



**Learning Algorithms For Micro-Doppler Radar Based Detection,
Classification and Imaging of Humans in Indoor Environments**

By

Shelly Vishwakarma

Under the supervision of **Shobha Sundar Ram**

Department of Electronics and Communication Engineering

Indraprastha Institute of Information Technology Delhi

New Delhi– 110020

February, 2020

**Learning Algorithms For Micro-Doppler Radar Based Detection,
Classification and Imaging of Humans in Indoor Environments**

By

Shelly Vishwakarma

A Thesis

submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy



Department of Electronics and Communication Engineering

Indraprastha Institute of Information Technology Delhi

New Delhi– 110020

February, 2020

Certificate

This is to certify that the thesis titled "*Learning Algorithms For Micro-Doppler Radar Based Detection, Classification and Imaging of Humans in Indoor Environments*" being submitted by *Shelly Vishwakarma* to the Indraprastha Institute of Information Technology Delhi, for the award of the degree of Doctor of Philosophy, is an original research work carried out by her under my supervision. In my opinion, the thesis has reached the standard fulfilling the requirements of the regulations relating to the degree.

The results contained in this thesis have not been submitted in part or full to any other university or institute for the award of any degree or diploma.

February, 2020

Dr. Shobha Sundar Ram

Assistant Professor,

Deptt. of Electronics Communication Engineering,

Indraprastha Institute of Information Technology Delhi-110020, India.

Declaration

This is certified that the thesis entitled "*Learning Algorithms For Micro-Doppler Radar Based Detection, Classification and Imaging of Humans in Indoor Environments*" being submitted by me to the Indraprastha Institute of Information Technology Delhi, for the award of degree of **Doctor of Philosophy**, is a bonafide work carried out by me. This research work has been carried out under the supervision of **Dr. Shobha Sundar Ram**. The study pertaining to this thesis has not been submitted in part or in full, to any other University or Institution for the award of any other degree.

February, 2020

Shelly Vishwakarma

PhD Student,

Deptt. of Electronics Communication Engineering,

Indraprastha Institute of Information Technology Delhi-110020, India.

Abstract

***Keywords*-Micro-Dopplers, indoor radar, target detection, classification, radar imaging, machine learning, sparse coding, dictionary learning, single-channel source separation, Denoising autoencoders, stochastic FDTD**

Indoor radars have been researched and developed to detect and monitor humans for applications that range from law enforcement, security and surveillance purposes to ubiquitous sensing applications such as smart home or occupancy detection, assisted living, and, bio-medical applications. These radars are typically phase synchronized to obtain Doppler information of human motions. Humans are non-rigid bodies. The movements of the arms and legs of the humans modulate the carrier frequency of the radar signal giving rise to micro-Doppler features in the radar returns. Research in micro-Dopplers has focused primarily on classifying different human activities, distinguishing between armed and unarmed personnel, and anomaly detection (such as fall detection). There are, however, some limitations associated with the current state-of-the-art methods using low complexity continuous wave (CW) micro-Doppler radars. This dissertation investigates signal processing, and machine learning assisted solutions for advancing the current state of the art techniques for mainly three tasks- target detection, target classification, and target imaging with indoor micro-Doppler radar.

Firstly, Doppler sensors are capable of detecting only a single dynamic target. However, indoor environments typically consist of multiple movers - humans, fans, and loudspeakers. When these multiple targets move simultaneously in the propagation channel, their radar backscatter interferes, resulting in distorted micro-Doppler signatures and poor classification accuracies. This limitation may be partially overcome at the cost of increased hardware complexity, but this would offset the inherent advantages of low cost, portable Doppler sensors. Instead, I focused on investigating signal processing solutions to detect multiple simultaneously moving targets. I presented a supervised dictionary learning ap-

proach to represent our micro-Doppler data. Since the resulting representations or dictionaries are customized or fine-tuned to the underlying data - as opposed to data independent transforms such as Fourier or wavelets - I hypothesize that they will have greater success in actually resolving the micro-Doppler signals. Superposed radar returns from multiple targets are resolved into individual components based on their sparse representations.

Secondly, in the current works, the training and testing of micro-Doppler signatures for classification have been carried out in the identical system and environmental conditions. However, these conditions may often be necessarily violated in real-world scenarios. For instance, situations may arise where the propagation channel or the presence of interference sources in the test site will permit only specific frequency bands of radar operation. These bands may differ from those used previously while training. In this dissertation, I examine the data-driven signal processing algorithms that demonstrate versatility in handling diversity in test and training data in real-life scenarios. I use customized dictionaries learned from micro-Doppler radar data gathered at different carrier frequencies to obtain sparse representations which are highly discriminative and characterize the target motion as opposed to the sensor parameters. These features are subsequently used for classifying test data from another distinct carrier frequency. Our experimental results show that the dictionary learning algorithms are capable of extracting meaningful representations of the micro-Dopplers despite the rich frequency diversity in the data.

Thirdly, there has been very limited research effort in imaging targets based on micro-Doppler radar returns due to the considerable variations that may exist in the indoor propagation environment. For instance, if the radar is deployed in through-wall settings, walls being dispersive and in-homogeneous mediums may introduce considerable distortions such as attenuation, delay and multipath to the radar returns, resulting in distorted radar images. In this dissertation, I focus on mitigating wall effects using a machine learning-based solution- denoising autoencoders- that does not require prior information of the wall parameters or room geometry. Instead, the method relies on the availability of a large volume of training radar data gathered in through-wall conditions and the corresponding clean data captured in line-of-sight conditions. I have validated the performance of the denoising solution for both static and dynamic human subjects.

In each of these cases, the signal processing and machine learning algorithms

are trained to handle diversity in human motion characteristics, radar system parameters, and propagation channel conditions. However, the performances of these machine learning algorithms are tied to a large volume of high-quality training data. Therefore, I gathered a highly curated data set of simulated and measured human micro-Dopplers in both line-of-sight conditions and through-wall conditions. I have also presented a computationally efficient method to model radar micro-Dopplers in indoor conditions by integrating the stochastic finite-difference time-domain (sFDTD) technique with the primitive based scattering center model of human radar returns. It captures diversity in the propagation environment using a single simulation.

Dedication

This thesis is dedicated to my husband, and my heavenly mother.

Acknowledgements

I would like to take this opportunity to express my sincere gratitude and appreciation to a number of incredible people in my life. Without their tremendous guidance, mentoring, help, care, and love, I could not have been what I am today.

First and foremost, I would like to thank my advisor, **Dr. Shobha Sundar Ram** for her consistent support, guidance, motivation, and encouragement over these years. She chose me as her Ph.D student when I knew nothing about research. She has been phenomenal in nurturing my research capabilities. I remember how she pushed me at the right times and guided me in all technical matters like improving my communication, and technical writing skills. On countless occasions, she boosted my confidence and have never failed to keep my optimism alive. I cannot thank her enough for the belief she showed in the moments of disappointments. She is probably one of the most energetic and passionate person I have ever seen. I admire many of her qualities, and no words can suffice my feelings of gratefulness towards her. I am fortunate to have her as my advisor.

I would also like to thank my supervisory committee members: Dr. Angshul Majumdar, and Dr. Saket Anand for their valuable advice, support and constructive feedbacks through these years. I have benefited from their knowledge and wisdom. The thesis has also benefited from comments and suggestions made by Dr. Sumit Roy, at University of Washington. I take this opportunity to thank him.

I would like to give a special thanks to my lab mate **Akash Deep Singh**, for working enthusiastically and meticulously with me over the years especially during hardware experiments. Countless number of brainstorming sessions with him have been very helpful and insightful. He has always been a extremely good friend, counsellor and co-advisor to me. He has helped me calm down in

number of occasions. I have had the pleasure to work with my other lab mates and friends: Vahini, Gaurav Duggal, Yoshana and Neeraj Pandey. Conducting research with them have been full of fun and excitement. They have not only been my collaborators, but have helped me broaden my horizons too. I very much appreciate Rahul Gupta and Vijay Sharma for their help and support during data collection that has made an invaluable contribution towards my PhD.

I must also express gratitude to all of my friends: Shikha, Neha Jain, Neelam Di, Ruchi Pandey, Dhananjay Kimothi, Ankita Deo, Divya Sitani, and Aditi Sharma for making my journey memorable. I have never felt time passing faster than when I am talking with them. I owe a special thanks to my best friend **Dr. Kavita Sharma**. I followed her footsteps during my entire Ph.D journey. I remember her calling every now and then for her excellent tips and tricks to tackle difficult situations, manage and organise life. She has always been there by my side during ups and downs of my life. No words can describe her friendship.

The person with the greatest indirect contribution to this work is **my mother**. I lost my father when I was a child. It was her whose sacrifices, strong determination, constant encouragement, support and unconditional love that made me the person who I am today. She made me a good human being, loving and caring. I will forever be grateful and blessed for this. I dedicate this Ph.D thesis to my heavenly mother. This doctorate is as much hers as it is mine. Life could not have been more painful then loosing her earthly presence. But, God has many ways to comfort those who are broken hearted. So, he comforted me by his angel, my soulmate, my dearest husband and my best friend, **Deepak**. Because of his understanding, love, care and support through the toughest moment of my life, I have started to embrace life again after the big loss. No words could ever express my appreciation. I thank God for enlightening my life with your presence. You are the best thing that has ever happened to me. I owe a special thanks to my family, my father-in-law, my mother-in-law and grandma. Their unconditional love, care and support have never let me feel the absence of my biological parents.

Finally, I would like to thank Indraprastha Institute of Information Technology-Delhi for providing excellent infrastructure and research environment. I would also like to thank Vivesvaraya Ph.D scheme for sponsoring my research.

Contents

Abstract	i
Dedication	iv
Acknowledgements	v
List of Figures	xi
List of Tables	xvi
List of Abbreviations	xix
1 Introduction	1
1.1 Multiple Target Detection: Using Low Complexity Narrowband Doppler Radars	5
1.2 Target Classification: Under Diverse Operating Conditions . . .	7
1.3 Target Imaging: In Through-Wall Scenarios	10
1.4 Research Objectives	12
1.5 List of Publications	13
1.5.1 Papers in refereed journals	13
1.5.2 Papers in conferences	14
1.6 Thesis Organisation	15

2	Introduction to Micro-Doppler Effect	18
2.1	Doppler effect	18
2.2	Micro-Doppler Effect	20
2.3	Experimental Data Collection	22
2.3.1	Simulation Data	23
2.3.2	Measurement Data	28
3	Detection of Multiple Movers Based on Single Channel Source Separation	31
3.1	Two Stage Implementation of Dictionary Learning	33
3.1.1	Sparse Coding Based Disaggregation	35
3.1.2	Sparse Representation based Classification	37
3.2	Simulation	38
3.2.1	Simulation Setup	39
3.2.2	Simulation Results and Analysis	39
3.3	Measurements	44
3.3.1	Measurement Data Collection	45
3.3.2	Measurement Results	50
3.4	Summary	58
4	Classification of Human Micro-Dopplers Across Multiple Carrier Frequencies	59
4.1	Theory	61
4.1.1	Synthesis Dictionary Learning (SDL)	61
4.1.2	Deep Dictionary Learning (DDL)	66
4.1.3	Analysis Dictionary Learning (ADL)	68
4.2	Simulation	72

4.2.1	Simulation Set Up	73
4.2.2	Simulation Results	75
4.3	Measurements	79
4.3.1	Measurement Set Up	79
4.3.2	Micro-Doppler Spectrograms Across Multiple Carriers .	81
4.3.3	Measurement Results and Analyses	85
4.4	Summary	98
5	Mitigation of Through-Wall Distortions in Frontal Radar Images	100
5.1	Theory	104
5.1.1	Training Stage	105
5.1.2	Test Stage	108
5.1.3	Computational Complexity	109
5.1.4	Metrics for evaluation	110
5.2	Simulation Method	112
5.2.1	Radar Signal Model	113
5.2.2	Stochastic Model of Through-Wall Propagation	115
5.2.3	Modeling of electromagnetic radar scatter from dynamic humans in through-wall scenarios	117
5.3	Simulation Results	118
5.3.1	Doppler enhanced frontal imaging of dynamic humans in through-wall scenarios	118
5.3.2	Results from denoising cluttered images	121
5.3.3	Impact of radar-target aspect angle	126
5.3.4	Aperiodic human crouch to run motion	129
5.4	Measurement Results	131
5.4.1	Measurement Data Collection	131

5.4.2	Measurement Results and Analyses	134
5.5	Discussion on Results	135
5.5.1	Computational Complexity Evaluation	135
5.5.2	Diversity of training data	137
5.5.3	Hyper-parameter Selection	138
5.6	Summary	141
6	Conclusion and Future Work	142
6.1	Contributions and Impact	142
6.1.1	Public database of radar micro-Doppler data	142
6.1.2	Representation of radar micro-Dopplers using customized dictionaries obtained from data-driven dictionary learning	143
6.1.3	Method to simulate large volumes of radar micro-Dopplers in indoor conditions	144
6.1.4	Clutter mitigation in radar images using denoising au- toencoders	145
6.2	Future Research Avenues	146
6.2.1	Diversity in radar systems, algorithms, and metrics . . .	146
6.2.2	Dictionary learning on higher-order data	146
6.2.3	RF Imaging on passive radar data	147
	References	148

List of Figures

2.1	Simulation setup configured to capture back-scattered signal from a simple rotating target	20
2.2	Simulation Models for Human Motion (a) Simple pendulum model, (b) Analytical models of walking motion derived from bio-mechanical experiments and (c) Computer animation models using Motion Capture (MOCAP) system for complex human motion	24
2.3	Room geometry of the simulation setup for stochastic FDTD. . .	27
2.4	Measurement setup of monostatic radar configured using Field-Fox Vector Network Analyzer and two horn antennas in line-of-sight scenario	29
3.1	Single channel source separation of radar signals from multiple targets.	36
3.2	(a) Point scatterer model of walking human, (b) Point scatterer model of rotating fan about y-axis, Spectrogram using STFT of (c) human walking towards radar, (d) rotating fan.	40
3.3	Geometry of the room.	41
3.4	(a) STFT of aggregate micro-Doppler returns from FH, and TF, (b) STFT of disaggregated micro-Doppler returns of FH and (c) TF respectively.	42

3.5	Experimental setup using FieldFox Vector Network Analyzer and two horn antennas as monostatic radar at 7.5 GHz for detecting one or more targets (human walking towards radar, human walking away from radar, human walking in the tangential direction, and table fan).	45
3.6	Doppler spectrogram of (a) human walking away from the radar, (b) human walking towards radar, (c) human walking tangentially before the radar and (d) rotating table fan generated with monostatic radar at 7.5GHz.	48
3.7	Doppler spectrogram of (a) aggregate micro-Doppler from a human walking towards the radar and a rotating table fan, (b) disaggregated micro-Doppler of human walking towards radar (indicating presence of target) (c) disaggregated micro-Doppler of human walking away from radar (indicating absence of target) (d) disaggregated return of rotating table fan (indicating presence of target).	51
3.8	Doppler spectrogram of (a) aggregate micro-Doppler from two humans, one walking towards the radar and one walking away from radar, (b) disaggregated micro-Doppler of human walking towards radar (indicating presence of target) (c) disaggregated micro-Doppler of human walking away from radar (indicating presence of target) (d) disaggregated return of rotating table fan (indicating absence of target).	52
3.9	Doppler spectrogram of (a) aggregate micro-Doppler from two humans, one walking towards the radar and one walking away from radar and a rotating table fan, (b) disaggregated micro-Doppler of human walking towards radar (indicating presence of target) (c) disaggregated micro-Doppler of human walking away from radar (indicating presence of target) (d) disaggregated return of rotating table fan (indicating presence of target).	53
3.10	Density of $\ Z\ $ when (a) FH is present and BH, TF are absent, (b) BH is present and FH, TF are absent, (c) FH and TF are present and only BH is absent and (d) FH and BH are present and only TF is absent.	54

4.1	Synthesis dictionary learning framework	62
4.2	Synthesis dictionary learning framework for multiclass classification	64
4.3	Deep dictionary learning framework	67
4.4	Analysis dictionary learning framework	69
4.5	Spectrogram using STFT of a simulated (a) human walking towards radar, (b) a human walking away from the radar radar and, (c) a table fan.	74
4.6	Classification accuracy of dictionary learning algorithms under noise.	78
4.7	Monostatic CW radar configured using a vector network analyzer and two linearly polarized horn antennas for five distinct carrier frequencies - 2.4, 3, 4, 4.5 and 5.8GHz	80
4.8	STFT spectrograms of (a, b) two walking humans - one walking towards the radar and the other walking away from the radar, (c, d) a human walking towards the radar, (e, f) a human boxing and (g, h) a rotating table fan. The figures in the left columns are generated from data collected at 2.4GHz and figures in the right columns are generated from data collected at 5.8GHz.	82
4.9	Variation of classification accuracy with (a) dictionary size for SDL, (b) depth in dictionary layers in DDL, (c) dictionary size for ADL and (d) Sparsity parameter (λ) for ADL.	89
4.10	Scatter plot to visualize (a) raw data, (b) handpicked features (PF) and (c) features extracted using DDL	92
5.1	Denoising Autoencoder	105
5.2	(a) During training stage, the autoencoder learns coefficients $\mathbf{W}_1$ and $\mathbf{W}_2$ from clean free space images ($\mathbf{Y}^{tr}$) and corrupted through-wall ($\hat{\mathbf{Y}}^{tr}$) images. (b) During test phase, corrupted through-wall images ($\hat{\mathbf{Y}}^{test}$) are denoised to obtain reconstructed images that resemble free space radar images ($\mathbf{Y}^{test}$).	111
5.3	Radar scattering model of human.	113

5.4	Room geometry in through-wall scenario (a) Dielectric (b) Reinforced wall and (c) Wall with air-gaps	116
5.5	(a)-(b) Magnitude response, (e)-(h) phase response at 7.5 GHz in freespace, through dielectric wall, reinforced dielectric wall, and dielectric wall with air-gaps respectively. (i)-(l) Corresponding simulated Doppler enhanced frontal image of a walking human.	119
5.6	Simulated Doppler enhanced frontal images over one walking stride of human motion, and the corresponding stick figures obtained using ground truth motion capture data.	121
5.7	Convergence curve of the denoising algorithm.	122
5.8	Group correlation across multiple frames	125
5.9	Simulated Doppler enhanced frontal images over one walking stride of human motion for different aspect angles	127
5.10	Simulated Doppler enhanced frontal images over a crouch to run motion of a human and the corresponding stick figures obtained using ground truth motion capture data	129
5.11	Simulated human (a) Group correlation across multiple frames and (b) SSIM variation for simulation results with respect to number of frames.	130
5.12	(a) Measurement setup in free space and measured range enhanced frontal image of a human subject in (b) free space, (c),(d) behind a glass wall, wood wall respectively , (e),(f) Denoised images in through wall scenarios-glass Wall, wood wall respectively using the presented denoising algorithm.	132
5.13	Computational time as a function of the number of nodes in the hidden layer for (left y-axis) training phase (right y-axis) test phase	137
5.14	SSIM variation for simulation results with respect to (a) number of frames, (b) number of nodes in the hidden layer for mapping functions-linear, tanh and sigmoid.	139

5.15 SSIM variation for measurement results with respect to the number of nodes in the hidden layer for mapping functions-linear, tanh and sigmoid for human subjects 140

List of Tables

3.1	Disaggregation algorithm for detection of multiple target classes	37
3.2	Classification accuracy of proposed method before and after disaggregation using five-fold cross validation.	44
3.3	Data Set Description	46
3.4	Classification of radar data across human walking towards radar (FH), human walking away from the radar (BH) and rotating table Fan (TF) Classes. here P is the predicted class and T is the true class	55
3.5	Detection based on disaggregation of data from human walking towards radar (FH), human walking away from radar (BH) and rotating table fan (TF)	57
4.1	Simulation Data Set Description	75
4.2	Description of test and training scenarios across multiple carrier frequencies	76
4.3	Classification results when training and test micro-Doppler data are simulated for different carrier frequencies for SNR= 10 dB .	76
4.4	Description of test and training scenarios across multiple carrier frequencies	80
4.5	Measurement Data Description	83
4.6	Comparison of average classification accuracy across multiple algorithms when training and test micro-Doppler data are gathered at same carrier frequencies	90

4.7	Comparison of average classification accuracy across multiple algorithms when training and test micro-Doppler data are gathered at different carrier frequencies	90
4.8	SDL results when training and test micro-Doppler data are gathered at different carrier frequencies	93
4.9	DDL results when training and test micro-Doppler data are gathered at different carrier frequencies	94
4.10	ADL results when training and test micro-Doppler data are gathered at different carrier frequencies	95
4.11	Comparison of overall classification accuracy and computational complexity	96
4.12	Comparison of theoretical computational cost	97
5.1	Simulated Radar Parameters	115
5.2	Denoising results between clean and corrupted Doppler enhanced frontal images for different through-wall conditions.SSIM:between corrupted and free space image before denoising (BD) and SSIM :between reconstructed and free space image after denoising (AD)	123
5.3	Denoising results between clean and corrupted Doppler enhanced frontal images for different through-wall conditions.NMSE:between corrupted and free space image before denoising (BD) and NMSE :between reconstructed and free space image after denoising (AD)	124
5.4	Denoising results between clean and corrupted images (captured behind reinforced wall) for different aspect angles. SSIM,NMSE:between corrupted and free space image before denoising (BD) and SSIM,NMSE: between reconstructed and free space image after denoising (AD)	128
5.5	Parameters of Real Radar Setup. *- values derived from available information	134

5.6 Denoising results between clean and corrupted measurement images of real humans for different through-wall conditions under varying percentage of training data. NMSE, SSIM :between corrupted and free space image before denoising (BD) and NMSE,SSIM: between reconstructed and free space image after denoising (AD) 136

List of Abbreviations

ADL	Analysis Dictionary Learning
ADMM	Alternating direction method of multipliers
COTS	commercial off the shelf components
CNN	Convolutional neural networks
CW	Continuous wave
DL	Dictionary Learning
DDL	Deep Dictionary Learning
FFT	Fast Fourier transform
FMCW	Frequency modulated continuous wave
FT	Fourier transform
FDTD	Finite difference time domain
ISM	Industrial, scientific and medical
kNN	k-nearest neighbors
LOS	Line-of-sight
MOCAP	Motion capture
PCA	Principal component analysis
ICA	Independent component analysis
RF	Radio frequency
RCS	Radar Cross Section
sFDTD	Stochastic finite difference time domain

SDL Synthesis Dictionary Learning

STFT Short-time Fourier transform

SVM Support vector machines

TF Time-frequency

TWRI Through-the-wall radar imaging

VNA Vector network analyzer

UWB Ultra-wide band

Chapter 1

Introduction

Over the last decade, research and development of radar-based indoor sensing technology have witnessed rapid growth. Radars offer many advantages compared to other sensors such as camera, infrared/thermal detectors, and x-ray sensors. Optical cameras are available at low prices, and automatic target recognition based on camera images have been extensively researched. However, they can not work in non-line-of-sight conditions as they have poor wall penetration capabilities. X-ray signals possess both good wall penetration and high-resolution imaging capabilities. However, the technology is expensive, and the radiation can be potentially harmful to the sensor operators. Infrared sensing results in a limited range. On the other hand, radars at microwave frequencies below X-band make a safe sensing alternative that can operate 24/7 and in non-line-of-sight scenarios. Radars are non-intrusive and typically easy to deploy using low-cost commercial off the shelf components (COTS) [1]. Humans are generally slow-moving bodies whose translational motions introduce a Doppler shift to the radar signals. Their micro-motions give rise to unique and

discriminative signatures in the joint time-frequency space popularly known as micro-Dopplers [2, 3]. The objective of indoor radar is to exploit these micro-Doppler signals to detect and monitor humans and other concealed objects for applications that range from law enforcement, security and surveillance purposes to ubiquitous sensing applications such as smart home or occupancy detection [4, 5], assisted living especially related to fall detection in elderly people [6–8], and, bio-medical applications for non-intrusively monitoring patients [9].

Micro-Dopplers have mostly been gathered by two types of coherent radars - narrowband and broadband. Broadband radars provide excellent downrange resolution. Some examples of wideband radars that have been used for gathering micro-Doppler data are [10, 11]. Alternatively, narrowband CW radars have been developed to detect dynamic targets based solely on their Doppler information. With the advancement in the radar hardware, there has been concurrent growth in radar signal processing techniques for interpreting the signals. Most of the current research have focused on using micro-Dopplers for classifying different human activities, human gesture recognition, anomaly detection, and distinguishing between armed and unarmed personnel. The algorithms have ranged from heuristic techniques [12, 13], to more sophisticated methods based on principal component analysis [14, 15], independent component analysis [16], empirical mode decomposition [17] and Hilbert Huang transform [18]. Heuristic methods involve the extraction of handcrafted features - such as torso Doppler and bandwidth of Doppler returns - from the spectrograms. Subse-

quently, studies have demonstrated that it may be far more efficient, in realistic scenarios, to use automatic feature extraction techniques to derive the most discriminating features from the radar data [19, 20]. In [21], the authors determined the most informative features in the micro-Doppler signatures for classification. Once the features are extracted, algorithms such as support vector machine [12, 17] and Bayesian classifier [22, 23] have been used for classification purposes.

With greater availability of digitally recorded data, almost unlimited memory capacity and ever-increasing processing speeds of computers, there is spurred growth in usage of deep learning assisted solutions in radar signal processing. More recently, the authors in [13, 24–28] used deep convolution neural networks (DCNN) which jointly learned informative features and classification boundaries without using an additional feature extraction algorithm.

Machine learning algorithms require humongous training databases - capturing the diversity of human motion characteristics, sensor parameters, and channel conditions - which are difficult to generate. Researchers have investigated many different methods to overcome the challenge of limited measurement databases. In [29, 30], researchers augmented limited experimental data gathered through measurements with simulated data using generative adversarial networks based on deep neural networks. Researchers also used deep neural network initialization methods for micro-Doppler classification with low training sample support in [24, 31]. They used a transfer learning technique where pre-trained networks from optical imagery (such as Alexnet, VGGNet, GoogleNet)

were trained with a limited radar data set. Finally, researchers in [5,25,26] used various data augmentation technique to generate larger databases synthetically through operations such as flipping, time-shifting, or using generative adversarial networks.

The essential functionality of the indoor radar in all these works has been the classification of humans based on their micro-Doppler signatures. There are, however, some limitations associated with the current methods using low complexity micro-Doppler radars:

1. First, the works have assumed that the propagation channel consists of only a single mover. When there are multiple movers, their returns interfere and are difficult to classify. Hence, the signals from multiple movers must first be disaggregated before the signal is classified.
2. Second, the classification algorithms have been trained and tested using the data captured in identical environmental and system conditions. However, in real-world scenarios, the algorithms must be capable of handling considerable diversity in channel conditions and sensor parameters.
3. There is very little literature on generating radar data from human returns in non-LOS conditions. This is due to the considerable variations that may exist in the propagation environment. For instance, if the radar is deployed in through-wall settings, walls being dispersive and in-homogeneous mediums may introduce considerable distortions such as attenuation, delay and multipath to the radar returns.

4. Finally, there has been a very limited effort in imaging targets in indoor conditions based on their micro-Doppler signals. This is because humans are spatially large dynamic targets with considerable variations in their pose and posture.

In this dissertation, I investigate signal processing, and machine learning assisted solutions for advancing the current state of the art techniques for mainly three tasks- target detection, target classification and target imaging with indoor micro-Doppler radar. The limitations of the existing literature and the presented solutions are briefly described in the following sections.

1.1 Multiple Target Detection: Using Low Complexity Narrowband Doppler Radars

Current studies have extensively focused on developing algorithms for the classification of different types of dynamic targets based on their micro-Dopplers. For example, studies have demonstrated the usefulness of micro-Doppler signatures for classifying different types of wind turbines [32, 33], ground moving vehicles [34–38] and even airborne targets such as drones [39–41]. More recently, micro-Dopplers have been extensively studied for human hand gesture recognition [42, 43]. However, all of these works are based on the assumption that the propagation channel consists of only *a single* dynamic target or target class. This condition is often violated in real-world scenarios. Indoor environments typically consist of multiple moving targets, such as humans, fans, and

speakers, while outdoors consists of multiple vehicles and pedestrians. When multiple targets are present simultaneously in the channel, the micro-Dopplers in the radar backscatter interfere, resulting in incorrect classification. Therefore, before classification is carried out, signals from multiple movers must be resolved or separated. Multiple target detection can be realized at the cost of increased hardware complexity by augmenting Doppler sensing with direction-of-arrival (DOA) processing using multiple receiver channel data to resolve targets in the Doppler-DOA space [44, 45]. Alternatively, in [11], authors used high range resolution capability of a frequency modulated CW radar to resolve multiple targets in the joint range-Doppler space.

I have retained the low cost and complexity of CW radar hardware with a single channel receiver. Instead of hardware augmentation, I have focused on examining the performance of signal processing based machine learning algorithms for *detection* of multiple simultaneously moving targets [46, 47]. Algorithmically, signal disaggregation is a single channel source separation problem rather than a classification problem. Single-channel source separation has been extensively researched in the speech community using data-independent transforms such as Discrete Cosine Transform (DCT), Fourier, and wavelets [48, 49]. There have been several research efforts into representing micro-Dopplers with traditional data-independent dictionaries such as short-time Fourier transform and wavelets [10, 12, 15, 17]. However, there is considerable overlap of radar micro-Dopplers of humans undergoing complex motions in the frequency domain due to which these data-independent representations are not suitable for

source separation. Secondly, when there are multiple targets, the selection of the appropriate tuning parameters for the data independent transforms becomes challenging. In this dissertation, I used a supervised dictionary learning method where a set of basis vectors or atoms are *learned* from training data to represent micro-Doppler from each class of dynamic targets. These dictionaries result in sparse representations of the radar signals when compared to classical data-independent dictionaries. Hence, these data-driven dictionaries can be used to resolve the superposed radar signals obtained from multiple targets into individual components. Subsequently, the presence of these targets is detected based on the strength of their corresponding disaggregated micro-Dopplers.

1.2 Target Classification: Under Diverse Operating Conditions

Since micro-Dopplers arise from the micro-motions of non-rigid targets, there has been significant focus on using the latest machine learning algorithms for *classifying* different human activities [12–21, 23, 50–67]. These include both periodic motions, such as the human gait [12, 50, 52–56], and non-periodic motions such as standing, sitting, arm motions or falling [7, 17, 57]. Most of the existing works utilize the time-frequency representations of the micro-Doppler returns to classify the data. Here, the short time window (also known as dwell time) is a critical parameter, which is usually heuristically selected to represent data from a specific type of dynamic target. Researches in [12, 64] used heuristic methods while [15] employed principle component analysis and independent component analysis [16] to extract features from time-frequency spectrograms

for classification. Empirical mode decomposition achieved classification accuracy up to 90% in [17]. All of these schemes extracted discriminating features from the micro-Doppler spectrograms.

More recently, there is an active thrust towards using deep learning assisted solutions to overcome the limitations of existing canonical approaches. Most of these works use deep convolution neural networks which jointly learn informative features and classification boundaries without using an additional feature extraction algorithm [13, 25–27]. In all of these works, the training and testing of micro-Doppler signatures for classification have been carried out in controlled laboratory conditions with fixed radar parameters (carrier frequency, position, transmitted waveform, etc.) and for fixed motion trajectories of the targets. However, these conditions may often be necessarily violated in real-world scenarios. The algorithms have not been tested in real-world conditions where there may be significant variations from the training conditions. For instance, through wall radars are mostly operated at 2.4GHz ISM band, which might interfere with other wireless networks such as microwave ovens, baby monitors, and the ubiquitous WiFi operating in the same band. These sources may directly interfere at the radar receiver or impinge upon the target and give rise to additional scatterings from the target.

Therefore, a certain degree of reconfigurability is desired of the radar system parameters to allow them to operate under unpredictable conditions arising from the presence of interference. The hardware reconfigurability must be supported by signal processing algorithms that can handle the diversity in the test

and training data. I examined three recent sparse coding-based techniques for classifying human micro-Doppler data [68, 69]. These are the synthesis dictionary learning (SDL) [70], the deep dictionary learning (DDL) [71] and the analysis dictionary learning (ADL) algorithms [72].

1. **SDL:** In SDL, I learn to express the training signals from each class using a linear combination of a few dictionary atoms. These dictionaries are fine-tuned to the underlying signals and are useful signatures for discerning the right target during classification.
2. **DDL:** In emerging learning methodologies, the depth of representation is perceived as a key aspect towards successful classification by capturing more detailed nuances of each class. Therefore, I extended the shallow dictionary learning, in SDL, into multiple levels, also known as deep dictionary learning (DDL). Here, the representation from each layer acts as an input to the subsequent layer. Each succeeding layer, thus, requires fewer features. The classification is carried out using the representation from the last layer and thus requires a much lower computation time than SDL.
3. **ADL:** The ADL is an alternative paradigm to the SDL. Here, a dictionary directly operates on the data to obtain its sparse representation. The key difference is that during the test phase, the sparse features required for classification are directly obtaining without requiring any inverse operation. As a result, the ADL involves the least computational cost and time.

1.3 Target Imaging: In Through-Wall Scenarios

Radar images, unlike camera images, consist of multiple scattering centers of a target convolved with point spread functions whose ambiguity/width is a function of a radar sensor parameter such as bandwidth (in case of range) or aperture size (in case of azimuth or elevation). Radar images of ground moving targets, airborne, space borne and water borne targets have been actively researched for decades by the radar community for automatic target recognition [73–75]. One of the least explored functionalities of micro-Doppler radar for indoor tracking is imaging of human activities. The reason is that humans are large three-dimensional dielectric bodies with tremendous variation in terms of pose and posture. Also, unlike air crafts and shifts, top view (range-azimuth) images of humans do not result in particularly informative images since humans are naturally larger along the length and breadth than width. Instead, frontal images (azimuth-elevation) resembling human perspectives may be more useful for indicating the nature of human activity. High-resolution frontal images require large planar antenna apertures at microwave frequencies. In [76,77], the authors explored using micro-Doppler processing to resolve multiple scattering centers on the human. Then their frontal positions were mapped using two-dimensional array processing. However, the radar images significantly deteriorate due to the presence of undesired clutter signals arising from the multipath in urban environments [45, 78]. Indoor clutter can be broadly categorized into *target independent static clutter*, *target dependent dynamic clutter* and *target independent*

dynamic clutter-

1. **Target independent static clutter** arise from the reflections off the wall - especially the front face in a through-wall scenario - ceiling, floor, and furniture. This type of clutter can cause significant challenges to the detection of still targets (such as still humans, concealed weapons, estimation of building layouts) using wideband radar signals. Considerable research efforts have been devoted to the study of methods to mitigate static clutter in wideband radars [79–81]. Static clutter is usually easy to eliminate when the objective is to detect moving targets. In the case of continuous-wave radars, these returns fall into the DC bin and can be easily removed through a baseband notch filter.
2. **Target independent dynamic clutter** arise from the reflections off other dynamic movers in the environment. In the case of human tracking, loudspeakers and fans (in indoor environments) and other ground moving vehicles and animals (in an outdoor environment) give rise to this type of clutter. I have already discussed the separation of Dopplers from multiple movers in the previous section.
3. **Target dependent dynamic clutter** arise from the interactions between the complex propagation environment and the moving target. This clutter is far more difficult to remove since the clutter is not independent of the target. In Doppler based tracking of moving humans, the performance of the radar detection is limited by the presence of dynamic clutter. Currently, there has

been limited research efforts made to eliminate dynamic *signal dependent* clutter due to the challenges in resolving the target returns from clutter especially when the knowledge of the exact propagation characteristics or analytical framework is unavailable.

This dissertation focuses on mitigating wall effects in through-wall frontal images of humans using denoising autoencoders based on machine learning [82, 83]. The algorithm does not require prior information about the wall parameters or room geometry. Instead, the method relies on the availability of a large volume of training radar images gathered in through-wall conditions and the corresponding clean images captured in line-of-sight conditions. The algorithm *learns* how to denoise or clean the corrupted signals using training data comprising of both corrupted and clean data. I demonstrate that the autoencoder can be used for removing signal-dependent clutter when no information or label of the through-wall scenario is assumed to be known during the test phase. Instead, the autoencoder is trained with a mixture of images gathered in diverse through-wall conditions.

1.4 Research Objectives

The research objectives of this dissertation are listed below.

1. To generate and publically release a curated data set of both simulated and measured micro-Dopplers of multiple humans of different periodic motions, in line-of-sight and through-wall scenarios. Currently, such data are

not easily available in the radar community, and hence this forms an important contribution of this dissertation.

2. This dissertation investigated signal processing solutions to enable the transition of micro-Doppler radar from the laboratory to real-life deployment by addressing the limitations of existing approaches. It examined signal processing algorithms for:

- (a) Single channel source separation of micro-Dopplers to enable detection of multiple targets using low complexity CW radars.
- (b) To classify micro-Doppler signatures when test and training conditions differ.
- (c) To mitigate target-dependent dynamic clutter arising from interactions between the moving human and the propagation channel using a machine learning-based denoising autoencoders- that does not require prior information of the wall parameters or room geometry.

1.5 List of Publications

The papers that contributed to this dissertation are the following.

1.5.1 Papers in refereed journals

1. **S. Vishwakarma** and S. S. Ram, “Mitigation of through-wall distortions of frontal radar images using denoising autoencoders,” *under revision IEEE TGRS*, 2019. The online version is available on arXiv preprint arXiv:1903.09451.

2. **S. Vishwakarma**, G. Duggal, K. V. Mishra, and S. S. Ram, “Doppler-resilient 802.11 ad-based ultra-short range automotive radar,” *under revision IEEE TAES*, 2019. The online version is available on arXiv preprint arXiv:1902.01306.
3. **S. Vishwakarma** and S. S. Ram, “Dictionary learning with low computational complexity for classification of human micro-Dopplers across multiple carrier frequencies,” *IEEE Access*, 6, 29793–29805.
4. **S. Vishwakarma** and S. S. Ram, “Detection of multiple movers based on single channel source separation of their micro-Dopplers,” *IEEE Transactions on Aerospace and Electronic Systems*, 54(1), 159–169.

1.5.2 Papers in conferences

1. **S. Vishwakarma**, N. Pandey, and S. S. Ram, “Clutter mitigation in range enhanced radar images using sparsity based denoising autoencoders,” *accepted in International Radar Conference*, IEEE, 2019.
2. **S. Vishwakarma** and S. S. Ram, “Mitigation of through-wall interference in radar images using denoising autoencoders,” *Radar Conference (RadarConf)*, IEEE, 2018, pp. 1543–1548.
3. A. D. Singh, S. S. Ram, and **S. Vishwakarma**, “Simulation of the radar cross-section of dynamic human motions using virtual reality data and ray tracing,” *RadarConference (RadarConf)*, IEEE, 2018, pp. 1555–1560.
4. **S. Vishwakarma** and S. S. Ram, “Dictionary learning for classification of

indoor micro-doppler signatures across multiple carriers,” in *Radar Conference (RadarConf)*, IEEE, 2017, pp. 0992–0997.

5. A. D. Singh, **S. Vishwakarma**, and S. S. Ram, “Co-channel interference between wifi and through-wall micro-Doppler radar,” in *Radar Conference (RadarConf)*, IEEE, 2017, pp. 1297–1302.
6. **S. Vishwakarma** and S. S. Ram, “Classification of multiple targets based on disaggregation of micro-doppler signatures,” in *Microwave Conference (APMC), 2016 Asia-Pacific IEEE*, 2016, pp. 1–4.

1.6 Thesis Organisation

This dissertation is organized as follows.

Notation: I use the following notation in this chapter. Matrices are written in capital bold letters while vectors and variables are written in normal letters.

- Chapter 2 introduces the basic concept of micro-Doppler phenomenon followed by a discussion of experimental data collection of radar micro-Doppler from humans. I present both the simulation and measurement methodologies to generate radar scatterings off the human subject in free space line-of-sight (LOS) conditions and through-wall conditions.
- In Chapter 3, a method based on the dictionary learning algorithm has been presented to facilitate the resolving of multiple moving targets in the same propagation channel based on single-channel source separation without

hardware augmentation of radar. The approach also facilitates the detection of targets across a wide dynamic range. In other words, it is able to detect weak micro-Dopplers in the presence of returns from a mover of strong radar cross-section.

- In Chapter 4, customized dictionaries are used to represent radar micro-Dopplers under diverse radar operating conditions (different carrier frequencies during training and test). The performance of data-driven dictionaries was benchmarked with data-independent dictionaries such as principal component, heuristic features, DCT coefficients, and cepstral coefficients. The data-driven dictionaries outperformed the data independent dictionaries in terms of their classification accuracy and computational time.
- In Chapter 5, a denoising autoencoder framework has been demonstrated to mitigate clutter and distortion in through-wall frontal images of both static and dynamic humans. The radar images of dynamic humans are simulated using Doppler-enhanced array processing while the images of the static humans are generated from measurement data using range-enhanced array processing. The autoencoder is implemented using an alternating direction method of multipliers approach to ensure convergence and fast training times. The presented approach required neither prior information of the wall characteristics nor any analytic framework to describe the wall propagation effects.

- Finally, Chapter 6 conclude this dissertation along with the suggested future research directions.

Chapter 2

Introduction to Micro-Doppler Effect

This chapter introduces the basic concept of Doppler effect followed by the micro-Doppler phenomenon. In the subsequent sections of the chapter, I outline the methods used in this dissertation for experimental data collection of radar micro-Doppler from humans. I discuss both the simulation and measurement methodologies to generate radar scatterings off the human subject in free space line-of-sight (LOS) and through-wall scenarios.

2.1 Doppler effect

Doppler radar is an electromagnetic system that transmits a CW signal and detects the shift in the carrier frequency of the reflected signal proportional to the relative velocity between the radar and the target. The Doppler effect is extensively used to distinguish dynamic targets from static clutter. Consider a

time-domain radar signal $s_i(t)$,

$$s_i(t) = ae^{j2\pi f_c t}, \quad (2.1)$$

where f_c is the transmitter frequency and a , denotes the signal amplitude which subsumes the transmitted power and the gain of the antenna. The time domain radar returns from a dynamic target, $s_r(t)$, are given by

$$s_r(t) = b(t) \exp \left(j2\pi f_c \left(t - 2\frac{r(t)}{c} \right) \right). \quad (2.2)$$

Here, $r(t)$ and $b(t)$ are the time-varying down-range and received signal strength from the target respectively and c is the speed of light. The radar cross-section of the target, the path loss and the gains of the radar antennas and transmitted power are subsumed in $b(t)$. If the target is moving with a radial velocity v , then

$$s_r(t) = b(t) \exp \left(j2\pi f_c \left(t - 2\frac{r_0 + vt}{c} \right) \right). \quad (2.3)$$

Then the baseband received signal sees a Doppler shifted frequency signal given by

$$f_D = -\frac{2}{\lambda_c} \frac{dr(t)}{dt} = -\frac{2v}{\lambda_c}, \quad (2.4)$$

where $\lambda_c = \frac{c}{f_c}$ is the wavelength of the transmitted radar signal. Now, if the target approaches the radar, the receiver will perceive a higher frequency resulting in positive Doppler shift. Conversely, if the target moves away from the radar, the Dopplers are negative.

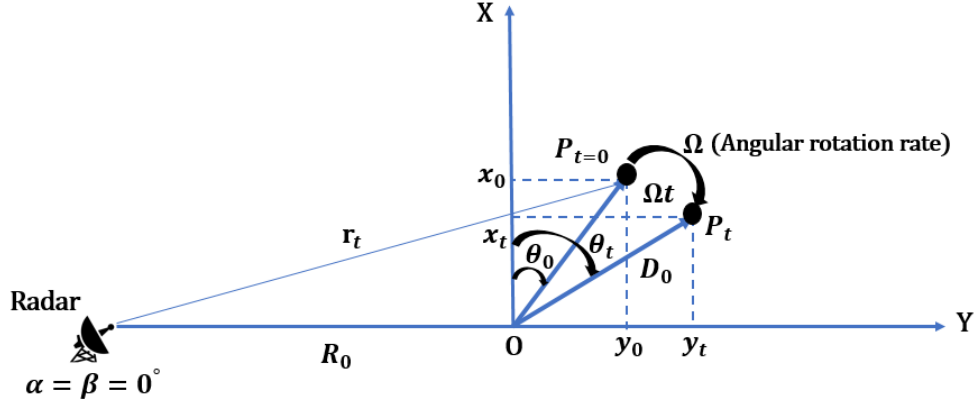


Figure 2.1: Simulation setup configured to capture back-scattered signal from a simple rotating target

2.2 Micro-Doppler Effect

Targets that undergo micro-motions such as rotation and vibration along with the bulk body translational motion, give rise to additional frequency modulations along the main Doppler called micro-Dopplers [2, 3]. I provide an example of the micro-Dopplers generated by a rotating target. Figure 2.1 shows a simple rotating target. Consider a point scatterer P lying at distance D_0 from a center point O and initial rotation angle of θ_0 from the positive X -axis at $t = 0$ seconds. The distance between radar and O is r_0 . The scatterer P rotates about O in the $X - Y$ plane with a rotation rate Ω . The elevation angle is 0° since both radar and rotating point scatterer P are assumed to lie on the same $X - Y$ plane. At t seconds, P rotates, resulting in radial range given by

$$r(t) = r_0 + vt + D_0 \sin \theta_0 \cos \Omega t + D_0 \cos \theta_0 \sin \Omega t. \quad (2.5)$$

Here, v is the radial velocity of scatterer P . The linear Doppler frequency shift of the scatterer P can be derived as

$$f_D = -\frac{2}{\lambda_c} \frac{dr(t)}{dt} = -\frac{2}{\lambda_c} [v + D_0(-\sin\theta_0 \sin\Omega t + \cos\theta_0 \cos\Omega t)]. \quad (2.6)$$

If the target possess only the rotational motion, the micro-Doppler frequency variations due to rotation can be captured by

$$f_{mD} = -\frac{2}{\lambda_c} D_0(-\sin\theta_0 \sin\Omega t + \cos\theta_0 \cos\Omega t). \quad (2.7)$$

Extensive studies have demonstrated that these micro-Doppler features can be exploited using various high-resolution time-frequency (TF) transforms to characterize the time-varying frequency content of a signal. Most of the existing works utilize the STFT based spectrograms of the micro-Doppler returns to represent the data [2].

$$\mathbf{STFT}(t, f) = \int x(\tau) h(t - \tau) e^{-j2\pi f\tau} d\tau \quad (2.8)$$

$$\mathbf{Spectrogram}(t, f) = |\mathbf{STFT}(t, f)|^2 \quad (2.9)$$

Here $h(t)$ is the short time window (also known as dwell time) that captures frequency content of the signal at different time instances. The duration of the window is often heuristically selected based on the nature of the target's motion. A shorter time window results in better time resolution but leads to poor frequency resolution and vice versa. Many alternative time-frequency represen-

tations have been studied to overcome the resolution limit of the STFT. I have, however, used spectrograms to represent micro-Doppler data in the dissertation.

2.3 Experimental Data Collection

Micro-Dopplers have been observed for different types of ground moving vehicles, humans and animals, wind turbines, helicopters, and other airborne targets as well as ships and underwater vehicles. In this thesis, I focus our discussion on human sensing with micro-Doppler radar. I investigate indoor micro-Doppler radar for three main applications- target classification, detection, and imaging of different human activities. The supporting signal processing algorithms are based on recently introduced machine learning assisted solutions. However, the performance of machine learning algorithms relies on the availability of large training databases gathered in a variety of scenarios. Therefore, I was motivated to gather large training databases of human micro-Dopplers in both free-space line-of-sight and indoor through-wall scenarios.

Experimental data can be gathered either by simulating the radar scatterings off human motions or through measurement data collection in laboratory conditions. Simulation data form a means for preliminary evaluation, for pinpointing cause and effect of radar phenomenology and for quickly generating large volumes of training data. Currently, I have excellent means for simulating human walking motions as a function of the height and velocity of the target, which forms the most important parameters that characterize the micro-Doppler spec-

tograms. Measurement data, on the other hand, are used for more detailed studies of the effect of hardware limitations and a more thorough evaluation of the performance of the algorithms in real-world scenarios. Both the simulation and measurement methodologies are detailed in the following two subsections.

2.3.1 Simulation Data

2.3.1.1 Radar Simulation of Humans in LOS

The earliest method of modeling human motion is shown in Figure 2.2(a). Here, the legs of the humans (which give rise to the maximum Dopplers) are modeled as a double pendulum structure [84]. However, this model have not been used in this thesis since it does not capture the variations due to the motion of other body parts such as hands and torso.

Boulic-Thalman model: The second method shown in Figure 2.2(b), is based on the Boulic-Thalman model that was derived from extensive bio-mechanical experiments [85]. The model provides a set of 12 equations that govern the motion trajectories of 17 different body parts as a function of three parameters - the height, the relative velocity of the subject and the starting phase of the leg motion. The advantage of this model is that one can generate a large variety of human walking motions by varying each of the parameters within specified bounds. However, the model has some limitations - First, it is a constant velocity model. Therefore, it does not describe the start of the motion; the acceleration turns or slowing down to halt. Second, it is an averaged model and hence does not show distinctions in human gait due to the subject's gender, age, mood, or

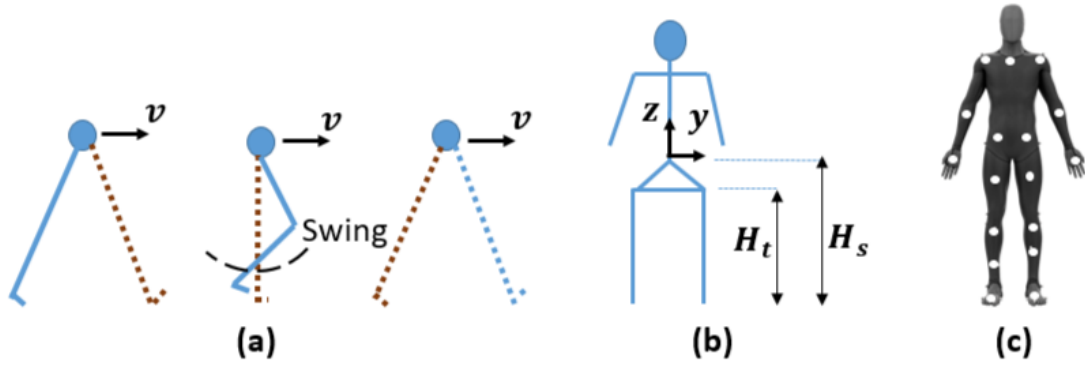


Figure 2.2: Simulation Models for Human Motion (a) Simple pendulum model, (b) Analytical models of walking motion derived from bio-mechanical experiments and (c) Computer animation models using Motion Capture (MOCAP) system for complex human motion

fitness. Also, the model cannot be extended to more complex motions such as jumping, skipping, and crawling.

Computer animation model: The third technique shown in Figure 2.2(c), use computer animation/motion capture data to model different human motions such as crawling, running and skipping [86–88]. In marker based motion capture technology several markers are placed on different parts of the live actor’s body to capture their three-dimensional positions at each time instant. Freely available databases of motion capture data from CMU, University of Pennsylvania and Ohio State are available at [89, 90]. The cheaper marker-less motion capture technology based on Microsoft’s Kinect is prone to high errors especially when the human is not front facing or during the turns [88]. In this thesis, I have used the analytical as well as marker based computer animation motion models of the human. These animation models are then combined with primitive based electromagnetic modeling to generate the radar scatterings off the humans. The primitive based model is computationally fast and simple to execute, especially compared to full-wave electromagnetic solvers. Here, different

parts of the human body such as - torso, arms, and legs are modelled as ellipsoids while the head as a sphere. The RCS of each b^{th} primitive (σ_b) is well characterized at microwave frequencies. If the scattering center of the primitive is assumed to be approximately at the center of the primitive $\vec{r}_b(t)$ at any time t , then the time domain radar returns of the human, $x(t)$, can be estimated by the complex sum of the returns of each of the body parts as shown in

$$x(t) = \sum_{b=1}^B \frac{\sqrt{\sigma_b(t)}}{r_b^2(t)} e^{-j \frac{2\pi f_c}{c} 2r_b(t)}. \quad (2.10)$$

2.3.1.2 Radar Simulation of Humans in Through-Wall Scenarios

All the existing simulation techniques simulate radar returns from humans moving in free space conditions. There is very little literature on simulating human radar returns in non-line-of-sight conditions due to the considerable variations that may exist in the propagation environment. For instance, if the radar is deployed in indoor settings, walls being dispersive and in-homogeneous mediums may introduce considerable distortions such as attenuation, delay and multipath to the radar returns. Additionally, multipath may arise from the ceiling, floor, and furniture. In [45,78,91], a deterministic model of through-wall propagation - derived from finite difference time domain (FDTD) simulations - was combined with human scattering models. However, these results provide the returns for a single environmental scenario. In real-world conditions, there are considerable variations in the electrical parameters (dielectric constant and conductivity) of the environment. Modeling each of these variants through an independent elec-

tromagnetic simulation would require tremendous computational resources. Instead, I present an approach to extend a single simulation framework to provide results for variations in the environment by incorporating stochasticity in the electrical parameters used in the FDTD equations through a stochastic FDTD (sFDTD) framework as described in [92].

The traditional FDTD models Maxwell's equations in the discrete-time domain over a discretized space. Electrical parameters such as the dielectric constant, conductivity, and magnetic permeability are specified for each spatial grid position. Electric and magnetic field equations are updated in each discrete time step in a leapfrog manner. In sFDTD, a normal distribution of the electrical parameters is considered for each spatial position. This results in a Gaussian distribution for the electric and magnetic field values across time and space. The sFDTD provides the mean and variance of the fields for every time step and for every grid point. While the process of updating of the mean electric field and magnetic field values is identical to that of the traditional FDTD, additional equations are incorporated to update the variance of the fields at each time step. The simulations have been carried out entirely in Matlab, through a custom developed software based on Sullivan's code [93].

In this thesis, the stochastic framework enables us to model the diversity in the indoor propagation effects due to the variations in the electrical characteristics. It is demonstrated with an example for through-wall propagation. Consider a two-dimensional simulation space shown in Figure 2.3, extending along the X and Z directions respectively (assuming the problem is invariant along Y height

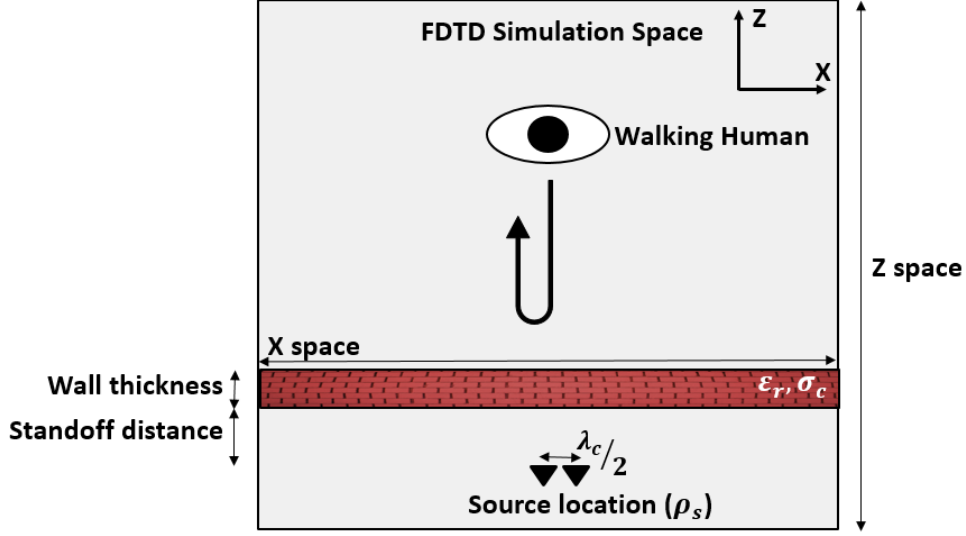


Figure 2.3: Room geometry of the simulation setup for stochastic FDTD.

axis). The space is bounded by a perfectly matched layer and divided into spatial grids of size of $1/10^{th}$ the wavelength λ_c of the excitation source located at $\vec{\rho}_s$. A suitable time step is chosen to satisfy the Courant-Friedrichs-Levy stability condition [94]. A dielectric wall of a certain thickness spans the space. Stochastic variations are introduced in the relative permittivity ϵ_r and conductivity σ_c of each grid point in the wall. Therefore, even a single layer dielectric wall is not truly homogeneous. Thus the model resembles real world conditions. For every point in space, $\vec{\rho}_p$, and at every time instant t , the sFDTD simulation gives the mean time-domain electric field $\mu_E(\vec{\rho}_p, \vec{\rho}_s, t)$ and its standard deviation $\sigma_E(\vec{\rho}_p, \vec{\rho}_s, t)$. One can use the Gaussian stochastic model to generate multiple samples (η) of time-domain electric field values $E(\vec{\rho}_p, \vec{\rho}_s, t, \eta) \sim \mathcal{N}(\mu_E, \sigma_E^2)$. The $E(\vec{\rho}_p, \vec{\rho}_s, t, \eta)$ is fast Fourier transformed to obtain the corresponding η samples of frequency domain wall transfer function between source and field, $H_{wall}(\vec{\rho}_p, \vec{\rho}_s, \eta)$.

The time-domain scattered returns from dynamic humans in a non-LOS through-wall scenarios can be obtained by modifying (2.10) to

$$x(t, \eta) = \sum_{b=1}^B \frac{\sqrt{\sigma_b(t)}}{r_b^2(t)} \times (Scale_{2D \rightarrow 3D} H_{wall}(\vec{r}_b(t), \vec{\rho}_s, \eta))^2. \quad (2.11)$$

Note that $H_{wall}(\vec{\rho}_p, \vec{\rho}_s, \eta)$, is the propagation factor from a two-dimensional source position $\vec{\rho}_s$ to field position $\vec{\rho}_b$ (projection of $\vec{r}_b$ in the two-dimensional space). Therefore, a suitable scaling operation is required to convert the two-dimensional propagation factor to three-dimensional transfer function as discussed in [91].

2.3.2 Measurement Data

Radar micro-Doppler data have been gathered with commercial sensors or using a variety of hardware platforms built using either- commercial off the shelf components (COTS), or software defined radios. The main aim is to enable enhanced developments in the framework of smart homes, healthcare and biomedical scenarios. [95], reported a radar-based technique for fall detection with application to elderly care. While researchers in [96], developed a portable frequency-modulated continuous-wave radar for indoor human tracking. In [97], an embedded multi sensor system was developed for safe point-to-point navigation of impaired users. Other low cost, low power, radars that have been used for measurement data collection are [3, 23, 98] and [53, 55, 56, 99] at Villanova University [100]. More recently, authors in [101], developed a software-defined Doppler radar for monitoring human breathing. Most of these radars

have been deployed in a monostatic configuration. However, micro-Dopplers are dependent on aspect angle of the target. Therefore, additional Doppler diversity has been investigated using bistatic or multistatic configurations [102]. Some of the commercially available radars that offer low-power, SDR development kits with tunable radar parameters are- Ancortek Radars [103], XeThru radar [104], Walabot Pro, a wideband(3.3-10.3GHz) 3D-programmable RF imaging sensor [105]. These SDR kits offer great flexibility for applications in industrial automation, healthcare monitoring, public safety, security and academic research. In this thesis, I have used the experimental set up shown in Figure

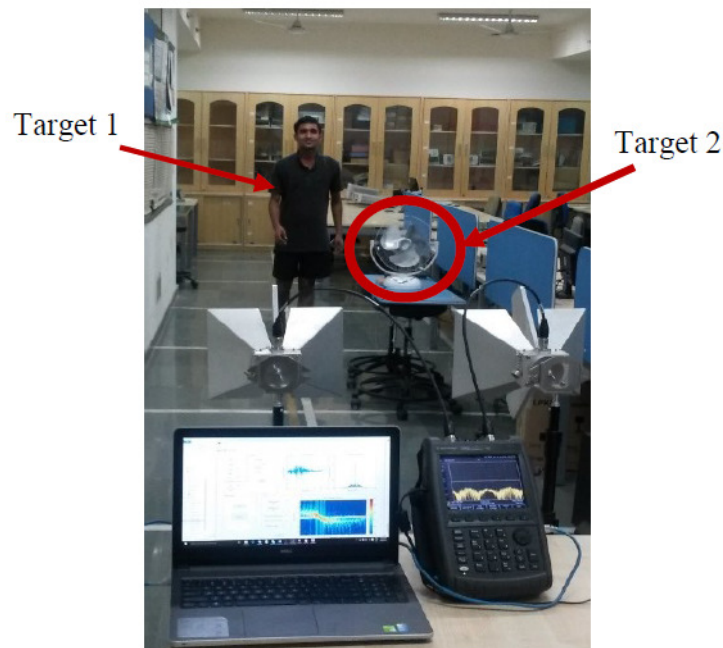


Figure 2.4: Measurement setup of monostatic radar configured using Field-Fox Vector Network Analyzer and two horn antennas in line-of-sight scenario

2.4, to generate measurement micro-Doppler data. The system comprise of two linearly polarized double-ridged waveguide horn antennas (HF907) rated up to 18 GHz, separated by a distance of 50cm, and a N9926A FieldFox vector net-

work analyzer (VNA) rated up to 14GHz. The VNA can be configured either in narrowband mode or broadband mode. Time-domain S_{21} measurements of the VNA capture the back-scattered radar returns from a test subject. The VNA can transmit up to a signal strength of $+3dBm$ and has an excellent dynamic range of $100dB$ with a noise floor at $-120dBm$. The system is portable. However, it has some limitations. The system allows for only one transmitter and one receiver. When the VNA is configured to the narrowband mode, it has a limited sampling frequency which is a function of sweep time. Hence, this system can only be used for detecting and tracking slow-moving targets at low carrier frequencies.

Chapter 3

Detection of Multiple Movers Based on Single Channel Source Separation

Indoor environments typically consist of multiple moving targets, such as humans, fans, loudspeakers, etc. Radars detect multiple targets by resolving them either along the down range (using wide bandwidths), Doppler or direction-of-arrival (using large antenna apertures). In the case of low complexity narrow-band Doppler radars (with single transmitter and receiver), multiple target detection is challenging. This is because when multiple targets are present simultaneously in the channel, their radar backscatter superimpose. As a result, the signatures are dominated by features arising from the strongest targets (based on their radar cross-section and proximity to the radar) with distortions arising from the weaker target returns. The distorted micro-Doppler signatures are then incorrectly classified. This is a serious limitation in most real-world scenarios since they often consist of multiple movers. Therefore all of the current research in micro-Doppler radars have assumed that the propagation channel consists of

only a single target.

In this chapter, a signal processing solution has been presented to detect multiple simultaneously moving targets by separating the aggregate backscatter signal into individual components from each target, thus retaining the low cost and complexity of CW radar. Algorithmically, signal disaggregation is a *single channel source separation* problem rather than a *classification* problem. The algorithm used in this chapter is based on sparse coding methods that were recently developed to separate the energy signals from multiple residential electrical appliances into their component signals [106]. Here, training data are used to learn a unique model or dictionary for the micro-Doppler data from each target class. Since dictionary learning is driven by the radar data, its atoms can be better tuned to match the underlying signal than traditional data-independent dictionaries such as Fourier or wavelets. Further, the dictionary parameters such as the short time window duration, in the case of Fourier, must be set based on the type of target. Hence, when there are multiple targets, the selection of the appropriate parameters for the successful representation of the signal becomes challenging. Dictionary learning presents an alternate method where a set of basis vectors or atoms (dictionaries, $\mathbf{D}$) are used to represent signals, $\mathbf{X}$, as shown in

$$\mathbf{X} = \mathbf{DZ}. \quad (3.1)$$

Here, each column of $\mathbf{X}$ indicates a time-domain measurement of the radar signal. These atoms can be tuned to the underlying signal and hence may yield slightly sparser representations of the signals, $\mathbf{Z}$, than the data independent

dictionaries. In essence, dictionary learning from training data, $\mathbf{X}$, involves minimizing the objective function, $J(\mathbf{D}, \mathbf{Z})$ as shown in

$$J(\mathbf{D}, \mathbf{Z}) = \min_{\mathbf{D}, \mathbf{Z}} \|\mathbf{X} - \mathbf{DZ}\|_F^2 + \lambda \|\mathbf{Z}\|_0. \quad (3.2)$$

Here, $\|\cdot\|_0$ is the l_0 norm that provides a count of the number of non-zero elements in coefficient matrix $\mathbf{Z}$. The optimization is done through a two staged procedure where atoms of dictionary and coefficients are updated through two alternating minimization operations as described below.

3.1 Two Stage Implementation of Dictionary Learning

N number of M dimensional training signals denoted by $\mathbf{X} = [x_1 \dots x_N]$, are considered where $x_n \in \mathbb{R}^{M \times 1}$. Dictionaries are learnt from $\mathbf{X}$ in a two staged iterative procedure. First, the dictionaries are initialized using randomly selected training signals such that $\mathbf{D} \in \mathbb{R}^{M \times P}$, is over complete (where the number of atoms P is greater than the signal dimension M).

1. **Sparse coding stage:** In order to ensure sparse representation of $\mathbf{X}$, the following constraint is imposed on $\mathbf{Z}$ as shown in

$$\mathbf{Z} = \min_{\mathbf{Z}} \|\mathbf{X} - \mathbf{DZ}\|_F^2 \quad \text{s.t.} \quad \|\mathbf{Z}\|_0 < \tau. \quad (3.3)$$

The sparsity level is controlled by parameter τ . Unfortunately l_0 -minimization is NP-hard [107]. Compressed sensing has demonstrated that the l_0 -norm

can be replaced with l_1 -norm, as shown in,

$$\mathbf{Z} = \min_{\mathbf{Z}} \|\mathbf{X} - \mathbf{DZ}\|_F^2 + \lambda \|\mathbf{Z}\|_1, \quad (3.4)$$

and still ensure sparsity. Here $\lambda \in R$ is a regularization parameter that controls the trade-off between the level of sparsity in $\mathbf{Z}$ and the error in data fitting. The formulation in (3.4) is a convex problem which is solved using the iterative soft thresholding algorithm (ISTA) algorithm suggested by [108].

2. **Dictionary update stage:** Once the sparse coefficients are obtained, the dictionary is updated using a least squares approach [109], as shown in

$$\mathbf{D} = \min_{\mathbf{D}} \|\mathbf{X} - \mathbf{DZ}\|_F^2 \quad \text{s.t.} \quad \|d_p\|_2^2 \leq 1, \quad \forall p = 1, 2, \dots, P. \quad (3.5)$$

Columns of dictionary are normalized to have norm less than unity. This two-staged process is iterated until $J(\mathbf{D}, \mathbf{Z})$ converges or reaches a very low tolerance level.

In each iteration of two-stages implementation of dictionary learning, we solve one simple linear least square problem and one iterative sparse-coding problem. Suppose that the training set $\mathbf{X}$ contains n number of m dimensional training signals, and the dictionary $\mathbf{D}$ contains k atoms with M dimensionality. Least square problems theoretically have closed form solutions in the form of pseudo inverse. The complexity of computing a pseudo inverse is $\mathcal{O}(mn^w)$, where $w < 2.37$, in practice it is conjectured to be $w = 2$. Therefore, the theo-

retical complexity per iteration is $\mathcal{O}(mn^2)$. For solving the sparse-coding problem, each iteration requires two matrix products and one thresholding. Suppose sparse coding algorithm is run for a fixed number of iterations (say t). The complexity of the solving l_1 is $\mathcal{O}(t(mk^2 + k^3))$. Therefore, the overall complexity per iteration of the algorithm becomes $\mathcal{O}(mn^2 + t(mk^2 + k^3))$.

If there are I target classes, the corresponding dictionaries, $\mathbf{D}_i$ are learnt for each i^{th} class using this procedure. Once learned, the dictionaries can be used either directly for classification as described in Section 3.1.1 or can be used to disaggregate or separate the superposed radar backscattered signals from multiple moving targets as described in Section 3.1.2. The limitation of the classification algorithm is that the aggregate signal is classified or assigned to just a single class. In contrast, the disaggregation algorithm enables the aggregate signal to be assigned across multiple classes.

3.1.1 Sparse Coding Based Disaggregation

If there are multiple movers in the propagation channel, the received signal at the radar receiver is the aggregate, $\mathbf{X}_{agg}$, from all of these targets as shown in Figure 3.1 and equation (3.6).

$$\mathbf{X}_{agg} = \sum_{i=1}^I \mathbf{X}_i \quad (3.6)$$

Here, $\mathbf{X}_i$ are the radar signals from i^{th} target class. The aim is to disaggregate $\mathbf{X}_{agg}$, into the constituent components $\mathbf{X}'_1 \dots \mathbf{X}'_I$ belonging to I different classes. First, the dictionaries from all the classes are combined together to form a set:

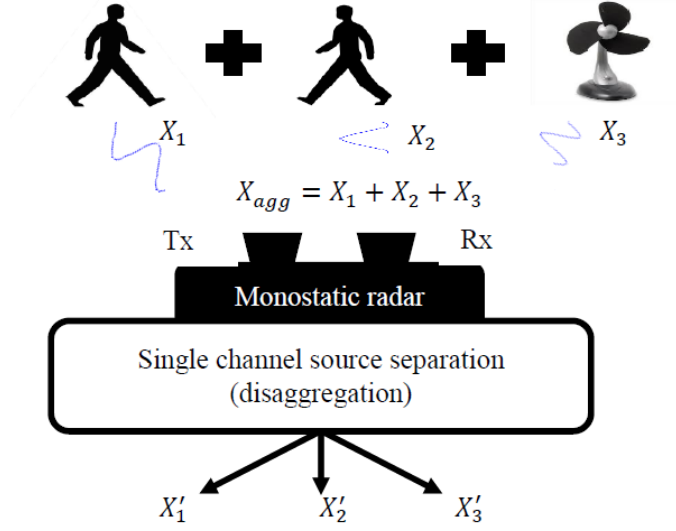


Figure 3.1: Single channel source separation of radar signals from multiple targets.

$\mathbf{D}_{(1:I)} = [\mathbf{D}_1 \dots \mathbf{D}_I]$. It is then solved for the sparsest solution, $\hat{\mathbf{Z}}_{1:I}$ for each class i as shown in

$$\hat{\mathbf{Z}}_{1:I} = \min_{\hat{\mathbf{Z}}_{1:I}} \left\| \mathbf{X}_{agg} - \mathbf{D}_{1:I} \hat{\mathbf{Z}}_{1:I} \right\|_F^2 + \lambda_1 \left\| \hat{\mathbf{Z}}_{1:I} \right\|_1. \quad (3.7)$$

The intuition behind this technique is as follows - If $\mathbf{D}_i$ is trained to reconstruct $\mathbf{X}_i$, then it will be able to reconstruct $\mathbf{X}'_i = \mathbf{D}_i \hat{\mathbf{Z}}_i$, which is the i^{th} portion of the aggregate signal better than any other $\mathbf{D}_j$ for $j \neq i$. The assumption here is that the dictionaries learnt for specific classes are discriminative and there is less coherence between inter-class dictionaries. An optional additional constraint can be imposed on the dictionary learning algorithm to ensure that the discrimination across dictionaries is maximized [110, 111]. In our case, this was found to not be necessary. Most of the active elements of $\hat{\mathbf{Z}}_{1:I}$ should be located in $\hat{\mathbf{Z}}_i$ where i is the class to which the signal belongs. Therefore, once the signal

has been disaggregated, the target i is detected if the strength, $\|\hat{\mathbf{Z}}_i\|_2$, is above a predefined threshold (γ_T). In other words, ideally, each target signal is expected to lie in its own class subspace and all the class sub-spaces are non-overlapping. The detailed algorithm is illustrated in Table 3.1.

Table 3.1: Disaggregation algorithm for detection of multiple target classes

Input: Training data matrix $\mathbf{X}_i^{Tr} \in \mathbf{R}^{M \times N}$, from I individual classes for $i = 1, 2, \dots, I$, $\lambda \in R$ is a regularization parameter.

Stage 1: Learning class specific dictionaries for $i = 1, 2 \dots I$

Loop until convergence

$$\{\mathbf{Z}_i\} = \min_{\mathbf{Z}_i} \|\mathbf{X}_i^{Tr} - \mathbf{D}_i \mathbf{Z}_i\|_F^2 + \lambda \|\mathbf{Z}_i\|_1$$

$$\{\mathbf{D}_i\} = \min_{\mathbf{D}_i} \|\mathbf{X}_i^{Tr} - \mathbf{D}_i \mathbf{Z}_i\|_F^2 \quad \text{s.t.} \quad \|d_{pi}\|_2^2 \leq 1$$

End

Stage 2: Disaggregation of $\mathbf{X}_{agg}$ using learnt dictionaries from I different classes concatenated together, $\mathbf{D}_{1:I} = [\mathbf{D}_1 \dots \mathbf{D}_I]$, $\lambda_1 \in R$ is a regularization parameter.

(a) Sparse Coding

$$\hat{\mathbf{Z}}_{1:I} = \min_{\hat{\mathbf{Z}}_{1:I}} \|\mathbf{X}_{agg} - \mathbf{D}_{1:I} \hat{\mathbf{Z}}_{1:I}\|_F^2 + \lambda_1 \|\hat{\mathbf{Z}}_{1:I}\|_1$$

(b) Detection based on strength of sparse coefficients $\hat{\mathbf{Z}}_i$, γ_T is the defined threshold

$$\hat{\mathbf{Z}}_i = \begin{cases} \hat{\mathbf{Z}}_i & \text{if } \|\hat{\mathbf{Z}}_i\|_2 \geq \gamma_T \\ 0 & \text{if } \|\hat{\mathbf{Z}}_i\|_2 < \gamma_T \end{cases}$$

(c) Reconstruction of disaggregated signal

$$\mathbf{X}'_i = \mathbf{D}_i \hat{\mathbf{Z}}_i$$

3.1.2 Sparse Representation based Classification

To classify the aggregate signal into one of the I classes, the learned dictionaries $\mathbf{D}_i$ can directly be used in the sparse representation-based classifier (SRC) [112].

The test signal will be assigned to the class, $\hat{i}$, having the least error amongst all class representations as shown in

$$\hat{i} = \min_i \left\| \mathbf{X}_{agg} - \mathbf{D}_i \hat{\mathbf{Z}}_i \right\|_2^2 \quad \forall i = 1, 2, \dots, I. \quad (3.8)$$

The classification algorithm is therefore not suited for detecting multiple targets that are simultaneously present in the channel. Note that in prior works the test signal was assumed to consist of radar backscatter from only a single target. Any backscatter from other moving objects in the background was simply treated as noise. In contrast, here, the test signal is considered to be the aggregate backscatter from multiple targets ($\mathbf{X}_{agg}$), and the objective is to detect all the movers.

The performance of sparse coding based dictionary learning (DL) algorithm is demonstrated with both simulation and measurement data.

3.2 Simulation

To demonstrate the application of DL based disaggregation technique on micro-Doppler signatures, I have simulated time domain returns from two classes of moving targets- human walking towards radar (FH) and a rotating table fan (TF) using the popular Boulic model, that is derived from bio-mechanical experimental data [85]. These simple human motion are considered over more complicated motions such as crawling or boxing due to the unavailability of suitable kinematic models. In the past, the radar returns of these motions have been sim-

ulated with motion capture data from existing animation databases [89,90]. But in those cases, the animation data were captured for only a very limited number of test subjects.

3.2.1 Simulation Setup

A Boulic model described in Section 2.3.1.1 is used to derive kinematics of human walking motion. A point scatterer model is used for modelling the targets backscatter. Here, the aggregate received signal is composed of returns from these 17 points representing the human body as shown in Figure 3.2(a). The joint time-frequency signature of the micro-Doppler returns is derived using short time frequency transform (STFT) with a Gaussian window of size 0.02sec as shown in Figure 3.2(c). The human walks along the XY plane and is upright along the Z axis. The rotating fan is modelled using a 3 point scatterer model, where the point scatterers are situated mid-way on 3 rectangular blades as shown in Figure 3.2(b). The scattered returns from the fan are a function of three parameters – the angular velocity of the motion, and the length and width of the blades. The fan is hung along the Z axis and the blades move circularly along the XY plane. The back-scattered signal is the superposition of returns from all 3 point scatterers as shown in Figure 3.2(d).

3.2.2 Simulation Results and Analysis

To validate the performance of the proposed algorithm, the radar returns of two targets- FH and TF, are simulated at a carrier frequency of 2.5 GHz. The dura-

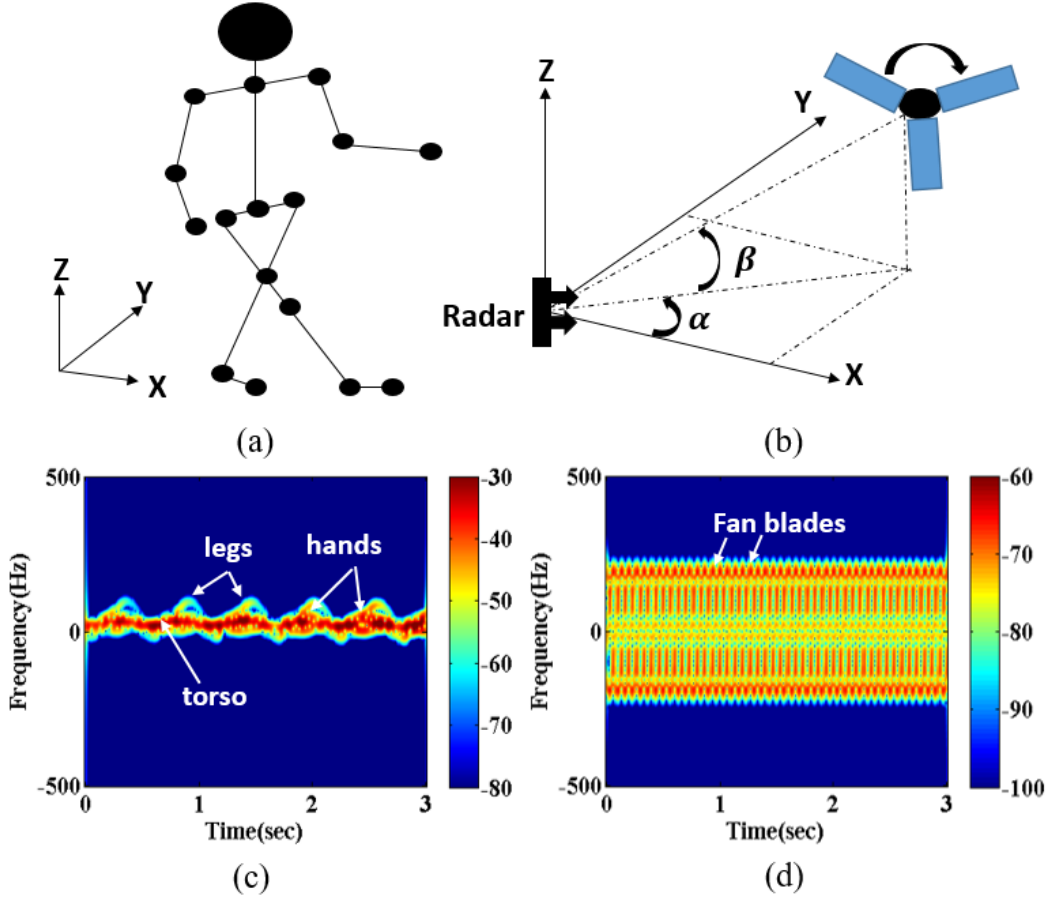


Figure 3.2: (a) Point scatterer model of walking human, (b) Point scatterer model of rotating fan about y-axis, Spectrogram using STFT of (c) human walking towards radar, (d) rotating fan.

tion of the signal is 1 second and a sampling frequency of 1 KHz is chosen to prevent aliasing. Therefore, each signal vector is of size $[1000 \times 1]$. The geometry of the room, the location of the monostatic radar and fan as well as the path of the human considered in the problem are depicted in Figure 3.3. The human parameters (height and relative velocity) and the fan parameters (blade length, width and angular velocity) are varied to simulate data for multiple unique cases. Human heights is varied from 1.6002m to 1.8288m and relative velocity ranging between 1 (H_t/sec) and 1.9 (H_t/sec), where H_t is the height from toe to hip as mentioned in [85]. Similarly, in the case of the fan, the blade lengths are varied between 0.60m to 0.80m, the blade widths between 0.09m to 0.12m and angular

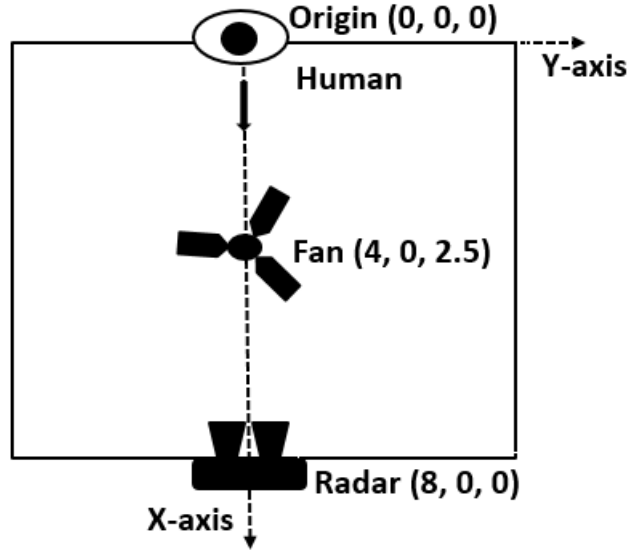


Figure 3.3: Geometry of the room.

velocities between 300 RPM to 500 RPM. Overall, 300 unique human and fan cases were generated out of which I used 80% of the data for training and 20% for the testing purposes with five-fold cross validation. Thus the training matrix for each class was of size $[1000 \times 240]$. From this, the techniques described in stage 1 of Table 3.1 is used to learn human and fan over-complete dictionaries of size $[1000 \times 4000]$ each. I tested our algorithm for disaggregation and classification using test data of size $[1000 \times 60]$, comprising of aggregated returns from both human and fan. The aggregated data is derived by simply superposing the time domain data from both the individual cases.

Then for each test case, the dictionaries of both classes were concatenated. Using the first stage of the algorithm described in Table 3.1, the discrete components belonging to each class under test are reconstructed. The algorithm was allowed to run for maximally 500 iterations in MATLAB on a 2.4GHz Intel processor.

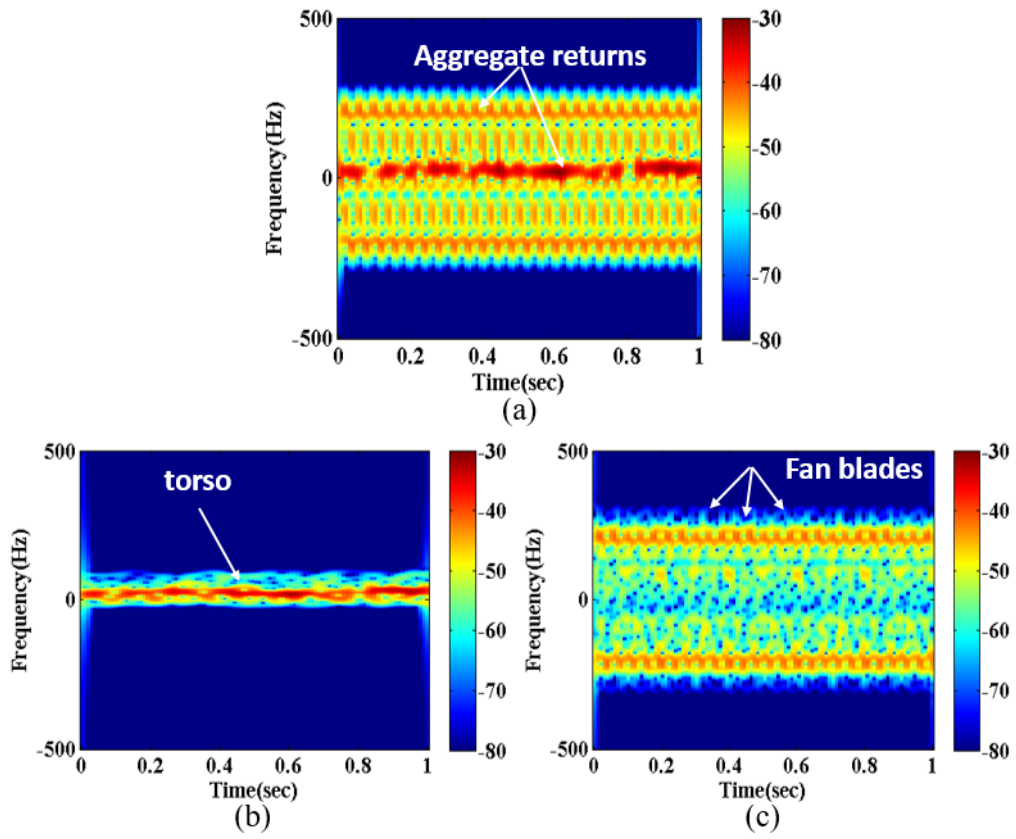


Figure 3.4: (a) STFT of aggregate micro-Doppler returns from FH, and TF, (b) STFT of disaggregated micro-Doppler returns of FH and (c) TF respectively.

Figure 3.4(a) shows the spectrogram corresponding to the aggregate signal from one test case comprising both the FH and TF returns. By comparing this figure to Figure 3.2(c) and (d), it is evident that the spectrogram shows features belonging to both the targets. Hence, it is quite likely that the classification algorithms will not be capable of correctly classifying the aggregated signal. In other words, the classifier is likely to detect the target with the stronger radar cross-section. A thresholding criterion was applied on the sparse coefficients obtained after disaggregation. These coefficients correspond to each class and hence the constituent signals can be reconstructed using these coefficients. Therefore, a target is detected only if the corresponding coefficients are above the threshold. The spectrograms generated from the disaggregated signals corresponding to the human and the fan are shown in Figure 3.4(b) and Figure 3.4(c) respectively. Both of these figures show significant similarity to Figure 3.2(c) and 3.2(d) respectively. There is error in the frequency regions of overlap between the two targets. However, some of the distinctive features (such as the frequency span and the periodicity) are retained in the disaggregated spectrograms. Therefore, the detection algorithms are likely to perform well on the disaggregated signals.

The aggregated test signals are classified before applying the proposed disaggregation algorithm and performance of classifier is evaluated. Next, the disaggregation processing is performed, and presented the separated constituents for target detection and classification. In both cases (before and after disaggregation), the SRC classifier is used.

In Table 3.2, the classification accuracies for three scenarios (each with 60

Table 3.2: Classification accuracy of proposed method before and after disaggregation using five-fold cross validation.

Cases	Classification	FH (%)	TF (%)
Single target-FH	Ground truth	100	0
	BD	100	0
	AD	100	0
Single target- TF	Ground truth	0	100
	BD	0	100
	AD	0	100
Two targets- FH+TF	Ground truth	100	100
	BD	100	0
	AD	100	100

test cases) have been presented. In the first scenario, only a single target, FH, is assumed to be present. In the second scenario, only TF is assumed to be present and in the third scenario, both FH and TF are assumed to be present. The ground truths reflect these three scenarios in the table. Then the classification results for before disaggregation (BD) and after disaggregation (AD) are presented. The results show that in the case of the single target, SRC classifier correctly identifies the targets using their corresponding dictionaries. However, in the case where two targets are present simultaneously, the algorithm detects FH but is not capable of identifying TF possibly since the RCS of the humans are greater than the fan. However, after disaggregation, both FH and TF are correctly identified.

3.3 Measurements

Next, the performance of the proposed algorithm is examined in real world scenarios using radar measurement data. Compared to the simulation setup, more complex target motion scenarios are considered.

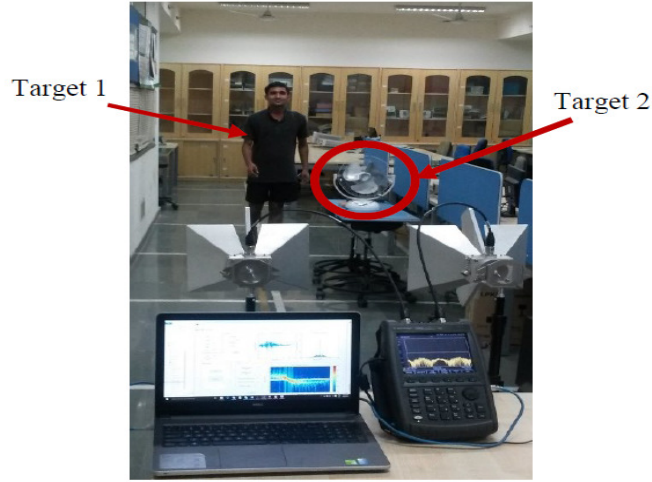


Figure 3.5: Experimental setup using FieldFox Vector Network Analyzer and two horn antennas as monostatic radar at 7.5 GHz for detecting one or more targets (human walking towards radar, human walking away from radar, human walking in the tangential direction, and table fan).

1. Human walking towards the radar (FH),
2. Human walking away from the radar (BH),
3. Human walking in tangential direction before the radar (SH)
4. Table fan (TF).

3.3.1 Measurement Data Collection

A monostatic CW radar shown in Figure 3.5 is used to collect measurement data from these four target classes. A CW radar configuration operating at 7.5 GHz gather measurement data from 40 human subjects (with varying gaits, heights and velocities) and a table fan (with different angular velocities, distances and orientations with respect to the radar). The human subjects moved roughly between 1m and 9m in front of the radar. The duration of each measurement is 2.7s with 1000 samples. These data are gathered in single, two, three and four

targets scenarios. The complete data set and its detailed description is available in [113]. The complete data set and its detailed description is presented in Table 3.3.

Table 3.3: Data Set Description

Target Scenario and Description	Target Parameters	Training Data	Test Data
Single target- FH, BH, SH	Number of subjects: 40 Target Heights: 5' to 6' Target Velocities: 0.6m/s to 1.5m/s Number of measurements: 200 (5 measurements from each of the 40 subjects)	150 (5 measurements from each of the 30 subjects)	50 (5 measurements from each of the 10 subjects)
Single target- TF	Number of fans: 1 Angular velocities: 1400 rpm, 2000 rpm and 2600 rpm Locations from the radar : 8 Number of measurements: 240 (10 measurements from each of the 24 cases)	150 (10 measurements from each of the 15 cases)	50 (10 measurements from each of the 5 cases)
Two targets- FH+TF, BH+TF, SH+TF	Number of humans: 5 Number of fans: 1 Locations of fan: 2 Number of measurements: 50 (5 measurements from each of the 5 x 2 cases)	0 (No training data)	50 (All measurement data are used for testing)
Two targets- FH+BH	Number of humans walking towards radar: 2 Number of humans walking away from radar: 2 Number of measurements: 40 (10 measurements from each of the 4 cases)	0 (No training data)	40 (All measurement data are used for testing)
Three targets- FH+BH+TF, FH+SH+TF, BH+SH+TF	Number of humans: 2 Number of fans: 1 Number of measurements: 20 (10 measurements from each of the 2 cases)	0 (No training data)	20 (All measurement data are used for testing)
Four targets- FH+BH+SH+TF	Number of humans: 3 Number of fans: 1 Number of measurements: 20 (10 measurements from each of the 2 cases)	0 (No training data)	20 (All measurement data are used for testing)

The joint time-frequency spectrograms generated from the baseband digi-

tized time-domain radar data by the application of short-time Fourier transform is shown in

$$\mathbf{STFT}(t, f) = \int x(\tau)h(t - \tau)e^{-j2\pi f\tau}d\tau \quad (3.9)$$

$$\mathbf{Spectrogram}(t, f) = |\mathbf{STFT}(t, f)|^2 \quad (3.10)$$

Here $h(t)$ is the short time window of duration 0.054 s.

Figure 3.6 shows the Doppler spectrograms generated from measured data gathered from a single target scenario using short time-frequency transform. When the human is walking away from the radar (BH), the Dopplers, as shown in Figure 3.6a, are mostly negative except for the backswing from the arms and legs. The strength of the Dopplers decreases with time as the distance from the radar increases. On the other hand, the Dopplers are mostly positive, in 3.6b, when the human is walking towards the radar (FH). The Doppler spread is directly a function of the velocity and height of the human since the feet give rise to the maximum absolute values of Dopplers. When the human is walking tangentially before the radar (SH), the Dopplers are less pronounced as shown in Figure 3.6c though their strengths are similar to the FH and BH cases. This is due to the low Doppler shift that arises from the target's cross line of sight motion with respect to the radar. The periodicity in the Dopplers corresponds to the gait of the human. Figure 3.6d shows the spectrogram due to the three rotating blades of the table fan (TF). The Dopplers here are a function of the blade length, the angular velocity, and the orientation of the fan. Due to low

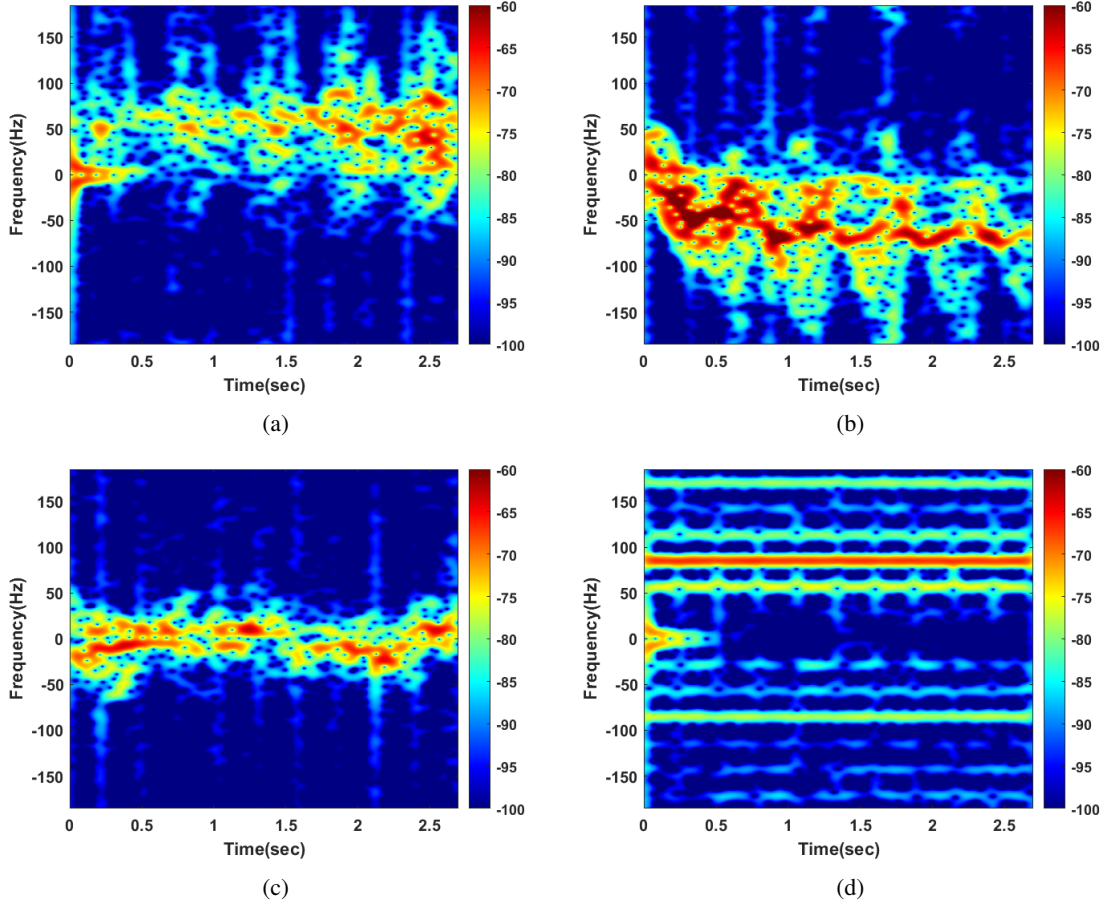


Figure 3.6: Doppler spectrogram of (a) human walking away from the radar, (b) human walking towards radar, (c) human walking tangentially before the radar and (d) rotating table fan generated with monostatic radar at 7.5GHz.

sampling frequency limits imposed by the VNA, aliasing can be observed in the spectrogram.

Training data:

75% of the measurement data gathered in the single target scenario for the four target classes ($\mathbf{X}_{FH}^{TR}$, $\mathbf{X}_{BH}^{TR}$, $\mathbf{X}_{SH}^{TR}$ and $\mathbf{X}_{TF}^{TR}$) are used for training purposes. As a result, the training signal matrix is of size $[1000 \times 150]$. During stage 1 of Table 3.1, class-specific dictionaries ($\mathbf{D}_{FH}$, $\mathbf{D}_{BH}$, $\mathbf{D}_{SH}$ and $\mathbf{D}_{TF}$) are learnt from the training data corresponding to the four targets. Dictionaries of size $[1000 \times 500]$ for all the target types, for four different values of $\lambda = 0.0001, 0.001, 0.01, 0.1$

are examined. Finally, the dictionaries with the least error of signal representation are adopted.

Test data:

Our training and test scenarios are not identical. They differ in the following key ways: first, the test data are gathered from a different set of subjects than the training data; second, the data, $\mathbf{X}_{agg}$, are gathered in two, three and four target scenarios (as opposed to the single target scenario for training data). In both the training and test cases, the humans and fans are placed in different locations and orientations with respect to the radar. Then $\mathbf{X}_{agg}$ and the concatenated dictionaries are used as input to the sparse coding based disaggregation algorithm described in stage.2 of Table 3.1 to obtain constituent signal components from different movers. A target i is detected if the strength of disaggregated coefficient, $||\hat{\mathbf{Z}}_i||_2$, is greater than a pre-defined threshold value for human targets and for table fan which is determined empirically from the noise floor of the measurements and the average radar cross-section of the targets. In the absence of prior works on disaggregation of micro-Doppler data, the performance of the algorithm is compared with the classification algorithm described in (3.8). Each test case is assigned to the class i that gives rise to the minimum residue between the test data $\mathbf{X}_{agg}$, and $\mathbf{D}_i, \mathbf{Z}_i$. Thus the classification algorithm can be used to identify the presence of only a single target in the channel. Strictly speaking, the disaggregation and classification algorithms have different objectives, and only the former is suited for the simultaneous detection of multiple targets.

3.3.2 Measurement Results

In this section, the performance of the disaggregation and classification algorithms are examined.

3.3.2.1 Doppler Spectrograms of Disaggregated Signals

The dictionary learning based algorithm is a supervised learning technique where the target classes must be known prior to detection. Multi-target class scenarios are considered where the objective is to detect the presence of one or more of the targets on the basis of disaggregation of backscatter signals. First, the Doppler spectrograms generated from the disaggregated components are compared qualitatively with the previously shown spectrograms in Figure 3.6.

In the first case, a two target scenario is considered where a human walking towards the radar (FH) and a table fan (TF) are present in the channel. The radar cross-section of the human is greater than that of the fan. The spectrogram in Figure 3.7a shows radar backscatter from both the targets with corresponding micro-Doppler features. The micro-Dopplers overlap at certain frequencies. If these aggregated data are directly used for classification, it is quite likely that the target will be assigned to class FH due to the strength of its returns. Figure 3.7.b d show the spectrograms generated from the disaggregated components and the advantages of the disaggregation algorithm. One can clearly observe targets belonging to the classes FH (Figure 3.7b) and TF (Figure 3.7d) while there is no target belonging to the class BH (Figure 4c) or SH (not shown). Note that the

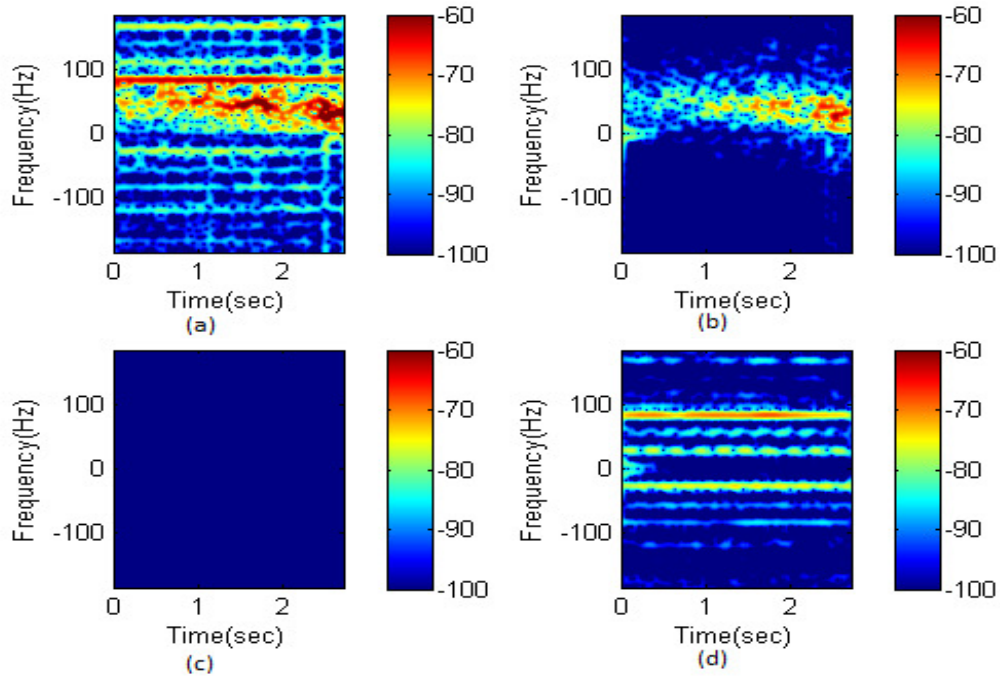


Figure 3.7: Doppler spectrogram of (a) aggregate micro-Doppler from a human walking towards the radar and a rotating table fan, (b) disaggregated micro-Doppler of human walking towards radar (indicating presence of target) (c) disaggregated micro-Doppler of human walking away from radar (indicating absence of target) (d) disaggregated return of rotating table fan (indicating presence of target).

strength of the disaggregated components (Figure 3.7a and Figure 3.7c) slightly differ from what is observed in the aggregate signal (Figure 3.7a). This may be attributed to the interference between the signals belonging to the different classes in the aggregate signal. Another two target scenario is when two humans are walking one towards and the other away from the radar (FH and BH). This is a more challenging scenario since the radar cross-sections of the two humans are comparable. The spectrogram shown in Figure 3.8a, clearly indicates both positive and negative micro-Dopplers emanating from the two targets. There is still some overlap between the micro-Dopplers from the back swing of the arms and legs. Due to the similarity in the strength of the returns from both the targets, the classification algorithm, based on these aggregate data, is quite likely

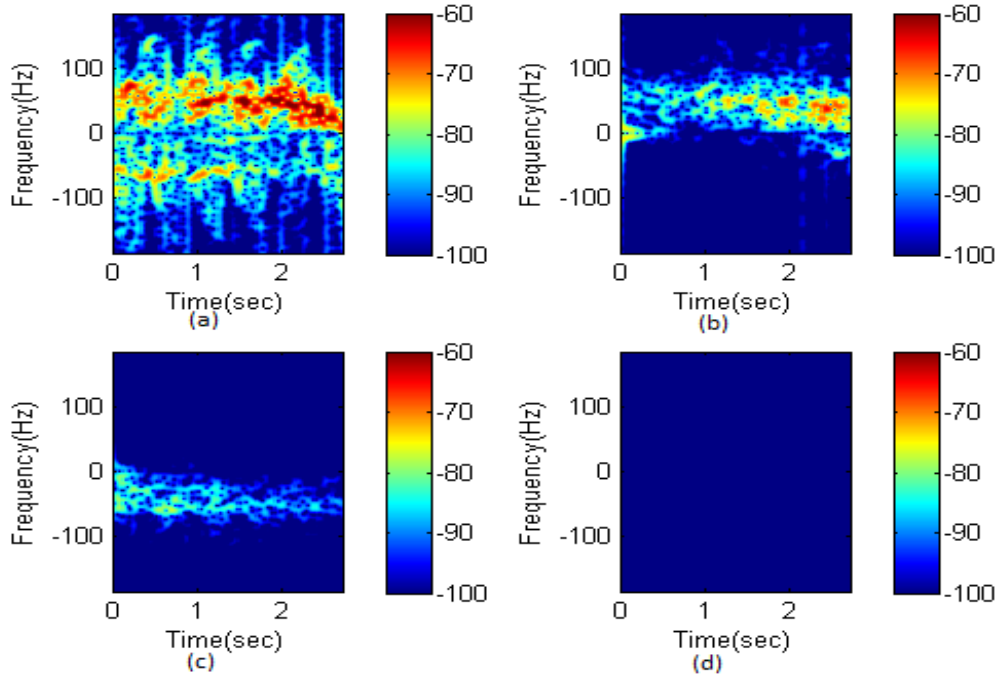


Figure 3.8: Doppler spectrogram of (a) aggregate micro-Doppler from two humans, one walking towards the radar and one walking away from radar, (b) disaggregated micro-Doppler of human walking towards radar (indicating presence of target) (c) disaggregated micro-Doppler of human walking away from radar (indicating presence of target) (d) disaggregated return of rotating table fan (indicating absence of target).

to be confused between classes FH and BH. The spectrograms from the disaggregated components, on the other hand, clearly indicate the presence of FH (Figure 3.8b), BH (Figure 3.8c) and the absence of TF (Figure 3.8d). The disaggregation algorithm, however, does seem to not correctly pick out the highest Dopplers arising from the motion of the feet. This may be due to the variation in the training data corresponding to humans of different heights. Finally, a three target scenario is considered where radar returns from three moving targets (FH, BH, and TF) are present. The spectrogram of aggregate signals, shown in Figure 3.9a, indicates that the aggregate data may be confused mostly between FH and BH since the returns from TF are weaker. In contrast, the spectrograms of the disaggregated components shown in Figure 3.9b, Figure 3.9c and Figure

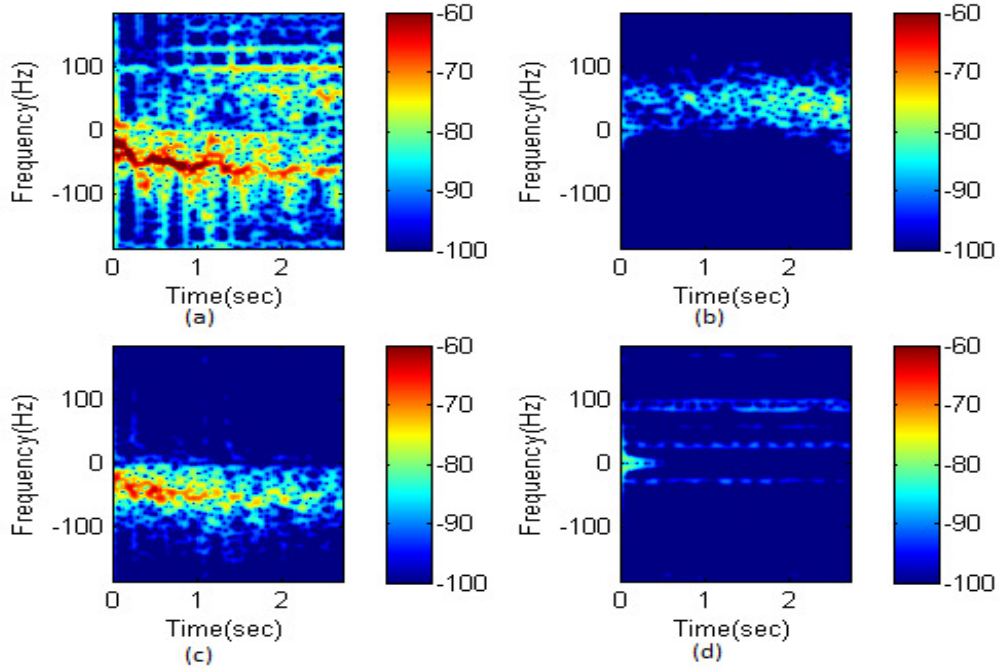


Figure 3.9: Doppler spectrogram of (a) aggregate micro-Doppler from two humans, one walking towards the radar and one walking away from radar and a rotating table fan, (b) disaggregated micro-Doppler of human walking towards radar (indicating presence of target) (c) disaggregated micro-Doppler of human walking away from radar (indicating presence of target) (d) disaggregated return of rotating table fan (indicating presence of target).

3.9d indicate that all three targets will be correctly detected.

The hypothesis is that most of the active elements of $\hat{\mathbf{Z}}_{1:I}$ should be located in $\hat{\mathbf{Z}}_i$ where i is the class to which the signal belonged and hence the vector is dense. Strength of sparse coefficients $Z = [Z_1 Z_2 Z_3]$ for some of the cases are shown in Figure 3.10. When the target is present, corresponding coefficients are dense with higher strength which are clearly visible in the spectrograms.

3.3.2.2 Classification results for SRC

The measurement data belonging to the single, two and three target scenarios are classified into one of three classes (FH, BH or TF) using the SRC algorithm discussed in Section 3.1.2. The results are presented in Table 3.4. In the single

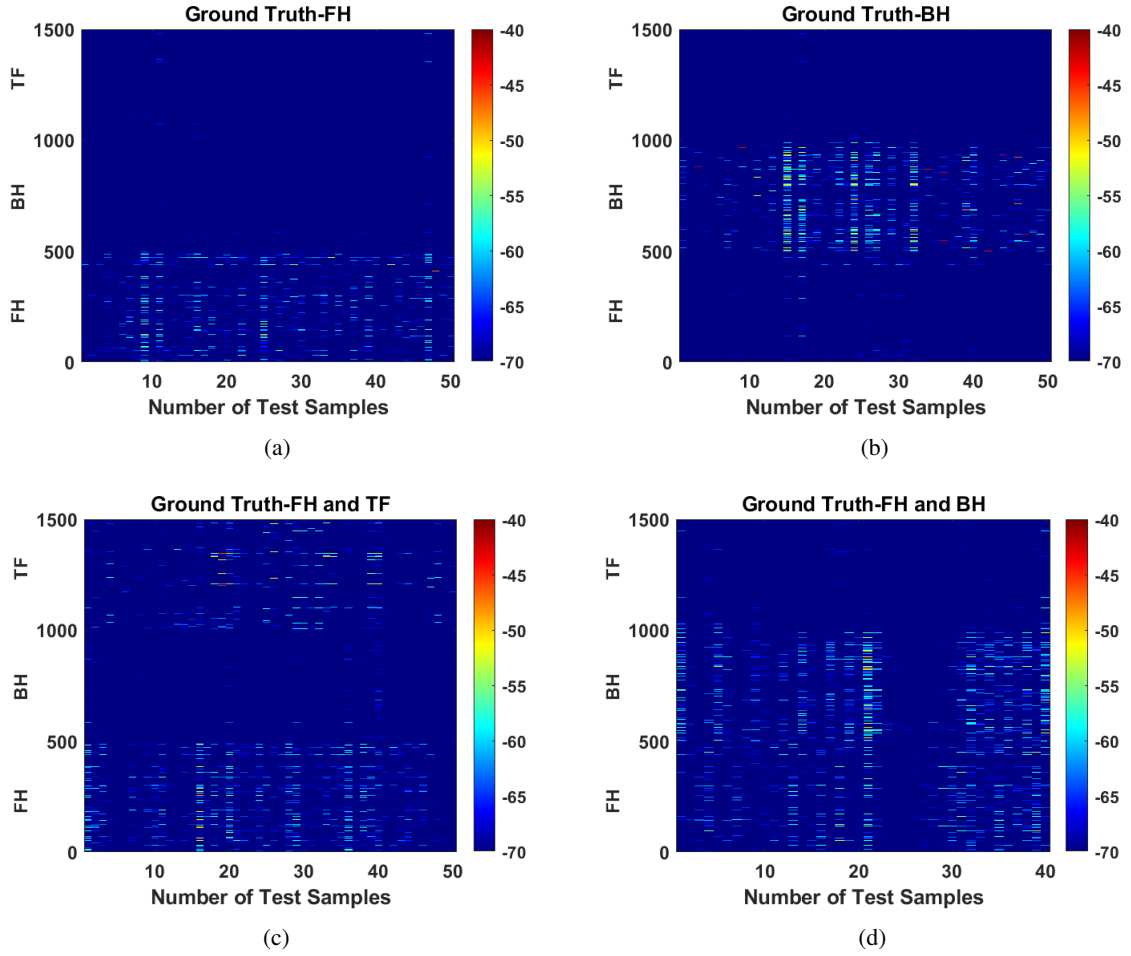


Figure 3.10: Density of $\|Z\|$ when (a) FH is present and BH, TF are absent, (b) BH is present and FH, TF are absent, (c) FH and TF are present and only BH is absent and (d) FH and BH are present and only TF is absent.

target scenario case (FH / BH / TF), the classifier most often picks the correct class for the target. The average accuracy for this scenario is 85%. This confirms the hypothesis that micro-Doppler signatures are useful tools for classifying moving targets. For comparison purposes, LC-KSVD algorithm reported in [114] is applied to the single target data and obtained 75% classification accuracy. The classifier is mostly confused between the FH and BH classes. Next, the performance of the algorithm is examined when multiple targets are present in the channel. The results for the two targets scenarios FH + TF and BH + TF, cases show that the classification performance is skewed towards the targets

with the stronger radar cross-section (humans). In other words, the returns from the weaker targets (TF) are treated as noise by the algorithm. On the other hand, when both the targets are humans (FH + BH), and therefore, of comparable radar cross-sections, the classifier pretty evenly distributes the cases between them. In the case of the three target scenario (FH + BH + TF), the classification accuracy is again skewed towards the stronger targets (humans) in favor of the weaker TF. These results show that the classifier can accurately detect the strongest target when all other moving objects in the background give rise to much weaker returns. The performance of the classifier deteriorates when there are multiple targets with comparable backscatter.

Table 3.4: Classification of radar data across human walking towards radar (FH), human walking away from the radar (BH) and rotating table Fan (TF) Classes. here P is the predicted class and T is the true class

Cases(T/P)	FH (%)	BH (%)	TF (%)
Single target-FH	92	8	0
Single target-BH	6	94	0
Single target-TF	20	10	70
Two targets-FH+TF	82	10	8
Two targets-BH+TF	6	84	10
Two targets-FH+BH	52.5	47.5	0
Three targets-FH+BH+TF	40	60	0

3.3.2.3 Detection results after disaggregation of micro-Doppler data

The disaggregation algorithm, unlike the classification algorithm, is meant for the detection of multiple targets. This time, the measurement data from four target class labels (FH, BH, SH and TF) are considered. The resultant detection accuracies for the single, two, three and four targets scenarios are presented in Table 3.5. An average true detection accuracy of 98.5% is achieved for the four

single target scenarios (FH / BH/ SH / TF) while the false alarm rate is 4.4%. When the two target scenarios is examined - the algorithm is able to detect both the targets in 88% of the FH + TF cases, 94% in BH + TF cases, 80% in FH + BH and 98% in SH + TF cases. This result demonstrates the usefulness of the disaggregation algorithm when compared to the classification algorithms for the detection of multiple targets. The algorithm was also able to detect the weak target (TF) in the presence of strong targets (FH, BH, and SH). It also detected two targets of comparable returns. However, the limitation is that the presented method has a high false alarm rate. This is because when two targets are present, there is a high probability of the overlap of the signals in the frequency domain.

The performance of the disaggregation algorithm is further investigated in the three target scenario. When three targets move simultaneously, it is almost impossible not to have micro-Doppler overlap. Despite this, the algorithm detects all three targets in 96% of the cases. Obviously, the humans are favored (more than 95% detection accuracy) since they have stronger radar cross-sections. Finally, our approach is tested in four target scenario where FH, BH, SH and TF moved simultaneously in the same channel. Here, detection accuracies for SH and TF are 100% where as for FH it is 87.5%. For the case of BH the accuracy dropped to 67.5%, this can be attributed to the weaker return signals in some of the BH cases in comparison to SH and FH and shadowing of BH by SH and FH.

Table 3.5: Detection based on disaggregation of data from human walking towards radar (FH), human walking away from radar (BH) and rotating table fan (TF)

Scenario(T/P)	True Detections (%)	Missed Detections (%)	False Alarms (%)
Single target-FH	FH:100 BH:NA TF:NA	FH:0 BH:NA TF:NA	FH:NA BH:4 TF:14
Single target-BH	FH:NA BH:100 TF:NA	FH:NA BH:0 TF:NA	FH:12 BH:NA TF:8
Single target-SH	FH:NA BH:NA TF:NA SH:100	FH:NA BH: NA TF:NA SH:0	FH:0 BH:0 TF:0 SH:NA
Single target-TF	FH:NA BH:NA TF:94	FH:NA BH: NA TF:6	FH:2 BH:0 TF:NA
Two targets-FH+TF	FH:88 BH:NA TF:94	FH:12 BH: NA TF:6	FH:NA BH:12 TF:NA
Two targets-BH+TF	FH:NA BH:94 TF:96	FH:NA BH:6 TF:4	FH:28 BH:NA TF:NA
Two targets-FH+BH	FH:87.5 BH:80 TF:NA	FH:12.5 BH:20 TF:NA	FH:NA BH:NA TF:52.5
Two targets-SH+TF	FH:NA BH:NA TF:98 SH:100	FH:NA BH: NA TF:2 SH:0	FH:22 BH:0 TF:NA SH:NA
Three targets-FH+BH+TF	FH:95 BH:90 TF:90	FH:5 BH:10 TF:10	FH:NA BH:NA TF:NA
Three targets-SH+FH+TF	FH:95 BH:NA TF:100 SH:100	FH:5 BH: NA TF:0 SH:0	FH:NA BH:0 TF:NA SH:NA
Three targets-SH+BH+TF	FH:NA BH:95 TF:100 SH:100	FH:NA BH: 5 TF:0 SH:0	FH:25 BH:NA TF:NA SH:NA
Four targets-FH+BH+TF+SH	FH:87.5 BH:67.5 TF:100 SH:100	FH:12.5 BH: 32.5 TF:0 SH:0	FH:NA BH:NA TF:NA SH:NA
Average	94	6	11

3.4 Summary

A supervised dictionary learning technique is used to represent micro-Doppler from dynamic targets. These dictionaries result in sparser representations of the radar signals than the classical data-independent dictionaries such as wavelets or Fourier. These data-driven dictionaries can be used to disaggregate the superposed radar signals obtained from multiple targets into individual components. This enables the detection of weak targets that would otherwise be missed in the presence of stronger returns.

The algorithm's performance has been evaluated using both the simulation and measurement data. Our results show a overall detection accuracy of 100% and 94% for simulated and measurement data respectively. However, the false alarm rate is higher in the case of measurement results which is 11%. Note that the false alarm values are not a function of the threshold selected in the algorithm and are instead due to the limitations in the dictionary learning algorithm. The discrimination capability of the dictionaries is governed by the degree of overlap in the micro-Dopplers. In all of the above cases, it is assumed that each target class (FH, BH, SH, or TF) consists of only a single target. The algorithm can be applied to disaggregate data where each class may comprise of several targets. However, there will be a need to incorporate additional complexity (hardware or software) to determine the actual number of targets within each class.

Chapter 4

Classification of Human Micro-Dopplers Across Multiple Carrier Frequencies

Over the last decade, classification of human motions based on their micro-Dopplers has attracted significant research focus for applications as diverse as law enforcement, indoor tracking, fall detection and assisted living [12, 13, 15, 17, 20, 51, 52, 55, 56, 58, 62]. However, in all of these cases, the radar system conditions and the propagation channel under training and test scenarios have been nearly identical. To demonstrate the actual use of radar micro-Dopplers in real-world scenarios, one needs to consider two factors - One, situations where the test conditions vary significantly from the training conditions. For instance, in indoor tracking, the presence of Wi-Fi or other wireless devices at specific bands (say 2.4 or 5.8GHz) may considerably interfere with the radar [115]. Alternatively, the walls being dispersive mediums may support certain frequencies over others. Therefore, a degree of reconfigurability or versatility is desired in the radar hardware parameters - such as the carrier frequency - which can be

achieved using software-defined radar platforms. Simultaneously, the classification algorithms need to use non-heuristic methods to derive features to handle the diversity in the training and test data. Second, the classification must occur real-time.

In the previous chapter, the performance of dictionary learning was studied for single-channel source separation of radar micro-Dopplers [46, 47]. In this chapter, a low computationally complex dictionary learning framework is presented for classifying different human motions using diverse multi-frequency training and test data. Here, three recent dictionary learning algorithms are investigated for classifying human micro-Doppler data gathered across multiple carrier frequencies. These are the synthesis dictionary learning (SDL) [70, 116, 117], the deep dictionary learning (DDL) [71] and the analysis dictionary learning (ADL) algorithms [72, 118–120]. In SDL, I learn to express the training signals from each class using a linear combination of a few dictionary atoms. These dictionaries are fine-tuned to the underlying signals and are useful signatures for discerning the right target during classification. The SDL uses a single layer sparse representation for each target class. In emerging learning methodologies, the depth of representation is perceived as a key aspect towards successful classification. Therefore, the shallow dictionary learning, in SDL, is extended into multiple levels also known as deep dictionary learning. Here, the representation from each layer acts as an input to the subsequent layer. Each succeeding layer, thus, requires fewer features. The classification is carried out using the representation from the last layer and thus requires a much lower com-

putation time than SDL. The ADL is an alternative paradigm to the SDL. Here, a dictionary directly operates on the data to obtain its sparse representation. The key difference is that during the test phase, the sparse features required for classification are directly obtaining without requiring any inverse operation. As a result, the ADL involves the least computational cost and time. All three algorithms, trained with data from a set of carrier frequencies, are used to classify test data from another distinct carrier. The performance of the presented data features is benchmarked with others such as LC-KSVD [114], PCA [14], physical features (PF) [12], DCT coefficients [121], and cepstral coefficients [122] in terms of their classification accuracy and computational time.

4.1 Theory

In this section, the synthesis, deep and analysis dictionary learning algorithms are presented for real-time classification of micro-Doppler data across multiple carriers.

4.1.1 Synthesis Dictionary Learning (SDL)

The objective of the algorithm is to first learn to represent the time-domain micro-Doppler signal, y_c , of c^{th} dynamic target with only few basis vectors, $\mathbf{D}_c$, and a sparse coefficient vector, z_c

$$y_c = \mathbf{D}_c z_c. \quad (4.1)$$

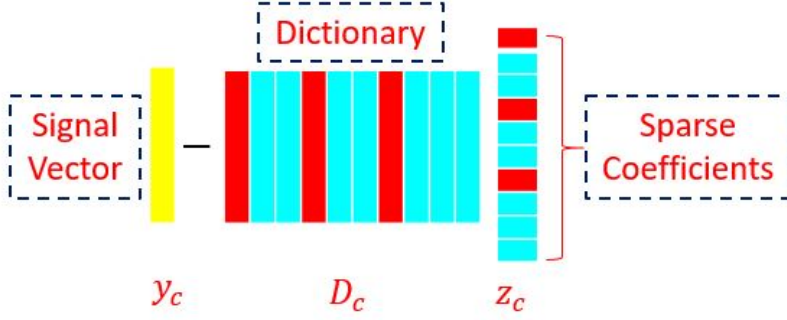


Figure 4.1: Synthesis dictionary learning framework

Subsequently, $\mathbf{D}_c$ will be used for classification purposes - both at the training and at the test stages. Figure 4.1 provides a pictorial representation of the equation. The red columns of dictionary matrix $\mathbf{D}_c$, are linearly combined with coefficient values, colored red in $\mathbf{z}_c$, to represent a signal vector y_c . As indicated in the figure, $\mathbf{z}_c$ is sparse. The learning problem in (4.1) is non-convex and solved using an alternating minimization technique. Consider a training data matrix $\mathbf{Y}_c \in \mathbb{R}^{N \times M}$. These are micro-Doppler measurements captured at multiple carrier frequencies for c^{th} target class. Each of the M columns represents an independent N -dimensional time-domain radar signal measurement. Learning the corresponding dictionary, $\mathbf{D}_c \in \mathbb{R}^{N \times K}$, and sparse code, $\mathbf{Z}_c \in \mathbb{R}^{K \times M}$, fundamentally involves minimizing the objective function, $J(\mathbf{D}_c, \mathbf{Z}_c)$, shown in

$$J(\mathbf{D}_c, \mathbf{Z}_c) = \min_{\mathbf{D}_c, \mathbf{Z}_c} \|\mathbf{Y}_c - \mathbf{D}_c \mathbf{Z}_c\|_F^2 \quad \text{s.t.} \quad \|\mathbf{Z}_c\|_0 \leq \tau. \quad (4.2)$$

Here, K defines the number of atoms in the dictionary $\mathbf{D}_c$ and τ is the sparsity level in $\mathbf{Z}_c$ which is controlled by the l_0 -norm to ensure sparse representation of $\mathbf{Y}_c$. l_0 -minimization is NP-hard [107]. $\mathbf{Z}_c$ can be updated using greedy matching pursuit algorithms such as orthogonal matching pursuit (OMP) [123, 124].

Alternatively, (5.3) can be relaxed to a higher order convex l_1 -minimization problem shown in

$$J(\mathbf{D}_c, \mathbf{Z}_c) = \min_{\mathbf{D}_c, \mathbf{Z}_c} \|\mathbf{Y}_c - \mathbf{D}_c \mathbf{Z}_c\|_F^2 + \lambda \|\mathbf{Z}_c\|_1. \quad (4.3)$$

$\lambda \in \Re$ is the regularization parameter that balances the trade-off between the data representation accuracy and sparsity. (4.3) is solved using a two-stage iterative procedure. First, the dictionary matrix is initialized using randomly selected signal vectors from the training data, $\mathbf{Y}_c$. Given $\mathbf{D}_c$, the coefficient matrix $\mathbf{Z}_c$ is updated using (4.4) which is known as Least Angle Shrinkage and Selection operator (LASSO) [125].

$$\mathbf{Z}_c = \min_{\mathbf{Z}_c} \|\mathbf{Y}_c - \mathbf{D}_c \mathbf{Z}_c\|_F^2 + \lambda \|\mathbf{Z}_c\|_1 \quad (4.4)$$

The literature has a plethora of techniques for solving a l_1 -minimization problem. In this dissertation, (4.4) is solved using the iterative soft-thresholding algorithm (ISTA) discussed in [126]. Once $\mathbf{Z}_c$ is obtained, estimating $\mathbf{D}_c$ reduces to a least squares problem [109] which is solved using

$$\mathbf{D}_c = \min_{\mathbf{D}_c} \|\mathbf{Y}_c - \mathbf{D}_c \mathbf{Z}_c\|_F^2 \quad \text{s.t.} \quad \|d_{c,m}\|_2^2 \leq 1 \quad \forall m = 1, 2, \dots, M. \quad (4.5)$$

At each iteration, the columns of the dictionary are normalized to have a unit norm to prevent scale ambiguities arising due to differences in strengths of the received signals from targets of varying radar cross-sections. I alternate between (4.4) and (4.5) till the algorithm converges.

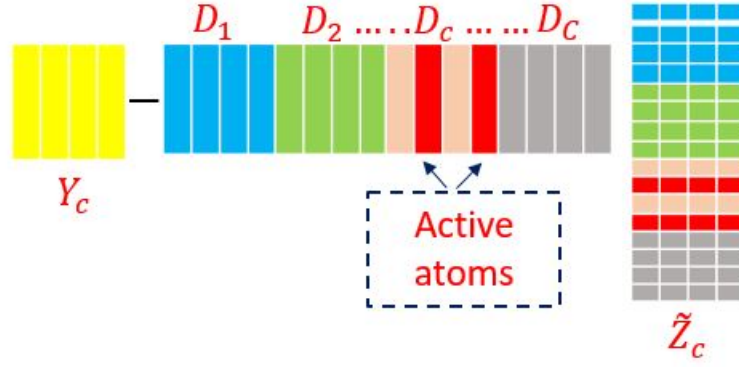


Figure 4.2: Synthesis dictionary learning framework for multiclass classification

Then, the dictionaries from all of the C classes are concatenated to form a single, over-complete $\mathbf{D} \in \mathbb{R}^{N \times KC}$ shown in

$$\mathbf{D} = [\mathbf{D}_1 \mathbf{D}_2 \mathbf{D}_3 \dots \mathbf{D}_C]. \quad (4.6)$$

This aggregate dictionary is used to generate the sparse coefficient matrix, $\tilde{\mathbf{Z}}_c \in \mathbb{R}^{KC \times M}$, for each class, c using

$$\tilde{\mathbf{Z}}_c = \min_{\tilde{\mathbf{Z}}_c} \left\| \mathbf{Y}_c - \mathbf{D} \tilde{\mathbf{Z}}_c \right\|_F^2 + \lambda \left\| \tilde{\mathbf{Z}}_c \right\|_1. \quad (4.7)$$

Note that $\tilde{\mathbf{Z}}_c$ is distinct from $\mathbf{Z}_c$ obtained from (4.4). It is hypothesized that if the target belongs to target class c , the representation must be a linear combination of few atoms of the corresponding dictionary $\mathbf{D}_c$. This is shown in Figure 4.2. Here the red columns show the active atoms of the aggregate dictionary and corresponding red rows in $\tilde{\mathbf{Z}}_c$ indicate the non-zero values of the coefficient matrix. Therefore, $\tilde{\mathbf{Z}}_c$ will mostly have significant values at similar positions for target class c while the remaining values will be either zero or negligibly small. In other words, $\tilde{\mathbf{Z}}_c$ will exhibit distinct row sparsity patterns for different

target classes c amid the cluster of coefficients belonging to all classes. The columns of sparse coefficient matrices, $\tilde{\mathbf{Z}}_c$, from all the target classes are used as input training features for a support vector machine (SVM) classifier with a linear kernel function (consistent across all the algorithms). The SVM is a popular classifier that has been extensively used for numerous machine learning applications. It defines boundaries to separate it from the other target classes and depends only on the class number c . Thus, the computational complexity of SVM is $\mathcal{O}(C)$.

Each radar test signal, y_{test} , is a single time-domain micro-Doppler measurement at a distinct carrier frequency from those used while training. I find the sparse features vector, $\tilde{z}_{test} \in \mathbb{R}^{KC \times 1}$, from y_{test} using $\mathbf{D}$ as shown

$$\tilde{z}_{test} = \min_{\tilde{z}_{test}} \|y_{test} - \mathbf{D}\tilde{z}_{test}\|_F^2 + \lambda \|\tilde{z}_{test}\|_1. \quad (4.8)$$

Similar to the training phase, the intuition here is that if the test sample y_{test} belongs to c class, the signal will be sparsely represented using few atoms of $\mathbf{D}_c$. As a result, the entries in the sparse coefficient vector, $\tilde{z}_{test}$, corresponding to the other classes, will be either zero or negligibly small. Meaning, the sparsity pattern in $\tilde{z}_{test}$ is most likely to resemble columns of $\tilde{\mathbf{Z}}_c$ among all the classes. Therefore, $\tilde{z}_{test}$, is classified by the SVM which is trained using $\tilde{\mathbf{Z}}_c$.

4.1.1.1 Computational Complexity of SDL

The computational complexity in each iteration of SDL shown in (4.9), consists of two parts- one simple linear least square problem and one iterative sparse-

coding problem.

$$J(\mathbf{D}_c, \mathbf{z}_c) = \min_{\mathbf{D}_c, \mathbf{z}_c} \|\mathbf{y}_c - \mathbf{D}_c \mathbf{z}_c\|_F^2 + \lambda \|\mathbf{z}_c\|_1. \quad (4.9)$$

The computational complexity of training a dictionary has already been computed in previous chapter which is $t\mathcal{O}(mn^2 + t1(mk^2 + k^3))$. Here, t is the number of iterations in training a dictionary and $t1$ is the number of iterations in sparse coding step. During the test phase, I solve the sparse-coding problem to extract the sparse features. Therefore, the computational complexity becomes $t2(\mathcal{O}(n^3))$, if we run the algorithm for $t2$, number of iterations.

4.1.2 Deep Dictionary Learning (DDL)

The SDL represents the micro-Doppler data through a single layer representation. Recent research on deep learning, mostly in neural networks, suggest that one can extract more fundamental or meaningful features through more profound representations of data. On similar lines, the single layer dictionary framework is extended to multiple layers of dictionaries [71]. The first step in the DDL algorithm is identical to the SDL algorithm. Training samples, $\mathbf{Y}_c$, are used to learn the dictionary and coefficient matrices, $\mathbf{D}_c^1$ and $\mathbf{Z}_c^1$, for the first level of DDL using the alternating minimization technique described in (4.4) - (4.5).

$$J(\mathbf{D}_c^1, \mathbf{Z}_c^1) = \min_{\mathbf{D}_c^1, \mathbf{Z}_c^1} \|\mathbf{Y}_c - \mathbf{D}_c^1 \mathbf{Z}_c^1\|_F^2 + \lambda_1 \|\mathbf{Z}_c^1\|_1 \quad (4.10)$$

For each successive n^{th} layer ($n = 2$ onwards to N total layers), the corresponding dictionary and coefficient matrices are learnt from the coefficients of

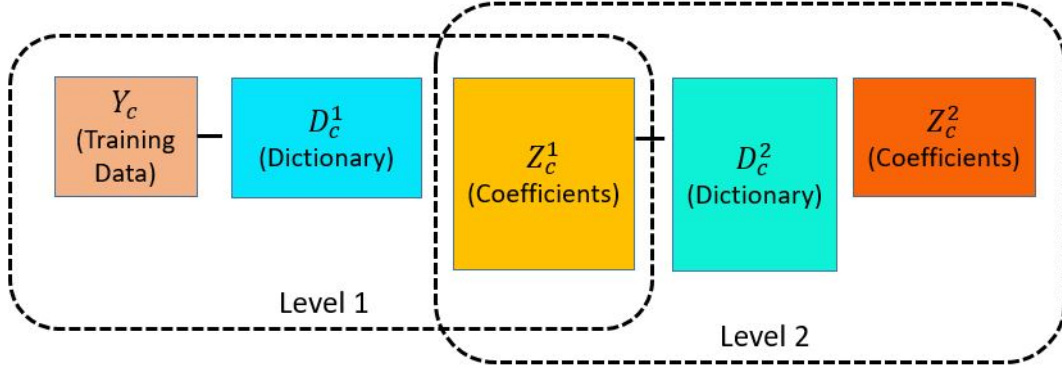


Figure 4.3: Deep dictionary learning framework

the previous layer such that $\mathbf{Z}_c^{n-1} = \mathbf{D}_c^n \mathbf{Z}_c^n$ as shown in Figure 4.3.

$$J(\mathbf{D}_c^n, \mathbf{Z}_c^n) = \min_{\mathbf{D}_c^n, \mathbf{Z}_c^n} \left\| \mathbf{Z}_c^{(n-1)} - \mathbf{D}_c^n \mathbf{Z}_c^n \right\|_F^2 + \lambda_n \|\mathbf{Z}_c^n\|_1 \quad (4.11)$$

The class dictionary, $\mathbf{D}_c$, is formed as a product of the multi-level dictionaries shown in

$$\mathbf{D}_c = \mathbf{D}_c^1 \times \mathbf{D}_c^2 \times \mathbf{D}_c^3 \dots \times \mathbf{D}_c^N. \quad (4.12)$$

The size of the class dictionary, $\mathbf{D}_c$, learned using an N -layer deep architecture is substantially reduced from the single layer class dictionary learned in SDL. In other words, as opting for deeper networks, fewer features are needed for representing the radar signals. The class dictionaries from multiple classes are concatenated to form $\mathbf{D}$ which is subsequently used for classification in a similar manner to (4.6) - (4.8). The sparse coefficient vector $\tilde{\mathbf{Z}}_c$ obtained from the training micro-Doppler data $\mathbf{Y}_c$, is used to train the SVM. Once the model is trained, $\tilde{z}_{test}$, is classified similarly to SDL.

4.1.2.1 Computational Complexity of DDL

In DDL, we usually extend single layer representation to have L -layered representation. Therefore, the total computational cost of training a dictionary in DDL becomes $L(\mathcal{O}(t(mn^2 + t1(mk^2 + k^3))))$, where L is the number of layers. The computation time in the training stage of DDL is greater than SDL due to the incorporation of these additional layers of learning. During the test phase, I solve for the sparse-coding problem to extract the sparse features similar to SDL with a computation cost of $\mathcal{O}(n^3)$. However, due to the reduced dimensionality of $\tilde{z}_{test}$ compared to SDL, the deep learning framework is faster during the test stage.

4.1.3 Analysis Dictionary Learning (ADL)

The algorithms discussed in Sections 4.1.1 and 4.1.2 belong to the synthesis framework where a signal y_c is synthesized with the linear combination of only a few atoms of the dictionary, leading to its sparse representation. An alternative generative framework - known as the analysis framework - is where a dictionary operates on y_c to generate its sparsest form $z_c = \mathbf{D}_c y_c$ as shown in Figure 4.4. Unlike SDL which focuses on classification based on sparsity (number and position of non-zero) patterns in the representation of the signals, here, the co-sparsity patterns (number of zeros) in z_c are utilized for distinguishing between multiple classes. The blue rows indicate these in the figure. The hypothesis here is that the signals belonging to different target classes will exhibit

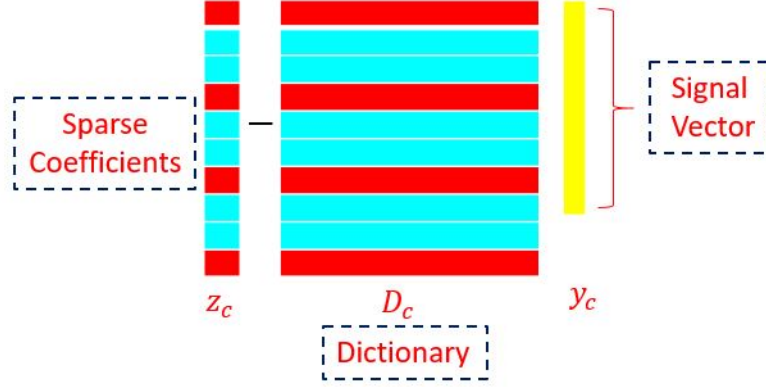


Figure 4.4: Analysis dictionary learning framework

unique co-sparsity patterns and thus belong to distinct subspaces from which they can be classified. The objective of ADL is to first learn a unique dictionary $\mathbf{D}_c \in \mathbb{R}^{P \times N}$ for each target class from the corresponding training data $\hat{\mathbf{Y}}_c \in \mathbb{R}^{N \times M}$. When $\mathbf{D}_c$ operates on $\hat{\mathbf{Y}}_c$, it produces sparse features $\mathbf{Z}_c = \mathbf{D}_c \hat{\mathbf{Y}}_c$. As mentioned before, M indicates the number of independent time-domain measurements consisting of N samples. To obtain high co-sparsity in $\mathbf{Z}_c$, the rows of $\mathbf{D}_c$ should exhibit high linear dependencies. Learning $\mathbf{D}_c$ involves the minimization of the following objective function -

$$J(\mathbf{D}_c, \hat{\mathbf{Y}}_c) = \min_{\mathbf{D}_c, \hat{\mathbf{Y}}_c} \left\| \mathbf{Y}_c - \hat{\mathbf{Y}}_c \right\|_F^2 \text{ s.t. } \left\| \mathbf{D}_c \hat{\mathbf{Y}}_c \right\|_0 \leq \tau \quad (4.13)$$

The l_0 constraint is relaxed to its nearest convex form by taking l_1 norm.

$$J(\mathbf{D}_c, \hat{\mathbf{Y}}_c) = \min_{\mathbf{D}_c, \hat{\mathbf{Y}}_c} \left\| \mathbf{Y}_c - \hat{\mathbf{Y}}_c \right\|_F^2 + \lambda \left\| \mathbf{D}_c \hat{\mathbf{Y}}_c \right\|_1 \quad (4.14)$$

(5.4) is solved using a variable splitting technique. To solve the problem more efficiently, a proxy variable $\mathbf{Z}_c$ is introduced such that the new objective function

becomes

$$J(\mathbf{D}_c, \hat{\mathbf{Y}}_c) = \min_{\mathbf{D}_c, \hat{\mathbf{Y}}_c} \left\| \mathbf{Y}_c - \hat{\mathbf{Y}}_c \right\|_F^2 + \lambda \left\| \mathbf{Z}_c \right\|_1 \quad \text{s.t. } \mathbf{Z}_c = \mathbf{D}_c \hat{\mathbf{Y}}_c \quad (4.15)$$

Here, solving exact Lagrangian is not desired. Therefore, the final objective function is reformulated as augmented Lagrangian [120] shown in (4.16)

$$\min_{\mathbf{D}_c, \hat{\mathbf{Y}}_c, \mathbf{Z}_c} \left\| \mathbf{Y}_c - \hat{\mathbf{Y}}_c \right\|_F^2 + \lambda \left\| \mathbf{Z}_c \right\|_1 + \mu \left\| \mathbf{Z}_c - \mathbf{D}_c \hat{\mathbf{Y}}_c \right\|_F^2 \quad \text{s.t. } \|d_{c,m}\|_2^2 \leq 1 \forall m = 1, 2, \dots, M \quad (4.16)$$

Here, λ acts as a regularizer trading off between representation error and sparsity. μ is a hyper-parameter that controls the equality between $\mathbf{Z}_c$ and its representation $\mathbf{D}_c \hat{\mathbf{Y}}_c$. For a high value of μ , equality is enforced. Otherwise, the constraint is relaxed. It is further ensured that each row of $\mathbf{D}_c$ is constrained to have a unit norm to prevent scale ambiguities due to variations in the signal strengths. Similar to synthesis framework, (5.4) is solved using a two-stage iterative framework. Here, $\mathbf{D}_c$ and $\hat{\mathbf{Y}}_c$, are initialized using randomly selected samples from the training data gathered at multiple carriers. The co-sparse coefficient matrix is obtained by $\mathbf{Z}_c = \mathbf{D}_c \hat{\mathbf{Y}}_c$. During the first stage of the iterative framework, $\mathbf{Z}_c$ and $\hat{\mathbf{Y}}_c$ are held constant and $\mathbf{D}_c$ is updated using least squares. In the second stage, I solve for $\hat{\mathbf{Y}}_c$ using the least squares as shown in

$$\min_{\hat{\mathbf{Y}}_c} \left\| \mathbf{Y}_c - \hat{\mathbf{Y}}_c \right\|_F^2 + \mu \left\| \mathbf{Z}_c - \mathbf{D}_c \hat{\mathbf{Y}}_c \right\|_F^2. \quad (4.17)$$

This is equivalent to solving the least squares problem defined in

$$\min_{\hat{\mathbf{Y}}_c} \left\| \begin{pmatrix} \mathbf{Y}_c \\ \sqrt{\mu} \mathbf{Z}_c \end{pmatrix} - \begin{pmatrix} \mathbf{I} \\ \sqrt{\mu} \mathbf{D}_c \end{pmatrix} \hat{\mathbf{Y}}_c \right\|_F^2. \quad (4.18)$$

Finally, $\mathbf{Z}_c$ is updated using the soft-thresholding method shown in

$$\mathbf{Z}_c = \text{soft}(\mathbf{D}_c \hat{\mathbf{Y}}_c, \gamma). \quad (4.19)$$

Here, $\gamma = \lambda/2\mu$, is the threshold that is selected for the co-sparse coefficient vector. The function above is defined as

$$\text{soft}(\mathbf{D}_c \hat{\mathbf{Y}}_c, \gamma) = \text{sign}(\mathbf{D}_c \hat{\mathbf{Y}}_c) \times \max(0, |\mathbf{D}_c \hat{\mathbf{Y}}_c| - \gamma) \quad (4.20)$$

The iterative procedure is continued until the objective function $J(\mathbf{D}_c, \hat{\mathbf{Y}}_c, \mathbf{Z}_c)$ converges to some local minimum. The class dictionaries from different target classes are then concatenated to form an aggregate dictionary $\mathbf{D}$. I train an SVM using the co-sparse features $\tilde{\mathbf{Z}}_c = \mathbf{D} \mathbf{Y}_c$, corresponding to each class. During the test phase, the SVM classifies the co-sparse feature vector $\tilde{z}_{test} = \mathbf{D} y_{test}$ of the test micro-Doppler data y_{test} .

4.1.3.1 Computational Complexity of ADL

During the first stage of the iterative framework of ADL, the dictionary is updated using least squares which takes $\mathcal{O}(mn^2)$ time. In the second stage, $\hat{\mathbf{Y}}_c$ is also updated using the least squares. The computational cost of this step is again $\mathcal{O}(mn^2)$. Finally, the sparse coefficient $\mathbf{z}_c$ is updated using a soft-thresholding

step which is $\mathcal{O}(n)$. Therefore, the overall complexity per iteration of the algorithm is $\mathcal{O}(mn^2) + \mathcal{O}(mn^2) + \mathcal{O}(n) = \mathcal{O}(mn^2) + \mathcal{O}(n)$. Note that the ADL is significantly faster in generating features at test time as compared to its synthesis counterpart as the feature generation in ADL involves only a simple product operation instead of the inverse operation in (4.8). This makes the algorithm more suited for real-time applications. The computational cost of test time is $\mathcal{O}(n)$.

4.2 Simulation

To demonstrate the application and effectiveness of dictionary learning based classification of micro-Doppler signatures, I have simulated time domain micro-Doppler returns from three classes of moving targets- human walking towards radar (FH), human walking away from radar (BH) and a rotating table fan (TF). I have used a point scatterer model for modelling the targets backscatter where it is assumed that the target is composed of few discrete scattering points and the superposition of the returns from these points are used to measure the aggregate reflected signal [2]. The radar returns of the human are obtained by combining a point scatterer distribution of the human with the Boulic-Thalmann model as described in detail in section 2.3.1.1. The model has some limitations - First, it is a constant velocity model. Therefore, it does not describe the start of the motion, the acceleration, turns or slowing down to halt. Second, it is an averaged model and hence does not show distinctions in human gait due to the subject's gender, age, mood or fitness. The model does not capture effects such as the

absorption, scattering, or shadowing between the different human body parts or background noise and clutter due to the propagation environment. Due to restrictions of this model, only two types of human motions - human walking towards the radar (FH) and human walking away from the radar (BH) are considered. In addition, I consider a source of clutter, the indoor table fan (TF), that gives rise to a distinct micro-Doppler signature due to the rotation of its blades. The dynamics of a rotating fan are described by its angular velocity and the length and width of the blades. The rotating fan (TF) is modelled using 3-point scatterers for each of its blades. The radar back-scattered returns from fan can be a function of number of parameters but, I have considered the following three parameters- angular velocity of rotational motion, blade length and blade width.

4.2.1 Simulation Set Up

To fully understand the underlying technique used for classification, I have performed extensive simulations. Here, one target is allowed to move in the channel at a time. The radar returns of each target type are simulated at multiple carrier frequencies- 2.4GHz, 4GHz, 5.8GHz, 7.5GHz and 10GHz. Here, the duration of each simulation is 1 second with a sampling frequency of 1KHz which is chosen in order to prevent aliasing across all carrier frequencies. To well describe the target kinematics and dynamics in joint time-frequency domain, I used short time frequency transform (STFT) with a Gaussian window of size 0.05sec. Spectrograms generated for simulated human walking towards radar

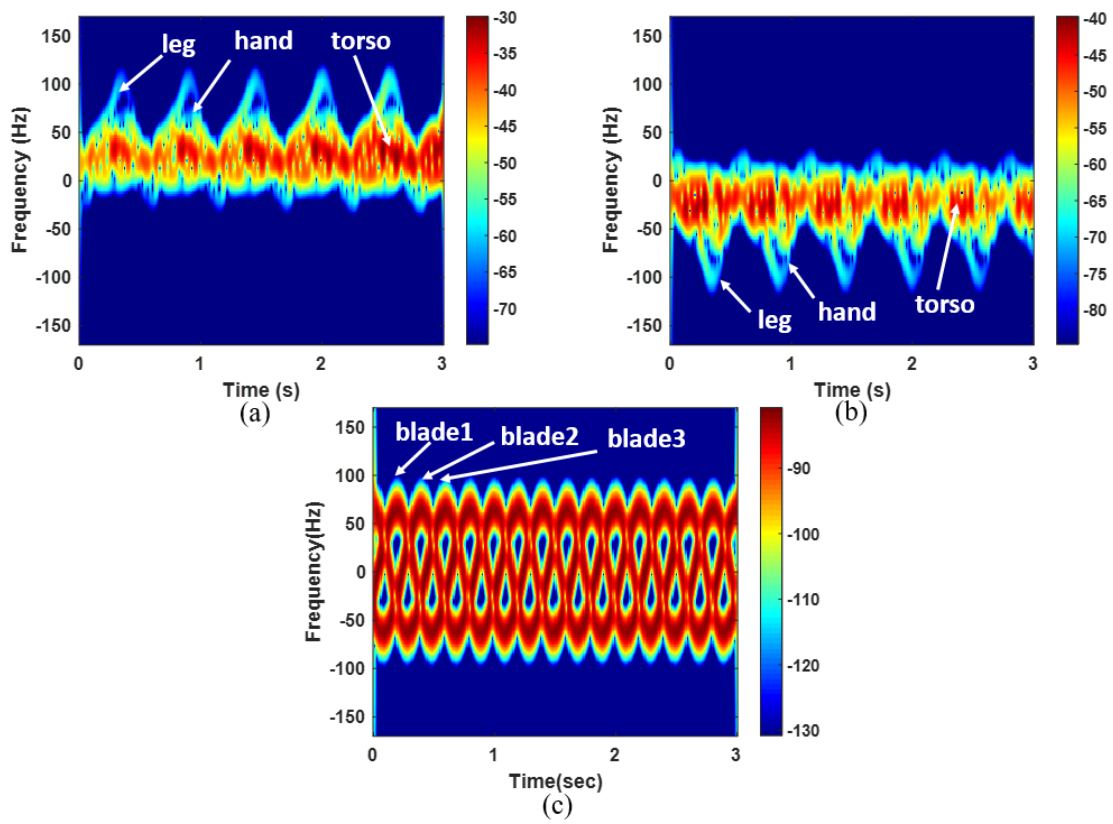


Figure 4.5: Spectrogram using STFT of a simulated (a) human walking towards radar, (b) a human walking away from the radar and, (c) a table fan.

Table 4.1: Simulation Data Set Description

Target Scenario	Target Parameters	Training Data	Test Data
Human walking towards radar (FH), Human walking away from radar (BH),	Target Heights: 5ft-6ft Target Velocities:0.8-1.7 m/s Number of measurements = 1800	1440 (360 measurements from each of 4 carrier frequencies)	360 (measurements from fifth carrier frequency)
Table fan (TF),	Number of blades: 3 Angular velocities: 200-400 rpm Length of the blades : 0.2-0.4m Width of blades: 0.14m-0.17m Number of measurements = 1800	1440 (360 measurements from each of 4 carrier frequencies)	360 (measurements from fifth carrier frequency)

(FH), human walking away from radar (BH) and a table fan (TF) at 4GHz are shown in Figure 4.5 where the Dopplers are positive when the target is approaching the radar and mostly negative as the distance between radar and target is increased. Table 4.1 summarizes the entire simulated data. I vary the human parameters (height and velocity) and the table fan parameters (blade length, width and angular velocity) to generate 360 simulations corresponding to each carrier frequency. To test the robustness of the algorithm under noisy conditions, the signal to noise ratio (SNR) is varied from -10dB to $+10\text{dB}$. data.

4.2.2 Simulation Results

I consider a training matrix of size $[1000 \times 1440]$ comprising data from 4 out of the 5 carrier frequencies each consisting of 1000 time-domain samples. The SDL technique described in Section 4.1.1, is used to learn under-complete class dictionaries each of size $[1000 \times 500]$. These dictionaries are then concatenated to form a single aggregate over-complete dictionary of size $[1000 \times 1500]$. The concatenated dictionary is used to generate a class specific sparse features matrix of size $[1500 \times 1440]$ which is used to train a support vector machine (SVM)

Table 4.2: Description of test and training scenarios across multiple carrier frequencies

Folds	Test Data Carrier Frequencies	Training Data Carrier Frequencies
Fold 1	2.4GHz	4, 5.8, 7.5, 10GHz
Fold 2	4GHz	2.4, 5.8, 7.5, 10GHz
Fold 3	5.8GHz	2.4, 4, 7.5, 10GHz
Fold 4	7.5GHz	2.4, 4, 5.8, 10GHz
Fold 5	10GHz	2.4, 4, 5.8, 7.5GHz

Table 4.3: Classification results when training and test micro-Doppler data are simulated for different carrier frequencies for SNR= 10 dB

Algorithms	Test Cases	Fold 1	Fold 2	Fold 3	Fold 4	Fold 5	Average Accuracy
SDL	FH	100	100	100	100	100	100
	BH	100	100	100	100	100	100
	TF	100	100	100	100	100	100
DDL	FH	100	100	100	100	100	100
	BH	100	100	100	100	100	100
	TF	100	100	100	100	100	100
ADL	FH	92.5	97.8	94.5	92.2	88.3	93.1
	BH	90.8	95.3	91.7	93	90.6	92.28
	TF	100	100	100	100	100	100
LC-KSVD	FH	11.4	77.8	86.4	90.8	98.6	73
	BH	0	69.8	23	42	1.4	27.2
	TF	99.1	18.9	95	13.7	50	55.34

classifier. It is hypothesized that radar signals from different classes will exhibit distinctive patterns in this sparse features matrix. In the test phase, a single $[1000 \times 1]$ micro-Doppler signal, corresponding to the fifth carrier frequency, is used to generate a sparse feature vector of size $[1500 \times 1]$. This vector is classified by the SVM as one of three classes- FH, BH or TF and the results are presented in Table.4.3.

In the DDL framework, a 3-layer deep network with the successive layers having dictionary sizes - $[1000 \times 500]$, $[500 \times 250]$ and $[250 \times 100]$ is learnt. As a result, the sparse feature matrix used for training the SVM is of size $[100 \times 1440]$

and the sparse feature vector used for testing is of size $[100 \times 1]$. In other words, the deep learning architecture requires fewer features for classification as we go down the hierarchy. This, in turn, reduces the computational complexity during the test phase. Finally, in ADL, a class dictionary of size $[500 \times 1000]$ is learnt that will operate on the training matrix. The concatenated dictionary from all three classes, of size $[1500 \times 1000]$, is used both for training and test. The size of the dictionaries is thus consistent across all the algorithms. The co-sparse feature matrix used for training the SVM is of size $[1500 \times 1440]$ and the co-sparse feature vector, used while testing, is of size $[1500 \times 1]$. I compare the performances of SDL, DDL and ADL with the LC-KSVD algorithm [127]. I fix two additional parameters in the LC-KSVD algorithm - the inter-class discrimination and class-wise sparsity regularizers. The variation of these parameters result in minor changes in the classification accuracy (around 2 to 3 %).

I select values close to those suggested by [114] that yield the best results. The results in Table 4.3 show the average classification accuracies across 5 folds as described in Table 4.2. The SDL and DDL show 100% accuracy across all folds while ADL give an average classification accuracy of 95.12%. These results demonstrate that the algorithms are very effective in extracting meaningful features corresponding to the target motion even when the data are gathered at different carrier frequencies. This facilitates the correct classification of micro-Doppler data even when there are substantial differences in the training and test scenarios. The LC-KSVD correctly recognizes FH and TF in 73% and 55.34% of the test cases. However, the average accuracy for the case of BH is only

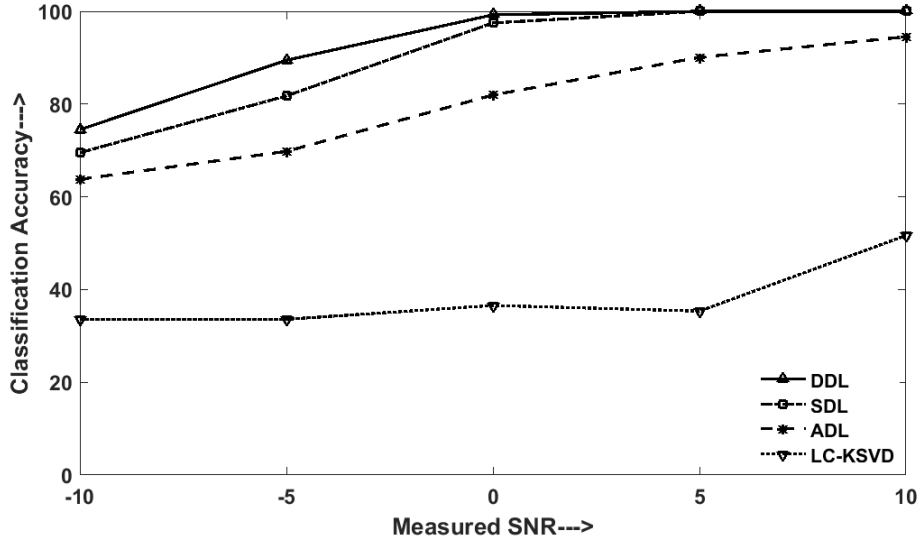


Figure 4.6: Classification accuracy of dictionary learning algorithms under noise.

27.2%. This result demonstrates that the radar returns for a human walking towards and away from the radar are not symmetrical. When the human is walking away from the radar, the micro-Dopplers spread is much lower than in the case of the forward motion due to restricted back swing of the leg compared to its forward swing. More significantly, the LC-KSVD algorithm is unable to obtain unique class dictionaries when the training data are gathered from multiple diverse carrier frequencies. A complete analysis of the performance of the LC-KSVD algorithm is provided with the measurement results. Figure 4.6 shows the average classification performance of the algorithms in the presence of Gaussian noise. The classification performs remains above 90% as long as the SNR is above 0dB.

4.3 Measurements

The simulation results present some useful insights on the performance of the algorithms. However, the simulation data do not capture the true diversity of data that a radar would encounter in a real scenario. Real data would be impacted by factors as varied as the radar hardware limitations, propagation environment effects, diversity of human behaviours even while performing similar motions and the interaction between the humans and the environment including other human subjects. In this section, I describe the experimental set up used to generate measurement micro-Doppler data at multiple carrier frequencies. I again used the same set up based on the hardware described in Chapter 2, Section 2.3.2. However, this time the radar set up is used for measurement data collection where only single dynamic target is considered in the propagation environment for the task of classification. All of the measurements are conducted in indoor, line-of-sight conditions.

4.3.1 Measurement Set Up

The VNA shown in Figure 4.7, is configured to the narrow band mode with a bandwidth of 10Hz and a center frequency of any one of five carrier frequencies - 2.4, 3, 4, 4.5 and 5.8 GHz. A set of lower carrier frequencies is considered to avoid aliasing issues since the sampling frequency is limited to 370Hz by the VNA hardware. The duration of each measurement is 27 seconds. The measured data are further segmented to 10 shorter signals, each of 2.7 seconds

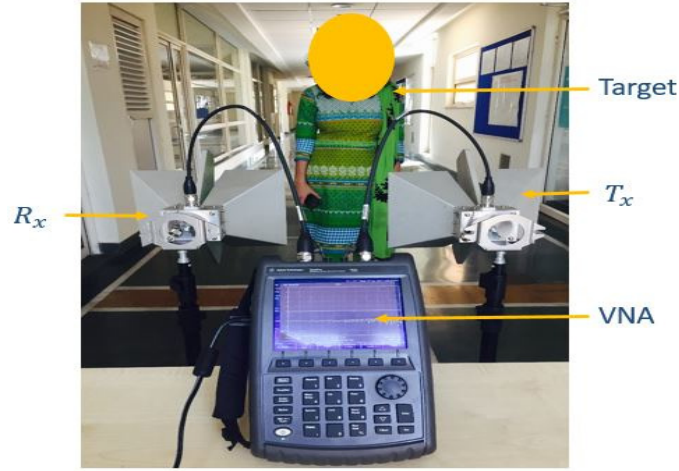


Figure 4.7: Monostatic CW radar configured using a vector network analyzer and two linearly polarized horn antennas for five distinct carrier frequencies - 2.4, 3, 4, 4.5 and 5.8GHz

Table 4.4: Description of test and training scenarios across multiple carrier frequencies

Folds	Test Data Carrier Frequencies	Training Data Carrier Frequencies
Fold 1	2.4GHz	3, 4, 4.5, 5.8GHz
Fold 2	3GHz	2.4, 4, 4.5, 5.8GHz
Fold 3	4GHz	2.4, 3, 4.5, 5.8GHz
Fold 4	4.5GHz	2.4, 3, 4, 5.8GHz
Fold 5	5.8GHz	2.4, 3, 4, 4.5GHz

duration.

Four distinct target classes are considered. Three target classes consist of human motions. The fourth target class is a table fan (TF) with rotating blades. I have specifically chosen this distinct fourth target class since a table fan gives rise to micro-Doppler signatures that can contribute to significant clutter in radar tracking of humans in indoor environments. The three human motion categories considered are - human walking in front of the radar (HW), two humans walking before the radar (TH) and a person standing and boxing his arms (HB). For each of these categories, I conducted measurements with 20 subjects of different heights, gait patterns, velocities, gender, and ages. These measurements

were repeated for each of the five different carrier frequencies mentioned above. The human motions were completely unrestricted. They, therefore, consist of motion transitions such as starting from rest, acceleration, turns, slowing down to halt, etc. Next, measurements were carried out with the table fan at different angular speeds, distances, and orientations with respect to the radar at the five carrier frequencies. There are therefore a total of 100 measurements for each motion category of which 80 measurements corresponding to 4 carrier frequencies are used for training the dictionary learning algorithms. The remaining data (20 measurements) corresponding to the fifth carrier frequency are used as test data. Table 4.5 summarizes the entire data collection. The complete data set and its detailed description is available in [128]. The targets move before the radar between 1m and 10m distance. If the human is assumed to have an average radar cross-section of $1m^2$, then the approximate dynamic range in the measurement data is 30dB (between -68 and -98 dBm). The radar cross-section of the table fan is much lower and can, therefore, result in weaker returns, sometimes close to the noise floor, especially when the fan is at an inclination away from the radar at 10m.

4.3.2 Micro-Doppler Spectrograms Across Multiple Carriers

The classification algorithms are posed with some unique challenges when the training and test measurement data are gathered under different conditions. I illustrate these challenges by presenting the micro-Doppler spectrograms for these motions for two carrier frequencies in Figure 4.8. The spectrograms are

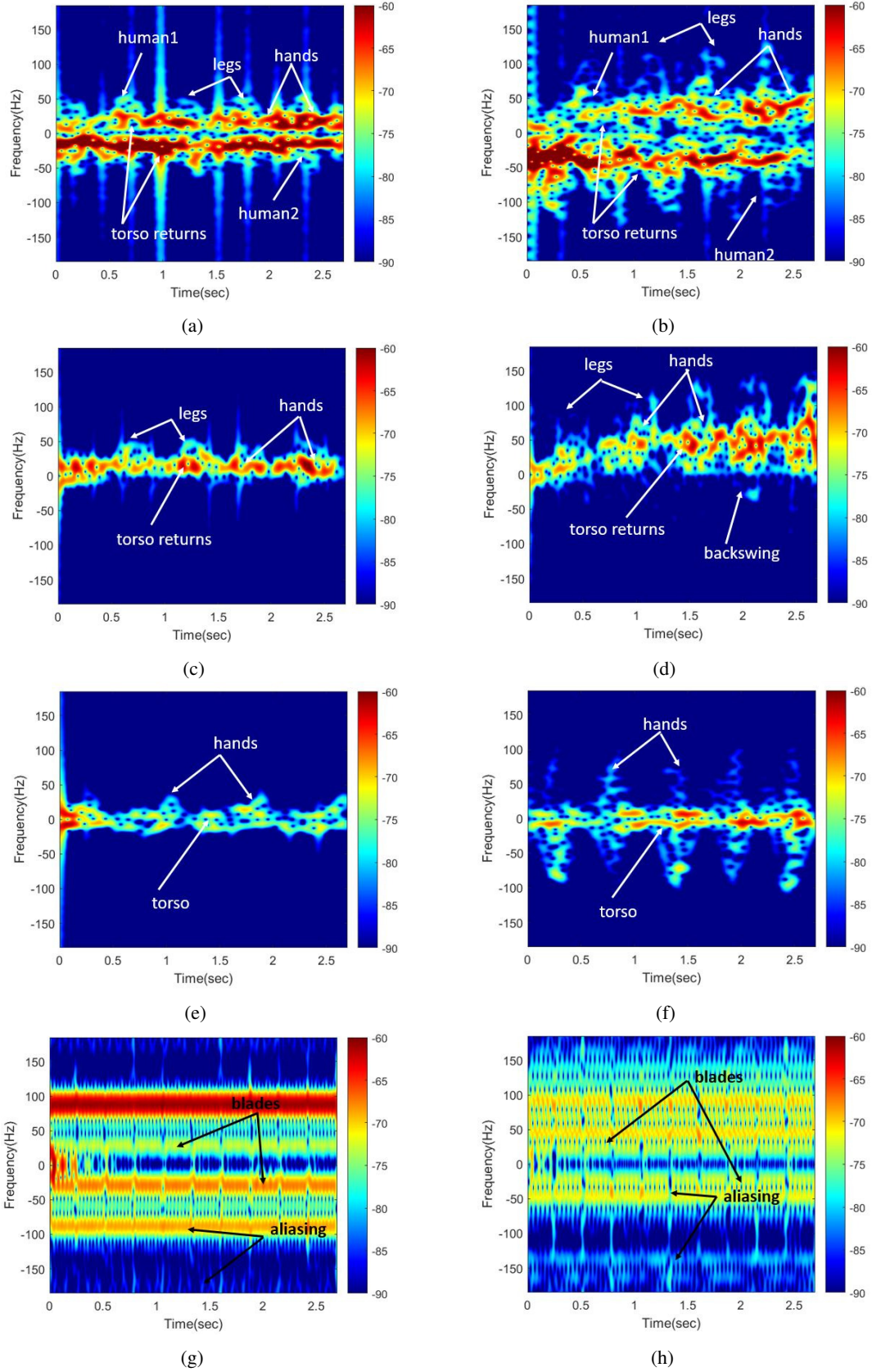


Figure 4.8: STFT spectrograms of (a, b) two walking humans - one walking towards the radar and the other walking away from the radar, (c, d) a human walking towards the radar, (e, f) a human boxing and (g, h) a rotating table fan. The figures in the left columns are generated from data collected at 2.4GHz and figures in the right columns are generated from data collected at 5.8GHz.

Table 4.5: Measurement Data Description

Target Scenario and Description	Target Parameters	Training Data	Test Data
Two humans walking (TH), Human boxing (HB), Human walking (HW)	Number of humans: 20 Target Heights: 5ft to 6ft Target Velocities: 0.6m/s to 1.5m/s Number of measurements = 100 (5 measurements from each of 20 subjects)	80 (5 measurements from each of 4 carrier frequencies)	20 (5 measurements from fifth carrier frequency)
Table fan (TF)	Number of fans: 1 Angular velocities: 1400 rpm, 2000 rpm Locations from radar : 10 Number of measurements = 100 (5 measurements from each of 20 cases)	80 (5 measurements from each of 4 carrier frequencies)	20 (5 measurements from fifth carrier frequency)

generated using the short-time Fourier transform with a dwell time of 0.05 seconds. The figures on the left column correspond to data generated at 2.4GHz while the figures from the right column correspond to 5.8GHz data. These are the lowest and highest carrier frequencies that have been selected for our experimental purposes. The first row - Figure 4.8(a) and Figure 4.8(b) - show the spectrograms corresponding to two humans walking before the radar. One human walks towards the radar, with mostly positive micro-Dopplers, while the second walks away from the radar, with mostly negative micro-Dopplers. The legs give rise to higher Dopplers than the arms or torso. These leg micro-Dopplers can occasionally be aliased to lower frequencies especially at 5.8GHz as seen in Figure 4.8(b). The next two figures, Figure 4.8(c) and Figure 4.8(d), correspond to the human walking before the radar. Again, in these figures, it is observed that the strongest returns arise from the torso while the limbs contribute to much weaker returns. This is especially evident in Figure 4.8(d). This is because of the lower cross-section of the limbs and due to shadowing of one limb by an-

other. Figure 4.8(e) and Figure 4.8(f) show the micro-Dopplers from a human standing still and boxing his arms. Here the torso Doppler is mostly around 0Hz since there is no translational motion of the human. Similarly, the legs do not have a distinct Doppler. Instead, both the positive and negative Dopplers observed are arising from the swinging motion of the boxer's arms. The last two figures, Figure 4.8(g) and Figure 4.8(h), are derived from the rotation of the three blades of the table fan. The rotation motion of the blades of the fan gives rise to micro-Dopplers that are a function of the number of blades, the angular speed of rotation, the orientation of the blades with respect to the radar and the length of the blades. Due to the low sampling frequency, the spectrograms of the table fans show significant aliasing across all the carrier frequencies.

These figures highlight the key challenges before the classification problem. The algorithm, trained with data gathered at particular carrier frequencies (say 2.4GHz to 4.5GHz data) must successfully classify data belonging to 5.8GHz. Some of the important factors to be taken into consideration are listed below.

1. The figures demonstrate that the Doppler spectrograms corresponding to high carrier frequency data (in the right column) show finer frequency resolution than those from low carrier frequency data (in the left column). This is particularly evident in the case of the three human motion classes.
2. The sampling frequency is identical for all the measurements. Therefore, there is a much higher probability of aliasing to occur at 5.8GHz (especially if the human moves at a high velocity) when compared to the lower carriers.

This problem of aliasing is particularly evident in the case of the table fan spectrograms as pointed out before. This limitation may give rise to some errors in classification.

3. There are differences in the micro-Doppler patterns within the same motion categories arising due to the variation in gait patterns across individuals (due to their height, weight, age, gender, fitness, and mood).
4. Measurement data have issues arising from shadowing of a target or parts of a target by the environment or the presence of other targets.
5. Even the measurement data gathered from a single individual, show micro-Doppler variations due to motion transitions.

4.3.3 Measurement Results and Analyses

The measurement data consists of four dynamic target classes- two humans walking (TH), human boxing while standing still (HB), a single human walking (HW) and a rotating table fan (TF). First, I consider the single carrier case. Here, the training and test micro-Dopplers are gathered at the same carrier frequency. Then I discuss the more challenging multi-carrier case. Here, the training features are obtained from micro-Doppler data from 4 out of 5 carrier frequencies while the test features are derived from the fifth carrier frequency.

I apply the three dictionary learning frameworks discussed in Section.4.1.1, 4.1.2, and 4.1.3 to the measurement data to study their effectiveness at classifying micro-Dopplers. I compare their performances with four non-DL based

feature extraction methods that have been used for micro-Doppler signature classification in recent literature - physical features (PF), DCT coefficients, cepstral features (CF), and principal components analysis. I also consider another DL based algorithm that was tested recently - the label consistent KSVD (LC-KSVD). The LC-KSVD is also based on the synthesis learning framework. Unlike the SDL, where class dictionaries are learned individually, the LC-KSVD learns multi-class dictionaries jointly with regularizers for class-wise sparsity and inter-class discrimination in its objective function. For a detailed description of LC-KSVD, I refer readers to [114]. The algorithms are run in MATLAB 2015b on an Intel(R) Core(TM) i7-5500U CPU running at 2.40 GHz; 16-GB RAM, Windows 10 (64 b).

4.3.3.1 Parameter selection for dictionary learning

A training matrix of size $[1000 \times 200]$ is considered for each target class by randomly repopulating the original measurements. Each signal vector has 1000 time domain samples over a duration of 2.7 seconds. The SDL technique described in Section 4.1.1, is used to learn under-complete class dictionaries each of size $[1000 \times K]$. These dictionaries from the four classes are then concatenated to form a single aggregate over-complete dictionary of size $[1000 \times 4K]$. The concatenated dictionary is used to generate a class-specific sparse features matrix of size $[4K \times 200]$, which is used to train a support vector machine (SVM) classifier. It is hypothesized that radar signals from different classes will exhibit distinctive patterns in this sparse features matrix. In the test phase,

a single $[1000 \times 1]$ micro-Doppler signal is used to generate a sparse feature vector of size $[4K \times 1]$. This vector is classified as one of the four classes by the SVM. Different sizes of class dictionaries, K , are examined and the results are presented in Figure 4.9(a). The results show that the algorithm is not very sensitive to the size of K provided the concatenated dictionaries from the four classes (of size $4K$) are over-complete. I choose $K = 500$.

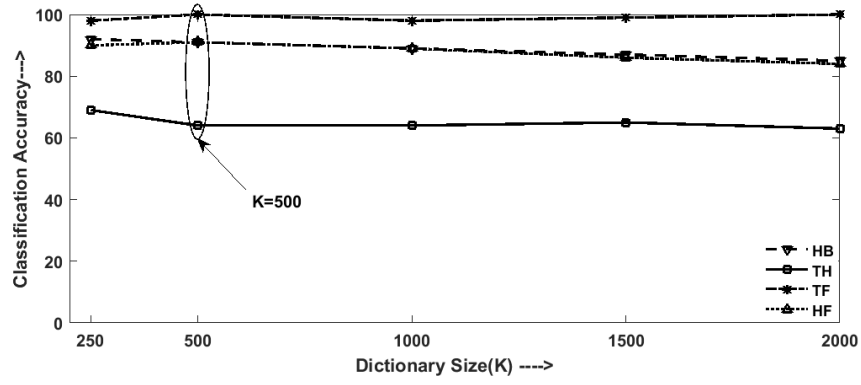
In the DDL framework, a N -layer deep network is learnt with each successive layer having a reduced dictionary size. I start with a dictionary size of $[1000 \times 500]$ to match that of SDL. I considered different depths as shown in Figure 4.9(b). Our results show that there is an improvement when $N = 3$ after which the advantage of abstraction is lost. Subsequent dictionary sizes are $[500 \times 250]$ and $[250 \times 100]$. The sparse feature matrix used for training the SVM is of size $[100 \times 200]$ and the sparse feature vector used for testing is of size $[100 \times 1]$. This size is considerably smaller than the SDL. Since the deep learning architecture requires fewer features for classification as we go down the hierarchy, the computational complexity during the test phase is reduced.

Finally, in ADL, I learn a class dictionary of size $[K \times 1000]$ that will operate on the training matrix. Again, I examine the effect of different values of K on the performance of the algorithm. The results are shown in Figure 4.9(c). The results show that the algorithm's performance is not very sensitive to K . I choose $K = 500$ to make it consistent with the other two dictionary algorithms. The concatenated dictionary from all four classes, of size $[2000 \times 1000]$, is used both for training and test. The size of the dictionaries is thus consistent across

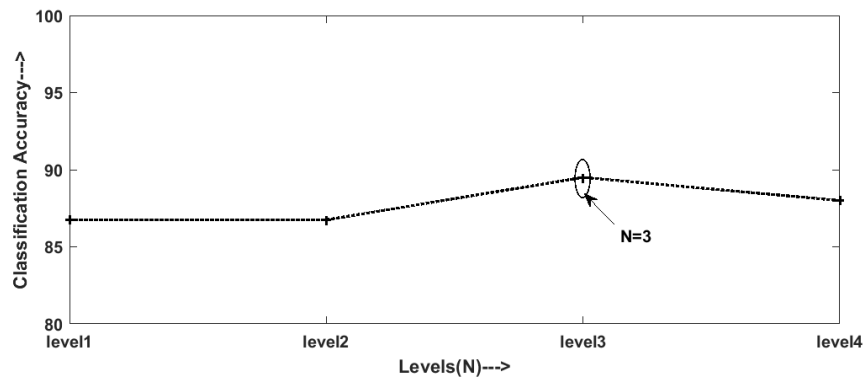
all the algorithms. The co-sparse feature matrix used for training the SVM is of size $[2000 \times 200]$ and the co-sparse feature vector, used while testing, is of size $[2000 \times 1]$. The choice of λ dictates the trade-off between representation and sparsity error. I show how the performance varies as a function of λ in Figure 4.9(d). I choose $\lambda = 0.001$ for the best results. I fix three additional parameters in the LC-KSVD algorithm - the inter-class discrimination that is the weights for label constraint term (α), weights for classification error term (β) and sparsity prior (S) regularizers. The variation of these parameters results in minor changes in the classification accuracy (around 3 to 4 %). I select values $\sqrt{\alpha} = 0.001$, $\sqrt{\beta} = 0.001$ and $S = 10$ that yield the best results.

4.3.3.2 Classification based on single-carrier data

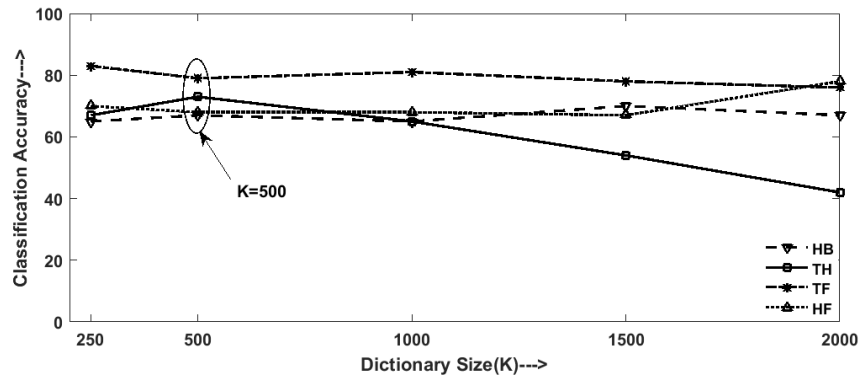
First, I present the classification results for the five-fold single-carrier frequency case in Table 4.6. The training and test for each fold are carried out on data from the same carrier. The classification accuracy for each entry in this table is obtained by the average across five folds. The results show 100% average classification success across five folds in all three synthesis algorithms (SDL, DDL, and LC-KSVD) and slightly lower performance for the ADL (average classification accuracy of 98.5%). The results compare favorably with other classification works on micro-Doppler data presented in literature (PF, CF, DCT, and PCA). The results show that the micro-Dopplers for these motions are *distinctive* and any conventional feature extraction technique is suitable for classification when I consider single carrier data.



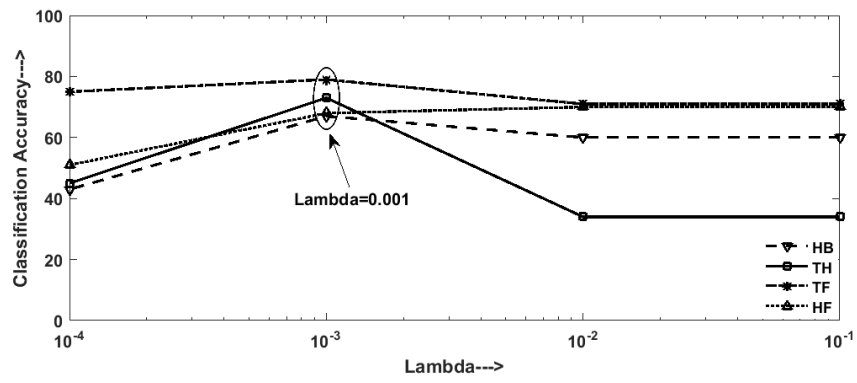
(a)



(b)



(c)



(d)

Figure 4.9: Variation of classification accuracy with (a) dictionary size for SDL, (b) depth in dictionary layers in DDL, (c) dictionary size for ADL and (d) Sparsity parameter (λ) for ADL.

Table 4.6: Comparison of average classification accuracy across multiple algorithms when training and test micro-Doppler data are gathered at same carrier frequencies

Cases	SDL	DDL	ADL	LC-KSVD	PF	CF	DCT	PCA
TH	100	100	100	100	95	100	100	100
HB	100	100	100	100	90	100	100	100
HW	100	100	94	100	80	100	100	100
TF	100	100	100	100	100	100	100	100
Average	100	100	98.5	100	91.25	100	100	100

Table 4.7: Comparison of average classification accuracy across multiple algorithms when training and test micro-Doppler data are gathered at different carrier frequencies

Cases	SDL	DDL	ADL	LC-KSVD	PF	CF	DCT	PCA
TH	67	68	73	13	89	38	69	58
HB	88	98	67	81	47	63	93	72
HF	93	97	68	25	52	40	56	59
TF	95	99	79	98	100	46	97	37
Average	85.75	90.5	71.5	54.25	72	46.75	78.75	56.5

4.3.3.3 Classification based on multiple carrier data

Next, I consider the more challenging scenario - when the training and test data are gathered at different carrier frequencies. I again consider five fold classification. Each fold now represents results for training data gathered from four out of five carrier frequencies and test data from the fifth (remaining) carrier. This is fully detailed in Table 4.4. The resulting classification accuracies are presented in Table 4.7. The results show a significant deterioration in the classification performance of all the algorithms in comparison to the single carrier case. This means that the micro-Dopplers from the multiple motions are no-longer distinctive. Among the non-DL algorithms, the DCT coefficients are most successful at classification. However, the performance is inferior to SDL and DDL algorithms.

To understand this further, I visualize the data reduced to a two-dimensional space using the t-distributed Stochastic Neighbor Embedding technique [129]. I consider the scatter plots of the raw data, the features extracted by DDL and the handpicked features extracted by a non-DL method (PF) as shown in Figure 4.10. Ideally, the four target class data must belong to four distinct clusters. However, this is not the case for the raw data in Figure 4.10(a). Figure 4.10(b), for PF case, shows a very distinct cluster for TF but the features from the three remaining motions overlap considerably. In the case of DDL, in Figure 4.10(c), The features from the four classes are mostly well separated - except for few cases corresponding to TH and HF. These result in the incorrect classifications seen in Table 4.9. I examine the performance of SDL, DDL, and ADL in greater detail in the following sections. SDL Table 4.8 shows the confusion matrix of the classification results for the SDL algorithm. Here, the row entries under the header "Test Cases" are the true class labels for the test micro-Doppler data and the column headers are the class labels to which the data are classified. The diagonal entries, therefore, indicate the correct classification results. The overall accuracy for HB, HW, and TF are superior to TH. TH is mostly confused with HW in all of the five folds. It may be because the radar returns from the second human may be much weaker than the first, in some instances, due to its greater distance from the radar or because the first subject shadows the second subject. There are also some instances when both the humans move in a synchronized manner with respect to the radar giving rise to overlapped micro-Doppler returns. TH is confused with HB to a lesser extent. This is likely

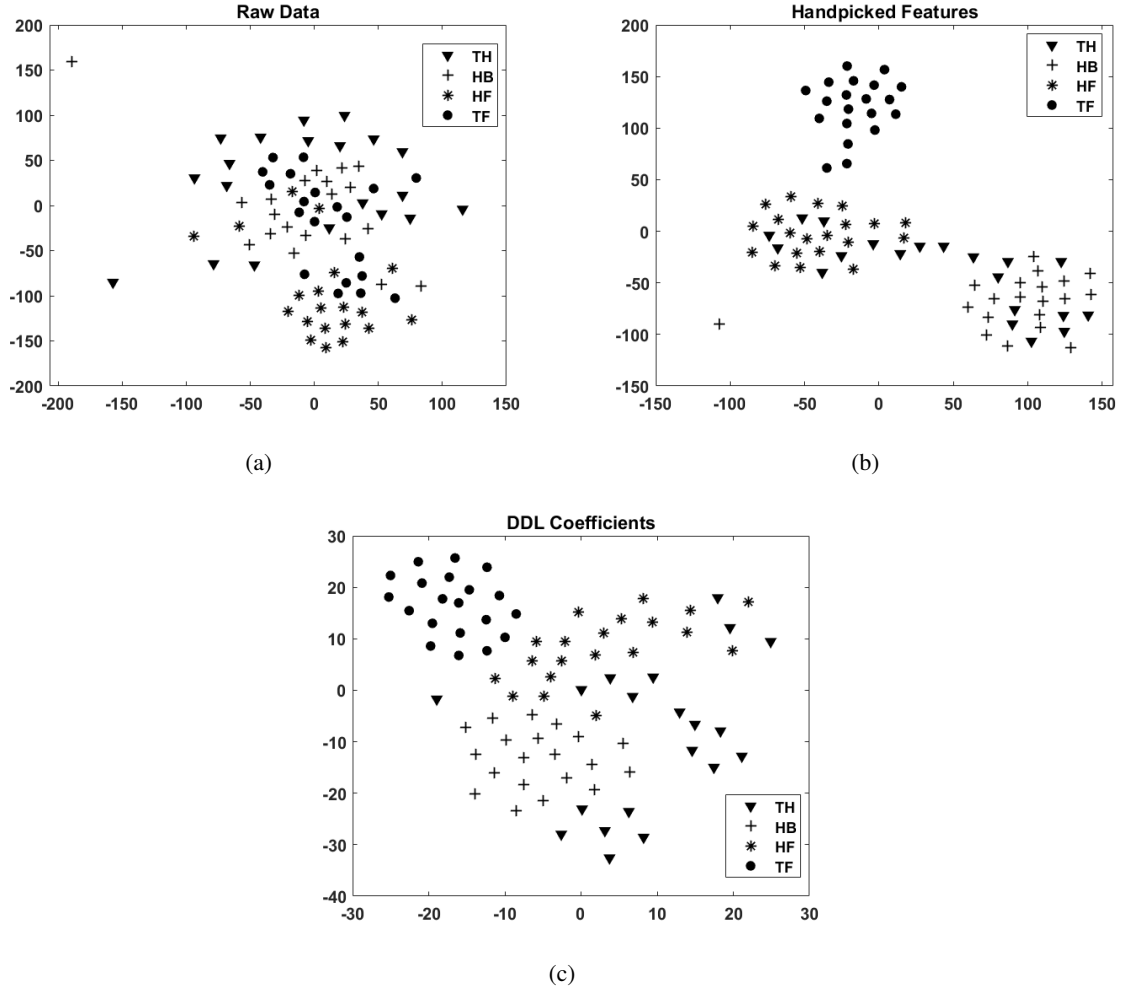


Figure 4.10: Scatter plot to visualize (a) raw data, (b) handpicked features (PF) and (c) features extracted using DDL

since both of these signatures show both positive and negative micro-Dopplers spreads. For the same carrier frequency, the micro-Doppler spread of the TH is higher than that of the HB due to the absence of micro-Dopplers from legs in the latter case. However, these spreads could be similar when data from different carrier frequencies are considered. This is why the confusion between the TH and HB did not occur in the single carrier frequency case while it occurs in the multi-carrier frequency case. HW and HB are at times confused. Here the confusion arises due to the back-swing motion of the arms and legs while walking that results in some negative Dopplers (when the human is walking

Table 4.8: SDL results when training and test micro-Doppler data are gathered at different carrier frequencies

Folds	Test cases	TH	HB	HW	TF
Fold 1	TH	75	10	15	0
	HB	0	70	25	5
	HW	0	20	80	0
	TF	0	0	5	95
Fold 2	TH	65	5	30	0
	HB	0	95	5	0
	HW	0	15	85	0
	TF	0	0	0	100
Fold 3	TH	45	10	45	0
	HB	0	90	10	0
	HW	0	0	100	0
	TF	0	0	0	100
Fold 4	TH	75	15	10	0
	HB	0	90	5	5
	HW	0	0	100	0
	TF	0	0	0	100
Fold 5	TH	75	0	25	0
	HB	5	95	0	0
	HW	0	0	100	0
	TF	0	0	20	80
Recall		67	88	93	95
Accuracy					

towards the radar). TF shows a poorer performance in the fifth fold. This can be attributed due to the similarity of micro-Doppler spreads from both cases due to aliasing of HW at high carrier frequencies.

DDL I present the results of the DDL algorithm in Table 4.9. The DDL shows an overall superiority to SDL across the classes - especially HB, HW, and TF. The TH shows the poorest performance and is confused mostly with HW and to a lesser extent with HB. Again, this poor performance can be attributed to the underlying challenge in distinguishing two targets with a radar system of limited dynamic range and frequency resolution. The superiority of the performance of DDL to SDL shows that the deeper representations lead to extraction of more

Table 4.9: DDL results when training and test micro-Doppler data are gathered at different carrier frequencies

Folds	Test cases	TH	HB	HW	TF
Fold 1	TH	75	15	10	0
	HB	0	100	0	0
	HW	0	0	100	0
	TF	0	0	0	100
Fold 2	TH	70	5	25	0
	HB	0	100	0	0
	HW	0	5	95	0
	TF	0	0	0	100
Fold 3	TH	55	10	30	5
	HB	0	95	5	0
	HW	0	0	100	0
	TF	0	0	0	100
Fold 4	TH	70	5	25	0
	HB	0	100	0	0
	HW	0	5	90	5
	TF	0	0	0	100
Fold 5	TH	70	0	30	0
	HB	5	95	0	0
	HW	0	0	100	0
	TF	0	0	5	95
Recall		68	98	97	99
Accuracy					

fundamental features from measurement data.

ADL The performance of the ADL is very poor compared to SDL and DDL across all classes except TH as seen in Table 4.10. It is an important observation since the TF class has hitherto been classified successfully (above 90%) by the other dictionary learning algorithms. The ADL results in this table mark a significant departure from the previous ADL results reported in Table 4.6 (measurement data with identical training and test scenarios). From these observations, it is inferred that the ADL algorithm is heavily impacted by the aliasing in the measurement data, both during training and testing. In particular, aliasing occurs for the TF data across all carriers. The human motions data, on the other

Table 4.10: ADL results when training and test micro-Doppler data are gathered at different carrier frequencies

Folds	Test cases	TH	HB	HW	TF
Fold 1	TH	70	10	20	0
	HB	0	55	25	20
	HW	10	25	60	5
	TF	0	0	5	95
Fold 2	TH	90	0	10	0
	HB	0	70	25	5
	HW	5	25	50	20
	TF	0	30	5	65
Fold 3	TH	75	10	15	0
	HB	0	75	20	5
	HW	0	0	60	40
	TF	0	0	0	100
Fold 4	TH	50	15	35	0
	HB	0	60	30	10
	HW	5	10	80	5
	TF	0	0	5	95
Fold 5	TH	80	5	15	0
	HB	20	75	5	0
	HW	10	0	90	0
	TF	0	0	60	40
Recall		73	66	68	79
Accuracy					

hand, are impacted more severely at the higher carrier frequencies. The performance of the ADL depends on the ability of the algorithm to generate a unique co-sparse representation of the data from each class. To ensure rich co-sparsity in the signal representation, the class dictionaries must exhibit a high degree of row-wise linear dependency. The algorithm fails to achieve this when there are overlaps in the micro-Doppler signal spreads due to aliasing. This problem does not occur in the single carrier case.

Table 4.11 also shows the computational time required during the test and training phases for all the algorithms. The training phase includes the time taken to learn the training features matrices for the SVM (for SDL, DDL, ADL, PF,

Table 4.11: Comparison of overall classification accuracy and computational complexity

Algorithm	Training time (seconds)	Test time (seconds)	Classification accuracy (%)
SDL	414.9	3.4	85.75
DDL	631.5	0.2	90.50
ADL	291.3	0.07	71.50
LC-KSVD	1815.3	0.2	50.25
PF	232.9	0.9	72
CF	2.96	0.4	46.75
DCT	4.15	0.76	78.75
PCA	3.1	0.47	56.5

CF, and DCT) as well as training the SVM. The training time for the DL algorithms is considerably higher than the non-DL methods since they involve learning class-specific dictionaries. In particular, the LC-KSVD takes the longest training time since it requires a l_0 -norm computation using a greedy approach. The SDL, DDL, and ADL algorithms use the faster l_1 -minimization operation instead. The table shows that the DDL algorithm takes longer than the SDL during the training phase since generating class-specific dictionaries includes learning at multiple layers. The ADL and PF take equivalent training time while CF, DCT, and PCA have the lowest training times as these use fixed dictionaries and do not involve any inverse operation. The training time, though an important consideration for practical deployment, does not factor into actual radar operation which depends only on the test time.

The test time includes the time taken to generate the test features and the time used by the SVM to classify these features in all the algorithms. The second step is mostly identical across all the algorithms and is approximately 0.01 seconds. It is observed that the ADL requires the lowest computational time. It

Table 4.12: Comparison of theoretical computational cost

Algorithm	Training Cost	Test Cost (Feature Extraction)
SDL	$t\mathcal{O}(mn^2 + t_1(mk^2 + k^3))$	$t_2\mathcal{O}(n^3)$
DDL	$L(t\mathcal{O}(mn^2 + t_1(mk^2 + k^3)))$	$t_2\mathcal{O}(n^3)$ (reduced feature space)
ADL	$\mathcal{O}(mn^2) + \mathcal{O}(n)$	$\mathcal{O}(n)$
PF	$\mathcal{O}(Tn^2\log(n))$	$\mathcal{O}(Tn\log(n))$
CF	$\mathcal{O}(n^2\log(n))$	$\mathcal{O}(n\log(n))$
DCT	$\mathcal{O}(n^3)$	$\mathcal{O}(n^2)$
PCA	$\mathcal{O}(p^2n + p^3)$	$\mathcal{O}(p^3)$

is because the test feature extraction in ADL uses a single matrix multiplication operation which is computationally much more straightforward than the matrix inversion operation in SDL and DDL (4.8). The DDL takes less time than SDL because the inversion operation in the case of DDL involves a matrix of reduced dimensionality due to the multi-layer dictionary synthesis. The baseline algorithms such as LC-KSVD, PF, CF, DCT, and PCA are also computationally inexpensive. The DDL is therefore comparable to the non-DL algorithms.

Next, I present a Table 4.12 showing a comparison of theoretical computational cost involved in training and testing each of the algorithms. The training and the test cost includes only the cost required to learn the training features matrices and test features respectively. Note that the ADL requires only a single linear operation $\mathcal{O}(n)$, during the test phase and thus has the lowest computational time unlike the other DL methods which involves inverse operations having complexity of the order of $\mathcal{O}(n^3)$. LC-KSVD uses l_0 -norm rather than l_1 minimization techniques. This greedy selection of the coefficients is substantially more expensive, growing in order of $\mathcal{O}(k^3)$. Note that this imposes a severe limitation when the number of non-zeros per atom k is significantly increased. The feature extraction using PCA involves two steps- co-variance

matrix computation ($\mathcal{O}(p^2n)$) and its eigen-value decomposition $\mathcal{O}(p^3)$, where n is the training samples, each represented with p features. So, the total time complexity of PCA is $\mathcal{O}(p^2n + p^3)$. The computational cost attached with CF, DCT are- $\mathcal{O}(n \log(n))$, $\mathcal{O}(n^2)$ per training sample respectively. In the case of PF, the features are extracted from micro-Doppler spectrograms whose generation consumes most of the training time. The computational cost involved in generating each spectrogram is $\mathcal{O}(Tn \log(n))$, where, T is the short time window.

In conclusion, the ADL algorithm offers some exceptional advantages regarding computational time and complexity. However, its performance, in our study, is limited by radar system issues such as the low sampling frequency. The CF, PCA also lend themselves to real-time operation - but they are not very successful when there are considerable variations between test and training data. The SDL and DDL algorithms, on the other hand, successfully learn unique dictionaries from multi-carrier data despite the system challenges. The DDL, in particular, is suited for real-time micro-Doppler classification due to its short testing time and high classification accuracy.

4.4 Summary

I demonstrate three sparse coding-based dictionary learning techniques - SDL, DDL, and ADL - to classify micro-Doppler data from dynamic indoor targets. These algorithms facilitate the representation of radar signals using unique basis

vectors even when the training data are gathered from multiple carrier frequencies. As a result, these class dictionaries successfully classify test data from a different carrier frequency from those used while training. This capability makes these algorithms suitable for re-configurable radar platforms for human tracking under diverse operating situations. The SDL and DDL use unique sparsity patterns while the ADL uses the unique co-sparsity patterns of the representations of the radar signals for classification. The computational complexity of the ADL, in the test phase, is much lower than the SDL and DDL. However, in our study, the ADL's performance was limited by radar system issues. As a result of using multi-layered dictionary learning in DDL, the algorithm has better classification accuracy and a much shorter computation time during testing in comparison to the single layer SDL thus making it an ideal candidate for real-world scenarios where low computational complexity and high accuracy are desired.

Chapter 5

Mitigation of Through-Wall Distortions in Frontal Radar Images

In the previous two chapters, the usefulness of representing human micro-Dopplers with customized data-dependent dictionaries is investigated for mainly two applications - *detection* [46, 47], and *classification* [68, 69], of human activities on the basis of their micro-Doppler returns. In this chapter, I discuss micro-Doppler based *through-the-wall radar imaging* (TWRI) for monitoring human activities in urban environments. There are varied applications for TWRI such as law enforcement, security, and surveillance, search and rescue, and indoor monitoring of the elderly [45, 57, 65, 87, 130, 131].

There are broadly two categories of through-the-wall radars: narrowband and broadband. Broadband radars provide excellent downrange resolutions to locate and resolve multiple targets as well as for estimating building layouts [132]. Alternatively, narrowband CW radars have been developed to detect dynamic targets based on their Doppler signatures [45, 57, 133]. Both of

these systems can be complemented with two-dimensional array processing to provide either range-enhanced frontal images or Doppler-enhanced frontal images [77, 78]. Frontal images of the humans provide informative signatures of their activities [76]. However, when the radars are deployed in through-wall scenarios, the quality of the radar images significantly deteriorate due to the through-wall propagation artifacts such as - attenuation, defocussing and multipath clutter [45, 77, 81, 134, 135].

Indoor clutter can be broadly categorized into *target independent static and dynamic clutter*, and *target dependent clutter*. *Target independent static clutter* arise from the reflections off the wall (especially the front face in a through-wall scenario), ceiling, floor, and furniture. Static clutter is easy to eliminate through filtering when the objective is to detect dynamic targets. The problem becomes more challenging in the context of detection of static and slow-moving targets. Authors in [136] assumed the availability of background data that could be coherently subtracted from the target measurements. Alternatively, sparsity-based multipath exploitation methods were explored in [137–139]. Here, the algorithm leveraged the orthogonality between the static clutter and the target scattering to mitigate the clutter. *Target independent dynamic clutter* arising from other moving objects in the environment can significantly interfere with Doppler signatures of moving targets. In Chapter 2, a method to segregate the Doppler returns from multiple targets is presented. This technique could be used for mitigating target-independent dynamic clutter. The third category is *target dependent clutter* that arises from the interactions of the target (static or

dynamic) and the complex propagation channel. As a result of refraction and multipath, the radar images are smeared, blurred, and there are shifts in the location of point scatterers in the images [77]. The authors in [140] and [141] used back-projection and sparsity-based change detection algorithms, respectively, to track slow-moving humans in the range-crossrange space in the presence of target-dependent clutter. Both these techniques, however, rely on the availability of accurate knowledge of the through-wall scenarios for detecting static targets. Alternatively, the authors in [79, 142] exploited the multipath (instead of suppressing the multipath) to improve the effective signal-to-clutter ratio (SCR) at the original target locations. They removed ghost artifacts by mapping the multipath ghosts to their true targets. Again the technique requires exact information of the room geometry and wall characteristics.

In this chapter, an alternative strategy, based on denoising autoencoders, is presented for recovering radar images corrupted by through-wall effects. An autoencoder is a neural network that extracts relevant features from the noisy input data for various tasks such as- dimensionality reduction and data denoising [143, 144]. Autoencoders have been widely used for applications such as anomaly detection, natural language processing, denoising and domain adaptation [143, 145–147]. Some preliminary results for clutter mitigation using autoencoders were presented in [82] where the nature of the type of through-wall scenario was assumed to be known during the test phase. The primary advantage of this technique is, however, that the autoencoders require neither prior information regarding the wall characteristics nor any kind of analytic frame-

work to describe the through-wall interference. Instead, the distorted radar signatures due to wall interference are treated as corrupt versions of ideal radar signatures obtained in free space conditions. The algorithm *learns* how to denoise or clean the corrupted signals using training data comprising of both corrupted and clean data. It is demonstrated, in this dissertation, that the autoencoder can be used for removing signal-dependent clutter when no information or label of the through-wall scenario is assumed to be known during the test phase. Instead, the autoencoder is trained with a mixture of images gathered in diverse through-wall conditions. Traditional autoencoders have been implemented using back-propagation algorithms such as- gradient descent [148], conjugate gradient descent [149] and steepest descent [144]. However, they have a very slow learning rate. This translates to long training times and, in some cases, the convergence may not be guaranteed. Instead, an alternating direction method of multipliers (ADMM) approach [150], is used where the complex convex optimization problem is broken into smaller sub-problems with closed-form solutions. Thus the convergence is guaranteed and training times are not very long.

The performance of the presented algorithm is tested on two types of radar images - Doppler enhanced and range enhanced frontal images. The *Doppler-enhanced frontal* images of *dynamic human motions* are generated from simulated narrowband radar data of human motions in through-wall environments using the techniques described in [77, 78]. I consider a variety of walls- a dielectric wall, a dielectric wall with metal reinforcements and one with air-gaps. The through-wall propagation phenomenology is modeled using finite-

difference time-domain (FDTD) techniques [151]. I introduce significant diversity in wall parameters such as dielectric constant and conductivity by incorporating stochasticity in the finite difference equations as suggested by [92]. This is a computationally, more efficient technique than running multiple FDTD simulations with varying wall parameters. The second set of images are *range-enhanced frontal images* captured of *static humans* using measurement data gathered with Walabot, a three dimensional programmable, wideband imaging radar [105]. During the training phase, the autoencoder is trained with a diverse mixture of data gathered from different through-wall scenarios. In the test phase, the network denoises the corrupted radar image without requiring any information of the type of wall or its parameters. Both the simulation and measurement results obtained from the autoencoder exhibit very low normalized mean square error and high structural similarity between the denoised reconstructed images and free space images.

5.1 Theory

Radar images deteriorate significantly due to distortions and multipath clutter signals introduced by through-wall environments. The images may be defocused, blurred, or smeared. Ghost targets may appear due to multipath. The objective, here, is to reconstruct clean radar images resembling free space images from corrupted through-wall images. I divide our denoising problem into two stages- training and the test stages.

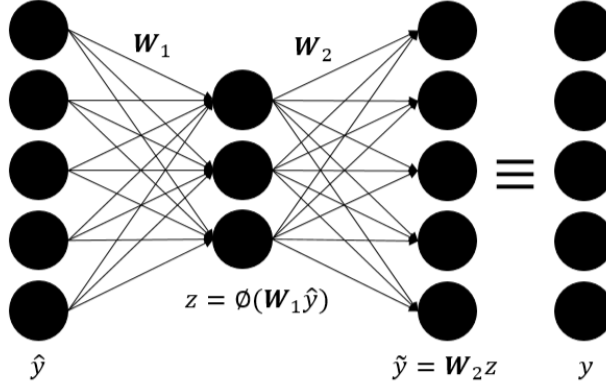


Figure 5.1: Denoising Autoencoder

5.1.1 Training Stage

A conventional denoising autoencoder shown in Figure 5.1, first corrupts the clean input data by adding stochastic Gaussian noise, then feeds the corresponding noisy version as input data to the next stage. In this work, the measurements in a through-wall are considered case as our noisy/corrupted data. The main difference is the non-Gaussian nature of interference. During the training stage, M radar images of the target are captured in free space. Each image of size $\mathfrak{R}^{I \times J}$ is vectorised to obtain $y^{tr} \in \mathfrak{R}^{N \times 1}$ where $N = I \times J$, is the total number of pixels in an image. M images are stacked as columns to form a data matrix $\mathbf{Y}^{tr} \in \mathfrak{R}^{N \times M}$. This exercise is repeated for the images captured in the corresponding through-wall scenarios to generate $\hat{\mathbf{Y}}^{tr}$ also of size $\mathfrak{R}^{N \times M}$. The autoencoder has primarily two stages- encoding and the decoding. In the encoding stage, the algorithm learns a latent/compressed representation $\mathbf{Z} \in \mathfrak{R}^{r \times M}$, of the input layer $\hat{\mathbf{Y}}^{tr}$ as shown in

$$\mathbf{Z} = \phi(\mathbf{W}_1 \hat{\mathbf{Y}}^{tr}). \quad (5.1)$$

Here, ϕ is the mapping function which can either be linear or nonlinear (such as sigmoid, tanh), $\mathbf{W}_1 \in \Re^{R \times N}$ is the corresponding weight matrix and R is the number of nodes in the hidden layer. Since the hidden layer, $\mathbf{Z}$, is the compressed representation of the input layer, $\hat{\mathbf{Y}}^{tr}$, it always has fewer nodes than the number of pixels ($R \ll N$). In the decoding stage, the algorithm maps $\mathbf{Z}$ back to obtain a reconstructed signal $\tilde{\mathbf{Y}}^{tr} = \mathbf{W}_2 \phi(\mathbf{W}_1 \hat{\mathbf{Y}})$ through weight matrix $\mathbf{W}_2 \in \Re^{N \times R}$ such that the error e ,

$$e = \left\| \mathbf{Y}^{tr} - \tilde{\mathbf{Y}}^{tr} \right\|_2^2, \quad (5.2)$$

between the reconstructed images and the free space images is minimized. Therefore the objective in the training stage is to learn weight matrices $\mathbf{W}_1$ and $\mathbf{W}_2$ so that the reconstructed images resemble free space images (instead of the corrupted through-wall images). The objective function,

$$J(\mathbf{W}_1, \mathbf{W}_2) = \min_{\mathbf{W}_1, \mathbf{W}_2} \left\| \mathbf{Y}^{tr} - \mathbf{W}_2 \phi(\mathbf{W}_1 \hat{\mathbf{Y}}^{tr}) \right\|_2^2, \quad (5.3)$$

can be solved in multiple ways - gradient descent, conjugate gradient descent, or steepest descent. In some of these ways, the error may become insignificant when back propagated. Additionally, these algorithms have a very slow learning rate. Instead, an alternating direction method of multipliers (ADMM) approach [150] is presented. Here, a simple variable separation technique is used to break the complex convex optimization problem into smaller sub-problems which have closed form solutions such that the convergence is guaranteed. The

objective function in (5.3) is reformulated to

$$J(\mathbf{W}_1, \mathbf{W}_2) = \min_{\mathbf{W}_1, \mathbf{W}_2} \|\mathbf{Y}^{tr} - \mathbf{W}_2 \mathbf{Z}\|_2^2 \quad s.t. \quad \mathbf{Z} = \phi(\mathbf{W}_1 \hat{\mathbf{Y}}^{tr}). \quad (5.4)$$

Since the formulation in (5.4) is a constrained optimization problem, it is relaxed using an augmented Lagrangian technique shown below

$$J(\mathbf{W}_1, \mathbf{W}_2, \mathbf{Z}) = \min_{\mathbf{W}_1, \mathbf{W}_2, \mathbf{Z}} \|\mathbf{Y}^{tr} - \mathbf{W}_2 \mathbf{Z}\|_2^2 + \mu^T (\mathbf{Z} - \phi(\mathbf{W}_1 \hat{\mathbf{Y}}^{tr})) + \lambda \|\mathbf{Z} - \phi(\mathbf{W}_1 \hat{\mathbf{Y}}^{tr})\|_2^2. \quad (5.5)$$

The Lagrangian term $\mu^T (\mathbf{Z} - \phi(\mathbf{W}_1 \hat{\mathbf{Y}}^{tr}))$ imposes equality at every iteration; this is too stringent a requirement in practice. The equality constraint is initially relaxed and enforced only during convergence. This assumption has been usually used in other research works such as [152, 153]. Therefore, the objective function is reformulated as shown in

$$J(\mathbf{W}_1, \mathbf{W}_2, \mathbf{Z}) = \min_{\mathbf{W}_1, \mathbf{W}_2, \mathbf{Z}} \|\mathbf{Y}^{tr} - \mathbf{W}_2 \mathbf{Z}\|_2^2 + \lambda \|\mathbf{Z} - \phi(\mathbf{W}_1 \hat{\mathbf{Y}}^{tr})\|_2^2. \quad (5.6)$$

Here, λ is the regularization parameter between the proxy variable $\mathbf{Z}$ and underlying representation $\phi(\mathbf{W}_1 \hat{\mathbf{Y}}^{tr})$. (5.6) is divided into a set of smaller sub problems as follows.

Problem1:

$$J(\mathbf{W}_1) = \min_{\mathbf{W}_1} \lambda \|\phi^{-1} \mathbf{Z} - \mathbf{W}_1 \hat{\mathbf{Y}}^{tr}\|_2^2. \quad (5.7)$$

Problem2:

$$J(\mathbf{W}_2) = \min_{\mathbf{W}_2} \|\mathbf{Y}^{tr} - \mathbf{W}_2 \mathbf{Z}\|_2^2. \quad (5.8)$$

Problem3:

$$\begin{aligned}
J(\mathbf{Z}) &= \min_{\mathbf{Z}} \|\mathbf{Y}^{tr} - \mathbf{W}_2 \mathbf{Z}\|_2^2 + \lambda \left\| \mathbf{Z} - \phi(\mathbf{W}_1 \hat{\mathbf{Y}}^{tr}) \right\|_2^2 \\
&= \min_{\mathbf{Z}} \left\| \begin{pmatrix} \mathbf{Y}^{tr} \\ \sqrt{\lambda} \phi(\mathbf{W}_1 \hat{\mathbf{Y}}^{tr}) \end{pmatrix} - \begin{pmatrix} \mathbf{W}_2 \\ \sqrt{\lambda} \mathbf{I} \end{pmatrix} \mathbf{Z} \right\|_2^2.
\end{aligned} \tag{5.9}$$

Sub problems in (5.7)-(5.9) are all simple least squares problems which already have a closed form solution [109]. At each iteration, the network weight $\mathbf{W}_1$, $\mathbf{W}_2$ and proxy variable $\mathbf{Z}$, are updated till the algorithm converges.

5.1.2 Test Stage

During test stage, P corrupted radar images are combined together to form $\hat{\mathbf{Y}}^{test} \in \Re^{N \times P}$ and passed through the autoencoder to obtain $\tilde{\mathbf{Y}} \in \Re^{N \times P}$. It is hypothesized that once the network is trained, the weight matrices $\mathbf{W}_1$ and $\mathbf{W}_2$ can be used to obtain a denoised form $\tilde{\mathbf{Y}}^{test}$ of the corrupted test data $\hat{\mathbf{Y}}^{test}$ as shown in Figure 5.2(b).

$$\tilde{\mathbf{Y}}^{test} = \mathbf{W}_2 \phi(\mathbf{W}_1 \hat{\mathbf{Y}}^{test}). \tag{5.10}$$

Note that the presented denoising algorithm is significantly faster in generating denoised images at test time as it involves only a simple product operation in (5.10). This makes the algorithm suitable for real-time applications.

5.1.3 Computational Complexity

5.1.3.1 Training Stage

During the training stage of denoising autoencoder network, we solve three sub-problems Problem1-Problem3. All these are simple least square problems. Theoretically they have a closed form solution in the form of $(X'X)^{-1}X'Y$. If X is an $(n \times m)$ matrix, one can solve it using the following steps-

1. Matrix multiplication $X'X$, complexity: $\mathcal{O}(n \times (m^2))$
2. Matrix inversion $A = (X'X)^{-1}$, complexity: $\mathcal{O}(m^3)$
3. Matrix multiplication $B = X'Y$, complexity: $\mathcal{O}(n \times (m^2))$
4. Matrix multiplication AB , complexity: $\mathcal{O}(m^3)$

If the network is trained for k number of iterations, the total theoretical computational cost becomes- $k(\mathcal{O}(n \times (m^2)) + \mathcal{O}(m^3) + \mathcal{O}(n \times (m^2)) + \mathcal{O}(m^3)) = k(\mathcal{O}(n \times (m^2)) + \mathcal{O}(m^3))$.

5.1.3.2 Test Phase

The presented denoising algorithm at test time involves only a simple product operation. Therefore, the computational complexity of test time becomes $\mathcal{O}(nm)$.

5.1.4 Metrics for evaluation

The effectiveness of the clutter mitigation algorithm is evaluated using two metrics- normalized mean square error (NMSE) and structural similarity index (SSIM). The image captured in free space are considered as the clean/ground truth image ($\mathbf{Y}^{test}$). The NMSE and SSIM is calculated between the through-wall image $\hat{\mathbf{Y}}^{test}$ and ground truth image before denoising (BD). Then the exercise is repeated after denoising (AD). In the second case, the NMSE and SSIM are calculated between the reconstructed/denoised image, $\tilde{\mathbf{Y}}^{test}$, and the ground truth image. The hypothesis, here, is that the NMSE and SSIM will improve after denoising.

The NMSE is computed between $\mathbf{Y}^{test}$ and $\hat{\mathbf{Y}}^{test}$ using

$$NMSE = \frac{\left\| \mathbf{Y}^{test} - \hat{\mathbf{Y}}^{test} \right\|_2^2}{\left\| \mathbf{Y}^{test} \right\|_2^2}. \quad (5.11)$$

NMSE is sensitive to the energy of absolute errors of all the pixels of an image. However, NMSE between two images may be low even if they have drastically different structural features [154]. SSIM [155], on the other hand, is a metric that provides information of the luminance (L), contrast (C) and structure difference (S), between the ground truth image $\mathbf{Y}^{test}$, and the test image $\hat{\mathbf{Y}}^{test}$. Its value should be 1 if the images are identical. The overall measurement metric becomes the multiplicative combination of three measures shown in

$$SSIM(\hat{\mathbf{Y}}, \mathbf{Y}) = [L(\hat{\mathbf{Y}}, \mathbf{Y})]^\alpha [C(\hat{\mathbf{Y}}, \mathbf{Y})]^\beta [S(\hat{\mathbf{Y}}, \mathbf{Y})]^\gamma. \quad (5.12)$$

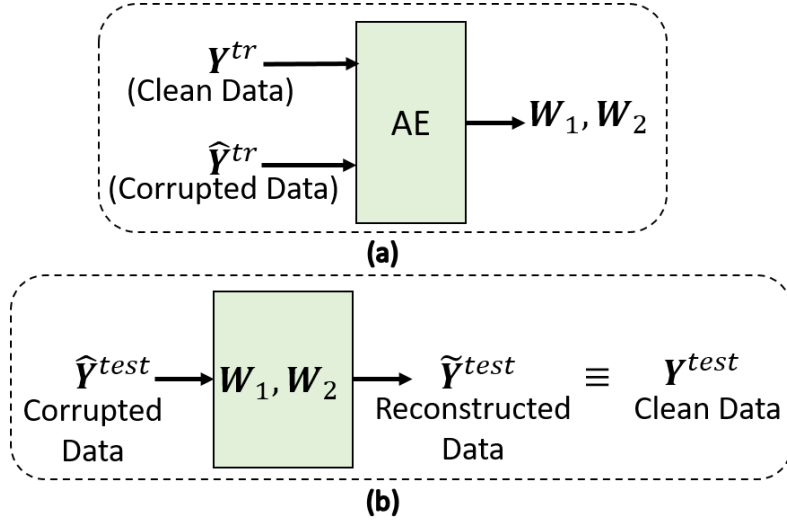


Figure 5.2: (a) During training stage, the autoencoder learns coefficients $\mathbf{W}_1$ and $\mathbf{W}_2$ from clean free space images ($\mathbf{Y}^{tr}$) and corrupted through-wall ($\hat{\mathbf{Y}}^{tr}$) images. (b) During test phase, corrupted through-wall images ($\hat{\mathbf{Y}}^{test}$) are denoised to obtain reconstructed images that resemble free space radar images ($\mathbf{Y}^{test}$).

I assume $\alpha = \beta = \gamma = 1$. The expressions for L, C, S are

$$L(\hat{\mathbf{Y}}, \mathbf{Y}) = \frac{2\mu_{\hat{\mathbf{Y}}}\mu_{\mathbf{Y}} + C_1}{\mu_{\hat{\mathbf{Y}}}^2 + \mu_{\mathbf{Y}}^2 + C_1}. \quad (5.13)$$

$$C(\hat{\mathbf{Y}}, \mathbf{Y}) = \frac{2\sigma_{\hat{\mathbf{Y}}}\sigma_{\mathbf{Y}} + C_2}{\sigma_{\hat{\mathbf{Y}}}^2 + \sigma_{\mathbf{Y}}^2 + C_2}. \quad (5.14)$$

$$S(\hat{\mathbf{Y}}, \mathbf{Y}) = \frac{\sigma_{\hat{\mathbf{Y}}\mathbf{Y}} + C_3}{\sigma_{\hat{\mathbf{Y}}}\sigma_{\mathbf{Y}} + C_3}. \quad (5.15)$$

Here, $\mu_{\mathbf{Y}}, \mu_{\hat{\mathbf{Y}}}, \sigma_{\mathbf{Y}}, \sigma_{\hat{\mathbf{Y}}}$ and $\sigma_{\hat{\mathbf{Y}}\mathbf{Y}}$ are the local means, standard deviations and the co-variance for the reference $\mathbf{Y}$ and test images $\hat{\mathbf{Y}}$ respectively. Assuming $C_3 = \frac{C_2}{2}$, the simplified index becomes

$$SSIM(\hat{\mathbf{Y}}, \mathbf{Y}) = \frac{(2\mu_{\hat{\mathbf{Y}}}\mu_{\mathbf{Y}} + C_1)(2\sigma_{\hat{\mathbf{Y}}\mathbf{Y}} + C_2)}{(\mu_{\hat{\mathbf{Y}}}^2 + \mu_{\mathbf{Y}}^2 + C_1)(\sigma_{\hat{\mathbf{Y}}}^2 + \sigma_{\mathbf{Y}}^2 + C_2)}. \quad (5.16)$$

I therefore conclude that two images can be regarded as similar only when both

NMSE is low and SSIM is close to 1.

5.2 Simulation Method

In this section I describe the simulation method to generate a large database of Doppler enhanced frontal images of humans in diverse through-wall conditions. I adopt the technique described in [77] and briefly describe its salient features in the following subsection. I only model the through-wall propagation phenomenology and do not consider multipath scattering from the ceiling, ground and lateral walls. The wall propagation phenomenology, modelled using finite difference time domain techniques (FDTD), and primitive based models of humans are hybridized to generate Doppler-enhanced frontal radar images. There may be considerable variations in the propagation conditions during training and test due to variations in the wall characteristics such as its dielectric constant and loss tangent. Modeling this diversity with independent FDTD simulations is computationally expensive. Therefore, I extend the simulation framework discussed in [87] by incorporating stochasticity in the propagation channel using the stochastic FDTD (sFDTD) technique suggested by [92]. The sFDTD method introduces statistical variations in the electrical properties of the medium. The results of the simulations provide the mean and the variance estimates of the time-domain electromagnetic fields at every point in the problem space from which numerous samples of the through-wall propagation can be generated. I describe these steps in greater detail in the following sections.

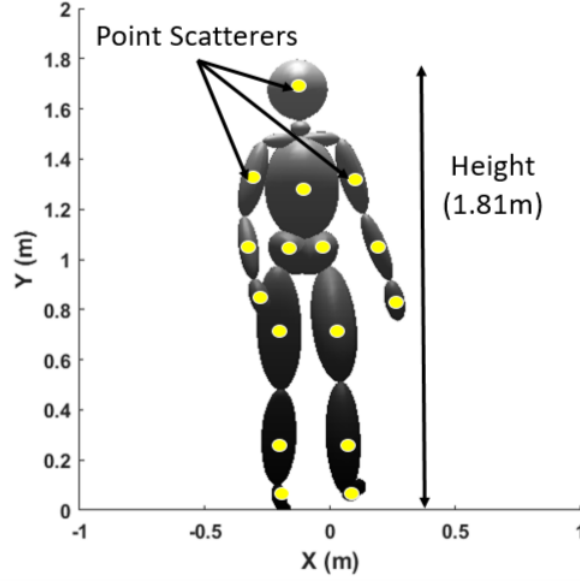


Figure 5.3: Radar scattering model of human.

5.2.1 Radar Signal Model

I consider a radar with a single CW transmitter of frequency f_c (wavelength λ_c). The receiver consists of a $P \times Q$ uniform planar array along the $X - Y$ plane with element spacing $\lambda_c/2$. I use motion capture data for realistically modeling complex human animation. I interpolate the data from the video frame rate to the radar sampling frequency. The data provides time-varying three-dimensional positions of B markers distributed over the entire human body - head, torso, arms, shoulders, and legs as shown in Figure 5.3. I assume that these markers correspond to point scatterers each of reflectivity a_b and located at spherical coordinate positions (r_b, θ_b, ϕ_b) where θ_b is the elevation from the XZ plane and ϕ_b is the azimuth with respect to the positive Z axis on the human body. Each body part moves with a Doppler frequency f_{D_b} . The baseband,

digitized radar data at each $(p, q)^{th}$ element is given by

$$\mathbf{X}_{p,q,n} = \sum_{b=1}^B a_b[n] e^{-j \frac{2\pi}{\lambda_c} \|\vec{r}_b[n] - \vec{r}_{p,q}\|_2^2}. \quad (5.17)$$

where n denotes the time sample and $\vec{r}_{p,q}$ is the position of the radar element. The data are processed with discrete short time Fourier transform across the time-domain using a sliding time window centered around τ and 2D Fourier beamforming across the array to obtain a time varying three-dimensional radar cube in terms of azimuth, elevation and Doppler, $\chi_{\phi,\theta,f_D}^\tau$, as shown in

$$\chi_{\phi,\theta,f_D}^\tau = F_{3D} \mathbf{X}_{p,q,n} = \sum_{b=1}^B a_b H_{3D} [\phi - \phi_b(\tau), \theta - \theta_b(\tau), f_D - f_{D_b}(\tau)]. \quad (5.18)$$

Here, $H_{3D}[\cdot]$ is the three-dimensional point spread function across the three domains and F_{3D} indicates the three-dimensional Fourier operator. From the radar cube, peak scatterers of strength $a_m^\tau[f_D]$ located at (ϕ_m, θ_m) at every Doppler bin f_D are identified. Doppler enhanced frontal images $\chi_{\phi,\theta}^\tau$ are generated for every τ by convolving the peak scatterers with a two-dimensional point spread function, $H_{2D}[\cdot]$ as shown in

$$\chi_{\phi,\theta}^\tau = \sum_{f_D} a_m^\tau(f_D) H_{2D} [\phi - \phi_m, \theta - \theta_m]. \quad (5.19)$$

The incorporation of the additional Doppler dimension enables us to resolve multiple scatterers of the human body along two spatial dimensions thus alleviating the need for larger arrays required for successful imaging. The radar signal model parameters used in this work are provided in Table 5.1.

Table 5.1: Simulated Radar Parameters

Radar Parameters	Values
Radar Type	Narrowband
Carrier frequency (f_c)	$7.5GHz$
Sampling frequency (f_s)	$1000Hz$
Integration time (T)	$0.8s$
Dwell time or short time (t_D)	$0.1s$
Maximum Doppler (f_{Dmax})	$\pm 500Hz$
Doppler resolution (Δf_D)	$10Hz$
Number of antenna elements ($P \times Q$)	10×10
Azimuth Beamwidth ($\Delta \phi_{azi}$)	10°
Azimuth Beamwidth ($\Delta \theta_{ele}$)	10°
Field of View (ϕ_{azi})	-90° to 90°
Field of View (θ_{azi})	-90° to 90°

5.2.2 Stochastic Model of Through-Wall Propagation

The formulation discussed above describes the radar images generated in free space conditions. In this section, I use through-wall propagation phenomenology described in Chapter 2, Section 2.3.1.2 to simulate human radar returns. As shown in Figure 5.4, I consider a 2D simulation space extending from -1m to 1m and 0m to 4m along the X and Z directions respectively (assuming the wall is invariant along Y height axis). The 2D simulation framework is chosen to reduce the computational complexity of the problem and because most walls show homogeneity along the height. In order to correspond to the radar signal model discussed earlier, I consider a uniform 10 element linear antenna array whose elements are spaced half wavelength apart. I independently simulate the excitation from each element of the array, located at $\vec{\rho}_p$, with a narrowband sinusoidal signal at f_c . The simulation space is bounded by a perfectly matched layer and divided into spatial grids of $\lambda_c/10$ size. I considered three different wall configurations - a homogeneous dielectric wall (Figure 5.4(a)), a wall rein-

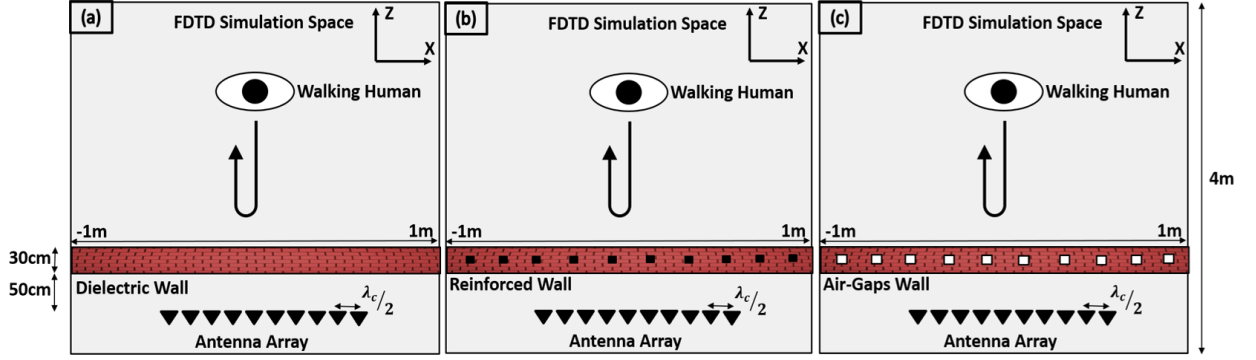


Figure 5.4: Room geometry in through-wall scenario (a) Dielectric (b) Reinforced wall and (c) Wall with air-gaps

forced with metal rods (Figure 5.4(b)), and a wall with air gaps (Figure 5.4(c)).

Each wall type is simulated independently. The dimensions of each wall is 2m x 30cm (X : -1m to 1m, Z : 1m to 1.3m). In all three cases, stochastic variations of 10% standard deviation are introduced in the relative permittivity

$\epsilon_r = 4$, and conductivity $\sigma_c = 0.001S/m$, of each grid point in the wall. There-

fore, even the single layer dielectric wall is not truly homogeneous. This is

done to mirror real world conditions. For every point in space, $\vec{\rho}_b$, and at every

time instant n , the sFDTD simulation gives the mean time-domain electric

field $\mu_E[\vec{\rho}_b, \vec{\rho}_p, n]$ and its standard deviation $\sigma_E[\vec{\rho}_b, \vec{\rho}_p, n]$. I use the Gaussian

stochastic model to generate 200 samples ($\eta = 1 \cdots 200$) of time-domain elec-

tric field values $E[\vec{\rho}_b, \vec{\rho}_p, n, \eta] \sim \mathcal{N}(\mu_E, \sigma_E^2)$. The $E[\vec{\rho}_b, \vec{\rho}_p, n, \eta]$ is fast Fourier

transformed to obtain the corresponding frequency domain wall transfer func-

tion $H_{wall}(\vec{\rho}_b, \vec{\rho}_p, \eta)$ at 7.5GHz. The simulations have been carried out entirely

in Matlab.

5.2.3 Modeling of electromagnetic radar scatter from dynamic humans in through-wall scenarios

I integrate the free space radar signal model with the through-wall propagation based on [45, 87]. I considered a scenario where the human is moving behind a wall before a radar. Since our FDTD simulation spatial extent is limited, the translational motion of the human is removed retaining only the dynamics of the swinging arms and legs. Therefore, the time-domain scattered returns at each (p, q) antenna element is obtained by hybridizing H_{wall} with the human scattering center model as shown in

$$\mathbf{X}_{p,q,n}[\eta] = \sum_{b=1}^B a_b \gamma_{2D \rightarrow 3D}(p, q, b) (H_{wall}[\vec{\rho}_b[n], \vec{\rho}_p, \eta])^2. \quad (5.20)$$

Each point scatterer on the human is projected from its three-dimensional position $\vec{r}_b$ to its two-dimensional counterpart $\vec{\rho}_b$ on the $X - Z$ plane. The term $H_{wall}[\vec{\rho}_b, \vec{\rho}_q, \eta]$ models the propagation from the source $\vec{\rho}_p$ to $\vec{\rho}_b$ and is generated from a stochastic realization (η) of the full wave electromagnetic solver. The scaling factor $\gamma_{2D \rightarrow 3D}$, given by (5.21), adjusts the phase for modeling three-dimensional physics from the two-dimensional simulation as described in [45].

$$\gamma_{2D \rightarrow 3D}(p, q, b) = e^{\frac{-j2\pi}{\lambda} [\|\vec{r}_b[n] - \vec{r}_{p,q}\|^2 - \|\vec{\rho}_b[n] - \vec{\rho}_p\|^2]}. \quad (5.21)$$

5.3 Simulation Results

First a single stride of a human walking motion from Sony Computer Entertainment America is considered. The duration of the motion is 0.8s which corresponds to one complete human stride. The human is walking away from the radar at an aspect angle of 180° . The Doppler-enhanced frontal images are generated for free space and the three wall types.

5.3.1 Doppler enhanced frontal imaging of dynamic humans in through-wall scenarios

The Doppler enhanced frontal images account for all the electromagnetic phenomenology introduced by the through-wall propagation conditions including attenuation, ringing and multipath.

To demonstrate the validity of our claim, I show the magnitude response for the four scenarios at a carrier frequency of 7.5 GHz and the corresponding simulated Doppler enhanced frontal image of a single frame of a walking human, at a radar-target aspect angle of 180° , in Figure 5.5. Some comments regarding the wall-target interaction phenomenology on the frontal images:

Free Space: Figure 5.5(a) shows that the magnitude response of the electric field decays as the distance from the source increases. The phase response in Figure 5.5(e) displays the circular wave-fronts emanating from an infinitely long line source excitation. Due to the undistorted phase response, we get the highest quality frontal images in the free space scenario. We can clearly discern both arms, legs and head of the human in the corresponding Doppler enhanced

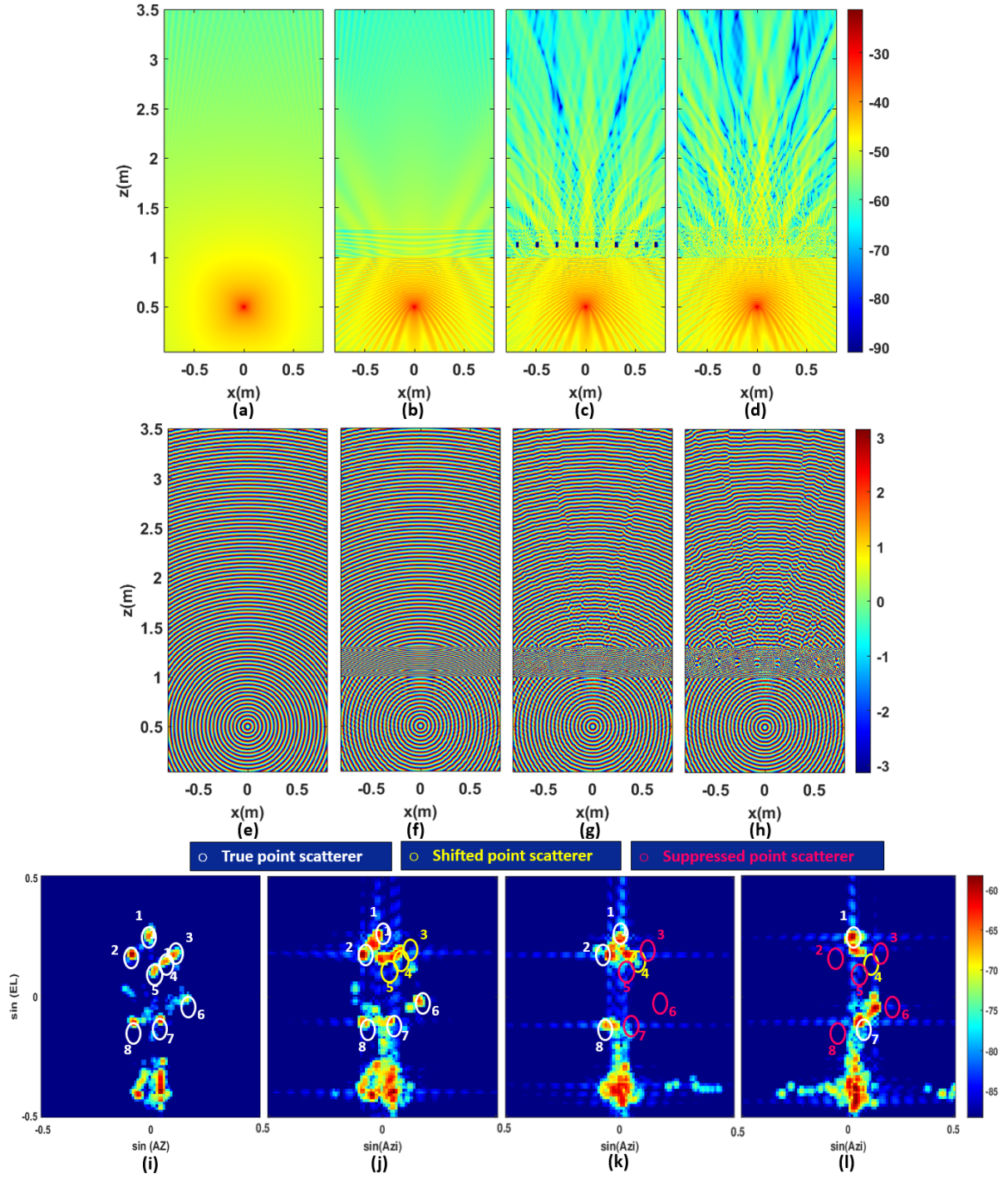


Figure 5.5: (a)-(b) Magnitude response, (e)-(h) phase response at 7.5 GHz in freespace, through dielectric wall, reinforced dielectric wall, and dielectric wall with air-gaps respectively. (i)-(l) Corresponding simulated Doppler enhanced frontal image of a walking human.

frontal image shown in Figure 5.5(i). The image falls within the $\pm 30^\circ$ field-of-view of the radar along elevation and azimuth. There is some smearing near the legs due to the limited resolution along azimuth and elevation of the array.

Dielectric Wall: The propagation of a signal through a homogeneous dielectric wall undergoes a two-way attenuation of approximately 12dB when compared to free space Figure 5.5(b). Hence, the strength of some of the peak scatterers in the frontal image Figure 5.5(f), become too weak to be visible on the same dynamic scale as that of free space case. This image clearly demonstrates the effect of through-wall attenuation on the images.

Reinforced Wall: Figure 5.5(c) and (g) shows that the inhomogeneity inside the wall causes multiple scattering that interferes destructively in some regions beyond the wall. As a result, the radar frontal images are significantly distorted. Additionally, positions of few scatterers get displaced along the azimuth direction due to refraction. Some of the point scatterers are not visible at all because these lie at regions of destructive interference shown in Figure 5.5(k).

Wall with Air Gaps: Finally, we see that the wall response is most severe in the case of the wall with air-gaps as visible in Figure 5.5(d) and (h). The phase response in Figure 5.5(h), shows that the multipath can interfere constructively or destructively in some regions leading to significant distortions. Therefore, the most significantly distorted frontal images are from this case, as can be seen in Figure 5.5(l).

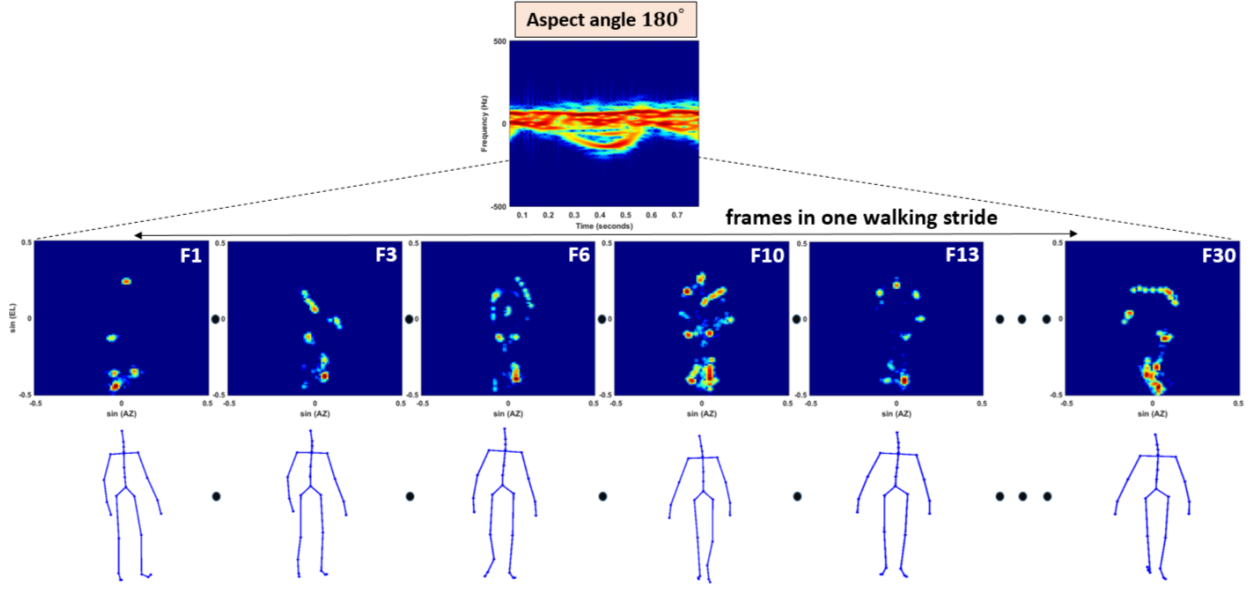


Figure 5.6: Simulated Doppler enhanced frontal images over one walking stride of human motion, and the corresponding stick figures obtained using ground truth motion capture data.

5.3.2 Results from denoising cluttered images

I selected 30 frames of a human subject walking away from the radar spanning a duration of 0.8s, as shown in Figure 5.6. The training data consists of images corresponding to different motion states within a stride, as shown in the stick figure models in Figure 5.6. The size of each image is $[92 \times 92]$. Corresponding to each of these frames, 200 distinct through-wall radar images are generated to capture statistical variations in the wall parameters. Each image is then vectorized to a column vector of size $[8464 \times 1]$, for final processing. Of the total set of 6000 images for each wall type, 80% images are used for training the autoencoder network and the remaining for the test.

The results are obtained by optimizing the number of nodes (R) in the hidden layer and the mapping function connecting the input and the hidden layer.

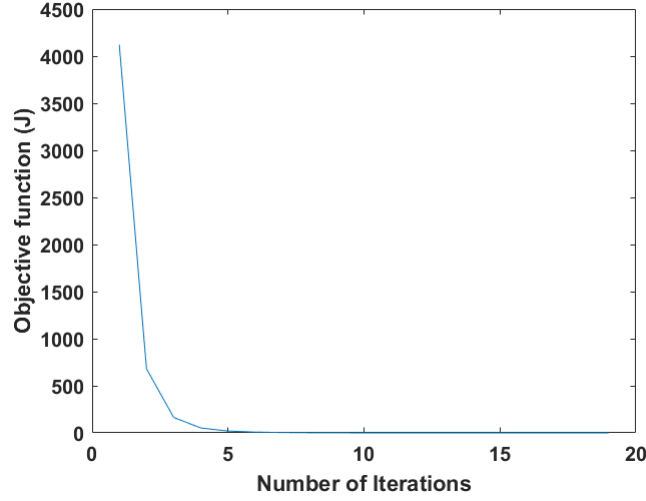


Figure 5.7: Convergence curve of the denoising algorithm.

I fixed the hidden layer dimension of the autoencoder network to be 500 and the mapping to be linear between input and the hidden layer. The choice of these parameters is presented in the discussion section 5.5. During training, the weight matrices $\mathbf{W}_1$ and $\mathbf{W}_2$ each of size $[500 \times 8464]$ and $[8464 \times 500]$ respectively are first randomly initialized. The weights are updated over successive iterations, as discussed in the previous section. Since equal importance is given to both encoding and decoding stage, the regularization parameter (λ) is chosen to be 1. I have shown the convergence of the objective function with the iterations of the learning algorithm. Figure 5.7 confirms that the convergence has reached. Once learned, the weights are used for the test.

Tables 5.2–5.3 show the results for the metrics, SSIM and NMSE, as a function of number of distinct frames of the human walking motion. I compare the metrics obtained from images generated *before denoising (BD)* with those obtained *after denoising (AD)*. To give readers a better understanding of the generality of the solution, I have studied the sensitivity of the performance of the

Table 5.2: Denoising results between clean and corrupted Doppler enhanced frontal images for different through-wall conditions. SSIM: between corrupted and free space image before denoising (BD) and SSIM :between reconstructed and free space image after denoising (AD)

Wall Scenario		Denoising Metric (SSIM)	Number of Frames				
			1	5	10	20	30
Train and test on same wall (Case 1)	Dielectric	BD	0.09	0.55	0.63	0.67	0.64
		AD	0.43	0.87	0.90	0.89	0.86
	Reinforced	BD	0.05	0.24	0.46	0.55	0.47
		AD	0.43	0.81	0.87	0.84	0.80
	Wall With Air-gaps	BD	0.03	0.01	0.34	0.51	0.43
		AD	0.42	0.82	0.83	0.82	0.77
Train and test on different walls (Case 2)		BD	0.05	0.24	0.46	0.55	0.46
		AD	0.41	0.43	0.55	0.08	0.01
Train on multiple walls (Case 3)		BD	0.04	0.26	0.43	0.52	0.45
		AD	0.19	0.81	0.84	0.83	0.78

algorithm to different wall types during training and test scenarios. I considered the following three training scenarios.

Case 1: Train and test on data from same wall:

First, I consider the scenario where the autoencoder is trained with data from a specific wall configuration and then subsequently tested on images generated from the same wall configuration. The training data set size, in this case, is $[8464 \times 4800]$. Note that even in the *same wall* case, there is diversity in the training and test data due to the statistical variations in the wall parameters as well as motion characteristics. Before denoising, the dielectric wall case has the lowest error when compared to the reinforced and air-gaps walls. This is because the quality of the radar images is a function of the phase and amplitude distortions introduced by the walls to the radar signals. Therefore, based on the magnitude and phase responses shown in Figure 5.5(c) and (d), it is observed the results deteriorate most in the case of the wall with air-gaps. The error be-

Table 5.3: Denoising results between clean and corrupted Doppler enhanced frontal images for different through-wall conditions. NMSE: between corrupted and free space image before denoising (BD) and NMSE :between reconstructed and free space image after denoising (AD)

Wall Scenario		Denoising Metric (NMSE)	Number of Frames				
			1	5	10	20	30
Train and test on same wall (Case 1)	Dielectric	BD	3.60	3.50	3.18	3.49	3.39
		AD	1.05	0.85	0.99	0.92	1.56
	Reinforced	BD	5.71	5.91	5.57	5.60	5.01
		AD	1.60	1.09	1.08	1.16	1.36
	Wall With Air-gaps	BD	4.42	4.01	3.50	3.25	2.86
		AD	1.14	1.18	1.36	1.32	1.32
Train on multiple walls (Case 3)		BD	5.01	4.49	4.07	4.13	3.78
		AD	0.71	1.16	1.26	1.60	1.58

tween reconstructed and the free space images drop significantly for all wall types after passing through the denoising network. The number of frames are varied from 1 to 30 to increase the diversity in the human motions. Since this is a continuous motion, there may be some slight correlation between images obtained from consecutive frames. However, this is not very evident from visual inspection of the images corresponding to the frames as seen in Figure 5.6. So, the group correlation index is mapped across the multiple frames as we increase the diversity of the training data in Figure 5.8. The figure shows that the group correlation increases until it reaches a plateau of about 10 frames. Hence, the denoising performance seems to improve when we increase from a single frame to 10 number of frames as the training data captures the diversity of motions in the Tables 5.2–5.3. Beyond this, the performance of the denoising algorithm slightly deteriorates due to the possible decorrelation between test and training data. However, the deterioration in the performance is very slight. The performance indicates that this algorithm is specifically suited for imaging continuous

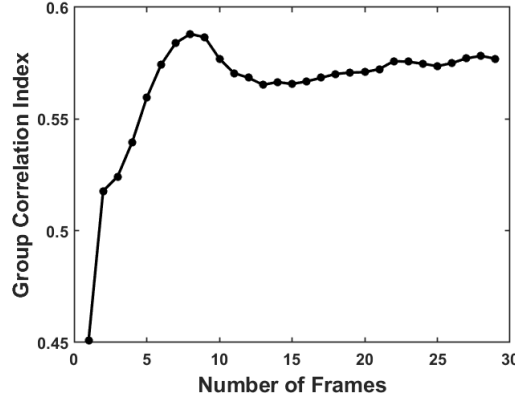


Figure 5.8: Group correlation across multiple frames

and periodic motions such as walking. Also, note that NMSE and SSIM do not behave in an identical manner for all the cases as they indicate different aspects of similarity of images. In practice, this case seems to be limited since in real-world scenarios. We may not know the type of wall available during the test phase. Therefore, I consider a significantly more challenging scenario where there is no information of the type of wall during the test phase.

Case 2: Train using data from a single wall type and test using data from a different wall type:

Next, I analyze the performance of the algorithms when the network is trained on images captured from a dielectric wall and tested with images captured from a reinforced wall type. The resulting denoising performance reported in Table 5.2 deteriorates considerably as the number of frames increases. The reason for the poor performance is because the nature of the clutter in the dielectric case is quite different from that of the reinforced wall case. The algorithm is unable to denoise the clutter due to the lack of diversity in the training data. Henceforth, I do not report results for this case.

Case 3: Train using data from multiple types of walls and test on data from a single wall:

To overcome the limitation of the previous case, I train the network on images captured from all three through-wall scenarios resulting in total training data set size of $[8464 \times 14400]$. Then, the data from any of these walls is randomly chosen for the test. This is a significantly more challenging scenario since no information about the type of wall is available during the test phase. Note that the images used in the test phase have not been used during training. Our algorithm is specifically suited for problems dealing with a great deal of diversity in the target and channel conditions (different wall scenarios). The results show very good performance (NMSE and SSIM) comparable to that of the same wall scenarios. Thus, the performance of the algorithms depends on the diversity of data provided while training.

5.3.3 Impact of radar-target aspect angle

In order to understand the generality of the presented denoising solution, I trained our autoencoder network with human radar images captured at different aspect angles with respect to radar line-of-sight conditions. I analyzed the performance of our algorithm for four aspect angles- 0° , 45° , 90° , 180° . The time-varying Doppler-enhanced frontal images generated in free space for 0° and 90° are shown in Figure 5.9. These can be compared to the frontal images for 180° that were shown earlier in Figure 5.6. Here 0° aspect angle means the person is walking towards the radar and 180° means the person is walking away

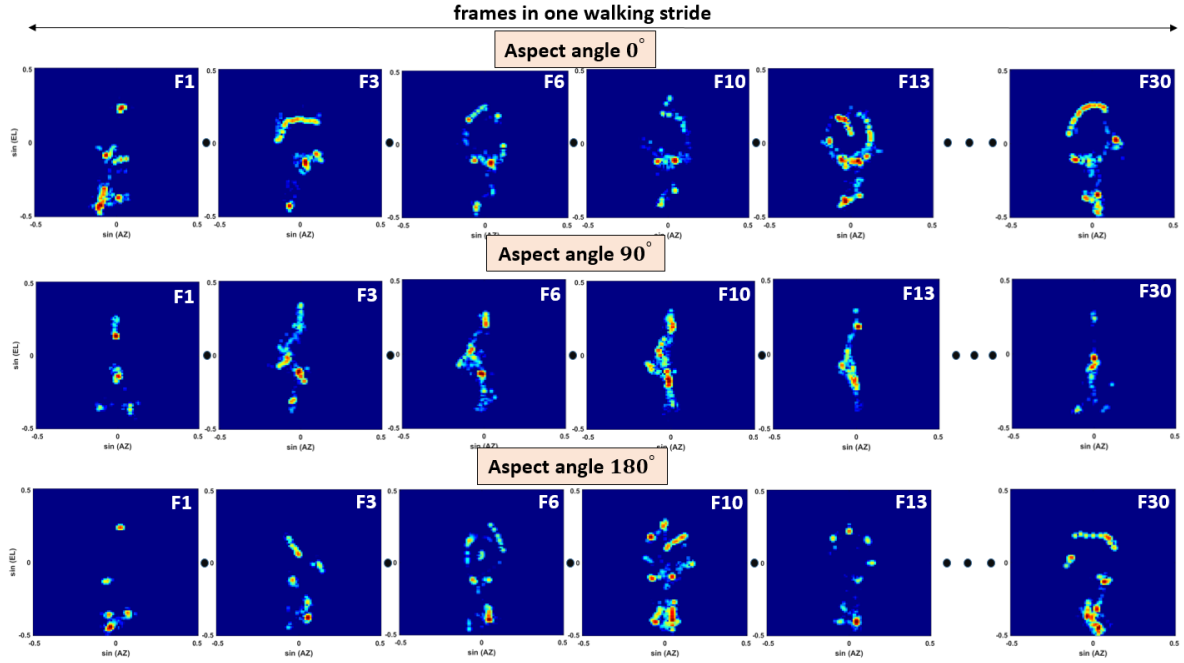


Figure 5.9: Simulated Doppler enhanced frontal images over one walking stride of human motion for different aspect angles

from the radar. Similarly 90° corresponds to the motion along the tangential direction to the radar. I studied the efficacy of the algorithm for reinforced wall which, as mentioned earlier, is one of the most complex walls. I tested the performance of the denoising autoencoder on two scenarios: In the first scenario, both the training and test data are gathered at the same aspect angle (identical training and test scenario); In the second scenario, I used images captured at multiple different aspect angles for both training and testing the autoencoder. Therefore, during test, the algorithm is not provided information of the aspect angle of the data. I used both NMSE and SSIM to measure the performance and report the results in Table 5.4. I observe highest error when the algorithm is trained with data captured at 90° aspect angle that is when the human walks in a direction tangential to the radar. This is most likely because of the inherent distortions in these frontal images due to the limited separation of point scatterers

Table 5.4: Denoising results between clean and corrupted images (captured behind reinforced wall) for different aspect angles. SSIM,NMSE:between corrupted and free space image before denoising (BD) and SSIM,NMSE: between reconstructed and free space image after denoising (AD)

Wall Scenario (Reinforced)			Denoising Metric	
			SSIM	NMSE
Aspect Angle	0°	BD	0.05	5.64
		AD	0.64	1.16
	45°	BD	0.36	4.01
		AD	0.73	1.02
	90°	BD	0.01	4.33
		AD	0.60	1.67
	180°	BD	0.46	5.01
		AD	0.80	1.23
	0°, 45°, 90°, 180°	BD	0.41	4.60
		AD	0.71	1.44

on the subject along the azimuth direction which can be clearly seen from Figure 5.9. Likewise, the Dopplers of the different point scatterers on the human body are not well resolved due to the tangential motion. The results reported for all the aspect angles show significant improvement after denoising. When data from multiple aspect angles is considered, the denoising significantly helps in reconstructing images close to free space images even when the algorithm is not provided any information of the exact aspect angle at which a person is walking. Therefore it can be inferred that the autoencoder is specifically suited for problems dealing with a great deal of diversity in the target and channel conditions. It can significantly denoise (i) images captured in similar and dissimilar- wall conditions as well as (ii) images captured at different aspect angles of the target provided there is sufficient diversity across training data.

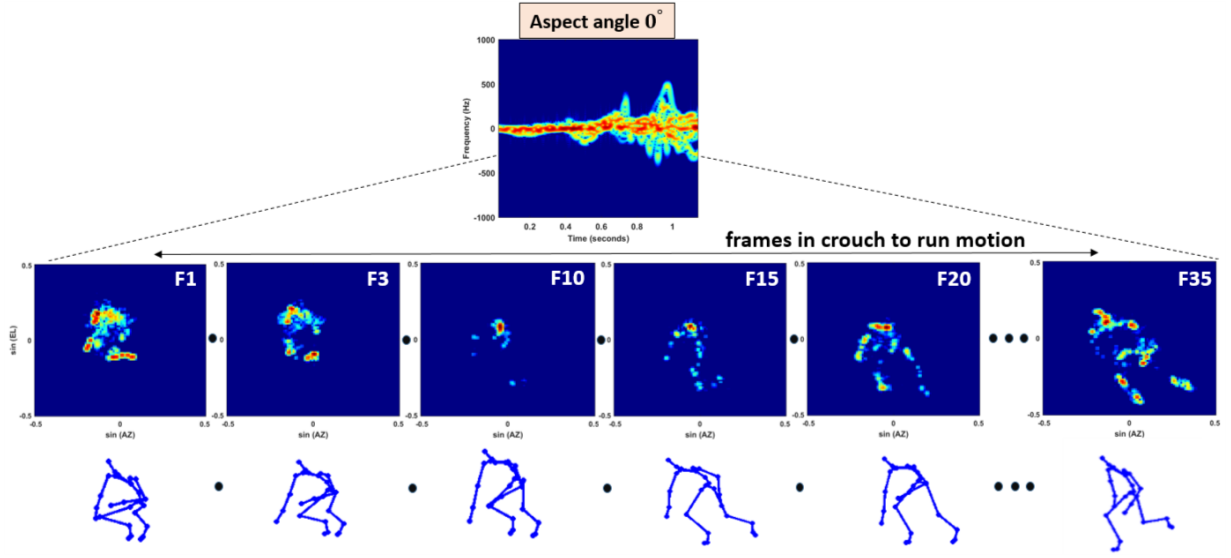


Figure 5.10: Simulated Doppler enhanced frontal images over a crouch to run motion of a human and the corresponding stick figures obtained using ground truth motion capture data

5.3.4 Aperiodic human crouch to run motion

I performed additional simulations to evaluate the performance of algorithm when human undergoes a non-periodic motion - when a human transitions from a crouch to a run towards the radar. Figure 5.10 shows the micro-Doppler signature of this motion. Since our FDTD simulation spatial extent is limited, the translational motion of the human is removed and only retained the dynamics of the swinging arms and legs. Here the torso velocity is mostly around 0 since there is no translational motion of the human. I considered 40 consecutive frames of a human subject over a duration of 1.2s. The transition between the motion states can be clearly seen from the frontal images and their stick figure counterparts. This motion is considerably more challenging than the simple human walking motion. As the number of frames increase, there is enough diversity in the motion which is also reflected in the group correlation index

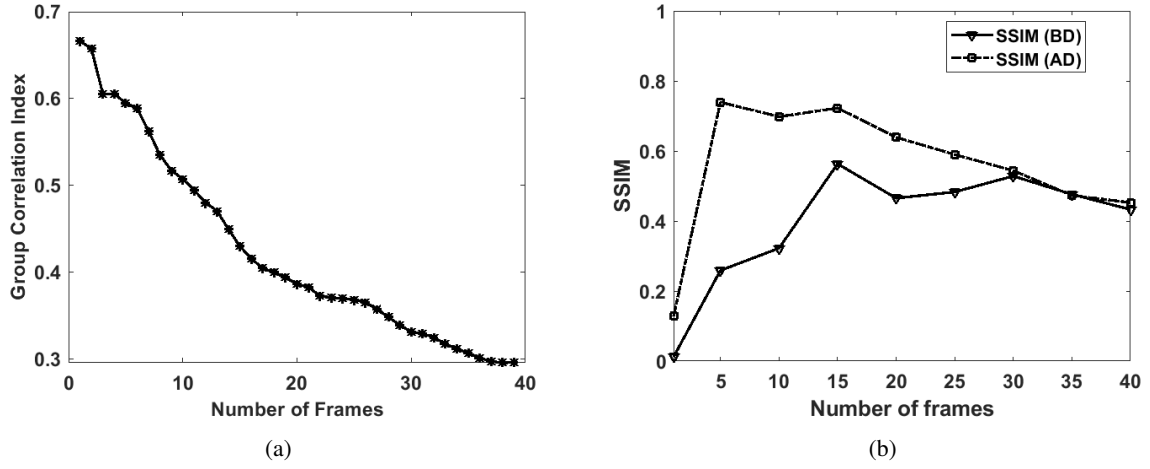


Figure 5.11: Simulated human (a) Group correlation across multiple frames and (b) SSIM variation for simulation results with respect to number of frames.

across multiple frames shown in Figure 5.11(a). I tested the performance of our algorithm on data set captured behind a reinforced wall (a complex wall). Figure 5.11(b) shows the denoising performance of the our denoising algorithm on the complex data set. The figure shows that the denoising algorithm (using the linear mapping function) results in significant improvement in the SSIM after denoising until approximately 15 frames. Beyond this, the performance of the denoising algorithm deteriorates due to the possible decorrelation between test and training data which is also reflected in Figure 5.11(a). Thus it can be concluded that the performance depends on how well the algorithm is trained to handle diversity in the test data.

5.4 Measurement Results

5.4.1 Measurement Data Collection

In this section, I evaluate the performance of our algorithm using wideband measurement data captured in both free space and through-wall conditions. The data is collected using Walabot Pro [105], a wideband(3.3-10.3GHz) 3D programmable RF imaging sensor. Walabot is a low power uncalibrated sensor with limited range in through-wall scenarios. It uses a 4×4 antenna array to illuminate the area in front of it to capture the back-scattered signals. It gives the ADC values of the voltage it detects. These values range from 0 to 255. This means that the the voltage level 5V represent 255 and anything less than 5V is a ratio between 5V and 255. The effective number of bits of Walabot are 8-bits. I converted these ADC values into voltage values. The hardware radar parameters are listed in Table 5.5. The range enhanced images are obtained in a manner similar to (5.19) by replacing the Doppler dimension with the range dimension. Here, the peak scatterers across all the range gates are coherently summed to obtain the frontal images of the targets. These steps are performed within the inbuilt processor in the sensor and we are provided with output range-enhanced images. The assumption here, is that the targets are still or slow moving. The clean measurement data consist of range enhanced frontal images of a human gathered in line-of-sight conditions in an environment mostly free of clutter. The through-wall measurement data comprise of images captured through a 2cm thick glass wall and a 3cm thick wooden wall. The reason for choosing

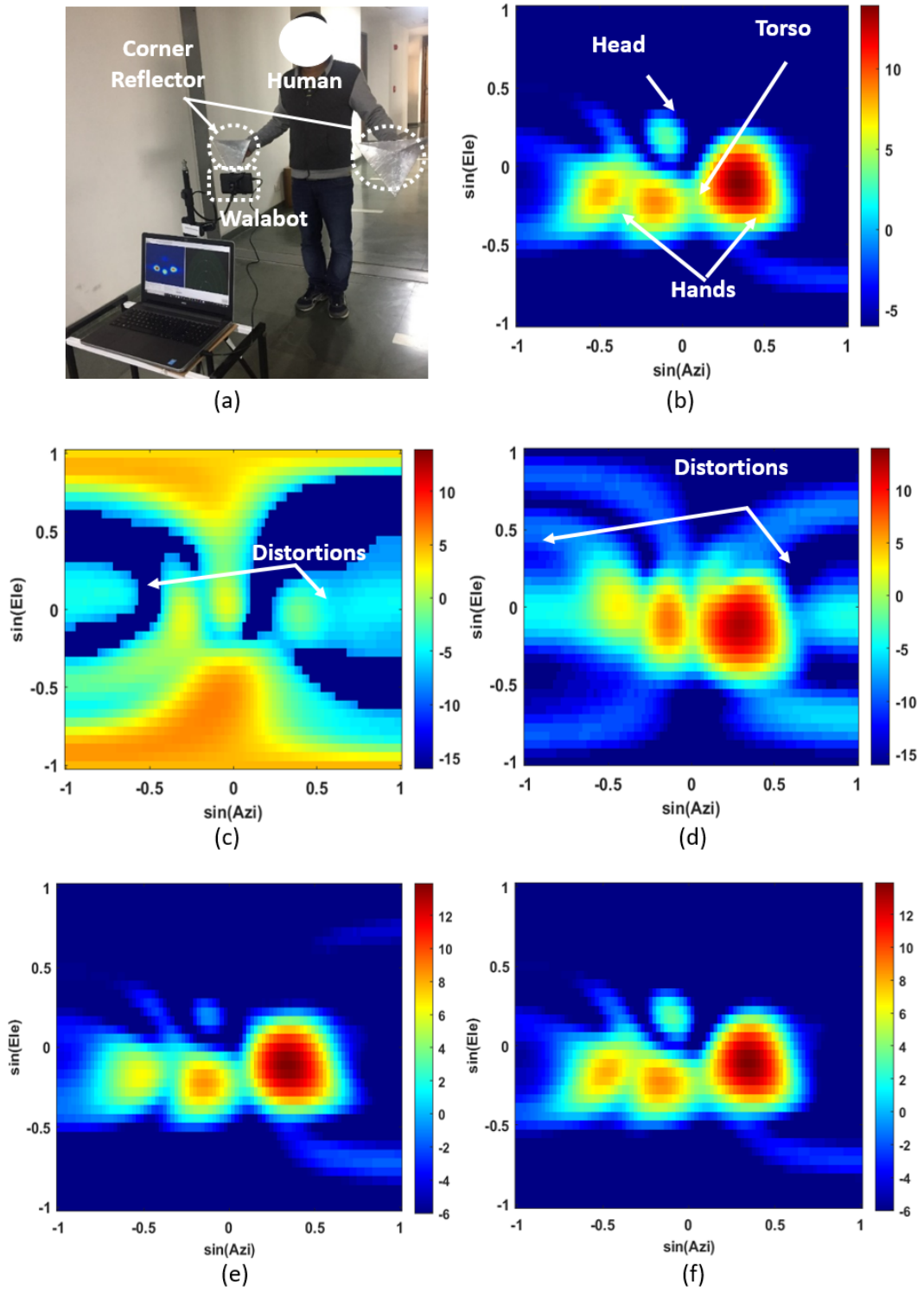


Figure 5.12: (a) Measurement setup in free space and measured range enhanced frontal image of a human subject in (b) free space, (c),(d) behind a glass wall, wood wall respectively , (e),(f) Denoised images in through wall scenarios-glass Wall, wood wall respectively using the presented denoising algorithm.

these through-wall conditions is that the Walabot (operating with a relatively low power at the relatively high frequency), have a very high attenuation for the more complex concrete walls. Please note that the signal propagation phenomenology may also vary for different wall parameters such as the dielectric constant, the conductivity, the thickness, and the width of the walls. The walls here at IIITD are 30cm thick walls which are made of brick or cement blocks held together with cement mortar and are often plastered with cement plaster on both sides. Walabot signals could not penetrate through these complex walls. Therefore, I considered simplistic glass and wooden walls rather than complex in-homogeneous through-wall scenarios. The subject stands in front of the radar at a standoff distance of 2m carrying two corner reflectors covered with aluminum tape to enhance the reflectivity from the hands as shown in Fig5.12(a). Therefore, the target is an extended target with multiple point scatterers. The experiments are performed on 4 human subjects of different heights and girth at different orientations (-45° to $+45^\circ$) with respect to the radar but always facing the radar. For each of these subjects, I captured 75 measurements resulting in a total of 300 images of which 80% are used for training and remaining for test. An example of the resultant radar image in free space is shown in Figure 5.12(b) where the torso, legs and two arms of the human can be clearly discerned. Radar images are corrupted when the measurements are gathered under different through-wall conditions. Some examples of the distortions are presented in Figure 5.12. Figure 5.12(c)-(d) correspond to images in through-glass wall and through-wood wall conditions respectively. These images are consid-

Table 5.5: Parameters of Real Radar Setup. *- values derived from available information

Radar Parameters	Values
Radar Type	Broadband
Bandwidth	$3.3GHz - 10.3GHz$
Maximum Range (R_{max})	$10m^*$
Range resolution (Δr)	$0.2m^*$
ADC	$8\ bit$
Number of antenna elements ($P \times Q$)	4×4
Azimuth Beamwidth ($\Delta\phi_{azi}$)	25°^*
Azimuth Beamwidth ($\Delta\theta_{ele}$)	25°^*
Field of View (ϕ_{azi})	$-90^\circ\ to\ 90^\circ^*$
Field of View (θ_{azi})	$-90^\circ\ to\ 90^\circ^*$

erably distorted due to the complex interaction between the wall and the target.

Each image of size $[91 \times 37]$ is vectorised to obtain a $[3367 \times 1]$ vector. Then all images are clubbed together to form a training data matrix of size $[3367 \times 240]$ and test data matrix of size $[3367 \times 60]$. Once trained, the weight matrices $\mathbf{W}_1$ and $\mathbf{W}_2$ are used to denoise the corrupted test images using the equation (5.10). Analogous to simulations, I examine the variation of denoising performance for the different number of nodes in the hidden layer, for different mapping functions (linear, non-linear- tanh and sigmoid) and for different proportions of training to test data.

5.4.2 Measurement Results and Analyses

I tested the performance of the denoising autoencoder for *same wall* and *different wall* scenarios. In the same wall scenario, both the training and test data are gathered from the same type of wall. In the different wall scenario, data from multiple walls are used for training the autoencoder which is subsequently

used for denoising images from any of the two walls. The reconstruction results are presented as a function of percentage of training data to test data in Table 5.6. These results have been obtained using an autoencoder where the hidden layer has 1500 nodes and the mapping function is sigmoid. The choice of these parameters are discussed in subsequent sections. The table shows SSIM and NMSE between the denoised radar images in through-wall and corresponding radar image gathered in free space conditions. I compare the metrics before denoising (BD) with those after denoising (AD). I observe that there is significant improvement in SSIM and reduction of NMSE after denoising. The performance improves as the percentage of training to test data increases for both the same wall and for different wall scenarios. In other words, the performance during test relies on adequate training data. The error for the different wall scenario is only slightly higher than the same wall scenario. This is the scenario when the test algorithm has no knowledge of the wall scenario. Note that in the case of the wideband measurements, I have not presented the result as a function of the number of frames. This is because, the targets are static and each measurement is independent with no correlation between them.

5.5 Discussion on Results

5.5.1 Computational Complexity Evaluation

The real-time performance of the algorithm relies on the test time and the test memory rather than training time. During test, I perform matrix multiplication

Table 5.6: Denoising results between clean and corrupted measurement images of real humans for different through-wall conditions under varying percentage of training data. NMSE, SSIM :between corrupted and free space image before denoising (BD) and NMSE,SSIM: between reconstructed and free space image after denoising (AD)

Wall Scenario			Denoising Metric								
			SSIM				NMSE				
			(% of Training Data)				(% of Training Data)				
			20	40	60	80	20	40	60	80	
Train and test on same wall	Glass Wall	BD	0.20	0.20	0.22	0.21	36.09	38.47	33.52	38.70	
		AD	0.53	0.70	0.88	0.97	7.81	5.10	4.68	3.69	
	Wood Wall	BD	0.22	0.22	0.23	0.23	28.21	29.04	27.51	33.76	
		AD	0.50	0.55	0.82	0.91	7.59	7.10	4.59	4.24	
Train on multiple walls			BD	0.23	0.22	0.23	0.22	29.00	28.63	28.43	29.49
			AD	0.46	0.59	0.70	0.89	8.40	7.39	6.19	4.97

operations of the trained weights $\mathbf{W}_1$ and $\mathbf{W}_2$ with test image $\hat{\mathbf{Y}}^{test}$. The sizes of the weight matrices and the image matrix are $R \times N$, $N \times R$ and $N \times 1$ respectively, where N denotes the number of pixels in the image and R denotes the number of hidden nodes in the autoencoder such that the number of nodes is always well below the number of pixels. The computational complexity therefore is $\mathcal{O}(RN)$. I ran our algorithm on Matlab 2015b, where all the variables were stored as 64 bit floats, with an Intel(R) Core(TM) i7-5500 processor running at 2.40 GHz. I report the test and training times of our algorithms as a function of the number of nodes of the hidden layers in Figure 5.13. Both the training and test times are higher when R is greater. The test time is significantly low even for the highest number of hidden nodes (1500). The computational memory in all of these cases was less than 500MB. Therefore, these test operations can be carried out in easily available processors such as Raspberry PI 3+ (with a 1GB RAM and 1.4GHz clock speed).

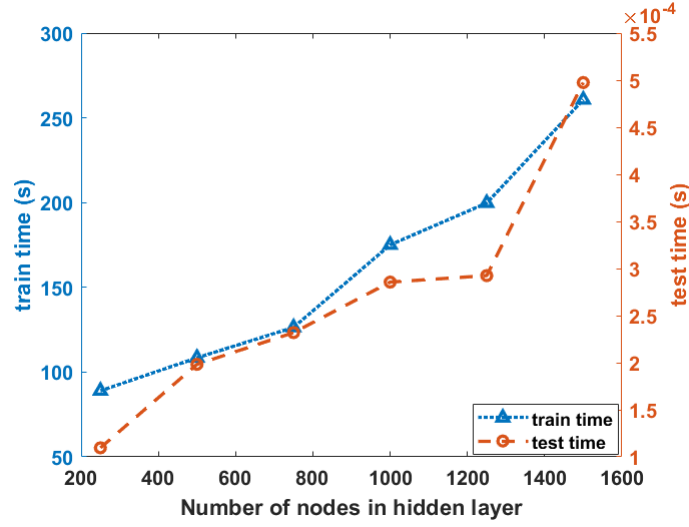


Figure 5.13: Computational time as a function of the number of nodes in the hidden layer for (left y-axis) training phase (right y-axis) test phase

5.5.2 Diversity of training data

The training data must be sufficiently large to handle the diversity of target conditions, channel conditions and any labeling errors between free space and through-wall images.

Diversity of target data: In our work, our autoencoder has been trained to handle the diversity in the size, shape, and orientation or aspect angle of the target with respect to the radar. In the case of dynamic motions, the correlation between consecutive frames facilitated in improving the denoising performance.

Diversity of channel data: Next, the approach does not require the knowledge of the exact wall conditions or analytical framework during the test phase. Instead, the algorithm was capable of denoising images obtained from diverse through-wall conditions.

Labelling errors between free-space and through-wall images: Finally, in

practice, it may be nearly impossible to gather correlated images in free space and through-wall conditions, especially for dynamic targets. For example, it may not be possible to replicate human motions in two different scenarios. Therefore, the algorithm must tolerate some degree of diversity in the motion characteristics during test and training phases. A sensitivity analysis of mismatch/labeling error between clean (free space) and the corrupted (through-wall) training images is not considered in this work. Generally, in machine learning scenarios, these algorithms are quite robust to reasonable random errors in the training set arising due to incorrectly labeled data. However, the algorithms are less robust to systematic errors when the samples are consistently incorrectly labeled.

5.5.3 Hyper-parameter Selection

I optimized the number of nodes in the hidden layer and mapping functions to obtain the results presented in the previous sections. First, I discuss the autoencoder used on the simulation data. I simulated narrowband time-domain returns at a carrier frequency of 7.5GHz with the gains of the antennas to be 10dBi, and the transmitted power at +30dBm. As a result, the maximum received signal strength from the human subject (at $R = 2.5\text{m}$) is -57dBm. Therefore, all the pixel values in the simulated images are negative. Figure 5.14a shows the performance for different mappings- linear, tanh, sigmoid - as a function of the number of frames. We observe that the linear and tanh mapping outperform the results obtained using sigmoid mapping functions. This is because both sigmoid

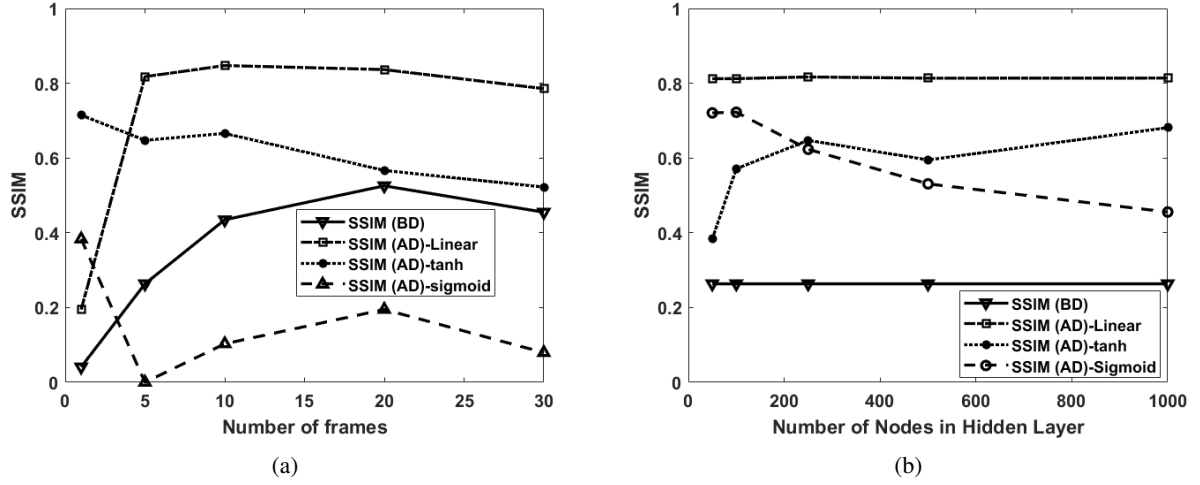


Figure 5.14: SSIM variation for simulation results with respect to (a) number of frames, (b) number of nodes in the hidden layer for mapping functions-linear, tanh and sigmoid.

and hyperbolic tangent (tanh) functions - are monotonically increasing functions that asymptote at $\pm\infty$. However, the tanh function is symmetric about the origin and produces outputs values between $[-1\ 1]$ while sigmoid function outputs are always positive $[0\ 1]$. The sigmoid mapping function is thus not suited to handle the negative signal strength in the radar images due to its asymptotic behavior. The linear mapping function is best suited for this case. Our studies also showed that using similar mapping functions in the encoder and decoder results in better performances when compared to the use of different mapping functions. Figure 5.14(b), show the variation of SSIM before and after denoising as a function of the number of nodes in the hidden layer. We observe that the performance converges when the number of nodes is approximately 500. Next, I discuss the autoencoder used on the measurement data of real humans in both line-of-sight and through-wall conditions. The measurement data are collected using the Walabot that has an ADC of 8 bits. The digitized data are calibrated to a voltage scale from 0 to 5V. In the radar images, the pixel values are con-

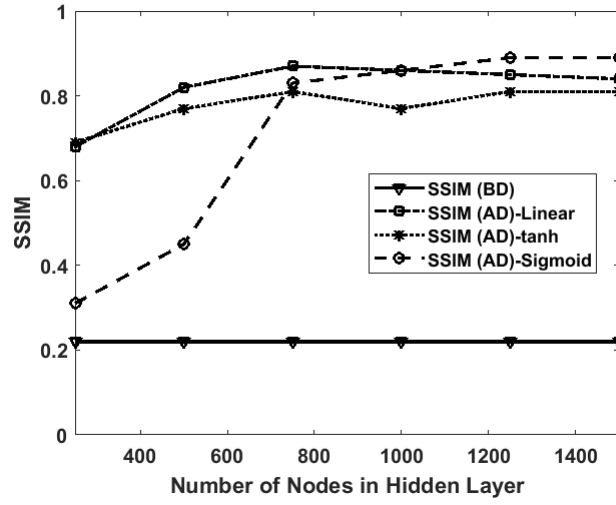


Figure 5.15: SSIM variation for measurement results with respect to the number of nodes in the hidden layer for mapping functions-linear, tanh and sigmoid for human subjects

verted to the logarithmic scale and hence consist of both negative and positive values. Due to the dynamic range of the pixel energy values, the sigmoid function is now able to handle the denoising, and got excellent results in Figure 5.15. The results are in perfect accordance with our previous hypothesis that the performance of these activation functions is sensitive to the signal strength and dynamic range of the pixels in the radar images. In fact, the non-linear mapping functions are slightly superior to the linear mapping function. The linear mapping function is able to handle the non-linearity in the wall response since the images are inherently sparse. The SSIM varies as a function of the number of nodes in the hidden layer for different mapping functions in Figure 5.15. The SSIM improves and tends towards 1 as we increase the number of nodes in the hidden layer to 1500.

5.6 Summary

I demonstrate the efficacy of the denoising autoencoder network at mitigating the distortions and clutter introduced by wall propagation on radar images of humans. The presented approach requires neither prior information of the wall characteristics nor any kind of analytic framework to describe the wall propagation effects. Instead, the algorithm relies on the availability of a huge training data set comprising of distorted radar images captured in diverse through-wall scenarios and the corresponding clean images in line-of-sight conditions. Once trained, the algorithm is capable of mitigating through-wall effects of similar walls though not necessarily identical walls. This capability makes this approach suitable for tracking humans under diverse propagation environments. I evaluated the performance of the algorithm on both static and dynamic targets. The radar images of dynamic humans were simulated using Doppler-enhanced array processing while the images of the static humans were generated from measurement data using range-enhanced array processing. Before denoising, the images were considerably distorted by through-wall propagation effects. Our algorithm showed that after denoising, the images were structurally similar to the free space images with the low mean square errors.

Chapter 6

Conclusion and Future Work

6.1 Contributions and Impact

The contributions of the dissertation are:

6.1.1 Public database of radar micro-Doppler data

An important contribution of my thesis is the release of a highly curated data set of simulated and measured human micro-Dopplers comprising:

- Single human walking scenarios of varying gaits, heights and velocities
- Multiple (two, three and four) humans walking together
- Humans performing different types of periodic motions in line-of-sight conditions
- Humans performing different types of periodic motions in through-wall conditions

Machine learning techniques have made major breakthroughs in various applications, and new algorithms appear every day. The performances of these algorithms are tied to the volume and diversity of high-quality training data. Unlike the vision and image processing and even the medical communities, there are currently a limited number of open databases of radar data. I hope that sharing our radar data will be useful for benchmarking future algorithms and reduce the expense and labor involved in data acquisition by other researchers.

6.1.2 Representation of radar micro-Dopplers using customized dictionaries obtained from data-driven dictionary learning

I demonstrated that highly sparse representations of micro-Doppler data can be generated based on data-driven dictionary learning approaches. Unlike data independent transforms such as Fourier and wavelets, these dictionaries are fine-tuned to data characteristics resulting in highly discriminative representations. These dictionaries can be used for a variety of applications similarly to those obtained from other transforms. I examined the advantages of data-driven dictionaries compared to traditional dictionaries for two applications in my thesis.

6.1.2.1 Single channel source separation of radar micro-Dopplers for multiple target detection

I demonstrated that the customized dictionaries can be exploited for detecting multiple dynamic targets in the propagation channel. The dictionary learning algorithm facilitated the resolving of multiple movers based on their sparse representations using a low complexity single channel CW radar. The algorithm

can be extended to outdoor automotive scenarios where the propagation environment might comprise of vehicles, pedestrians, and animals whose returns must be separated before they are individually classified. Similarly, in inverse synthetic aperture radar (ISAR) imaging involving range-Doppler processing, micro-motions of components of a target (such as the wheels of an automobile) may give rise to distortions that significantly distort the ISAR images. Here, single channel source separation of micro-Dopplers may enable the mitigation of such distortions.

6.1.2.2 Classification of radar targets using dictionary learning

I found that the data-driven dictionaries outperformed the data independent dictionaries in terms of their classification accuracy and achieved near real-time radar data processing when the test and training micro-Doppler radar data were gathered at different carrier frequencies. This is because these sparse representations of the data are highly discriminative and characterize the target motion as opposed to the sensor parameters. Hence, these types of dictionaries can be used to support reconfigurable radar hardware where sensor parameters are changed during deployment in response to unforeseen channel conditions.

6.1.3 Method to simulate large volumes of radar micro-Dopplers in indoor conditions

I have presented a computationally efficient method to model radar micro-Dopplers in indoor conditions by integrating the stochastic finite-difference time-domain (sFDTD) technique with the primitive based scattering center model of human

radar returns. I introduced significant diversity in the electrical parameters of static indoor scatterers by incorporating stochasticity in a single sFDTD simulation instead of running multiple FDTD simulations.

6.1.4 Clutter mitigation in radar images using denoising autoencoders

I demonstrated a denoising autoencoder network for mitigating through-wall propagation artifacts - such as attenuation, defocussing and multipath clutter - on micro-Doppler based radar images. The algorithm does not require analytic models to represent these complex wall behaviors or prior information of the wall parameters or room geometry. Instead, the algorithm relies on a large volume of training data that are diverse in terms of target conditions (the diversity in the size, shape and orientation or aspect angle of the target with respect to the radar) and channel conditions. The denoising autoencoder algorithm has been implemented using an alternating direction method of multipliers approach to ensure convergence and fast training times. Though the method was demonstrated for indoor through-wall images, I hypothesize that these techniques can be used to mitigate distortions in other types of radar images where large training databases are available (such as the automotive radar scenarios) since the method does not rely on physics-based modeling and instead on data driven learning based inferences.

6.2 Future Research Avenues

Some future avenues for research on the micro-Doppler radar are provided below.

6.2.1 Diversity in radar systems, algorithms, and metrics

Radar micro-Doppler data are currently gathered with a variety of hardware platforms both in research labs and with commercial sensors. There are considerable variations in the system parameters such as carrier frequency, transmitted power, waveform, and polarization. The hardware is supported by significantly different signal processing and machine learning-based algorithms, each fine-tuned for achieving specific objectives. Unlike other radar systems, micro-Doppler radar hardware and software have not been uniformly benchmarked against a common data set. In fact, there are currently no common metrics for evaluating the system performance for micro-Doppler based sensing of humans. It would, therefore, be desirable to derive the radar operating characteristics (ROC) curves - the probability of detection and false alarm of human sensing based on micro-Doppler data.

6.2.2 Dictionary learning on higher-order data

I presented an approach to detect a limited number of dynamic movers based on the sparse representations of their micro-Dopplers. However, when the number of target classes becomes very large, the discriminativeness of dictio-

naries gathered from Doppler-only data decreases. It would be interesting to investigate the dictionary learning framework in conjunction with higher-order radar data which include range-Doppler ambiguity maps, Doppler-direction-of-arrival (DOA) ambiguity maps, and three-dimensional radar data cubes (range, Doppler and DOA).

6.2.3 RF Imaging on passive radar data

Micro-Doppler radars have been observed with both active and passive radar sensors. Owing to the sudden rise in transmitters of opportunity, passive sensing has attracted great attention for indoor tracking and monitoring applications. Passive sensing leads to low power consumption and lighter construction due to its receive only nature. However, there is very limited research into imaging targets using passive signals. Interesting future research direction would be to investigate radar image of indoor movers using passive radar sensing techniques.

References

- [1] K. Ranney, K. Gallagher, D. Galanos, A. Hedden, R. Cutitta, S. Freeman, C. Dietlein, B. Kirk, and R. Narayanan, “Software-defined radar: recent experiments and results,” in *Radar Sensor Technology XXII*, vol. 10633. International Society for Optics and Photonics, 2018, p. 106331E.
- [2] V. C. Chen, F. Li, S.-S. Ho, and H. Wechsler, “Analysis of micro-doppler signatures,” *IEE Proceedings-Radar, Sonar and Navigation*, vol. 150, no. 4, pp. 271–276, 2003.
- [3] —, “Micro-doppler effect in radar: phenomenon, model, and simulation study,” *IEEE Transactions on Aerospace and electronic systems*, vol. 42, no. 1, pp. 2–21, 2006.
- [4] M. G. Amin, A. Ravisankar, and R. G. Guendel, “Rf sensing for continuous monitoring of human activities for home consumer applications,” in *Big Data: Learning, Analytics, and Applications*, vol. 10989. International Society for Optics and Photonics, 2019, p. 1098907.
- [5] S. Zubeyde Gurbuz and M. G. Amin, “Radar-based human-motion recognition with deep learning: Promising applications for indoor monitoring,”

- in *Signal Processing Magazine*. IEEE, 2019.
- [6] B. Erol and M. G. Amin, “Radar data cube analysis for fall detection,” in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 2446–2450.
- [7] J. Le Kernec, F. Fioranelli, C. Ding, H. Zhao, L. Sun, H. Hong, J. Loran-del, and O. Romain, “Radar signal processing for sensing in assisted living: The challenges associated with real-time implementation of emerging algorithms,” *IEEE Signal Processing Magazine*, vol. 36, no. 4, pp. 29–41, 2019.
- [8] D. L. Hall, T. D. Ridder, and R. M. Narayanan, “Abnormal gait detection and classification using micro-doppler radar signatures,” in *Radar Sensor Technology XXIII*, vol. 11003. International Society for Optics and Photonics, 2019, p. 110030Q.
- [9] F. Fioranelli, H. Li, J. Le Kernec, V. Busin, N. Jonsson, G. King, M. Tomlinson, and L. Viora, “Radar-based evaluation of lameness detection in ruminants: preliminary results,” in *2019 IEEE MTT-S International Microwave Biomedical Conference (IMBioC)*, vol. 1. IEEE, 2019, pp. 1–4.
- [10] T. Thayaparan, S. Abrol, E. Riseborough, L. Stankovic, D. Lamothe, and G. Duff, “Analysis of radar micro-doppler signatures from experimental helicopter and human data,” *IET Radar, Sonar & Navigation*, vol. 1, no. 4, pp. 289–299, 2007.

- [11] Z. A. Cammenga, G. E. Smith, and C. J. Baker, “High range resolution micro-doppler analysis,” in *SPIE Defense+ Security*. International Society for Optics and Photonics, 2015, pp. 94 611G–94 611G.
- [12] Y. Kim and H. Ling, “Human activity classification based on micro-doppler signatures using a support vector machine,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 47, no. 5, pp. 1328–1337, 2009.
- [13] Y. Kim and T. Moon, “Human detection and activity classification based on micro-doppler signatures using deep convolutional neural networks,” *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 1, pp. 8–12, 2016.
- [14] J. Zabalza, C. Clemente, G. Di Caterina, J. Ren, J. J. Soraghan, and S. Marshall, “Robust pca micro-doppler classification using svm on embedded systems,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 3, pp. 2304–2310, 2014.
- [15] F. H. C. Tivive, S. L. Phung, and A. Bouzerdoun, “Classification of micro-doppler signatures of human motions using log-gabor filters,” *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1188–1195, 2015.
- [16] V. C. Chen, “Spatial and temporal independent component analysis of micro-doppler features,” in *Radar Conference, 2005 IEEE International*. IEEE, 2005, pp. 348–353.

- [17] D. Fairchild and R. Narayanan, "Classification of human motions using empirical mode decomposition of human micro-doppler signatures," *IET Radar, Sonar & Navigation*, vol. 8, no. 5, pp. 425–434, 2014.
- [18] C. Cai, W. Liu, J. S. Fu, and Y. Lu, "Radar micro-doppler signature analysis with hht," *IEEE Transactions on Aerospace and Electronic systems*, vol. 46, no. 2, pp. 929–938, 2010.
- [19] Y. Kim, S. Ha, and J. Kwon, "Human detection using doppler radar based on physical characteristics of targets," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 2, pp. 289–293, 2015.
- [20] F. Fioranelli, M. Ritchie, and H. Griffiths, "Classification of unarmed/armed personnel using the netrad multistatic radar for micro-doppler and singular value decomposition features," *IEEE Geoscience and Remote Sensing Letters*, vol. 12, no. 9, pp. 1933–1937, 2015.
- [21] B. Tekeli, S. Z. Gurbuz, and M. Yuksel, "Information-theoretic feature selection for human micro-doppler signature classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 5, pp. 2749–2762, 2016.
- [22] J. A. Nanzer and R. L. Rogers, "Bayesian classification of humans and vehicles using micro-doppler signals from a scanning-beam radar," *IEEE Microwave and Wireless Components Letters*, vol. 19, no. 5, pp. 338–340, 2009.

- [23] R. Ricci and A. Balleri, “Recognition of humans based on radar micro-doppler shape spectrum features,” *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1216–1223, 2015.
- [24] M. S. Seyfioğlu, S. Z. Gürbüz, A. M. Özbayoğlu, and M. Yüksel, “Deep learning of micro-doppler features for aided and unaided gait recognition,” in *Radar Conference (RadarConf), 2017 IEEE*. IEEE, 2017, pp. 1125–1130.
- [25] M. S. Seyfioğlu, A. M. Özbayoğlu, and S. Z. Gürbüz, “Deep convolutional autoencoder for radar-based classification of similar aided and unaided human activities,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 4, pp. 1709–1723, 2018.
- [26] M. S. Seyfioğlu, B. Erol, S. Z. Gurbuz, and M. G. Amin, “Diversified radar micro-doppler simulations as training data for deep residual neural networks,” in *2018 IEEE Radar Conference (RadarConf18)*. IEEE, 2018, pp. 0612–0617.
- [27] Z. Chen, G. Li, F. Fioranelli, and H. Griffiths, “Personnel recognition and gait classification based on multistatic micro-doppler signatures using deep convolutional neural networks,” *IEEE Geoscience and Remote Sensing Letters*, vol. 15, no. 5, pp. 669–673, 2018.
- [28] Q. Chen, Y. Liu, F. Fioranelli, M. Ritchie, B. Tan, and K. Chetty, “Dopnet: A deep convolutional neural network to recognize armed and unarmed

- human targets,” *IEEE Sensors Journal*, vol. 19, no. 11, pp. 4160–4172, 2019.
- [29] G. S. Z. Erol, Baris, , and M. Amin, “Gan-based synthetic radar micro-doppler augmentations for improved human activity recognition,” in *Radar Conference (RadarConf), 2019 IEEE*. IEEE, 2019.
- [30] B. Erol, S. Z. Gurbuz, and M. G. Amin, “Gan-based synthetic radar micro-doppler augmentations for improved human activity recognition,” in *2019 IEEE Radar Conference (RadarConf)*. IEEE, 2019, pp. 1–5.
- [31] E. Al Hadhrami, M. Al Mufti, B. Taha, and N. Werghi, “Transfer learning with convolutional neural networks for moving target classification with micro-doppler radar spectrograms,” in *2018 International Conference on Artificial Intelligence and Big Data (ICAIBD)*. IEEE, 2018, pp. 148–154.
- [32] R. Nepal, J. Cai, and Z. Yan, “Micro-doppler radar signature identification within wind turbine clutter based on short-cpi airborne radar observations,” *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1268–1275, 2015.
- [33] W. Al-Mashhadani, A. Brown, L. Danoon, C. Horne, R. Palamà, H. Griffiths, J. Patel, and F. Fioranelli, “Measurements and modelling of radar signatures of large wind turbine using multiple sensors,” in *2018 IEEE Radar Conference (RadarConf18)*. IEEE, 2018, pp. 1389–1394.

- [34] Z. Slavik and K. V. Mishra, “Phenomenological modeling of millimeter-wave automotive radar,” in *URSI Asia-Pacific Radio Science Conference*, 2019, in press.
- [35] K. V. Mishra, M. R. Bhavani Shankar, V. Koivunen, B. Ottersten, and S. A. Vorobyov, “Toward millimeter wave joint radar communications: A signal processing perspective,” *IEEE Signal Processing Magazine*, 2019, in press.
- [36] A. Ghaleb, L. Vignaud, and J. Nicolas, “Micro-doppler analysis of wheels and pedestrians in isar imaging,” *IET Signal Processing*, vol. 2, no. 3, pp. 301–311, 2008.
- [37] Y. Li, L. Du, and H. Liu, “Hierarchical classification of moving vehicles based on empirical mode decomposition of micro-doppler signatures,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 5, pp. 3001–3013, 2013.
- [38] S. Vishwakarma, G. Duggal, K. V. Mishra, and S. S. a. Ram, “Doppler-resilient 802.11 ad-based ultra-short range automotive radar,” *arXiv preprint arXiv:1902.01306*, 2019 .
- [39] P. Molchanov, K. Egiazarian, J. Astola, A. Totsky, S. Leshchenko, and M. P. Jarabo-Amores, “Classification of aircraft using micro-doppler bicoherence-based features,” *IEEE Transactions on Aerospace and Electronic systems*, vol. 50, no. 2, pp. 1455–1467, 2014.

- [40] M. Ritchie, F. Fioranelli, H. Borrión, and H. Griffiths, “Multistatic micro-doppler radar feature extraction for classification of unloaded/loaded micro-drones,” *IET Radar, Sonar & Navigation*, vol. 11, no. 1, pp. 116–124, 2016.
- [41] J. S. Patel, C. Al-Ameri, F. Fioranelli, and D. Anderson, “Multi-time frequency analysis and classification of a micro drone carrying payloads using multistatic radar,” *Journal of Engineering*, 2019.
- [42] Y. Kim and B. Toomajian, “Application of doppler radar for the recognition of hand gestures using optimized deep convolutional neural networks,” in *2017 11th European Conference on Antennas and Propagation (EUCAP)*. IEEE, 2017, pp. 1258–1260.
- [43] X. Liang, H. Li, W. Wang, Y. Liu, R. Ghannam, F. Fioranelli, and H. Heidari, “Fusion of wearable and contactless sensors for intelligent gesture recognition,” *Advanced Intelligent Systems*, p. 1900088, 2019.
- [44] A. Lin and H. Ling, “Three-dimensional tracking of humans using very low-complexity radar,” *Electronics Letters*, vol. 42, no. 18, pp. 1062–1064, 2006.
- [45] S. S. Ram and H. Ling, “Through-wall tracking of human movers using joint doppler and array processing,” *IEEE Geoscience and Remote Sensing Letters*, vol. 5, no. 3, pp. 537–541, 2008.

- [46] S. Vishwakarma and S. S. Ram, “Classification of multiple targets based on disaggregation of micro-doppler signatures,” in *Microwave Conference (APMC), 2016 Asia-Pacific*. IEEE, 2016, pp. 1–4.
- [47] ———, “Detection of multiple movers based on single channel source separation of their micro-dopplers,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 54, no. 1, pp. 159–169, 2017 .
- [48] P. K. M. Krishna and K. Ramaswamy, “Single channel speech separation based on empirical mode decomposition and hilbert transform,” *IET Signal Processing*, vol. 11, no. 5, pp. 579–586, 2017.
- [49] P. Mowlae, J. Kulmer, J. Stahl, and F. Mayer, *Single Channel Phase-Aware Signal Processing in Speech Communication: Theory and Practice*. Wiley Online Library, 2017.
- [50] V. C. Chen and H. Ling, *Time-frequency transforms for radar imaging and signal analysis*. Artech House, 2002.
- [51] Z. Zhang, P. O. Pouliquen, A. Waxman, and A. G. Andreou, “Acoustic micro-doppler radar for human gait imaging,” *The Journal of the Acoustical Society of America*, vol. 121, no. 3, pp. EL110–EL113, 2007.
- [52] B. G. Mobasser and M. G. Amin, “A time-frequency classifier for human gait recognition,” in *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 2009, pp. 730 628–730 628.

- [53] I. Orović, S. Stanković, and M. Amin, “A new approach for classification of human gait based on time-frequency feature representations,” *Signal Processing*, vol. 91, no. 6, pp. 1448–1456, 2011.
- [54] J. Park, J. T. Johnson, N. Majurec, M. Frankford, K. Stewart, G. E. Smith, and L. Westbrook, “Simulation and analysis of polarimetric radar signatures of human gaits,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 3, pp. 2164–2175, 2014.
- [55] M. G. Amin, F. Ahmad, Y. D. Zhang, and B. Boashash, “Human gait recognition with cane assistive device using quadratic time–frequency distributions,” *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1224–1230, 2015.
- [56] B. Lyonnet, C. Ioana, and M. G. Amin, “Human gait classification using microdoppler time-frequency signal representations,” in *Radar Conference, 2010 IEEE*. IEEE, 2010, pp. 915–919.
- [57] R. M. Narayanan, M. C. Shastry, P.-H. Chen, and M. Levi, “Through-the-wall detection of stationary human targets using doppler radar,” *Progress In Electromagnetics Research B*, vol. 20, pp. 147–166, 2010.
- [58] F. Fioranelli, M. Ritchie, and H. Griffiths, “Aspect angle dependence and multistatic data fusion for micro-doppler classification of armed/unarmed personnel,” *IET Radar, Sonar & Navigation*, vol. 9, no. 9, pp. 1231–1239, 2015.

- [59] J. Bryan, J. Kwon, N. Lee, and Y. Kim, "Application of ultra-wide band radar for classification of human activities," *IET Radar, Sonar & Navigation*, vol. 6, no. 3, pp. 172–179, 2012.
- [60] M. G. Anderson and R. L. Rogers, "Micro-doppler analysis of multiple frequency continuous wave radar signatures," in *Defense and Security Symposium*. International Society for Optics and Photonics, 2007, pp. 65 470A–65 470A.
- [61] S. Björklund, H. Petersson, A. Nezirovic, M. B. Guldogan, and F. Gustafsson, "Millimeter-wave radar micro-doppler signatures of human motion," in *Radar Symposium (IRS), 2011 Proceedings International*. IEEE, 2011, pp. 167–174.
- [62] A. Balleri, K. Chetty, and K. Woodbridge, "Classification of personnel targets by acoustic micro-doppler signatures," *IET radar, sonar & navigation*, vol. 5, no. 9, pp. 943–951, 2011.
- [63] D. P. Fairchild and R. M. Narayanan, "Micro-doppler radar classification of human motions under various training scenarios," in *SPIE Defense, Security, and Sensing*. International Society for Optics and Photonics, 2013, pp. 873 407–873 407.
- [64] T. Thayaparan, L. Stanković, and I. Djurović, "Micro-doppler-based target detection and feature extraction in indoor and outdoor environments," *Journal of the Franklin Institute*, vol. 345, no. 6, pp. 700–722, 2008.

- [65] C.-P. Lai, R. Narayanan, Q. Ruan, and A. Davydov, “Hilbert-huang transform analysis of human activities using through-wall noise and noise-like radar,” *IET Radar, Sonar & Navigation*, vol. 2, no. 4, pp. 244–255, 2008.
- [66] O. R. Fogle and B. D. Rigling, “Micro-range/micro-doppler decomposition of human radar signatures,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 48, no. 4, pp. 3058–3072, 2012.
- [67] S. Bjorklund, H. Petersson, and G. Hendeby, “Features for micro-doppler based activity classification,” *IET radar, sonar & navigation*, vol. 9, no. 9, pp. 1181–1187, 2015.
- [68] S. Vishwakarma and S. S. Ram, “Dictionary learning for classification of indoor micro-doppler signatures across multiple carriers,” in *Radar Conference (RadarConf), 2017 IEEE*. IEEE, 2017, pp. 0992–0997.
- [69] —, “Dictionary learning with low computational complexity for classification of human micro-dopplers across multiple carrier frequencies,” *IEEE Access*, vol. 6, pp. 29 793–29 805, 2018.
- [70] M. Aharon, M. Elad, and A. Bruckstein, “k-svd: An algorithm for designing overcomplete dictionaries for sparse representation,” *IEEE Transactions on signal processing*, vol. 54, no. 11, pp. 4311–4322, 2006.
- [71] S. Tariyal, A. Majumdar, R. Singh, and M. Vatsa, “Deep dictionary learning,” *IEEE Access*, 2016.

- [72] S. Shekhar, V. M. Patel, and R. Chellappa, “Analysis sparse coding models for image-based classification,” in *Image Processing (ICIP), 2014 IEEE International Conference on*. IEEE, 2014, pp. 5207–5211.
- [73] F. Ahmad, M. G. Amin, and S. A. Kassam, “Synthetic aperture beamformer for imaging through a dielectric wall,” *IEEE transactions on aerospace and electronic systems*, vol. 41, no. 1, pp. 271–283, 2005.
- [74] A. Lin and H. Ling, “Doppler and direction-of-arrival (ddoa) radar for multiple-mover sensing,” *IEEE transactions on aerospace and electronic systems*, vol. 43, no. 4, 2007.
- [75] R. Perry, R. Dipietro, and R. Fante, “Sar imaging of moving targets,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 35, no. 1, pp. 188–200, 1999.
- [76] S. S. Ram, “Doppler enhanced frontal radar images of multiple human activities,” in *Radar Conference (RadarCon), 2015 IEEE*. IEEE, 2015, pp. 1166–1171.
- [77] S. S. Ram and A. Majumdar, “High-resolution radar imaging of moving humans using doppler processing and compressed sensing,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 51, no. 2, pp. 1279–1287, 2015.
- [78] —, “Through-wall propagation effects on doppler-enhanced frontal radar images of humans,” in *Radar Conference (RadarConf), 2016 IEEE*. IEEE, 2016, pp. 1–6.

- [79] P. Setlur, G. Alli, and L. Nuzzo, “Multipath exploitation in through-wall radar imaging via point spread functions,” *IEEE Transactions on Image Processing*, vol. 22, no. 12, pp. 4571–4586, 2013.
- [80] Q. Tan, H. Leung, Y. Song, and T. Wang, “Multipath ghost suppression for through-the-wall radar,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 3, pp. 2284–2292, 2014.
- [81] M. Leigsnering, F. Ahmad, M. Amin, and A. Zoubir, “Multipath exploitation in through-the-wall radar imaging using sparse reconstruction,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 50, no. 2, pp. 920–939, 2014.
- [82] S. Vishwakarma, V. Ummalaneni, M. S. Iqbal, A. Majumdar, and S. S. Ram, “Mitigation of through-wall interference in radar images using denoising autoencoders,” in *Radar Conference (RadarConf18), 2018 IEEE*. IEEE, 2018, pp. 1543–1548.
- [83] S. Vishwakarma and S. S. Ram, “Mitigation of through-wall distortions of frontal radar images using denoising autoencoders,” *Accepted for publication in IEEE TGRS. Online Available arXiv preprint arXiv:1903.09451*, 2019 .
- [84] G. E. Smith, K. Woodbridge, and C. J. Baker, “Multistatic micro-doppler signature of personnel,” in *Radar Conference, 2008. RADAR’08. IEEE*. IEEE, 2008, pp. 1–6.

- [85] R. Boulic, N. M. Thalmann, and D. Thalmann, “A global human walking model with real-time kinematic personification,” *The visual computer*, vol. 6, no. 6, pp. 344–358, 1990.
- [86] S. S. Ram and H. Ling, “Simulation of human microdopplers using computer animation data,” in *2008 IEEE Radar Conference*. IEEE, 2008, pp. 1–6.
- [87] S. S. Ram, C. Christianson, Y. Kim, and H. Ling, “Simulation and analysis of human micro-dopplers in through-wall environments,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 48, no. 4, pp. 2015–2023, April 2010.
- [88] A. D. Singh, S. S. Ram, and S. Vishwakarma, “Simulation of the radar cross-section of dynamic human motions using virtual reality data and ray tracing,” in *2018 IEEE Radar Conference (RadarConf18)*. IEEE, 2018, pp. 1555–1560 .
- [89] “Cmu graphics lab motion capture database,” online accessed on 24 December 2019. [Online]. Available: <http://mocap.cs.cmu.edu/>
- [90] “Accad motion capture lab database,” online accessed on 24 December 2019. [Online]. Available: <https://accad.osu.edu/research/motion-lab/mocap-system-and-data/>
- [91] S. S. Ram, C. Christianson, Y. Kim, and H. Ling, “Simulation and analysis of human micro-dopplers in through-wall environments,” *IEEE Trans-*

- actions on Geoscience and remote sensing*, vol. 48, no. 4, pp. 2015–2023, 2010.
- [92] S. M. Smith and C. Furse, “Stochastic fdtd for analysis of statistical variation in electromagnetic fields,” *IEEE Transactions on Antennas and Propagation*, vol. 60, no. 7, pp. 3343–3350, 2012.
- [93] D. M. Sullivan, *Electromagnetic simulation using the FDTD method*. John Wiley & Sons, 2013.
- [94] A. Taflove and S. C. Hagness, *Computational electrodynamics: the finite-difference time-domain method*. Artech house, 2005.
- [95] Q. Wu, Y. D. Zhang, W. Tao, and M. G. Amin, “Radar-based fall detection based on doppler time–frequency signatures for assisted living,” *IET Radar, Sonar & Navigation*, vol. 9, no. 2, pp. 164–172, 2015.
- [96] Z. Peng, J.-M. Muñoz-Ferreras, Y. Tang, R. Gómez-García, and C. Li, “Portable coherent frequency-modulated continuous-wave radar for indoor human tracking,” in *2016 IEEE Topical Conference on Biomedical Wireless Technologies, Networks, and Sensing Systems (BioWireleSS)*. IEEE, 2016, pp. 36–38.
- [97] A. Mancini, E. Frontoni, and P. Zingaretti, “Embedded multisensor system for safe point-to-point navigation of impaired users,” *IEEE Transactions on Intelligent Transportation Systems*, vol. 16, no. 6, pp. 3543–3555, 2015.

- [98] M. Otero, “Application of a continuous wave radar for human gait recognition,” in *Signal Processing, Sensor Fusion, and Target Recognition XIV*, vol. 5809. International Society for Optics and Photonics, 2005, pp. 538–549.
- [99] P. Setlur, M. Amin, and F. Ahmad, “Urban target classifications using time-frequency micro-doppler signatures,” in *Signal Processing and Its Applications, 2007. ISSPA 2007. 9th International Symposium on*. IEEE, 2007, pp. 1–4.
- [100] “Radar systems at villanova university,” online accessed on 24 December 2019. [Online]. Available: <https://www1.villanova.edu/villanova/engineering/research/Advanced-Communications/facilities/rillab.html>
- [101] S. Costanzo, “Software-defined doppler radar sensor for human breathing detection,” *Sensors*, vol. 19, no. 14, p. 3085, 2019.
- [102] T. Derham, S. Doughty, K. Woodbridge, and C. J. Baker, “Realisation and evaluation of a low cost netted radar system,” in *Radar, 2006. CIE’06. International Conference on*. IEEE, 2006, pp. 1–4.
- [103] “Ancortek radar,” online accessed on 24 December 2019. [Online]. Available: <https://ancortek.com/>
- [104] “Xethru radar,” online accessed on 24 December 2019. [Online]. Available: <https://www.xethru.com/>

- [105] “Walabot,” online accessed on 18 January 2020. [Online]. Available: https://cdn.sparkfun.com/assets/learn_tutorials/7/2/4/walabot-tech-brief-416.pdf
- [106] J. Z. Kolter, S. Batra, and A. Y. Ng, “Energy disaggregation via discriminative sparse coding,” in *Advances in Neural Information Processing Systems*, 2010, pp. 1153–1161.
- [107] B. K. Natarajan, “Sparse approximate solutions to linear systems,” *SIAM journal on computing*, vol. 24, no. 2, pp. 227–234, 1995.
- [108] I. W. Selesnick, “Sparse signal restoration,” *Connexions*, pp. 1–13, 2009.
- [109] C. M. Bishop, “Pattern recognition,” *Machine Learning*, vol. 128, pp. 1–58, 2006.
- [110] I. Ramirez, P. Sprechmann, and G. Sapiro, “Classification and clustering via dictionary learning with structured incoherence and shared features,” in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 3501–3508.
- [111] M. Yang, L. Zhang, X. Feng, and D. Zhang, “Fisher discrimination dictionary learning for sparse representation,” in *Computer Vision (ICCV), 2011 IEEE International Conference on*. IEEE, 2011, pp. 543–550.
- [112] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, “Robust face recognition via sparse representation,” *IEEE transactions on pattern analysis and machine intelligence*, vol. 31, no. 2, pp. 210–227, 2009.

- [113] “Multiple target radar database collected at iiitd,” online accessed on 24 December 2019. [Online]. Available: <https://bit.ly/2zkWGrc>
- [114] F. K. Coutts, D. Gaglione, C. Clemente, G. Li, I. K. Proudler, and J. J. Soraghan, “Label consistent k-svd for sparse micro-doppler classification,” in *Digital Signal Processing (DSP), 2015 IEEE International Conference on*. IEEE, 2015, pp. 90–94.
- [115] A. D. Singh, S. Vishwakarma, and S. S. Ram, “Co-channel interference between wifi and through-wall micro-doppler radar,” in *Radar Conference (RadarConf), 2017 IEEE*. IEEE, 2017, pp. 1297–1302 .
- [116] K. Engan, S. O. Aase, and J. H. Husoy, “Method of optimal directions for frame design,” in *Acoustics, Speech, and Signal Processing, 1999. Proceedings., 1999 IEEE International Conference on*, vol. 5. IEEE, 1999, pp. 2443–2446.
- [117] Q. Zhang and B. Li, “Discriminative k-svd for dictionary learning in face recognition,” in *Computer Vision and Pattern Recognition (CVPR), 2010 IEEE Conference on*. IEEE, 2010, pp. 2691–2698.
- [118] R. Rubinstein, T. Peleg, and M. Elad, “Analysis k-svd: A dictionary-learning algorithm for the analysis sparse model,” *IEEE Transactions on Signal Processing*, vol. 61, no. 3, pp. 661–677, 2013.
- [119] S. Nam, M. E. Davies, M. Elad, and R. Gribonval, “Cospase analysis modeling-uniqueness and algorithms,” in *Acoustics, Speech and Signal*

- Processing (ICASSP), 2011 IEEE International Conference on.* IEEE, 2011, pp. 5804–5807.
- [120] S. Singh and A. Majumdar, “Analysis co-sparse coding for energy disaggregation,” *IEEE Transactions on Smart Grid*, 2017.
 - [121] P. Molchanov, J. Astola, K. Egiazarian, and A. Totsky, “Ground moving target classification by using dct coefficients extracted from micro-doppler radar signatures and artificial neuron network,” in *Microwaves, Radar and Remote Sensing Symposium (MRRS), 2011.* IEEE, 2011, pp. 173–176.
 - [122] E. Hughes, M. Lewis, and E. Reid, “The application of speech recognition techniques to radar target doppler recognition: A case study,” 2006.
 - [123] J. A. Tropp and A. C. Gilbert, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Transactions on information theory*, vol. 53, no. 12, pp. 4655–4666, 2007.
 - [124] T. T. Cai and L. Wang, “Orthogonal matching pursuit for sparse signal recovery with noise,” *IEEE Transactions on Information Theory*, vol. 57, no. 7, pp. 4680–4688, 2011.
 - [125] R. Tibshirani, “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.
 - [126] I. W. Selesnick, “Sparse signal restoration,” in *Connexions*, 2006.

- [127] Z. Jiang, Z. Lin, and L. S. Davis, “Label consistent k-svd: Learning a discriminative dictionary for recognition,” *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 35, no. 11, pp. 2651–2664, 2013.
- [128] “Radar database collected at iitd,” online accessed on 24 December 2019. [Online]. Available: <https://bit.ly/2L8W4KT>
- [129] L. v. d. Maaten and G. Hinton, “Visualizing data using t-sne,” *Journal of machine learning research*, vol. 9, no. Nov, pp. 2579–2605, 2008.
- [130] M. G. Amin, *Through-the-wall radar imaging*. CRC press, 2016.
- [131] B. Liu, M. Jian, Z. Lu, and R. Chen, “Indoor monitoring human movements using dual-receiver radar,” in *2017 IEEE Radar Conference (RadarConf)*. IEEE, 2017, pp. 0520–0523.
- [132] Y. Jia, G. Cui, L. Kong, and X. Yang, “Multichannel and multiview imaging approach to building layout determination of through-wall radar,” *IEEE Geoscience and Remote Sensing Letters*, vol. 11, no. 5, pp. 970–974, 2014.
- [133] C. Clemente, A. Balleri, K. Woodbridge, and J. J. Soraghan, “Developments in target micro-doppler signatures analysis: radar imaging, ultrasound and through-the-wall radar,” *EURASIP Journal on Advances in Signal Processing*, vol. 2013, no. 1, p. 47, 2013.
- [134] F. Ahmad, M. G. Amin, and G. Mandapati, “Autofocusing of through-the-wall radar imagery under unknown wall characteristics,” *IEEE transactions on image processing*, vol. 16, no. 7, pp. 1785–1795, 2007.

- [135] F. H. C. Tivive, A. Bouzerdoun, and M. G. Amin, “A subspace projection approach for wall clutter mitigation in through-the-wall radar imaging,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 4, pp. 2108–2122, 2014.
- [136] F. Ahmad and M. G. Amin, “Multi-location wideband synthetic aperture imaging for urban sensing applications,” *Journal of the Franklin Institute*, vol. 345, no. 6, pp. 618–639, 2008.
- [137] M. Leigsnering, *Sparsity-Based Multipath Exploitation for Through-the-Wall Radar Imaging*. Springer, 2018.
- [138] V. H. Tang, A. Bouzerdoun, and S. L. Phung, “Multipolarization through-wall radar imaging using low-rank and jointly-sparse representations,” *IEEE Transactions on Image Processing*, vol. 27, no. 4, pp. 1763–1776, 2017.
- [139] V. H. Tang, S. L. Phung, F. H. C. Tivive, and A. Bouzerdoun, “A sparse bayesian learning approach for through-wall radar imaging of stationary targets,” *IEEE Transactions on Aerospace and Electronic Systems*, vol. 53, no. 5, pp. 2485–2501, 2017.
- [140] A. Martone, K. Ranney, and R. Innocenti, “Through-the-wall detection of slow-moving personnel,” in *Radar Sensor Technology XIII*, vol. 7308. International Society for Optics and Photonics, 2009, p. 73080Q.

- [141] F. Ahmad and M. G. Amin, “Through-the-wall human motion indication using sparsity-driven change detection,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 2, pp. 881–890, 2013.
- [142] P. Setlur, M. Amin, and F. Ahmad, “Multipath model and exploitation in through-the-wall and urban radar sensing,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 49, no. 10, pp. 4021–4034, 2011.
- [143] P. Vincent, H. Larochelle, Y. Bengio, and P.-A. Manzagol, “Extracting and composing robust features with denoising autoencoders,” in *Proceedings of the 25th international conference on Machine learning*. ACM, 2008, pp. 1096–1103.
- [144] G. E. Hinton and R. R. Salakhutdinov, “Reducing the dimensionality of data with neural networks,” *science*, vol. 313, no. 5786, pp. 504–507, 2006.
- [145] M. Sakurada and T. Yairi, “Anomaly detection using autoencoders with nonlinear dimensionality reduction,” in *Proceedings of the MLSDA 2014 2nd Workshop on Machine Learning for Sensory Data Analysis*. ACM, 2014, p. 4.
- [146] R. Socher, E. H. Huang, J. Pennin, C. D. Manning, and A. Y. Ng, “Dynamic pooling and unfolding recursive autoencoders for paraphrase detection,” in *Advances in neural information processing systems*, 2011, pp. 801–809.

- [147] M. Chen, Z. Xu, K. Weinberger, and F. Sha, “Marginalized denoising autoencoders for domain adaptation,” *arXiv preprint arXiv:1206.4683*, 2012.
- [148] B. Du, W. Xiong, J. Wu, L. Zhang, L. Zhang, and D. Tao, “Stacked convolutional denoising auto-encoders for feature representation,” *IEEE transactions on cybernetics*, vol. 47, no. 4, pp. 1017–1027, 2017.
- [149] I. Sutskever, J. Martens, G. Dahl, and G. Hinton, “On the importance of initialization and momentum in deep learning,” in *International conference on machine learning*, 2013, pp. 1139–1147.
- [150] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, “Distributed optimization and statistical learning via the alternating direction method of multipliers,” *Foundations and Trends® in Machine Learning*, vol. 3, no. 1, pp. 1–122, 2011.
- [151] K. Yee, “Numerical solution of initial boundary value problems involving maxwell’s equations in isotropic media,” *IEEE Transactions on antennas and propagation*, vol. 14, no. 3, pp. 302–307, 1966.
- [152] J. Mehta, K. Gupta, A. Gogna, A. Majumdar, and S. Anand, “Stacked robust autoencoder for classification,” in *International Conference on Neural Information Processing*. Springer, 2016, pp. 600–607.
- [153] A. Majumdar, “Graph structured autoencoder,” *Neural Networks*, vol. 106, pp. 271 – 280, 2018.

- [154] Z. Wang and A. C. Bovik, “A universal image quality index,” *IEEE signal processing letters*, vol. 9, no. 3, pp. 81–84, 2002.
- [155] A. Hore and D. Ziou, “Image quality metrics: Psnr vs. ssim,” in *2010 20th International Conference on Pattern Recognition*. IEEE, 2010, pp. 2366–2369.