equation 5.19.
- Via.$Stimulate\_other\_candidate\_antibody()$ other candidate antibodies will be stimulated by highest performing antibody using equation 5.18.
- $Calculate\_concentration()$ calculate the total concentration of the antibodies present in the network by equation 5.16.
- $Calculate\_time\_to\_apoptosis()$ calculates the remaining life of an antibody using equation 5.17.
- $Purge_antibody()$purge an antibody from the network if its lifetime expires or its action value is less than β.
- $Calculate\_concentration()$ It calculates the total concentration of the antibodies present in the network using equation 5.16.
- $update\_affinity\_repertoire(R)$ It updates the affinities of the antibodies in the repertoire using equation 5.20.
- $Insert\_antibody()$ It insert a new eligible antibody from the repertoire if number of antibodies in the current network is less than maximum capacity of the network.

For the execution of the above mentioned algorithm, it is assumed that initially, two data points for each antibody are buffered to initiate the mechanism and the upper limit (Max_{num}) on the number of antibodies that can be part of the present in the network at any given instant of time is set. After that, whenever there is an antigen attack, antibodies forming the network are stimulated or suppressed using equations5.14, 5.18 and 5.19, which in turn depend on the affinity values of respective antibodies calculated by equation5.11. If the activation value α, for a

Algorithm 4 Immuno-inspired online feature selection algorithm

```
for all Antigen (Ag) attacks do
  for all Antibody in the current immune network do
    calc_affinity(Ab, Ag)
    Antigenic_stimulate(incomingdata)
    if Reward from the environment = 1 then
      Stimulate_best_antibody(affinity)
      Suppress_other_candidate_antibody(affinity)
      Append Ag to the Abs data.
    else if Penalty from the environment = 1 then
      Suppress_best_antibody(affinity)
      Stimulate_other_candidate_antibody(affinity)
    end if
    Calculate_concentration(affinity)
    Calculate_time_to_apoptosis(activationvalue)
    if (α_i < β) OR (H(I) ≤ 0) then
      Purge_antibody(timetoapoptosis)
    end if
    Calculate_concentration(affinity)
    update_affinity_repertoire(R)
    if Numberofantibodiesinthenetwork < Max_num then
      Insert_antibody()
    end if
  end for
end for
```

given antibody Ab is lower than the defined threshold β, then the Ab is removed from the network and an antibody which is present in the repertoire having highest affinity for the network will enter the network.

The above algorithm will converge when only the last few antibodies remain alive and perform equally well in the network while the rest in the repertoire are the ones that have been purged.

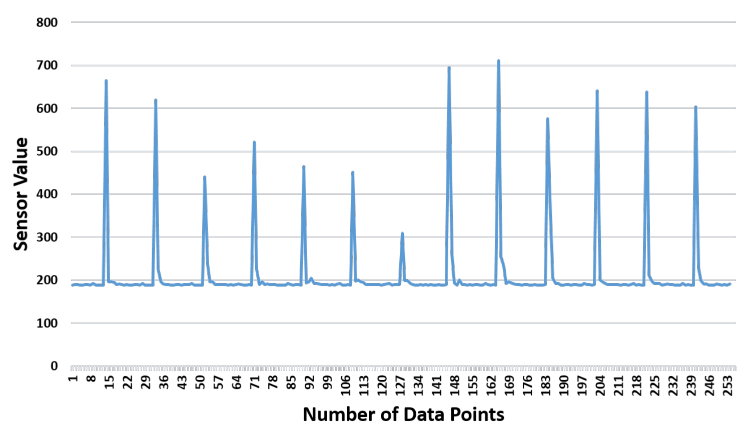
5.7 Experiments

In an IoT scenario where devices are equipped with multiple sensors and numerous features can be extracted from each sensor, heavy computation is required. For instance, if k number of sensors are to be used and from each sensor m number of features can be extracted, then the total number of features to be considered for computation would be $k * m$. Some of the features could be redundant and thus can be removed without affecting the system accuracy. Reducing the number of features cuts the computational complexity. This can be done by selecting the better suited features for each sensor. For experimentation, we have taken the task of human detection which plays a vital role in range of applications such as smart homes, traffic management, surveillance, etc. Human detection refers to the classification of human beings from non-human things or objects. Considering the privacy as of main concern, we have used two sensors for human detection in an indoor environment [133], viz. an Ultrasonic sensor (US) and a Pyro-Infrared (PIR) sensor. The aim of the experiment was to find the reduced set of features for both the US as well as the PIR sensor without degrading the performance of classifier. In this section, we first describe the experimental set up followed by the experiments conducted and results. To check the stability of the proposed approach, we have also compared our proposed approach with conventionally used feature selection methods such as subset analysis, the ranking algorithm , the GA wrapper and Principle Component Analysis (PCA).

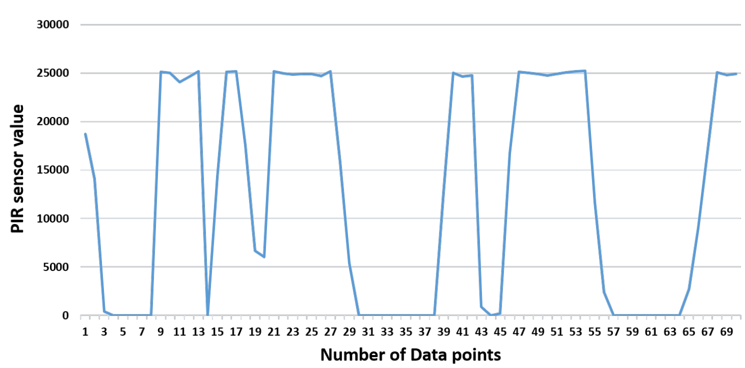
5.7.1 Experimental Set up

In order to differentiate human beings from non-humans entering a room, an analog ultrasonic sensor and a PIR sensor were placed at a height of 72 cm from the ground level within the room. Since the proposed approach is supervised in nature, a camera was mounted nearby so as to confirm the detection of the human/non-human. If both the system and the camera detected a human a reward was provided; else a penalty was inflicted. The detection uses analog signals sensed by both the ultrasonic and PIR sensors.

Figure 5.7 shows a snapshot of the raw signals obtained from these sensors when a human being moves past them.



(a)



(b)

Figure 5.7: Raw Sensor Signal (a) Ultrasonic Sensor Signal (b) PIR Sensor Signal

To analyse these raw signals, feature selection was performed for both sensors simultaneously using these signals. To start with, for each sensor, 8 features, described by the equations 5.21 through 5.28, were considered as candidate features. This amounted to a total number of 255 antibodies in the repertoire per sensor. Initially, each antibody was assumed to have an affinity value equal to zero. The mechanism of feature selection was carried out concurrently on the data received from both the sensors. Unlike other online feature selection mechanisms, prior data sets for training were not used. In order to extract the features defined by equation 5.21 through equation 5.28, data from each sensor was separately buffered for 4 sec and considered to be an instance for each sensor of the human detection problem. These two instances constituted two separate antigens which were processed by two separate processes running simultaneously, one for the US and the other for the PIR.

$$Maximum = Max\left(x_i\right)_{i=1}^{N} \tag{5.21}$$

$$Median = median\left(x_i\right)_{i=1}^{N} \tag{5.22}$$

$$Mean = \frac{\Sigma_{i=1}^{N} x_i}{N} \tag{5.23}$$

$$Kurtosis = N\frac{\Sigma_{i=1}^{N}\left(x_i - average\right)^4}{\left(\Sigma_{x=1}^{N}\left(x_i - average\right)^2\right)^2} \tag{5.24}$$

$$Energy = \Sigma_{i=1}^{N}\left(x_i * x_i\right) \tag{5.25}$$

$$CrestFactor = \frac{\frac{1}{2}\left(Maximum - Minimum\right)}{RootMeanSquare} \tag{5.26}$$

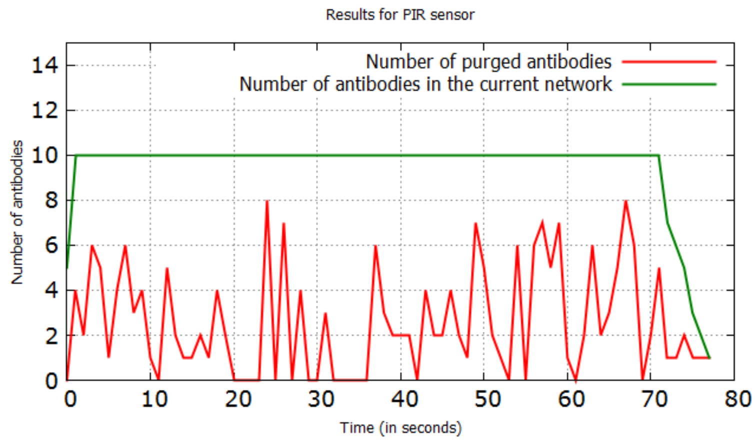
$$RootMeanSquare = \sqrt{\frac{1}{N}\Sigma_{i=1}^{N} x_i * x_i} \tag{5.27}$$

$$StandardDeviation = \sqrt{\frac{1}{N-1}\Sigma_{i=1}^{N}\left(x_i - Mean\right)^2} \tag{5.28}$$

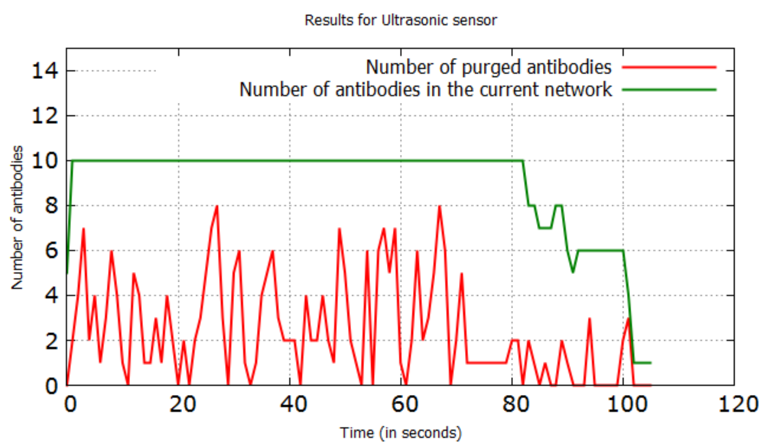
where, N is the number of data points in the problem instance and a_i represents a single data point of the problem instance. The network initially had 5 antibodies in the network which were selected randomly bounded with a condition that each feature should be a part of at-least one of the selected antibody (Ab). In this case subset of features associated with these five antibodies were:

1. Maximum, Energy, Crest Factor
2. Maximum, Energy, Kurtosis
3. Maximum, Median, Mean, Standard Deviation
4. Maximum, Mean, Kurtosis, Energy
5. Median, Median, Kurtosis, Energy, Standard Deviation

The following values were used in the experiments carried out: $\rho = 1000$,$\gamma = 1$, $\beta = -5.00$ and $max_{num} = 10$, $T_{max} = 10$. Repeated experiments were performed to realize the proposed approach and the corresponding results were obtained. The winning feature subsets form the better suited subset of features for human detection problem in the given indoor environment. Once, the system converged to the better subset of features, the selected antibody was used for classification without taking any feedback from camera. For comparison of the proposed approach, classification was also performed using the widely used Support Vector Machine (SVM) [136]. One-class SVM classifier is applied both with the candidate set of features and also with the reduced feature set (obtained after applying the iFS). To compare the proposed approach, experiments were performed using the greedy ranking algorithm [48], the GA wrapper [69] and Principle Component Analysis (PCA) [77] to select the subset of features coupled with an SVM for classification. In case of GA, the population size was taken to be twenty and a single point crossover and a bit flip mutation strategy was used.



(a)



(b)

Figure 5.8: Number of antibodies which are purged out of the network along with total number of antibodies in the current network for (a) PIR Sensor (b) Ultrasonic Sensor

Inference 1: The graph in figure 5.8 shows the variation in the number of antibodies discarded with respect to time for both US as well as PIR sensor data along with the number of antibodies populating the network. It can be seen that several antibodies die with time indicating the removal of a combination of features with every dying antibody. Only a few survive at the end which constitute the set of better suited features. It can also be inferred that different kinds of data can have different converging times. For the PIR sensor data the convergence is faster as compared to that of the US sensor. For both the sensors, the number of antibodies populating the network decreases over time indicating that the approach filters out the more suited ones from the given set.

Inference 2: As features may behave differently for different data instances and so is the convergence rate of an applied algorithm to select the features. It is also clear from the graph in figure 5.16 that PIR sensor data converges fast as compare to US sensor data.

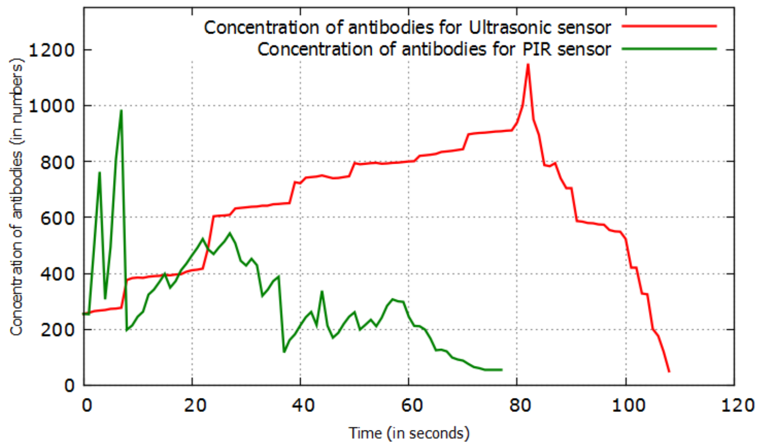


Figure 5.9: Varying Net Concentration of the antibodies for PIR Sensor and Ultrasonic Sensor

Inference 3: The graph in figure 5.9 shows the variations in net concentrations of antibodies. In the initial stages the concentration of antibodies increases to high values while it reduces towards the end. This is so because over time, the better performing antibodies tend to have higher concentrations while the concentrations of the rest dwindle till they eventually die off. The network is thus left with only the better ones which constitute a very limited number of antibodies, thereby lowering the net concentration.

Inference 4: The experiments performed reveal that for the US sensor data, the features - Energy and the Crest Factor emerge as the better suited features whereas Energy alone emerges likewise for the PIR data. Out of the initially selected 16 features for both the sensors, the approach reduces the feature set to a mere 3 saving in terms of the computation involved. The graph in figure 5.16 shows how, in the latter part of the experiments, the concentrations of the better performing antibodies increase while those of the others decrease.

Inference 5: Table 5.6 depicts the classifications results. The SVM classifier was used for classification of human from non-humans. The SVM classifier was used with all the candidate features and with the selected features obtained using iFS. It is clear from the results shown that the accuracy of the SVM classifier improves in terms of enhanced True Positive Rate (TPR) and reduced False Positive rate (FPR), when the reduced subset of features obtained using the iFS is used. Once the features are selected, the clusters obtained from the selected features (after applying iFS) are used for classification. It can be inferred from this table that the clusters obtained using iFS without feedback gives better classification results as compared to the SVM classifier.

Inference 6: Results in table 5.7 and table 5.8 shows the classification results for each of the sensors used. The results show the dominance of proposed method over other widely used feature selection methods viz. Greedy ranking, GA wrapper and PCA.

Table 5.6: Classification Results

	SVM without feature selection	SVM with feature Selection using AIS	AIS based classification
True Positive Rate (TPR)	83.26	95.26	99.06
False Positive Rate (FPR)	0.172	0	0.002

Table 5.7: Classification Results for ultrasonic sensor data with SVM

Feature Selection Algorithm	Maximum Performance	Number of selected features
Greedy ranking	79	3
GA wrapper	92	4
PCA	79	5
iFS	99	3

Table 5.8: Classification Results for PIR sensor data with SVM

Feature Selection Algorithm	Maximum Performance	Number of selected features
Greedy ranking	80.5	3
GA wrapper	93.8	2
PCA	85	5
iFS	99	1

5.8 Chapter Summary

When the dimensionality of the data is very high, solving an FSP is crucial. In addition, feature selection plays an important role in making a system adaptive, especially when data changes with a change in environment. Features extracted previously might not be able to perform equally well in the changed environment. With a reduced number of features, the processing time also decreases. In this chapter, two different approaches for feature selection have been proposed. One is a Reward and Penalty based approach and another one is an immuno-inspired online feature selection mechanism.

Reward and Penalty based approach is a simple but effective approach presented in this chapter. To test the efficacy of the approach, it is implemented using a mobile robot. The robot needs to classify human beings from other objects in an environment using a set of sensors and a Support Vector Machine-based classifier. As the robot moves from one environment to the other, the set of features that were used initially, may not prove to be suitable for use in the new environment. With the proposed approach, the set of suitable features are identified on-the-fly for the new environment and a low false positive rate has been achieved. Further, the approach will also be suitable for other applications based on the Internet of Things (IoT) that attempt to realize smart environments by embedding numerous sensors in the environment. The aim of such applications is to perform the desired task with reduced human intervention. These applications, typically running in resource-constrained devices, demand the processing of streaming data with constrained memory space and time. Thereby, any redundancy in features can increase the computational time as well as consume memory space.

"Success is not the key to happiness. Happiness is the key to success. If you love what you are doing, you will be successful."

Albert Schweitzer (1875 – 1965)
Alsatian theologian, organist and physician

6

Unobtrusive and Pervasive Monitoring of Geriatric Subjects for Early Screening of Mild Cognitive Impairment : A case study of Human Sensing

In the current era of the Internet of Things (IoT), living environments are getting smarter with advancing sensor technology. To monitor the activities of human beings, different kind of sensors are being embedded in the infrastructure. This chapter discusses a case study of non-intrusive human sensing. Human behavior has been studied for a plethora of applications such as elderly care, digital health, assisted living, device automation, security, etc. This chapter discusses one of such applications where elders activities are monitored health. Though the primary focus of such monitoring has been to find trends in the physical health of the subject, recent studies have indicated that the inferences can also be used for a study on cognition.

Healthy aging and access to convenient health care are of significant concern for the geriatric population because the primary marker for on-boarding a neurological patient for engagement is old-age. Hence, cognitive impairment becomes a primary focus of geriatric care. Instrumented elderly care homes, providing ambient assisted living (AAL) use sensors to monitor the activities of daily living (ADL) of users. MCI with recollection complains and deficits are consistently to have most likely progression to dementia, concretely of the Alzheimer's variety. However, differentiating between normal aging and the development of MCI there-in is a research challenge. MCI affects the behavior of the human being; thus this deviation in behavior can be a key for early detection of dementia and symptoms thereof [104]. In this chapter, the use of unobtrusive, non-contact ADL sensors for early detection of Mild Cognitive Impairment in the geriatric population is discussed. The feasibility of using deep learning techniques to make such inferences has been explained.

6.1 Detection of Deviation in the Routine of a Human Being

In this section, an approach to detect early symptoms of MCI is explained which is dependent on the deviation in the daily routine of a human being. The primary goal of the proposed approach is not to detect an unhealthy person as a healthy one. However, it is believed that if a healthy person is suspected of having the symptoms of mild cognitive then going to medico for a checkup is not a destructive process, provided the count of such false detection is not more than 20%. Figure 6.1 represents an overview of the proposed framework.

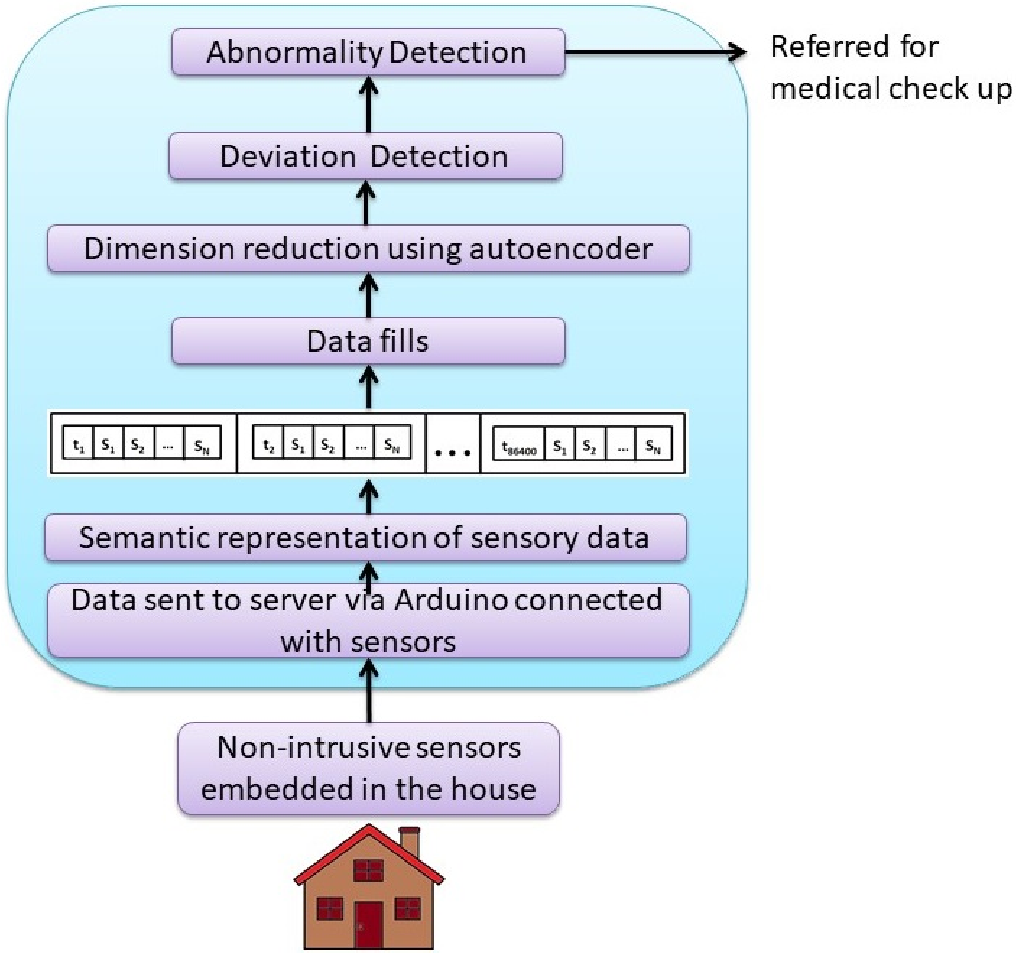


Figure 6.1: Overview of the proposed framework

As shown in figure 6.1 multiple sensors are embedded in the house at different places which are connected to the server via Arduino boards to monitor the activity of a senior person. Data is unlabelled data. Therefore labels of activities are inferred on the basis of sensors physical position and their readings. For example "if the output of the PIR sensor located at the door of the bathroom is one then, it is inferred that person is either entering or exiting the bathroom." Data is stored with timestamp with a frequency of per second.

6.1.1 Semantic Representation of Sensory Data

To monitor the routine of a target human being multiple heterogeneous sensors are embedded in his living infrastructure. Along with the sensor replication, sensor status, sensor position, remaining battery life (used to power up the sensor), current date and time are additionally buffered. Sensor replication of a sensor at a given instant of time is its output value. Sensor status is stored as OK if it is sensing (betokens it is getting powered up) else sensor status is NOK. Sensor position gives information about the physical location of the sensor. With the avail of remaining battery life, the power consumed by the total system utilized for pervasive computing can be calculated. For the experimental purpose, setup is done at 50 geriatric homes in Asia Pacific geography to monitor the health and wellness of the elderly people. Data is labeled on the basis of sensor replication, sensor status, and sensor position. For example "if on 1-1-2017 at 1:00:07 sensor position = bed, sensor replication = 1 and sensor status = OK, then it means on 1-1-2017 at 1:00:07 subject is on the bed and activity is labeled as resting/sleeping." For the analysis of human behavior, data is represented in a vector form. There are two types of vectors, one is a long vector and the second one is a short vector as shown in figure 6.2. To explain the long vector and short vector, let $S_1, S_2, .., S_N$ represents the sensor replication of N number of sensors respectively and $t_1, t_2, ...t_{86400}$ represents the time stamp in seconds associated with sensor values to represent a $24hrs(86400seconds)$ routine. Long vector represents a 24 hrs data of the given subject. This long vector comprises of multiple short vectors. A short vector represents the activity of a human for one second. This vector has the value of sensor replication along with the associated timestamp. Therefore x_{th}long vector for i_{th} subject can be represented as $\mathcal{V}_x^i) = (t_1, S_1, S_2, ..S_N), (t_2, S_1, S_2, ...S_N), .., (t_{86400}, S_1, S_2, .., S_N)$ where,$(t_j, S_1, S_2, ...S_N)$ represents a single short vector at the j_{th} second of a particular day.

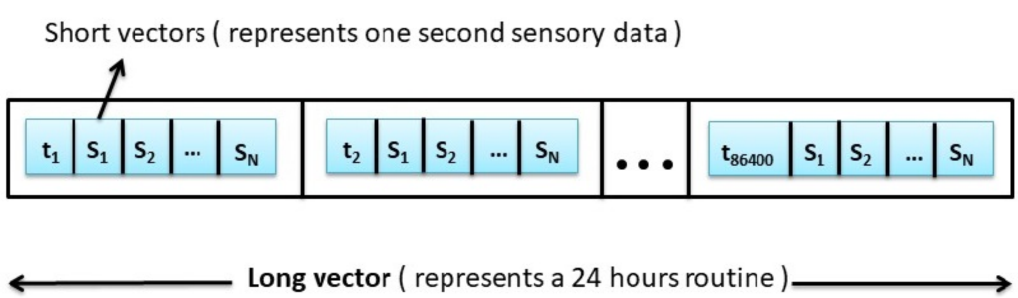


Figure 6.2: Data representation in the form of vectors

6.1.2 Data fills

It is a probable case that sensors can go faulty at any instance of time, leading to gaps in the data stored. To find the missing values in the stored data, an attribute is defined designated as sensor status. Sensor status is set to "Ok" if Arduino connected with sensor receives data and also gateway receives the data from Arduino else sensor status is set to "Nok". When at a given timestamp sensor status is "Nok", then corresponding to that sensor replication is zero. Figure 6.3 shows a few instances of data stored. This causes gaps/wrong sensory value in the data. Therefore, there are gaps in the corresponding long vectors. These gaps need to be filled for the perpetual routine of the human being.

Date	Time	Sensor Status	Sensor Replication	Location of sensor	Sensor type
2017:12:01	00:00:37	Ok	1	door	PIR
2017:12:01	00:00:37	Nok	0	bed	Vibration
2017:12:01	00:00:37	Ok	23	bed_room	temperature
2017:12:01	00:00:37	Ok	0	kitchen	PIR

Figure 6.3: A snapshot of the data stored

In this chapter, to fill the data gaps four prediction techniques are explored which can be listed as:

- Using Hidden Markov Model (HMM)
- Using average of last one week
- Using average of last one month average of the same day
- Using Recurrent Neural Network (RNN)

Using HMM

HMM is one of the most commonly used techniques to predict the activities in advance. Given a last one week data as input HMM model[13] is used to predict the next 24 hrs routine.

Using average of last one week

For this technique, it is assumed that a person follows a similar routine every day. Therefore, next day presaged data is an average of sensory replications over last one week.

Using average of last one month average of the same day

Unlike, the previous technique, it assumes that a person follows almost similar routine on the same day of every week. Therefore, instead of averaging over last one week, data for a particular day of the week is an average of sensory data of the same day over last one month.

Using RNN

Recurrent Neural Network(RNN) [6] predicts the next state of the system on the basis of precedent two states of the system. Hence, using precedent two vectors, next day data is predicted As there are more than one techniques that can be applied to fill the missing data; therefore, the performance of different methods need to be compared to select the best-suited one. With all the four prior mentioned techniques routine of the next day(which has data gaps) is predicted. In the database only a few sensor readings are missing, so for performance measure, an extent of the similarity is being quantified between the anticipated routine and the vector with gaps. The similarity between two vectors is calculated using cosine similarity. While quantifying the similarity extend the sensory value of the sensors showing the status as "Nok" are ignored. Higher the similarity between the two vectors, better will be the performance of the prognostication technique.

6.1.3 Dimension Reduction

A long vector represents the 24 hrs routine of a person which further comprises of 86400 short vectors. Thus, if $N(N is very high value)$ number of sensors are embedded in a house then the size of a long vector will be $84600 * (N + 1)$ (sensor replications from N number of sensors for every second along with the timestamp). As dimensions of the long vectors are very high for manual interpretation of the daily routine of a human being, therefore, the dimension of the data needs to be reduced for further analyses. As per prior research, dimensions of a vector can be reduced by increasing the window size. It infers that instead of considering every second data, data can be considered for a window of five or more minute. However, there is a probability of data loss. So, to avoid the data loss sensory data is considered for every second. Autoencoder [61] is one of the neural network based models which is used by researchers to reduce the dimensions of the data also. Auto-encoder model is fed with long vectors to get the data in a condensed format (reduced vectors) as an output. Auto-encoder model has 3 hidden layers. Data that is taken for further analysis is the output of the shortest hidden layer of the auto-encoder as shown in figure6.6.

6.1.4 Deviation Detection

According to medical studies, during the early phase of MCI, a person tends to forget the random things which reflect in his daily routine activities. Therefore, if the routine of a person is observed continuously then changes in the behavior of a person can be examined using sensory data. Detection of deviation in the behavior of a person is an essential aspect of analyzing the health conditions of a person. When data is represented in condensed form using auto-encoder, then it is expected that similar long vector (represents the one day routine of a person) will have similar representations. If the similarity between two adjacent vectors is less than a particular threshold (θ), then it triggers an alarm for further investigations.

Threshold (θ) calculation

The threshold is a deciding factor for the health and wellness detection of a person. Therefore, any random value cannot be assigned to θ. It needs to evolve on the basis of the routine of the person himself. The initial value of the θ is set to the measure of the extent of the similarity between the first two long vectors. The extent of similarity is measured on the basis of normalized(on a scale of 0-10) value of the difference of two adjacent vectors. For the next comparison threshold range is set to $(2 - \theta_{previous} + 2)$. Current extent of similarity is stored as $\theta_{current}$ and accordingly $\theta_{previous}$ is updated as $(\theta_{previous} + \alpha_c * (\theta_{current} - \theta_{previous}))$. If there is a sudden jump (greater than 1) in the value of $\theta_{previous}$ then an alarm is raised indicating the deviation in the behavior of the human being. In the case of the sudden jump, changes are not reflected in the value of $\theta_{previous}$. However, α_c is set equal to 0.01 to ensure the slow evolution of θ with time and is also capable enough to detect the diversions.

6.1.5 Abnormality detection

Once, the diversion is detected in the behavior of the human being, then the long vector corresponding to the current reduced vector is analyzed and compared with the previous adjacent long vector. Accordingly, the start time and end time of the deviated behavior is stored. Short vectors are analyzed for the duration of deviated behavior(between the start time and end time of the deviation). If some random activities are performed repeatedly in this duration and duration of each activity is very less as compared to the usual duration of that particular activity then a flag is raised to mark the suspected case of abnormal behavior. For example, "if a person wants to go to the kitchen to cook something and he wants some ingredient for cooking in the kitchen which is placed in the bedroom and went to the bedroom but forgets the reason to enter in the bedroom, again come to the kitchen and then goes to the washroom, again comes to the kitchen and goes to the hall". This shows that a person is forgetting things and is confused. Another example is " a person goes to the washroom to take a bath but forgets the reason for entering the bathroom so come back immediately."

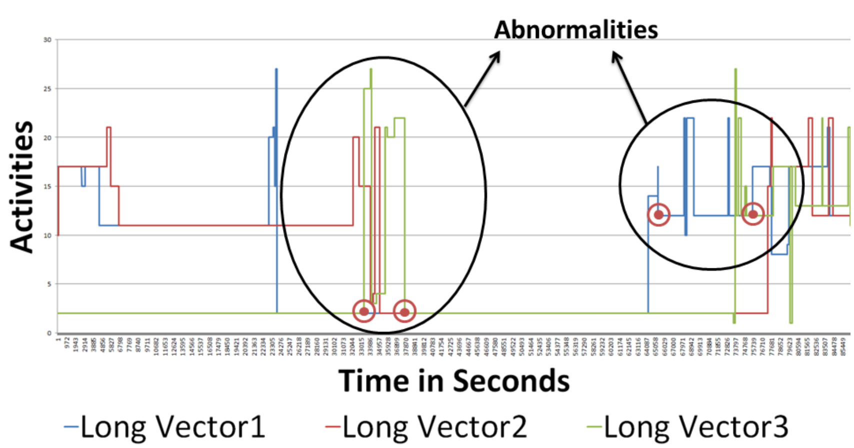


Figure 6.4: Showing abnormality in the behavior

Methodology	**Accuracy**
HMM model	79
average of last one week	85
average of last one month average of the same day	89.5
RNN	92

Table 6.1: Accuracy measures of different methods used to fill data gaps

Figure 6.4 shows a case of diversion where the start time of deviation and end time of the diversion are labeled. If the cases of repeated, random activities and reduced duration of daily activities are encountered continuously over a month then the person is suggested to go for a medical check-up.

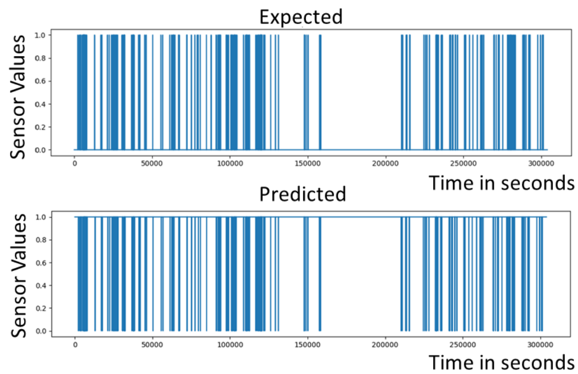
6.2 Experiments and Results

For experiment purpose 14 different kinds of non-intrusive sensors were embedded in the old age homes of 50 subjects. Data was buffered continuously for six months. Whenever sensors status was "NOK" a gap was there in the corresponding long vectors (as defined in semantic representation of data section). As mentioned earlier, a long vector is used to represent the $24hrs$ routine of a subject. As per the requirement of the proposed methodology, these gaps need to be filled. Different methodologies as mentioned in the data fills section were used. Performance of these methodologies was calculated on the basis of the extent of similarity between the predicted vector$\mathcal{V}_{predicted}$ and the existing vector $\mathcal{V}_{actual}$(with data gaps). On comparing the performance of the different method, better one was selected to fill the data gaps of long vectors. Table 6.1 depicts the performance of the 4 methodologies in terms of accuracy. Accuracy can be calculated by using the equation 6.1

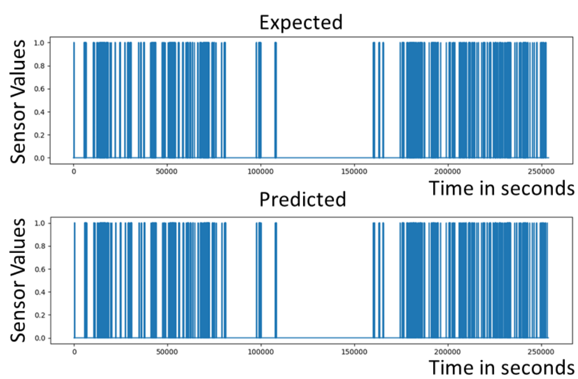
$$Accuracy = \frac{\Sigma_1^{N_v} 100 - (\mathcal{V}_{predicted} - \mathcal{V}_{actual})}{N_v} \tag{6.1}$$

Where, N_v represents the total number of long vectors with data gaps.

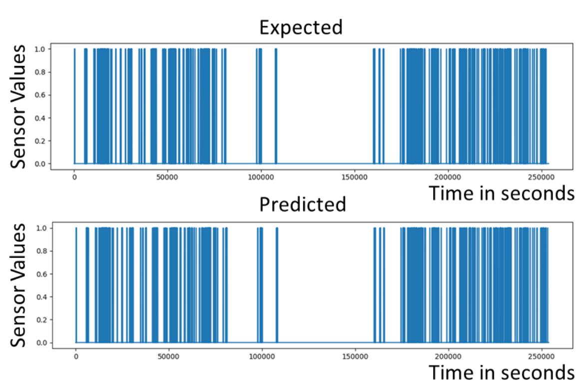
As it can be inferred from the table 6.1 that RNN performed better than other methods in terms of accuracy. Therefore, RNN was used to fill the data gaps. Graphs in figure 6.5 shows the expected value and predicted value (using RNN) of the bed sensor(figure 6.5a), the door sensor of kitchen(figure 6.5b) and the bathroom sensor(figure 6.5c). On the basis of expected values and predicted values of sensors, the difference in the routine of two consecutive days (which was defined in terms of long vectors) was calculated. Once, missing data was filled, long vectors were given as an input to the auto-encoder [61]. As to train the entire system, data of six months is not enough consequently, a new data set was engendered based on the available data set. Considering the fact of varying timings of sunrise and sunset around the year, growing age of elders and incremental global warming, the routine of the subject may shift by some seconds and additionally, accordingly optimum temperature requisites may change. Therefore, the new data set was created with a slight continuous shift in the routine and changing requirements of temperature. While engendering the new data set arbitrariness of the data was maintained to evade the biasing. Training of the neural network was performed in batches and cross-validation was performed to avoid the over fitting of the data. Thus the new data had training data of 150 months. Long vectors representing the routine of each subject for 150 months acted as input to the auto-encoders. The output of this auto-encoder was reduced vectors which represents the same routines but, in condensed form as shown in figure 6.6.



(a) Predicted and actual value of bed sensor



(b) Predicted and actual value of kitchen door sensor



(c) Predicted and actual value of bathroom sensor

Figure 6.5: Predicted (using RNN) and Actual sensor values with respect respect to time stamp on x-axis

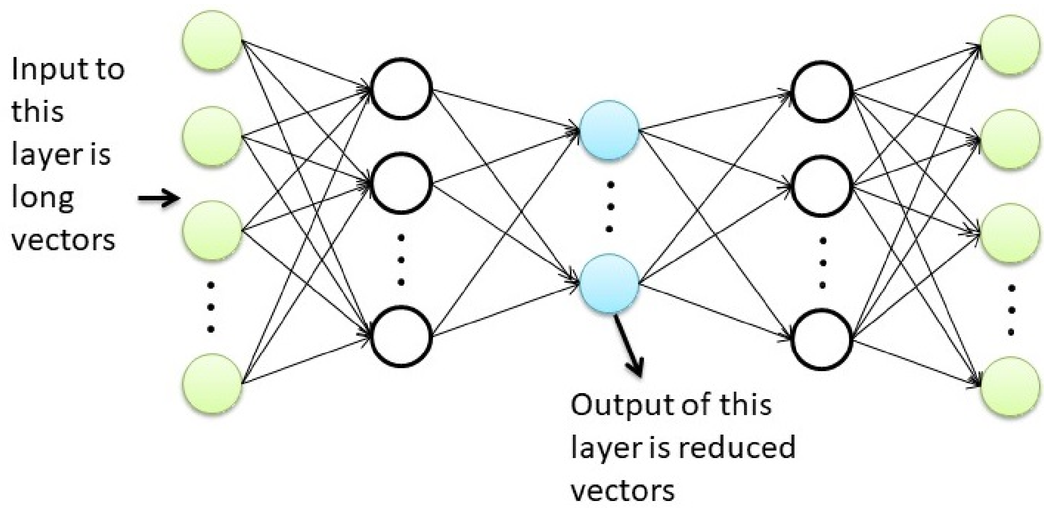


Figure 6.6: Figure showing the input and output with a dummy auto-encoder

After that, deviation in routine was calculated on the basis of similarity between the adjacent reduced vectors. Threshold (θ) was calculated in the same way as mentioned in the deviation detection section. If the similarity between two consecutive reduced vectors was below the defined value of θ, an alarm was triggered. Based on the detected deviation corresponding long vector was investigated for abnormality in the behavior and tagged as an abnormal vector. The difference of this abnormal vector was calculated with its precedent long vector. Based on the calculated difference, along with the duration, the start time and end time of deviated behavior were inferred. As mentioned earlier, if there is randomness in the activities or activity duration is reduced as compared to the usual duration of the same then it is an indication of abnormal behavior of the subject. If this kind of behavior continues over a month, then the subject is alarmed and advised to go for medication. Routines of 50 subjects were analyzed for the detection of early symptoms of MCI. As it can be depicted from the numbers shown in table ?? that out of subjects, an alarm is triggered for 20 subjects for abnormality check. However, the abnormality is finally detected in 10 subjects.

Total Number of targets	50
Alarm triggered for abnormality Check	20
Abnormality Detected	10

Table 6.2: Accuracy measures of different methods used to fill data gaps

Graph in figure 6.7 shows the average routine of 50 subjects. For most of the cases, it varies between 0 to 4 which is normal. However, as it can be depicted that for some subjects such 13, 17, etc. deviation is more than 6, which are a suspected case of illness and referred for medication. Subjects showing deviation in the range of (4-7) are boundary cases and observed more carefully in future.

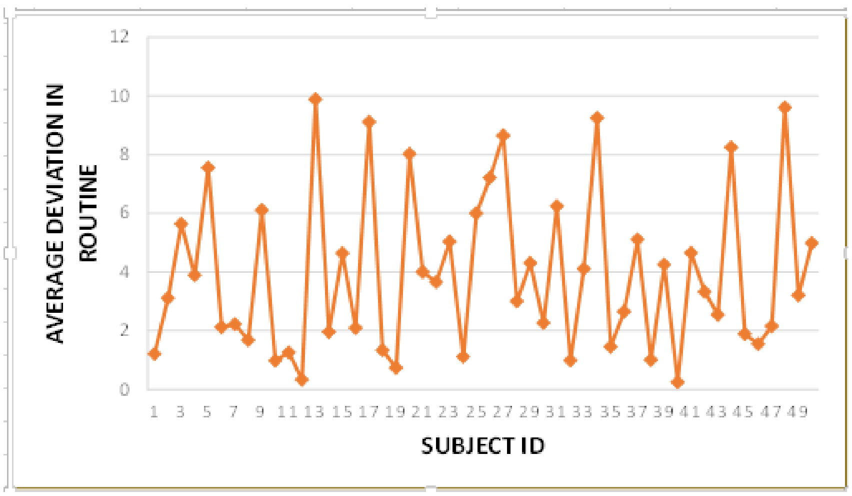


Figure 6.7: Average Deviations in routines of 50 subjects

6.2.1 Classifier Design

Now, so far we can differentiate between a healthy aging person and suspected case of illness. Therefore, this data can be used to train a two class classifier. One class is normal growing person and another class is of deviated behaviors. A 2-layer deep RNN is designed with ReLu function for hidden layers and a sigmoid function for the output layer. Figure 6.8 shows the overview of the framework used for the classifier. As, it can be inferred from the figure that reduced features of all the subjects are given as input to train the classifier. This classifier is tested for 10 healthy subjects which are correctly classified. This chapter presents the initial results only due to limitation of time. However,it opens up the scope of future research for testing the build classifier more rigorously .

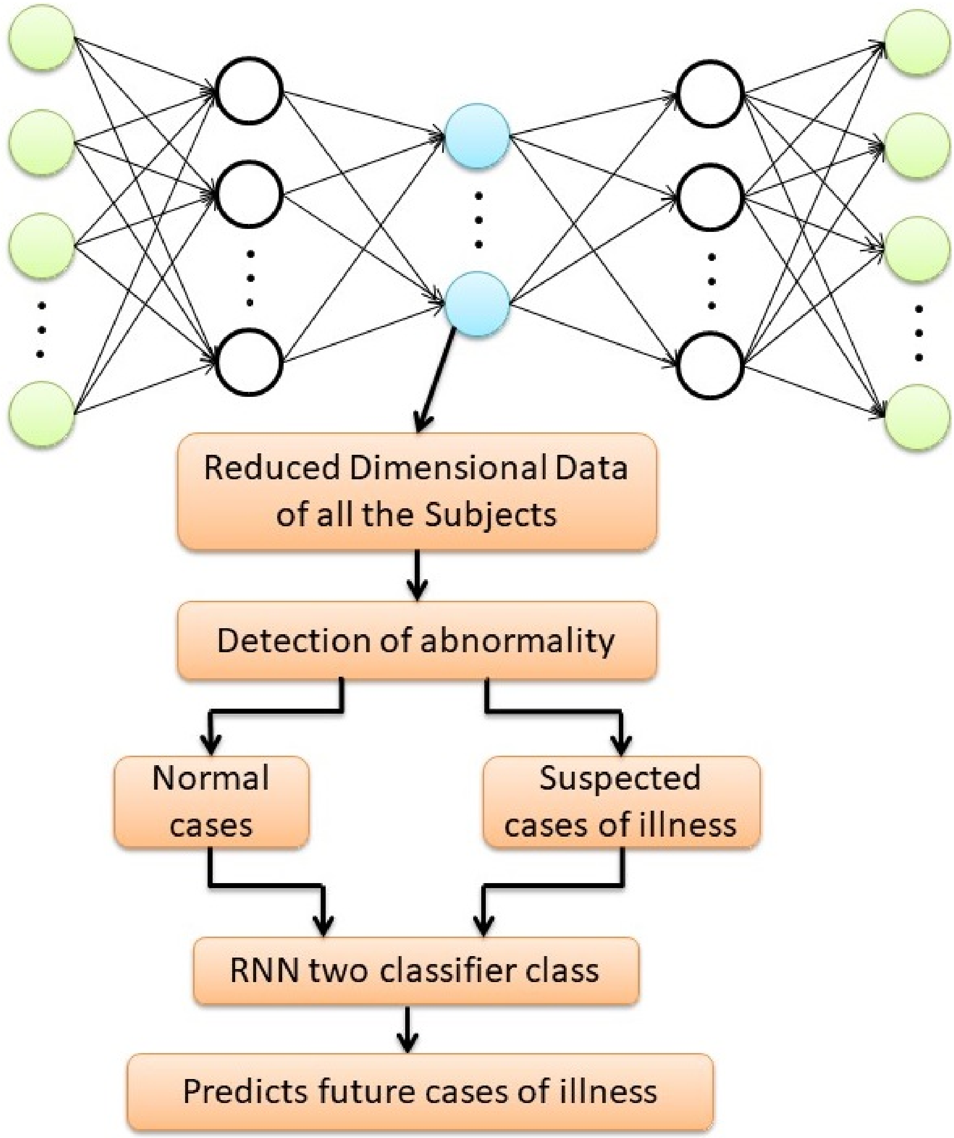


Figure 6.8: Strategy of making the classifier

6.3 Chapter Summary

In this chapter, the feasibility of early detection of MCI using coarse-grained pervasive sensing has been discussed. The proposed approach is an outcome of the analysis of the spatio-temporal data of human activities. Certain challenges such as missing data, vector representation are also discussed in this chapter. Different techniques to handle missing data have also described and tested experimentally. It can be concluded from the obtained results that the LSTM based approach predicts the missing values more accurately. In order to detect early symptoms of dementia the approach proposed in this chapter makes use of the deviations in the behavior of a human being. The results show that a longitudinal study of behavior can shed important insights into the onset and progression of MCI in the elderly population. The methodology can be applied only when a sufficient amount of training data is available. What can be done if training data is not available? Can knowledge be transferred from a place where sufficient amount of data is available, to add smartness to a place without collection of training time. The next chapter attempts to address these issues.

"Don't say you don't have enough time. You have exactly the same number of hours per day that were given to Helen Keller, Pasteur, Michelangelo, Mother Teresa, Leonardo da Vinci, Thomas Jefferson, and Albert Einstein."

H. Jackson Brown Jr. (1991 – 1994)
American author

7

Transfer Learning in the domain of Smart Homes

With advancements in sensor technology the environments within human habitats are becoming smarter. They are now in a position to cater to a variety of needs thus enhancing living standards. Since human beings are dynamic in nature, a solution to add smartness to his living space needs to be tailored to suit his/her specific requirements. This calls for methodologies that can understand, analyze and learn the behavior of the individual. Machine learning techniques form one such solution. However they require a large amount of training data to learn from. Collection and labeling such data can be an enormously time consuming task. As discussed in the previous chapter, the training of a model to detect early symptoms of dementia required six months of data. Collecting such data for each individual is virtually impossible. This chapter thus leverages the use of Transfer learning in the domain of smart homes. *Transfer Learning (TL)* is a concept inspired by the brain's ability to perform tasks, that have never been encountered earlier, using the information learnt from previous tasks.

TL becomes useful when either no data is available or when the quantum of data is not enough for ML techniques to be used. In ML the knowledge obtained from a source domain is utilized in the target domain. This requires one to calculate the similarity between the source and target domains. Multiple parameters have been used to calculate the same. This chapter discusses the utility TL in the domain of smart homes. Using different scenarios, the experiments performed and their corresponding outcomes have been discussed herein to prove the efficacy of the proposed approach.

7.1 Problem Definition

Let Sr denotes source domain and Tg denotes target domain. Sr is well equipped with multiple sensors placed at different locations. The data received from different embedded sensors is used to train a particular algorithm which is implemented to predict the future routine of a person. S has sufficient amount of training data which is collected over months. However, Tg is also equipped with similar sensors but its layout and its resident is different from Sr. Also, Tg lacks in amount of training data. Now, the challenge is to find a way in which the knowledge gained in Sr can be reused in Tg. In other words how the leanings can be transferred from the source domain to the target domain. Knowledge can be values of parameters, machine learning model, raw data, processed data, etc. Let Tg_{prm} denotes the parameters in the target domain and Sr_{prm} denotes the parameters in source domain. So, the problem can be formally defined as :

$$Tg_{prm}\alpha_r Sr_{prm} \tag{7.1}$$

where, α_r represents the relation between parameters of the target domain and the source domain. Knowledge of Sr_{prm} is available in source domain. Challenge is to find the relation between source domain and target domain that is α_r and then using that similarity to determine the value of Tg_{prm} by leveraging the knowledge of Sr_{prm}.

7.2 Proposed approach

Conventional intelligent systems that are designed to cater the needs of a human beings, requires information related to different activities being performed by the human. An activity recognition algorithm relies on training data and yet need improvisation to perform well under diverse circumstances. Labelling a data set consumes a lot of man hours and it is difficult to get the substantial amount of labeled data in every possible scenario. To resolve the mentioned issue researchers are focusing on designing a generalized similarity matrix to find relation between two different data sets and to perform transfer learning. In transfer learning, knowledge obtained from a source domain is transferred to a target domain. Transfer learning is about finding the relevant source data set. One of the challenges in Transfer Learning is to find a compatible source data-set. When this source data-set has been recognized, its relating model parameters can be utilized for transfer learning. However, in case of smart houses, structural likenesses of the living spaces and number of residents can be a useful criteria. The structural likeliness can be measured by maintaining a count of number of sensors, living room, kitchen, washroom, number of smart devices, floor map, etc. Even though limiting the scope to activity recognition, it is unfeasible to calculate all the possible differences between source domain and target domain. In the domain of pattern recognition and behavior analysis of a human being, there can be differences across time, people, devices, data sampling rate, sensor modalities, etc. These differences needs to be considered while calculating similarity between source domain and target domain. Unlike conventional approaches, the proposed approach in this paper does not require labelled data thus saves time and effort required to label the data.

7.2.1 Data Representation

To collect the data for human activities, sensors are embedded in the living space of human beings. Sensory data is stored and labelled manually. For the proposed approach, routine of a human being is represented in a vector form. Let, a vector $\mathcal{V}$ represents the one day routine of a human being. A routine can be defined as the sequence of sensory data collected from various embedded sensors over a period of 24 hours. A vector $\mathcal{V}$ can be divided into subvectors. A subvector consists of data received from all the embedded sensors at a given time instance 't'. If data is stored in seconds then it can be inferred that V has a sequence of $24{*}60{*}60$ subvectors. The dimension of a subvector is equal to the number of embedded sensors in the living space of a human being. Figure 7.1 shows the structure of a vector and subvectors.

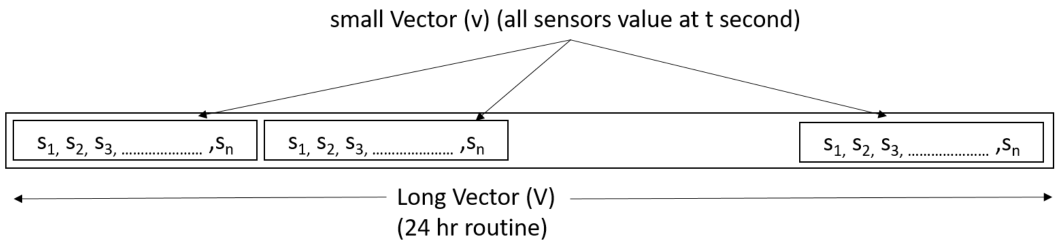


Figure 7.1: Vector representation of data

7.2.2 Similarity Matrix

Transfer Learning is about finding a matching source domain. Depending on the similarity between source domain and the target domain, performance in the target domain using TL can be improved to high extends. Depending on the choice of source domain, transfer learn can affect the performance of the target domain negatively. In the domain of smart homes where apart from the data patterns and its features, sensors types, their modalities, their physical placement and the way in which a human being performs an activity, it becomes a challenge to find the similarity between a source domain and a target domain. For obvious reasons, smart homes are tailored to the needs of individual human being. Being a dynamic creature, every human being has a different way to perform a given activity which is difficult to capture and finding a relation with other human being. Several parameters [26] [103] are defined by researchers to calculate the similarity between the source domain and the target domain. Solutions being proposed by researchers [26] are dependent on the amount of labeled data in the source domain which further add complexity to the problem. Unlike conventional approaches, the proposed approach in this paper is independent of labeled data because it relies the vector representation of the sensory data as explained in the data representation section. Following are the certain parameters that are explained to calculate the similarity between a source domain and a target domain.

Knowledge can be transferred in two ways: 1. Inter-House transfer 2. Intra-House transfer

7.2.3 Inter House Knowledge Transfer

Inter house knowledge transfer can be referred as to transfer the knowledge of the activities of a resident of a house to train a model built for another person residing in another house. However, one of the challenges in Transfer Learning is to find a compatible source data-set. When this source data-set has been recognized, its relating model (LSTM) can be utilized for transfer learning. However, in case of smart houses, structural likenesses of the living spaces and number of residents can be a useful criteria. The structural likeliness can be measures by maintaining a count of number of sensors, living room, kitchen, washroom, number of smart devices, floor map, etc. Even though limiting the scope to activity recognition, it is not feasible to calculate all the possible differences between source domain and target domain. Figure 7.2 shows a scenario where knowledge is transferred from a house to another one.

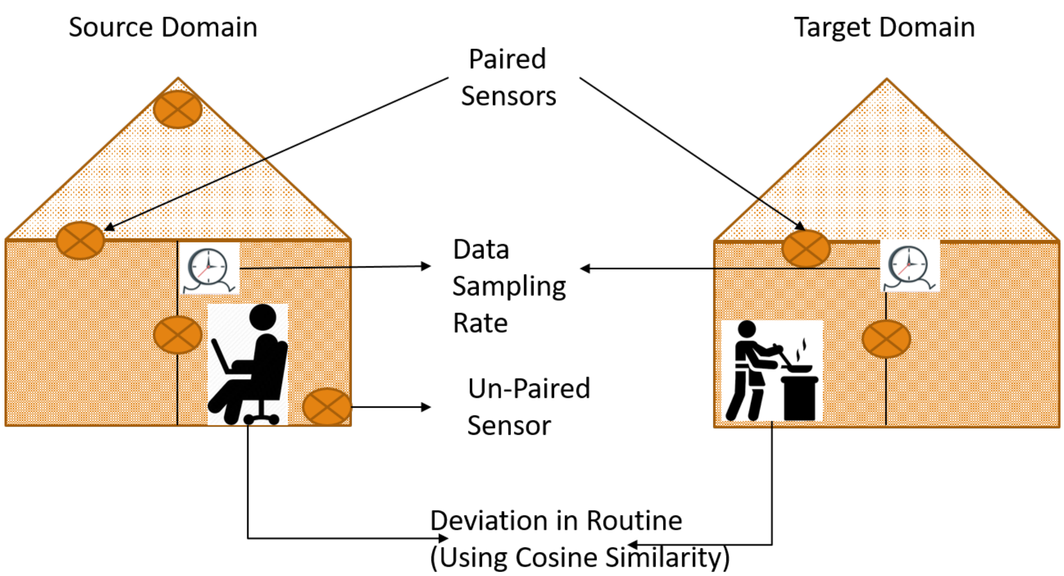


Figure 7.2: Transfer learning in inter-house scenario

Sensor Modality and Physical Space (α_1)

Sensor modalities is one of the essential factor to be considered for transfer-learning techniques. Some techniques may be generalized to sensor modalities, but some techniques are too specific for sensor modalities depending on the application. One of such application is activity recognition where difference in sensor modalities infer the differences between source and target domain. This in turn effect the knowledge that is transferred from the source domain to the target domain. Thus, physical settings of a space is important for the domain of activity recognition. To enumerate the differences between source and target domain in terms of sensor modality and physical space settings we define a term *paired sensors*. Two sensors S_i in target domain and S_j in source domain, are said to be *Paired* iff they have the same modality and possess the same physical settings. For example there is a PIR sensor (S_1) embedded at the door of kitchen in the source domain to detect the movement in that area. Similarly there is a PIR or similar sensor (which can sense the motion of a human) S_2 at the door of kitchen in the target domain. Sensor in the source domain as well as in the target domain is located at the entrance of kitchen and serve the same purpose that is detection of human (movement) in that particular area. Therefore as per definition of paired sensors S_1 and S_2 are paired sensors.

Enumeration of the difference between source and target domain in terms of sensor modality and physical space is dependent on :

1. **Number of paired sensors in target domain:**$|\chi_t| - |\chi_{s_i}|$ Higher the number of paired sensors in target domain, high will be the similarity between target domain and source domain.

2. **Number of unpaired sensors in target domain and source domain:** Corresponding to the unpaired sensors in the source domain there is no sensor in the target domain to capture the similar data. Therefore, learning from the data of unpaired sensors in source domain cannot be utilized in target domain. Therefore higher the number of unpaired sensors in both source and target domain, lesser will be the similarity between source and target domain.

3. **Difference in the total number and sensors places in source domain and target domain:** Total count of sensors and their placement are important to find the most compatible source domain for transfer learning. The difference in the placement of sensors reflects the difference in the data collected for the same activity. Thus, lowering the similarity between source domain and the target domain.

Combining the above points all together α_1 is calculated by equation 7.2.

$$\alpha_1 = \frac{|(\chi_t \cap \chi_{s_i}|}{|\chi_t|} - \frac{|\chi_t - \chi_{s_i}| + |\chi_{s_i} - \chi_t|}{|\chi_t|} + \frac{|\chi_t|}{(|\chi_t| - |\chi_{s_i}|)} \tag{7.2}$$

where $|(\chi_t \cap \chi_{s_r}|$:Number of paired sensors in target domain.

$|\chi_t - \chi_{s_r}| + |\chi_{s_r} - \chi_t|$:Number of unpaired sensors in target and source domain.

$|\chi_t|$: Total number of sensors in target domain.

$|\chi_t| - |\chi_{s_r}|$: Difference in the number of sensors in target domain and number of sensors in source domain.

Number of residents and Data Sampling Rate (α_2)

TL can be performed between multi-resident and single resident spaces and between multi-resident spaces. In these cases data sampling rate is an important factor to be considered to calculate the difference between source domain and target domain. If the sampling frequency of the source and target domain matches, then cosine similarity between the routine of the residents of the source domain and target domain is calculated. Difference in the sampling rate of a source domain and a target domain. To normalize the difference below is the equation to caluculate the value of α_3. Corresponding to calculated cosine similarity, α_2 can be calculated as:

$$\alpha_2 = \frac{\sum fq_{s_r}^{s_k} - \sum fq_{fq_t}^{s_k}}{\sum fq_{s_r}^{s_k}} \tag{7.3}$$

Where $fq_{s_r}^{s_k}$ is the sampling frequency of k^{th} sensor in the source domain and $fq_{f_t}^{s_k}$ is the sampling frequency of k^{th} in the target domain.

Deviation in routines α_3

Conventional machine learning algorithms use either supervised or unsupervised machine learning. However, in case of TL in the domain of activity recognition and prediction, data can be either labeled or unlabeled. With the labeled data , relationship between two instances can be learned which is difficult to learn with unlabelled data. In case of unlabelled data α_3 is defined, which calculates the average deviation in the routine of a human being in source domain. On the basis of average deviation between daily routines this factor helps to improve the build model.

$$\alpha_3 = \frac{1}{N}\sum Cosine(R_{s_{r_i}})(R_{s_{r_j}}) \tag{7.4}$$

7.2.4 Intra House Transfer Learning (α_4)

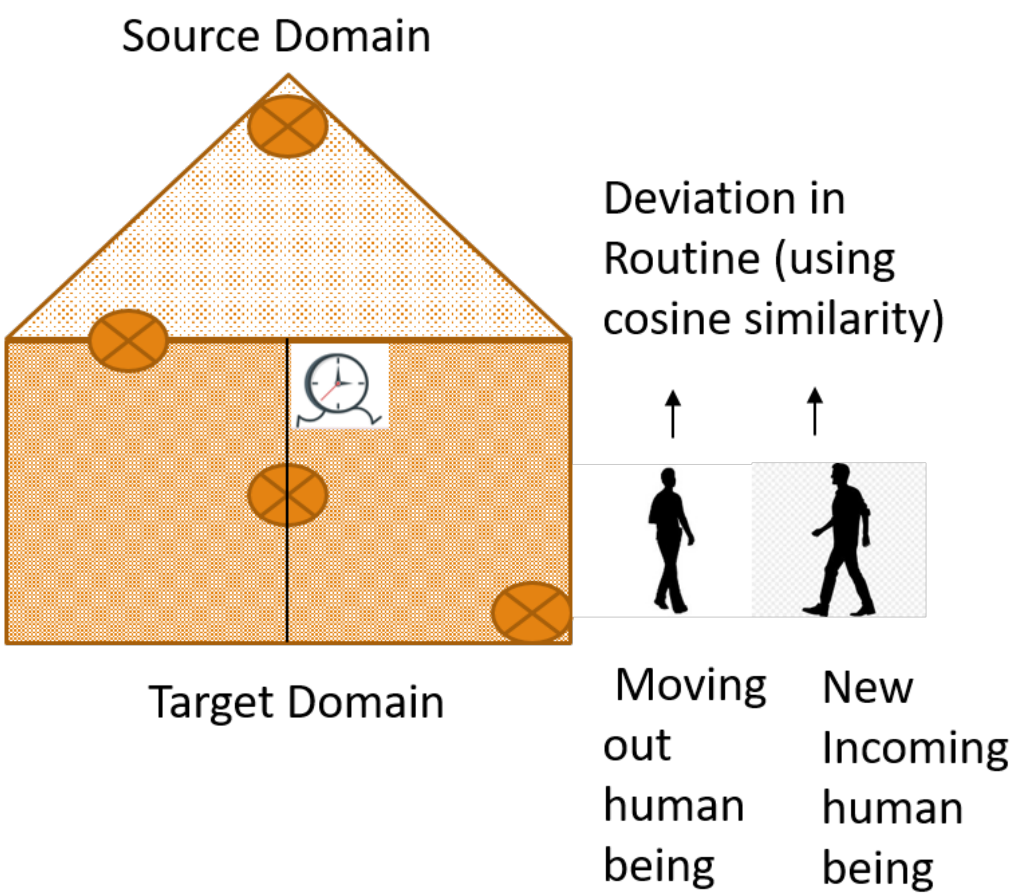


Figure 7.3: Transfer learning in intra-house scenario

In a house, it is one of the probable cases that a person moves out of the house and a new person enters as shown in figure 7.3. In this case, data for the previous resident is available and there is no data for the new resident. In this particular scenario, number of sensors, data sampling rate, physical settings are same except the resident. This scenario refers to the the case of intra house Transfer Learning where the knowledge of the old resident can be leveraged to build a model for the new resident. Below is the equation to calculate α_4:

$$\alpha_4 = W_{old} \tag{7.5}$$

Where W_{old} is the weight matrix of the model trained for the old resident.

Also, in case of multi-resident scenario, the knowledge can be transferred from its old residents to support a new entry in the house. In that particular scenario, below is the equation to calculate α_4

$$\alpha_4 = (\frac{1}{n-1}\sum_{i=1}^{(n-1)} \theta_{ij}^{'}) \tag{7.6}$$

7.2.5 Relation between source and target domain

After, calculation of $alpha_1, alpha_2, alpha_3, alpha_4$, weights for the model to built LSTM for a new comer can be defined as:

$$\begin{aligned}\alpha_{new} = \frac{1}{2}(\frac{1}{3}(\alpha_1 + \alpha_2 + \alpha_3)(W_x^{S_r}) \\ + \alpha_4(W_x^T))\end{aligned} \tag{7.7}$$

7.3 Data Collection

To collect the data, non-intrusive sensors are embedded in the living space of 50 oldage subjects. The sensors include PIR, Vibration sensor, Temperature and humidity sensor, water sensor, gas sensor, ultrasonic sensor, touch sensor, etc. In total 25 sensors are embedded to capture the daily routine of a human being. Sensors are placed at different locations to capture the data of different activities such as bed room, living room, kitchen, bathroom, etc. As shown in figure 7.4 data from different sensors is stored along with time. Data is stored every second. A complete set up for data collection is explained in figure 7.5. Sensors are connected with an arduino board to collect the data. Via gateway data is sent to the IoT middleware for the further analysis.

Time	PIR_bed1	PIR_Kitchen	Heat_Kitchen	PIR_Living	PIR_bathroom	PIR_mainentrance	US_Readi	Vibration_	PIR_bed2
7:10:22	1		20	0	1	0	56	25	0
7:10:23	1		20	0	1	0	56	20	0
7:10:24	0		20	0	1	0	56	0	0
7:10:25	0		80	0	1	0	67	0	0
7:10:26	0		80	0	1	0	76	0	0
7:10:27	0		100	0	1	0	86	0	0
7:10:28	0		100	0	1	1	100	0	0
7:10:29	0		120	0	1	1	255	0	0
7:10:30	1		120	0	1	1	255	0	0
7:10:31	1		120	0	1	0	255	0	0
7:10:32	1		120	0	1	0	255	0	0
7:10:33	0		120	0	1	0	255	0	0
7:10:34	0		121	0	1	0	255	0	0

Figure 7.4: A snapshot of data

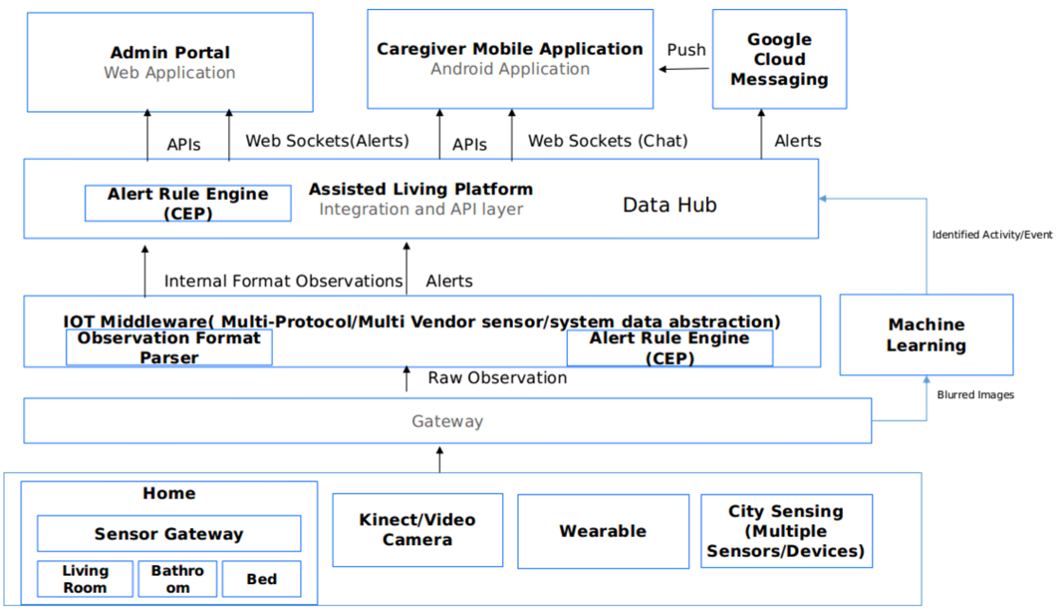


Figure 7.5: Sensor Set-up

7.4 Experiments and Results

To know the efficacy of the proposed approach, different experiments were conducted. Experiments were performed for different scenarios. In every scenario, routine of a human being was defined in a structured manner as defined in the data representation section. The data dimension is too high, so to reduce the dimension encoder decoder was utilized [24]. For every human being in source domain a LSTM [45] model was built. Six months data was used to train the LSTM. Accuracy in the target domain is calculated from the data of three months. Different possible cases were considered to perform the experiments:

7.4.1 Case 1

For the first set of experiments single resident houses were considered. We considered 20 houses for first set of experiments. First 10 houses were tagged as source domains and the another 10 houses were tagged as target domains. For each target house, similarity parameters as defined in the proposed approach were calculated. Based on the similarity parameters, knowledge is transferred from source domain to the target domain. To quantify the results, the cosine similarity is calculated between predicted routines and the actual routine.

Target Domain

Source Domain

	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
S1	76.48	26.07	26.35	0.93	11.2	26.95	13.54	14.21	28.31	66.77
S2	49.25	75.16	14.82	81.18	82.12	12.5	67.35	29	61.93	61.55
S3	5.82	12.48	84.53	15.72	62.17	22.01	86.07	49.13	9.46	85.08
S4	81.15	80.81	81.85	4.46	48.06	13.21	47.74	86.59	65.69	19.9
S5	64.88	14.07	65.73	80.94	6.7	27.87	23.74	64.29	18.04	64.17
S6	20.23	6.59	25.06	19.78	82.99	82.56	12.5	48.08	24.2	24.06
S7	81.13	68.47	60.8	9.82	77.82	22.18	4.75	10.67	49.75	65.6
S8	65.99	26.5	3	61.1	29.13	46.64	67.41	45.06	12.95	5.49
S9	16.91	25.13	24.31	81.96	48.05	28.47	64.77	78.25	26.04	74.82
S10	48.23	14.03	61.78	65.51	18.29	65.51	15.58	28.73	79.26	11.78

Figure 7.6: Average accuracy calculated for every pair of source domain and target domain

Table 7.6 shows the average of the accuracy calculated for every possible pair of a source domain and a target domain. In source domain and the target domain, only single resident houses were considered. It can be depicted from the table that approximately above 80, accuracy can be achieved by using the concept of transfer learning. In the table, some low numbers are also present which agree with the statement that transfer learning can also affect performance negatively. It can be inferred from the table, for domain $T1$, $S4$ is most matching source domain and $S3$ is the domain which donot match at all. Low number in table 7.6 for any combination of a source and target shows the lack of similarity between them.

7.4.2 Case 2

For the second set of experiments both single resident houses and multi-resident houses were considered. For this set of experiments, two multi resident houses were tagged as source domain and 10 single resident houses were tagged as target domain. Each multi-resident house had data of 3 members. Similarity parameters were calculated between the each member of the multi-resident house and the target domain houses. Results are quantifying in the same way as mentioned in the case 1.

Target Domain

Source Domain	T1	T2	T3	T4	T5	T6	T7	T8	T9	T10
S1M1	73.30731	9.610309	62.61693	84.64065	82.74095	15.92107	28.81665	64.64988	8.906789	7.624435
S1M2	17.06282	37.8543	61.95128	70.5779	61.35627	43.89165	69.96214	67.74358	61.41534	46.9376
S1M3	18.60097	17.70722	70.38774	32.24387	76.78658	55.84707	82.17945	8.37364	47.27428	42.43239
S2M1	13.64158	7.86047	2.34441	1.604179	4.98424	17.65348	63.69389	70.9792	25.60029	68.27088
S2M2	20.53806	41.13188	4.98995	71.03995	3.48378	21.13465	26.65539	14.22077	29.47934	58.28119
S2M3	69.50329	5.95727	19.92665	14.06362	44.56252	63.63871	55.90245	80.20133	37.99771	31.68551

Figure 7.7: Average accuracy calculated for multi-resident source domain and single-resident target domain

Table 7.7 shows the accuracy results obtained for scenarios where knowledge is transferred from a multi-resident house to a single resident house. In the table 7.7, for each target domain, maximum and minimum accuracy achieved are highlighted. Accuracy is calculated for every member in a multi-resident house. Similar to the case 1, in this case also there are low numbers which indicate that knowledge can be transferred. For example, for target $T2$, knowledge of member number 2 from house number 2 cannot be utilized.

7.4.3 Case 3

Third set of experiments is opposite to case 2 experiments. For this set of experiments also both muti-resident and single resident houses were considered. But for this set of experiments, single resident houses were tagged as source domain and multi-resident houses were tagged as target domain. Similar to case 2, 2 multi-resident houses were considered and 10 single resident houses were considered. Each multi-resident house needed data for 3 members. For each member of the multi-resident houses, knowledge is transferred from single resident house. For each member in the house, results are quantified using cosine similarity between the actual routine vectors and the predicted routine vectors.

Target Domain

Source Domain	S1M1	S1M2	S1M3	S2M1	S2M2	S2M3
T1	46.93862	63.73833	49.01944	59.93795	23.33763	7.930122
T2	2.25788	28.73875	29.60722	31.71052	43.45623	26.45292
T3	52.35889	22.0797	14.75856	15.11803	20.97263	62.27874
T4	11.73032	4.6316	69.46603	46.18287	48.0228	60.85795
T5	2.67548	43.63824	20.25874	31.90634	58.12904	3.37439
T6	32.94206	1.445989	39.57631	53.10977	23.69626	3.57257
T7	38.58801	51.2638	0.9948	24.05173	8.95955	55.62141
T8	41.37693	24.67901	1.25677	67.42847	3.81904	49.39017
T9	66.83326	35.21875	34.53293	52.45906	21.83953	46.25598
T10	48.99494	57.13905	28.57748	1.18259	54.27336	13.15018

Figure 7.8: Average accuracy calculated where source is single resident and target is multi-resident house

Values presented in table 7.8 shows the average accuracy obtained while transferring knowledge from single resident house to multi-resident houses.

7.4.4 Case 4

Fourth and the final set of experiments multi-resident houses were considered. For this set of experiments, to collect the data of human activities in smart houses, sensors were embedded in 4 smart-house set ups. In small houses 25 sensors and in large houses 35 sensors were embedded. A small house has one bedroom, one living room, one kitchen and one washroom. A large house has two bedrooms, one living room, one reading room, one kitchen and two washroom. The sensors were embedded at different places to capture the different activities of a human being(s). The sensors that were embedded were PIR, US, Vibration sensor, temperature and humidity sensor, water sensor, gas sensor, touch sensor. Sensory Data is buffered at every second. Embedded sensors are non-intrusive to respect the privacy concerns of residents. Data was collected continuously for nine months. In house1 and house 2 data was collected for 4 persons and in house 3 and 4 data was collected for 3 persons. However, one of the member in the house 2 and 4 entered in the house from 180^{th} day onwards. For this set of experiments, house 1 and house 2 were tagged as source domain and house 3 and house 4 were tagged as target domain. Similarity parameters were calculated so as to transfer the knowledge from source domain to the target domain so as to predict the routine of a new entry in the target domain. Knowledge is transferred from different members of the houses in source domain for the new entry in the target domain. Results are quantified in the same as of previous cases.

Target Domain

Source Domain	T1	T2
S1M1	51.67718	60.346783
S1M2	7.184336	6.74416
S1M3	48.20476	23.43874
S1M4	39.92432	26.76426
S2M1	62.45561	40.62514
S2M2	59.32822	50.78214
S2M3	2.37728	4.98166
S2M4	25.49385	57.19945

Figure 7.9: Average accuracy calculated where source and target both are multi-resident houses

Table 7.9 shows the results obtained for transfer learning from multi-resident house to multi-resident house. For each of the target domain the source domain whose knowledge gives maximum accuracy and minimum or low accuracy.

7.5 Chapter Summary

The concept of Transfer Learning was proposed in order to mimic the capability of the human brain to learn new things from old knowledge. TL is useful in cases when there is little or no data available. However, it can work only when the source and target domains are similar. In this chapter, the use of TL in the domain of smart homes has been explored. Earlier work conducted by researchers relied on the amount of labeled data in a source domain. To overcome this, an approach which uses unlabeled data in source domain has been proposed herein. Different parameters have been used to calculate the similarity between the source and target domains. To prove the efficacy of the proposed approach, different experiments were performed.

Experiments include cases where knowledge is transferred from single resident houses to single resident houses, single resident to multi resident houses and vice versa and multi-resident houses to multi-resident houses. However, the usage of transfer learning between different application areas such as smart hospitals, smart offices, smart homes is left as future work.

"Study Nature, Love Nature, Stay Close To Nature, It Will Never Fail You."

Frank Lloyd Wright (1867 – 1959)
American architect

8

Conclusions and Avenues for Future Research

With advancements in sensor technology, ambient environments are getting more keenly intellectual so as to avail the living standards of human beings. Sensors are constantly creeping in human environments and habitats thus improving both quality and security. In the current era of the Internet of Things(IoT) where human beings also participate in the loops, a technique that can reliably sense their presence is mandatory. State-of-the-art smart living environments have the goal of being ubiquitous and pervasive with minimal obstruction to the subject. In this respect, device-free and passive technologies are increasingly being employed. Multiple sensors such as Ultrasonic, Pyroelectric Infra Red (PIR), camera, etc. are being embedded in scenarios such as home automation [48], elderly care [67], security [31], etc. With emerging machine learning techniques, signals from sensors are being analyzed for human detection in indoor scenarios. This thesis addresses the importance and challenges in human sensing. The thesis follows a bottom up approach. The first half of the thesis introduces the methods to improve the accuracy of human sensing followed by the deployment of human sensing to analyse the human behaviour for health and wellness. This chapter presents a summary of the contributions made and discusses their applications envisioned in the real world. It concludes with the avenues for future research and directions for their extension.

8.1 Summary of the Thesis

This thesis is aimed at non-intrusive human sensing. It is envisaged that the contributions made in this thesis are applied to scenarios wherein non-intrusive human sensing is needed. Such applications include assisted living, elderly care, digital health, human robot interaction, pedestrian safety, border security. The success of all these applications are dependent on the accuracy of human detection. In this direction, the first contribution (Chapter 3) , focuses on the sensor selection to sense human presence. To sense the human presence, an Ultrasonic (US) and a (PIR) sensor were separately used to differentiate human beings from non-human things. However, as discussed in chapter 2 every sensor has its own inherent limitation(s). One of the feasible solutions to overcome this limitation is to use the multiple sensors. Therefore, to improve the accuracy of human sensing, an algorithm to combine the data from both these sensors is proposed in the thesis. The proposed approach was tested by using a combination of US and PIR sensors in an indoor environment. From the results obtained it can be inferred that using combination of sensors better accuracy for human sensing can be achieved than methods incorporating only single sensors. This contribution confirms the fact that a sensor is susceptible to failure under certain scenarios and one viable solution to overcome this is to use other complementary sensors.This laid the foundation of the next contribution (Chapter 4), thus presents a multi-modal human sensing approach that facilitates autonomous sensor selection based on the changes in the environment. The proposed approach made use of eliminating features as well as decision making features. On the basis of eliminating features, sensory data from which information regarding human sensing cannot be extracted were ignored. From the remaining sensors data, decision making features were extracted for further processing. Experiments were carried out using a camera, a PIR and US sensors in different environments. The results obtained using this approach prove its efficacy. An underlying model for human detection is expected to adapt and perform well in terms of accuracy and detection time. The number of features that can be extracted from the raw signals forms one of the parameters that define the computational and time complexity of a given system. Therefore, the selection of the minimum number of features useful in correctly detecting a human being, is a non-trivial issue. An approach to find such features from a given set of features constitutes the third contribution (Chapter 5). In this chapter two different approaches for feature selection are explained. One is reward and penalty based feature selection and another is Immuno-inspired online feature selection mechanism.This contribution applies the proposed feature-selection approach to the data received via US and PIR sensors for human sensing. A comparison of results obtained by using a classifier with and without using feature selection, has been presented. Results indicate that a change in the environment causes the set of selected features to also change. The approach has been substantiated by experiments carried out in the real world.

Behaviour profiling of a human depends on the success of human sensing. Different sensors can be embedded in the living space of human beings so as to track his/her activities. The behavior of a person needs to be analyzed from the data obtained. The next contribution (Chapter 6) describes techniques to detect the early symptoms of cognitive impairment in human beings. Instrumented elderly care homes, providing ambient assisted living (AAL) use sensors for monitoring the activities of daily living (ADL) of the users. In the fourth contribution, the use of unobtrusive, non-contact ADL sensors for early detection of Mild Cognitive Impairment in a geriatric population has been explored. This approach make use of deviation in the routine of a human being. The feasibility of using deep learning techniques to make such inferences has been shown. The proposed approach is tested for elderly people staying in old-age homes. Further, this information is used to design a classifier to predict the future cases of illness.

Delivering accurate and helpful information on the following action to be performed by a person is an essential factor in pervasive computing. The research of activity detection in a single occupant space is mature but it requires ample amount of training data *apriori*. However, data collection is a time consuming. The next contribution in the thesis, deals with an approach using transfer learning in case when there is insufficient or no data. To add flexibility to a system, the concept of transfer learning has been explored to support the new entry of a person in a space which is shared by several people. TL can however, be used only when the similarity between a source domain and a target domain is high. TL can have both positive as well as negative effects. Experiments performed considering multiple combinations of source and target domains prove the utility of proposed approach.

8.2 Future Research Avenues

The work reported in this thesis provide ample scope and promulgate several clear directions for future research endeavors. This thesis in no way discourages intrusive human sensing but highlights the benefits of using non-intrusive human sensing and how it can be achieved. Further due to the dynamic behavior of human beings, this thesis does not claim to have solved the human sensing problem in totally. The problem still has several unanswered questions which open the gates for future research.

The proposed mechanism for feature selection in chapter 5 when used for human detection based feature classification on incoming data streams from a US and PIR sensor provided results comparable and superior to that obtained in conventional methods. Since the Artificial Immune Network evolves over time, addition of new antibodies (subsets or new features) to the repertoire and hence to the network, could also be accommodated. This on-the-fly addition of new features can thus facilitate scalability especially in cases where the incoming data is highly heterogeneous as in case of networked scenarios such as an IoT. Data from a set of sensors could become heterogeneous especially when the environment around it changes. New features can be added by adding new antibodies in the repertoire so that the network can evolve them over the resulting period. One may thus add a new sensor resulting in new features being added in the evolution of the network. Using the same technique one may also remove a sensor which has proven to be less useful or redundant thereby saving on both power and computation. In future, the study of the effects of on-the-fly addition and deletion of new feature subsets, on the classification of heterogeneous data to make such immune networks scalable and adaptable to varying environmental conditions, can further improve the accuracy of ML algorithms.

With the massive growth in sensor technology, human behavior analysis has proved its importance in the domain of digital health. In chapter 6 a case study of non-intrusive human sensing has been discussed to detect the early symptoms of

MCI. The analysis has been performed using single resident elders houses. However, it will be interesting to analyze how temporary visitors data could be handled. Extending this work to other age groups to track their health and wellness in a multi-resident space also needs to be investigated. Whether deep learning methods can be combined with statistical methods to analyze the spatio-temporal data of human activities is another aspect that could be explored.

Chapter 7 discusses the utility of the concept of transfer learning in the domain of smart homes. However, can knowledge be transferred among different smart spaces for example between a smart office and a smart home? Can learning be performed over the different modality of sensors? Answers to these are left for future research.

Overall it is felt that the time has come to look into and exploit non-traditional approaches to develop algorithms that are scalable, robust, adaptable, and learn and evolve continuously during their lifetime while also coping up with the dynamism introduced by both human beings and the environment.

Bibliography

[1] M. A. A. Akhoundi and E. Valavi. Multi-sensor fuzzy data fusion using sensors with different characteristics. *arXiv preprint arXiv:1010.6096*, 2010.

[2] W. M. Alobaidi, E. A. Alkuam, H. M. Al-Rizzo, and E. Sandgren. Applications of ultrasonic techniques in oil and gas pipeline industries: a review. *American Journal of Operations Research*, 5(04):274, 2015.

[3] M. Alwan, P. J. Rajendran, S. Kell, D. Mack, S. Dalal, M. Wolfe, and R. Felder. A smart and passive floor-vibration based fall detector for elderly. In *Information and Communication Technologies, 2006. ICTTA'06. 2nd*, volume 1, pages 1003–1007. IEEE, 2006.

[4] A. Avci, S. Bosch, M. Marin-Perianu, R. Marin-Perianu, and P. Havinga. Activity recognition using inertial sensing for healthcare, wellbeing and sports applications: A survey. In *Architecture of computing systems (ARCS), 2010 23rd international conference on*, pages 1–10. VDE, 2010.

[5] I. Aydin, M. Karakose, and E. Akin. A multi-objective artificial immune algorithm for parameter optimization in support vector machine. *Applied Soft Computing*, 11(1):120–129, 2011.

[6] I. Aydin, M. Karakose, and E. Akin. A multi-objective artificial immune algorithm for parameter optimization in support vector machine. *Applied Soft Computing*, 11(1):120–129, 2011.

[7] Y.-W. Bai, C.-C. Cheng, and Z.-L. Xie. Use of a time-variation ultrasonic signal and pir sensors to enhance the sensing reliability of an embedded surveillance system. In *2013 26th IEEE Canadian Conference on Electrical and Computer Engineering (CCECE)*, pages 1–6. IEEE, 2013.

[8] R. D. Baruah and P. Angelov. Evolving local means method for clustering of streaming data. In *Fuzzy Systems (FUZZ-IEEE), 2012 IEEE International Conference on*, pages 1–8. IEEE, 2012.

[9] N. Bellotto and H. Hu. Multisensor-based human detection and tracking for mobile service robots. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 39(1):167–181, 2009.

[10] M. E. Berglund, J. Duvall, and L. E. Dunne. A survey of the historical scope and current trends of wearable technology applications. In *Proceedings of the 2016 ACM International Symposium on Wearable Computers*, pages 40–43. ACM, 2016.

[11] A. Brooks and S. Williams. Tracking people with networks of heterogeneous sensors. In *Proceedings of the Australasian Conference on Robotics and Automation*, pages 1–7. Citeseer, 2003.

[12] S. F. M. Burnet et al. *The clonal selection theory of acquired immunity*, volume 3. Vanderbilt University Press Nashville, 1959.

[13] S. F. M. Burnet et al. *The clonal selection theory of acquired immunity*, volume 3. Vanderbilt University Press Nashville, 1959.

[14] M. Buscema. Back propagation neural networks. *Substance use & misuse*, 33(2):233–270, 1998.

[15] L. M. Candanedo and V. Feldheim. Accurate occupancy detection of an office room from light, temperature, humidity and co2 measurements using statistical learning models. *Energy and Buildings*, 112:28–39, 2016.

[16] J.-L. Chabert, É. Barbin, et al. *A history of algorithms: from the pebble to the microchip*, volume 23. Springer, 1999.

[17] G. Chandrashekar and F. Sahin. A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1):16–28, 2014.

[18] G. Chandrashekar and F. Sahin. A survey on feature selection methods. *Computers & Electrical Engineering*, 40(1):16–28, 2014.

[19] R. Chattopadhyay, N. C. Krishnan, and S. Panchanathan. Topology preserving domain adaptation for addressing subject based variability in semg signal. In *2011 AAAI Spring symposium series*, 2011.

[20] L. Chen, H. Wei, and J. Ferryman. A survey of human motion analysis using depth imagery. *Pattern Recognition Letters*, 34(15):1995–2006, 2013.

[21] Y. Chen and Y. Rui. Real-time speaker tracking using particle filter sensor fusion. *Proceedings of the IEEE*, 92(3):485–494, 2004.

[22] Y. Cheng. Mean shift, mode seeking, and clustering. *IEEE transactions on pattern analysis and machine intelligence*, 17(8):790–799, 1995.

[23] E. H. Chi, P. Pirolli, K. Chen, and J. Pitkow. Using information scent to model user information needs and actions and the web. In *Proceedings of the SIGCHI conference on Human factors in computing systems*, pages 490–497. ACM, 2001.

[24] K. Cho, B. Van Merriënboer, D. Bahdanau, and Y. Bengio. On the properties of neural machine translation: Encoder-decoder approaches. *arXiv preprint arXiv:1409.1259*, 2014.

[25] D. Comaniciu and P. Meer. Mean shift: A robust approach toward feature space analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (5):603–619, 2002.

[26] D. Cook, K. D. Feuz, and N. C. Krishnan. Transfer learning for activity recognition: A survey. *Knowledge and information systems*, 36(3):537–556, 2013.

[27] E. Cornelis, E. Gorus, I. Beyer, I. Bautmans, and P. De Vriendt. Early diagnosis of mild cognitive impairment and mild dementia through basic and instrumental activities of daily living: Development of a new evaluation tool. *PLoS medicine*, 14(3):e1002250, 2017.

[28] L. Da Xu, W. He, and S. Li. Internet of things in industries: A survey. *IEEE Transactions on industrial informatics*, 10(4):2233–2243, 2014.

[29] R. Damarla and D. Ufford. Personnel detection using ground sensors. In *Unattended Ground, Sea, and Air Sensor Technologies and Applications IX*, volume 6562, page 656205. International Society for Optics and Photonics, 2007.

[30] M. Dash and H. Liu. Feature selection for classification. *Intelligent data analysis*, 1(1-4):131–156, 1997.

[31] M. Dash and H. Liu. Feature selection for classification. *Intelligent data analysis*, 1(1-4):131–156, 1997.

[32] L. N. de Castro and F. J. Von Zuben. ainet: an artificial immune network for data analysis. *Data mining: a heuristic approach*, 1:231–259, 2001.

[33] K. G. Derpanis. Mean shift clustering. *Lecture Notes. http://www. cse. yorku. ca/˜ kosta/CompVis_Notes/mean_shift. pdf*, 2005.

[34] T. Diethe, N. Twomey, and P. Flach. Active transfer learning for activity recognition. In *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*, 2016.

[35] N. Eagle and A. S. Pentland. Eigenbehaviors: Identifying structure in routine. *Behavioral Ecology and Sociobiology*, 63(7):1057–1066, 2009.

[36] I. A. Essa. Ubiquitous sensing for smart and aware environments. *IEEE personal communications*, 7(5):47–49, 2000.

[37] A. Fleury, M. Vacher, and N. Noury. Svm-based multimodal classification of activities of daily living in health smart homes: sensors, algorithms, and first experimental results. *IEEE transactions on information technology in biomedicine*, 14(2):274–283, 2010.

[38] S. Flores, J. Geiß, and M. Vossiek. An ultrasonic sensor network for high-quality range-bearing-based indoor positioning. In *2016 IEEE/ION Position, Location and Navigation Symposium (PLANS)*, pages 572–576. IEEE, 2016.

[39] G. Forman. An extensive empirical study of feature selection metrics for text classification. *Journal of machine learning research*, 3(Mar):1289–1305, 2003.

[40] A. Gandomi and M. Haider. Beyond the hype: Big data concepts, methods, and analytics. *International Journal of Information Management*, 35(2):137–144, 2015.

[41] G. Gao, Y. Zhang, Z. Yu, and D. Guanghuang. Data fusion and multi-fault classification based on support vector machines. In *9th Joint International Conference on Information Sciences (JCIS-06)*. Atlantis Press, 2006.

[42] W. Gao, H. Ai, and S. Lao. Adaptive contour features in oriented granular space for human detection and segmentation. In *2009 IEEE Conference on Computer Vision and Pattern Recognition*, pages 1786–1793. IEEE, 2009.

[43] S. Gauthier, B. Reisberg, M. Zaudig, R. C. Petersen, K. Ritchie, K. Broich, S. Belleville, H. Brodaty, D. Bennett, H. Chertkow, et al. Mild cognitive impairment. *The lancet*, 367(9518):1262–1270, 2006.

[44] D. M. Gavrila. The visual analysis of human movement: A survey. *Computer vision and image understanding*, 73(1):82–98, 1999.

[45] F. A. Gers, J. Schmidhuber, and F. Cummins. Learning to forget: Continual prediction with lstm. 1999.

[46] T. Gu, Z. Wu, X. Tao, H. K. Pung, and J. Lu. epsicar: An emerging patterns based approach to sequential, interleaved and concurrent activity recognition. In *2009 IEEE International Conference on Pervasive Computing and Communications*, pages 1–9. IEEE, 2009.

[47] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182, 2003.

[48] I. Guyon and A. Elisseeff. An introduction to variable and feature selection. *Journal of machine learning research*, 3(Mar):1157–1182, 2003.

[49] K. N. Ha, K. C. Lee, and S. Lee. Development of pir sensor based indoor location detection system for smart home. In *2006 SICE-ICASE International Joint Conference*, pages 2162–2167. IEEE, 2006.

[50] H. Hachiya, M. Sugiyama, and N. Ueda. Importance-weighted least-squares probabilistic classifier for covariate shift adaptation with application to human activity recognition. *Neurocomputing*, 80:93–101, 2012.

[51] Q. Hao, F. Hu, and Y. Xiao. Multiple human tracking and identification with wireless distributed pyroelectric sensor systems. *IEEE Systems Journal*, 3(4):428–439, 2009.

[52] M. S. Hashemipour and S. A. Soleimani. Artificial immune system based on adaptive clonal selection for feature selection and parameters optimisation of support vector machines. *Connection Science*, 28(1):47–62, 2016.

[53] K. Hashimoto, N. Yoshiike, and K. Morinaka. Human occupancy detection method and system for implementing the same, Dec. 30 1997. US Patent 5,703,367.

[54] H. M. Hondori, M. Khademi, and C. V. Lopes. Monitoring intake gestures using sensor fusion (microsoft kinect and inertial sensors) for smart home tele-rehab setting. In *2012 1st Annual IEEE Healthcare Innovation Conference*, 2012.

[55] D. H. Hu and Q. Yang. Transfer learning for activity recognition via sensor mapping. In *Twenty-second international joint conference on artificial intelligence*, 2011.

[56] D. H. Hu, V. W. Zheng, and Q. Yang. Cross-domain activity recognition via transfer learning. *Pervasive and Mobile Computing*, 7(3):344–358, 2011.

[57] C.-L. Huang and C.-J. Wang. A ga-based feature selection and parameters optimizationfor support vector machines. *Expert Systems with applications*, 31(2):231–240, 2006.

[58] C. Ionescu, D. Papava, V. Olaru, and C. Sminchisescu. Human3. 6m: Large scale datasets and predictive methods for 3d human sensing in natural environments. *IEEE transactions on pattern analysis and machine intelligence*, 36(7):1325–1339, 2014.

[59] N. K. Jerne. Towards a network theory of the immune system. In *Annales d'immunologie*, volume 125, page 373, 1974.

[60] N. K. Jerne. Towards a network theory of the immune system. In *Annales d'immunologie*, volume 125, page 373, 1974.

[61] N. K. Jerne. Towards a network theory of the immune system. In *Annales d'immunologie*, volume 125, page 373, 1974.

[62] S. S. Jha, K. Shrivastava, and S. B. Nair. On emulating real-world distributed intelligence using mobile agent based localized idiotypic networks. In *Mining Intelligence and Knowledge Exploration*, pages 487–498. Springer, 2013.

[63] X. Jin, S. Gupta, A. Ray, and T. Damarla. Multimodal sensor fusion for personnel detection. In *Information Fusion (FUSION), 2011 Proceedings of the 14th International Conference on*, pages 1–8. IEEE, 2011.

[64] A. Kailas, V. Cecchi, and A. Mukherjee. A survey of communications and networking technologies for energy management in buildings and home automation. *Journal of Computer Networks and Communications*, 2012, 2012.

[65] P. K. Kikiras, J. N. Avaritsiotis, et al. Unattended ground sensor network for force protection. *Journal of Battlefield Technology*, 7(3):29, 2004.

[66] K. Kira and L. A. Rendell. The feature selection problem: Traditional methods and a new algorithm. In *AAAI*, volume 2, pages 129–134, 1992.

[67] K. Kira and L. A. Rendell. The feature selection problem: Traditional methods and a new algorithm. In *AAAI*, volume 2, pages 129–134, 1992.

[68] L. Kleeman and R. Kuc. An optimal sonar array for target localization and classification. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, pages 3130–3135. IEEE, 1994.

[69] R. Kohavi and G. H. John. Wrappers for feature subset selection. *Artificial intelligence*, 97(1-2):273–324, 1997.

[70] K. Kong and M. Tomizuka. A gait monitoring system based on air pressure sensors embedded in a shoe. *IEEE/ASME Transactions on mechatronics*, 14(3):358–370, 2009.

[71] B. Kosucu, N. Hernández, and T. Olugbade. Opportunistic sensing and detection of mild cognitive impairment. In *Proceedings of the 7th International Conference on PErvasive Technologies Related to Assistive Environments*, page 67. ACM, 2014.

[72] N. C. Krishnan, P. Lade, and S. Panchanathan. Activity gesture spotting using a threshold model based on adaptive boosting. In *2010 IEEE International Conference on Multimedia and Expo*, pages 155–160. IEEE, 2010.

[73] R. Kuo, M. Shieh, J. Zhang, and K. Chen. The application of an artificial immune system-based back-propagation neural network with feature selection to an rfid positioning system. *Robotics and Computer-Integrated Manufacturing*, 29(6):431–438, 2013.

[74] M. Kurz, G. Hölzl, A. Ferscha, A. Calatroni, D. Roggen, and G. Tröster. Real-time transfer and evaluation of activity recognition capabilities in an opportunistic system. *machine learning*, 1(7):8, 2011.

[75] M. Längkvist, L. Karlsson, and A. Loutfi. A review of unsupervised feature learning and deep learning for time-series modeling. *Pattern Recognition Letters*, 42:11–24, 2014.

[76] O. D. Lara and M. A. Labrador. A survey on human activity recognition using wearable sensors. *IEEE Communications Surveys and Tutorials*, 15(3):1192–1209, 2013.

[77] J. Li and R. R. Linear. Principal component analysis, 2014.

[78] Q. Liang and J. M. Mendel. An introduction to type-2 tsk fuzzy logic systems. In *FUZZ-IEEE'99. 1999 IEEE International Fuzzy Systems. Conference Proceedings (Cat. No. 99CH36315)*, volume 3, pages 1534–1539. IEEE, 1999.

[79] M. Lim, J. Choi, D. Kim, and S. Park. A smart medication prompting system and context reasoning in home environments. In *2008 Fourth International*

Conference on Networked Computing and Advanced Information Management, volume 1, pages 115–118. IEEE, 2008.

[80] J. Llata, E. G. Sarabia, and J. Oria. Fuzzy expert system with double knowledge base for ultrasonic classification. *Expert Systems with Applications*, 20(4):347–355, 2001.

[81] A. Lotfi, C. Langensiepen, S. M. Mahmoud, and M. J. Akhlaghinia. Smart homes for the elderly dementia sufferers: identification and prediction of abnormal behaviour. *Journal of ambient intelligence and humanized computing*, 3(3):205–218, 2012.

[82] C. Lu and L. Fu. Robust location-aware activity recognition using wireless sensor network in an attentive home. *IEEE Transactions on Automation Science and Engineering*, 6(4):598–609, 2009.

[83] X.-Y. Lu, X.-Q. Chu, M.-H. Chen, P.-C. Chang, and S.-H. Chen. Artificial immune network with feature selection for bank term deposit recommendation. *Journal of Intelligent Information Systems*, 47(2):267–285, 2016.

[84] S. Malviya, D. Mittal, and A. Birle. Agriculture multi sensor data fusion and analysis system. *measurements*, 4(5), 2015.

[85] I. V. Maslov and I. Gertner. Multi-sensor fusion: an evolutionary algorithm approach. *Information Fusion*, 7(3):304–330, 2006.

[86] M. M. Masud, Q. Chen, L. Khan, C. C. Aggarwal, J. Gao, J. Han, A. Srivastava, and N. C. Oza. Classification and adaptive novel class detection of feature-evolving data streams. *IEEE Transactions on Knowledge and Data Engineering*, 25(7):1484–1497, July 2013.

[87] D. A. Matthews, M. R. Lerner, D. G. De Vorchik, S. Sechrest, S. Zou, and B. P. Anderson. Search tool using multiple different search engine types across different data sets, Apr. 26 2016. US Patent 9,323,867.

[88] N. Mattsson, H. Zetterberg, O. Hansson, N. Andreasen, L. Parnetti, M. Jonsson, S.-K. Herukka, W. M. van der Flier, M. A. Blankenstein, M. Ewers, et al.

Csf biomarkers and incipient alzheimer disease in patients with mild cognitive impairment. *Jama*, 302(4):385–393, 2009.

[89] O. Mazhelis. One-class classifiers: a review and analysis of suitability in the context of mobile-masquerader detection. *South African Computer Journal*, 2006(36):29–48, 2006.

[90] Z. Miao, W. Ji, Y. Xu, and J. Yang. Human face classification using ultrasonic sonar imaging. *Japanese Journal of Applied Physics*, 48(7S):07GC11, 2009.

[91] T. B. Moeslund and E. Granum. A survey of computer vision-based human motion capture. *Computer vision and image understanding*, 81(3):231–268, 2001.

[92] M. Moghavvemi and L. C. Seng. Pyroelectric infrared sensor for intruder detection. In *2004 IEEE Region 10 Conference TENCON 2004.*, volume 500, pages 656–659. IEEE, 2004.

[93] V. G. Motti and K. Caine. Users privacy concerns about wearables. In *International Conference on Financial Cryptography and Data Security*, pages 231–244. Springer, 2015.

[94] T. Murakita, T. Ikeda, and H. Ishiguro. Human tracking using floor sensors based on the markov chain monte carlo method. In *Pattern Recognition, 2004. ICPR 2004. Proceedings of the 17th International Conference on*, volume 4, pages 917–920. IEEE, 2004.

[95] R. R. Murphy. Dempster-shafer theory for sensor fusion in autonomous mobile robots. *IEEE Transactions on Robotics and Automation*, 14(2):197–206, 1998.

[96] C. Nadee and K. Chamnongthai. Multi sensor system for automatic fall detection. In *2015 Asia-Pacific Signal and Information Processing Association Annual Summit and Conference (APSIPA)*, pages 930–933. IEEE, 2015.

[97] T. S. Nikhil Sivanandan and S. B. Nair. Immuno-inspired behaviour adaptation in multi-robot systems. *IEEE International Conference on Systems, Man and Cybernetics, SMC-2016*, 2016.

[98] S. J. Pan, J. T. Kwok, Q. Yang, and J. J. Pan. Adaptive localization in a dynamic wifi environment through multi-view learning. In *AAAI*, pages 1108–1113, 2007.

[99] S. J. Pan, I. W. Tsang, J. T. Kwok, and Q. Yang. Domain adaptation via transfer component analysis. *IEEE Transactions on Neural Networks*, 22(2):199–210, 2011.

[100] S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Transactions on knowledge and data engineering*, 22(10):1345–1359, 2010.

[101] J.-T. Park and J.-B. Song. Sensor fusion-based exploration in home environments using information, driving and localization gains. *Applied Soft Computing*, 36:70–86, 2015.

[102] P. Pastor, L. Righetti, M. Kalakrishnan, and S. Schaal. Online movement adaptation based on previous sensor experiences. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 365–371, 2011.

[103] D. N. Perkins, G. Salomon, et al. Transfer of learning. *International encyclopedia of education*, 2:6452–6457, 1992.

[104] R. C. Petersen, R. Doody, A. Kurz, R. C. Mohs, J. C. Morris, P. V. Rabins, K. Ritchie, M. Rossor, L. Thal, and B. Winblad. Current concepts in mild cognitive impairment. *Archives of neurology*, 58(12):1985–1992, 2001.

[105] L. Piwek, D. A. Ellis, S. Andrews, and A. Joinson. The rise of consumer health wearables: promises and barriers. *PLoS Medicine*, 13(2):e1001953, 2016.

[106] D. N. Politis and J. P. Romano. Multivariate density estimation with general flat-top kernels of infinite order. *Journal of Multivariate Analysis*, 68(1):1–25, 1999.

[107] M. Popescu, B. Hotrabhavananda, M. Moore, and M. Skubic. Vampir-an automatic fall detection system using a vertical pir sensor array. In *2012 6th International Conference on Pervasive Computing Technologies for Healthcare (PervasiveHealth) and Workshops*, pages 163–166. IEEE, 2012.

[108] A. Priyono, M. Ridwan, A. J. Alias, R. A. O. Rahmat, A. Hassan, and M. A. M. Ali. Generation of fuzzy rules with subtractive clustering. *Jurnal Teknologi*, 43(1):143–153, 2012.

[109] J. Randi and L. Corno. Theory into practice: A matter of transfer. *Theory into practice*, 46(4):334–342, 2007.

[110] P. Rashidi and D. J. Cook. Transferring learned activities in smart environments. In *Intelligent Environments*, pages 185–192, 2009.

[111] P. Rashidi and D. J. Cook. Multi home transfer learning for resident activity discovery and recognition. *KDD Knowledge Discovery from Sensor Data*, pages 56–63, 2010.

[112] P. Rashidi and D. J. Cook. Activity knowledge transfer in smart environments. *Pervasive and Mobile Computing*, 7(3):331–343, 2011.

[113] R. Rawassizadeh, B. A. Price, and M. Petre. Wearables: Has the age of smartwatches finally arrived? *Communications of the ACM*, 58(1):45–47, 2015.

[114] D. Riboni, C. Bettini, G. Civitarese, Z. H. Janjua, and R. Helaoui. Fine-grained recognition of abnormal behaviors for early detection of mild cognitive impairment. In *2015 IEEE International Conference on Pervasive Computing and Communications (PerCom)*, pages 149–154. IEEE, 2015.

[115] J. K. Rosenblatt and D. Payton. A fine-grained alternative to the subsumption architecture for mobile robot control. In *Proceedings of the IEEE/INNS international joint conference on neural networks*, volume 2, pages 317–324, 1989.

[116] J. Sasiadek and Q. Wang. Sensor fusion based on fuzzy kalman filtering for autonomous robot vehicle. In *Robotics and Automation, 1999. Proceedings. 1999 IEEE International Conference on*, volume 4, pages 2970–2975. IEEE, 1999.

[117] J. Sasiadek and Q. Wang. Fuzzy adaptive kalman filtering for ins/gps data fusion and accurate positioning. *IFAC Proceedings Volumes*, 34(15):410–415, 2001.

[118] J. Z. Sasiadek. Sensor fusion. *Annual Reviews in Control*, 26(2):203–228, 2002.

[119] T. Semwal and S. Nair. Agpi: Agents on raspberry pi. *Electronics*, 5(4):72, 2016.

[120] T. Semwal, P. Yenigalla, G. Mathur, and S. B. Nair. A practitioners' guide to transfer learning for text classification using convolutional neural networks. In *Proceedings of the 2018 SIAM International Conference on Data Mining*, pages 513–521. SIAM, 2018.

[121] S. Seneviratne, Y. Hu, T. Nguyen, G. Lan, S. Khalifa, K. Thilakarathna, M. Hassan, and A. Seneviratne. A survey of wearable devices and challenges. *IEEE Communications Surveys & Tutorials*, 19(4):2573–2620, 2017.

[122] X. Shao, H. Zhao, K. Nakamura, K. Katabira, R. Shibasaki, and Y. Nakagawa. Detection and tracking of multiple pedestrians by using laser range scanners. In *2007 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 2174–2179. IEEE, 2007.

[123] S. Sharma, R. D. Baruah, and S. B. Nair. A reward and penalty based approach for online feature selection. In *2017 3rd IEEE International Conference on Cybernetics (CYBCONF)*, pages 1–6, June 2017.

[124] B. Song, H. Choi, and H. S. Lee. Surveillance tracking system using passive infrared motion sensors in wireless sensor network. In *2008 International Conference on Information Networking*, pages 1–5. IEEE, 2008.

[125] L. Song, A. Smola, A. Gretton, K. M. Borgwardt, and J. Bedo. Supervised feature selection via dependence estimation. In *Proceedings of the 24th international conference on Machine learning*, pages 823–830. ACM, 2007.

[126] Sonia and A. Ghose. Analysis of analog signal from pir-sensors for human analytics in an iotized environment. In *2018 3rd International Conference On*

Internet of Things: Smart Innovation and Usages (IoT-SIU), pages 1–6, Feb 2018.

[127] Sonia and A. Ghose. Unobtrusive and pervasive monitoring of geriatric subjects for early screening of mild cognitive impairment. In *2018 IEEE International Conference on Pervasive Computing and Communications Workshops (PerCom Workshops)*, pages 179–184, March 2018.

[128] Sonia, S. B. Nair, and R. D. Baruah. An immune-inspired online feature selection mechanism. In *2017 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2123–2128, Oct 2017.

[129] Sonia, M. Singh, R. D. Baruah, and S. B. Nair. A voting-based sensor fusion approach for human presence detection. In A. Basu, S. Das, P. Horain, and S. Bhattacharya, editors, *Intelligent Human Computer Interaction*, pages 195–206, Cham, 2017. Springer International Publishing.

[130] Sonia, A. M. Tripathi, R. D. Baruah, and S. B. Nair. Ultrasonic sensor-based human detector using one-class classifiers. In *2015 IEEE International Conference on Evolving and Adaptive Intelligent Systems (EAIS)*, pages 1–6, Dec 2015.

[131] Sonia, A. M. Tripathi, R. D. Baruah, and S. B. Nair. Ultrasonic sensor-based human detector using one-class classifiers. In *Evolving and Adaptive Intelligent Systems (EAIS), 2015 IEEE International Conference on*, pages 1–6, Dec 2015.

[132] Sonia, P. Vij, and R. D. Baruah. An experience of multi-sensors robot for adaptive human sensing. In *IEEE Symposium Series on Computational Intelligence*. IEEE, 2018.

[133] R. D. B. S. B. N. Sonia, Manish Singh. A voting-based sensor fusion approach for human presence detection. 2016.

[134] R. D. S. B. N. Sonia, Manish Singh. A voting-based sensor fusion approach for human presence de- tection. In *Intelligent Human Computer Interaction(IHCI),206*. springer, 2016.

[135] L. Spinello and K. O. Arras. People detection in rgb-d data. In *2011 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 3838–3843. IEEE, 2011.

[136] J. A. Suykens and J. Vandewalle. Least squares support vector machine classifiers. *Neural processing letters*, 9(3):293–300, 1999.

[137] T. Teixeira, G. Dublon, and A. Savvides. A survey of human-sensing: Methods for detecting presence, count, location, track, and identity. *ACM Computing Surveys*, 5(1):59–69, 2010.

[138] J. Timmis, M. Neal, and J. Hunt. An artificial immune system for data analysis. *Biosystems*, 55(1):143–150, 2000.

[139] A. K. Troyer, K. J. Murphy, N. D. Anderson, M. Moscovitch, and F. I. Craik. Changing everyday memory behaviour in amnestic mild cognitive impairment: a randomised controlled trial. *Neuropsychological rehabilitation*, 18(1):65–88, 2008.

[140] P. Urwyler, R. Stucki, L. Rampa, R. Müri, U. P. Mosimann, and T. Nef. Cognitive impairment categorized in community-dwelling older adults with and without dementia using in-home sensors that recognise activities of daily living. *Scientific reports*, 7:42084, 2017.

[141] H. van der Kooij, R. Jacobs, B. Koopman, and F. van der Helm. An adaptive model of sensory integration in a dynamic environment applied to human stance control. *Biological cybernetics*, 84(2):103–115, 2001.

[142] F. Wahl, M. Milenkovic, and O. Amft. A distributed pir-based approach for estimating people count in office environments. In *2012 IEEE 15th International Conference on Computational Science and Engineering*, pages 640–647. IEEE, 2012.

[143] J. Wang, P. Zhao, S. C. H. Hoi, and R. Jin. Online feature selection and its applications. *IEEE Transactions on Knowledge and Data Engineering*, 26(3):698–710, March 2014.

[144] L. Wang, T. Tan, H. Ning, and W. Hu. Silhouette analysis-based gait recognition for human identification. *IEEE transactions on pattern analysis and machine intelligence*, 25(12):1505–1518, 2003.

[145] Y. Watanabe, A. Ishiguro, and Y. Uchikawa. Decentralized behavior arbitration mechanism for autonomous mobile robot using immune network. In *Artificial immune systems and their applications*, pages 187–209. Springer, 1999.

[146] G. Welch, G. Bishop, et al. An introduction to the kalman filter. 1995.

[147] L. Xia and J. Aggarwal. Spatio-temporal depth cuboid similarity feature for activity recognition using depth camera. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 2834–2841, 2013.

[148] L. Xia, C.-C. Chen, and J. K. Aggarwal. Human detection using depth information by kinect. In *CVPR 2011 workshops*, pages 15–22. IEEE, 2011.

[149] M.-T. Yang and S.-Y. Huang. Appearance-based multimodal human tracking and identification for healthcare in the digital home. *Sensors*, 14(8):14253–14277, 2014.

[150] A. Yazar, F. Erden, and A. E. Cetin. Multi-sensor ambient assisted living system for fall detection. In *Proceedings of the IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP14)*, pages 1–3, 2014.

[151] J. Yun and S.-S. Lee. Human movement detection and identification using pyroelectric infrared sensors. *Sensors*, 14(5):8057–8081, 2014.

[152] P. Zappi, E. Farella, and L. Benini. Tracking motion direction and distance with pyroelectric ir sensors. *IEEE Sensors Journal*, 10(9):1486–1494, 2010.

[153] Z. Zhang. Microsoft kinect sensor and its effect. *IEEE multimedia*, 19(2):4–10, 2012.

[154] Z. Zhao, Y. Chen, J. Liu, and M. Liu. Cross-mobile elm based activity recognition. *International Journal of Engineering and Industries*, 1(1):30–38, 2010.

[155] V. W. Zheng, S. J. Pan, Q. Yang, and J. J. Pan. Transferring multi-device localization models using latent multi-task learning. In *AAAI*, volume 8, pages 1427–1432, 2008.

[156] C. Zhu and W. Sheng. Multi-sensor fusion for human daily activity recognition in robot-assisted living. In *Proceedings of the 4th ACM/IEEE international conference on Human robot interaction*, pages 303–304. ACM, 2009.

[157] Z. Zhu, Y.-S. Ong, and M. Dash. Wrapper–filter feature selection algorithm using a memetic framework. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(1):70–76, 2007.

[158] Z. Zhu, Y.-S. Ong, and M. Dash. Wrapper–filter feature selection algorithm using a memetic framework. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 37(1):70–76, 2007.

[159] Y. Zigel, D. Litvak, and I. Gannot. A method for automatic fall detection of elderly people using floor vibrations and soundproof of concept on human mimicking doll falls. *IEEE Transactions on Biomedical Engineering*, 56(12):2858–2867, 2009.

[160] M. Zocchi. Ultrasonic liposculpturing. *Aesthetic plastic surgery*, 16(4):287–298, 1992.