Thèse de doctorat
présentée à
l'UNIVERSITE POLYTECHNIQUE HAUTS-DE-FRANCE

N° d'ordre: 19/06

Pour l'obtention du grade de
Docteur en Sciences
Spécialité: **ELECTRONIQUE**

par:
Rahmad SADLI

Ecole doctorale :
Sciences Pour l'Ingénieur (ED SPI 072)

Equipe de recherche, Laboratoire :
Institut d'Electronique, de Micro-Electronique et de Nanotechnologie - Département d'Opto-Acousto-Electronique (IEMN DOAE - UMR 8520)
Institut Français des Sciences et Technologies des Transports de l'Aménagement et des Réseaux (Laboratoire Électronique Ondes et Signaux pour les Transport IFSTTAR/LEOST)

Etude et développement d'un dispositif routier d'anticollision basé sur un radar ultra large bande pour la détection et l'identification notamment des usagers vulnérables

Soutenue le 12/03/2019, devant le jury composé de:

Président du jury:
Mr. Abdenour HADID. Professeur à l'Université de Oulu, Finland

Rapporteurs:
Mr. Vahid MEGHDADI. Professeur à l'Université de Limoges - ENSIL
Mr. Anthony GHIOTTO. MCF-HDR à IMS, Bordeaux INP

Examinateur
Mme. Maria-Gabriella DI BENEDETTO. Professeur à l'Université La Sapienza, Rome

Directeur de thèse:
Mme. Atika RIVENQ MENHAJ, Professeur à IEMN-DOAE, UPHF, Valenciennes

Co-directeur de thèse:
Mr. Charles TATKEU, Directeur de Recherche à IFSTTAR, Villeneuve d'Ascq

Encadrant:
Mr. Yassin EL HILLALI, MCF-HDR à IEMN-DOAE, UPHF, Valenciennes

Invité
Mr. Amir NAKIB, MCF HDR à Université Paris-Est, Paris

POLYTECHNIQUE HAUTS-DE-FRANCE UNIVERSITY

Doctoral School of Engineering Science

Doctoral Thesis
Speciality: **Electronique**

Study and Development of a Road Collision Avoidance System Based on Ultra Wide-Band Radar for Obstacles Detection and Identification Dedicated to Vulnerable Road Users

by:
Rahmad SADLI

Prof. Vahid MEGHDADI (Rapporteur)
Dr. Anthony GHIOTTO (Rapporteur)
Mme. Prof. Maria-Gabriella DI BENEDETTO (Examiner)
Prof. Abdenour HADID (Examiner)
Mme. Prof. Atika RIVENQ MENHAJ (Supervisor)
Dr. Charles TATKEU (Co-Supervisor)
Dr. Yassin EL HILLALI(Co-Supervisor)
Dr. Amir NAKIB (Invited Member)

March 12, 2019

À ma mère
À la mémoire de mon père
À ma femme
À mes filles
et
À mes frères et mes sœurs

Aknowledgements

This PhD thesis has been supported by the **French National Agency for Research (ANR)** through the project **CYCLOPE**. I gratefully acknowledge the support of this institution. I also thank **European Union**, the **French Government** and the **Region Hauts-de-France** who have supported this work through **ELSAT 2020** project (Ecomobility, Logistics, Security and Adaptability in Transport to 2020).

Thanks to **IEMN-DOAE** and **IFSTTAR-COSYS-LEOST** for welcoming me during my PhD thesis.

I am very grateful to my PhD thesis advisors, **Mr. Yassin EL HILLALI**, Associate Professor at the University of Polytechnique Hauts-de-France, **Prof. Atika RIVENQ MENHAJ**, Professor at the University of Polytechnique Haut-de-France, and **Mr. Charles TATKEU**, Research Director at the IFSTTAR, Villeneuve d'Ascq, for their guidance and encouragement.

I would like also to thank **Prof. Jamal ASSAAD**, Director of IEMN-DOAE laboratory, and **his staff members** as well for allowing me to join the IEMN-DOAE researchers and to access the highly cost Lab materials.

Special thanks to the committee members: **Prof. Vahid MEGHDADI**, **Mr. Anthony GHIOTTO**, **Prof. Maria-Gabriella DI BENEDETTO**, **Prof. Abdenour HADID**, and **Mr. Amir NAKIB** for their time to review this manuscript and for their remarks to enrich this PhD work.

I am indebted to **my mother** for all of the love and support she has given throughout my life time. I would not be where I am today without her. She is a single most influential person in my life.

A very special thanks to my beloved wife, **Andika Vebrina**, who is always beside me. Her prayers and her supports make me strong to do all of this.

I dedicate this PhD thesis to my little girls, **Fatimah** and **Arifa**. I love you so much.

Special thanks also to all my brothers and sisters, **Bang Din**, **Akak**, **Bang Lah**, **Fanda**, **Nyanyak**, **Lala**, **Dek Yan**, and **Dek Si**. Thanks for supporting and praying for me all the time.

I would like to thank also all of **my colleagues** in **Polytechnique of Aceh**, **my fellow Indonesian Students** in Valenciennes. Their supports, their helps, and their advises are invaluable to me.

Finally, thanks to **all my friends** at IEMN-DOAE and IFSTTAR for being a part of my journey and making it memorable and unforgettable.

Résumé | Abstract

Résumé Dans ce travail de thèse, nous présentons nos travaux qui portent sur l'identification des cibles en général par un radar Ultra-Large Bande (ULB) et en particulier l'identification des cibles dont la surface équivalente radar (SER) est faible telles que les piétons et les cyclistes. Ce travail se décompose en deux parties principales, la détection et la reconnaissance.

Dans la première approche du processus de détection, nous avons proposé et étudié un détecteur de radar ULB robuste qui fonctionne avec des données radar 1-D (A-scan) à une dimension. Il exploite la combinaison des statistiques d'ordres supérieurs (HOS) et du détecteur de seuil automatique connu sous le nom de CA-CFAR pour Cell-Averaging Constant False Alarm Rate. Cette combinaison est effectuée en appliquant d'abord le HOS sur le signal reçu afin de supprimer une grande partie du bruit. Puis, après avoir éliminé le bruit du signal radar reçu, nous implémentons le détecteur de seuil automatique CA-CFAR. Ainsi, cette combinaison permet de disposer d'un détecteur de radar ULB à seuil automatique robuste. Afin d'améliorer le taux de détection et aller plus loin dans le traitement, nous avons évalué l'approche des données radar 2-D (B-Scan) à deux dimensions. Dans un premier temps, nous avons proposé une nouvelle méthode de suppression du bruit, qui fonctionne sur des données B-Scan. Il s'agit d'une combinaison de WSD (Wavelet Shrinkage Denoising) et de HOS. Pour évaluer les performances de cette méthode, nous avons fait une étude comparative avec d'autres techniques de suppression du bruit telles que l'analyse en composantes principales (PCA-Principal Component Analysis), la décomposition en valeurs singulières (SVD-Singular Value Decomposition), la WSD, et la HOS. Les rapports signal à bruit -SNR- des résultats finaux montrent que les performances de la combinaison WSD et HOS sont meilleures que celles des autres méthodes rencontrées dans la littérature.

A la phase de reconnaissance, nous avons exploité les données des deux approches à 1-D et à 2-D obtenues à partir du procédé de détection.

Dans la première approche à 1-D, les techniques SVM (Support Vector Machines) et le DBN (Deep Belief Networks) sont utilisées et évaluées pour identifier la cible en se basant sur la signature radar. Les résultats obtenus montrent que la technique SVM donne de bonnes performances pour le système proposé où le taux de reconnaissance global moyen atteint 96,24%, soit respectivement 96,23%, 95,25% et 97,23% pour le cycliste, le piéton et la voiture. Dans la seconde approche à 1-D, les performances de différents types d'architectures DBN composées de différentes couches ont été évaluées et comparées. Nous avons constaté que l'architecture du réseau DBN avec quatre couches cachées est meilleure et la précision totale moyenne peut atteindre 97,80%. Ce résultat montre que les performances obtenues avec le DBN sont meilleures que celles obtenues avec le SVM (96,24%) pour ce système de reconnaissance de cible utilisant un radar ULB.

Dans l'approche bidimensionnelle (2-D), le réseau de neurones convolutifs (CNN) a été utilisé et évalué. Nous avons proposé trois architectures de CNN. La première est le modèle modifié d'Alexnet, la seconde est une architecture avec les couches de convolution arborescentes et une couche entièrement connectée, et la troisième est une architecture avec les cinq couches de convolution et deux couches entièrement connectées. Après comparaison et évaluation des performances de ces trois architectures proposées nous avons constaté que la troisième architecture offre de bonnes performances par rapport aux autres propositions avec une précision totale moyenne qui peut atteindre 99,59%.

Enfin, nous avons effectué une étude comparative des performances obtenues avec le CNN, DBN et SVM. Les résultats montrent que CNN a les meilleures performances en termes de précision par rapport à DBN et SVM. Cela signifie que l'utilisation de CNN dans les données radar bidimensionnels permet de classer correctement les cibles radar ULB notamment pour les cibles à faible SER et SNR telles que les cyclistes ou les piétons.

Mots-clés CA-CFAR, CNN, détection des cyclistes et piétons, Deep Belief Networks, DBN, HOS, identification des cyclistes et piétons, radar Ultra-Large Bande, radar ULB, Réseau de Neurones Convolutifs, Statistiques d'Ordres Supérieurs, Support Vector Machines, SVM, usagers vulnérables, Wavelet Shrinkage Denoising, WSD.

Abstract In this thesis work, we focused on the study and development of a system identification using UWB-Ultra-Wide-Band short range radar to detect the objects and particularly the vulnerable road users (VRUs) that have low RCS-Radar Cross Section- such as cyclist and pedestrian. This work is composed of two stages i.e. detection and recognition.

In the first approach of detection stage, we have proposed and studied a robust UWB radar detector that works on one dimension 1-D radar data (A-scan). It relies on a combination of Higher Order Statistics (HOS) and the well-known CA-CFAR (Cell-Averaging Constant False Alarm Rate) detector. This combination is performed by firstly applying the HOS to the received radar signal in order to suppress the noise. After eliminating the noise of the received radar signal, we apply the CA-CFAR detector. By doing this combination, we finally have an UWB radar detector which is robust against the noise and works with the adaptive threshold.

In order to enhance the detection performance, we have evaluated the approach of using two dimensions 2-D (B-Scan) radar data. In this 2-D radar approach, we proposed a new method of noise suppression, which works on this B-Scan data. The proposed method is a combination of WSD (Wavelet Shrinkage Denoising) and HOS. To evaluate the performance of this method, we performed a comparative study with the other noise removal methods in literature including Principal Component Analysis (PCA), Singular Value Decomposition (SVD), WSD and HOS. The Signal-to-Noise Ratio (SNR) of the final result has been computed to compare the effectiveness of individual noise removal techniques. It is observed

that a combination of WSD and HOS has better capability to remove the noise compared to that of the other applied techniques in the literature; especially it is found that it allows to distinguish efficiency the pedestrian and cyclist over the noise and clutters whereas other techniques are not showing significant result.

In the recognition phase, we have exploited the data from the two approaches 1-D and 2-D, obtained from the detection method.

In the first 1-D approach, Support Vector Machines (SVM) and Deep Belief Networks (DBN) have been used and evaluated to identify the target based on the radar signature. The results show that the SVM gives good performances for the proposed system where the total recognition accuracy rate could achieve up to 96,24%.

In the second approach of this 1-D radar data, the performance of several DBN architectures compose of different layers have been evaluated and compared. We realised that the DBN architecture with four hidden layers performs better than those of with two or three hidden layers. The results show also that this architecture achieves up to 97.80% of accuracy. This result also proves that the performance of DBN is better than that of SVM (96.24%) in the case of UWB radar target recognition system using 1-D radar signature.

In the 2-D approach, the Convolutional Neural Network (CNN) has been exploited and evaluated. In this work, we have proposed and investigated three CNN architectures. The first architecture is the modified of Alexnet model, the second is an architecture with three convolutional layers and one fully connected layer, and the third is an architecture with five convolutional layers and two fully connected layers. The performance of these proposed architectures have been evaluated and compared. We found that the third architecture has a good performance where it achieves up to 99.59% of accuracy.

Finally, we compared the performances obtained using CNN, DBN and SVM. The results show that CNN gives a better result in terms of accuracy compared to that of DBN and SVM. It allows to classify correctly the UWB radar targets like cyclist and pedestrian.

Keywords CA-CFAR, Convolutional Neural Networks, CNN, cyclists and pedestrians detection, cyclists and pedestrians identification, Deep Belief Networks, DBN, Higher Order Statistics, HOS, Support Vector Machines, SVM, Ultra-Wide Band radar, UWB radar, vulnerable users, Wavelet Shrinkage Denoising, WSD

Contents

List of Figures

List of Tables

List of Acronyms

ACAS/FOT	Advanced Collision Avoidance System/Field Operational Test
ACS	Anti-Collision System
ANN	Artificial Neural Network
ATIS	Advanced Traveler Information Service
ATMS	Advanced Traffic Management System
AVCS	Advanced Vehicle Control and Safety
AWGN	Additive White Gaussian Noise
BS	Base Station
BSS	Blind Source Separation
CA-CFAR	Cell-Averaging Constant False Alarm Rate
CMS	Collision Mitigation System
CNN	Convolution Neural Network
CYCLOPE	Cyclo Protection Electronically
DBN	Deep Belief Network
DSRC	Dedicated Short Range Communication
CEM	Electromagnetic compatibility
ETC	Electronic toll collection
EU	European Union
FCC	Federal Communications Commission
FCW	Forward Collision Warning
FMCW	Frequency Modulated Continuous Wave
FN	False Negative
FP	False Positive
FSK-CW	Frequency Shift-Keying-Continuous Wave
GPS	Global Positioning System
HOS	Higher Order Statistics
ITS	Intelligent Transport System
MTI	Moving Target Indication
MBWA	Mobile Broadband Wireless Access
MCMC	Markov Chain Monte Carlo
NC	Negative Condition
ODD	Operational Design Domain

OEDR	Object and Event Detection and Response
PC	Personal Condition
PCA	Principal Component Analysis
PROSPECT	Proactive Safety for Pedestrians and Cyclists
RBF	Radial Based Function
RCS	Radar Cross-Section
ReLU	Rectified Linear Unit
ROC	Receiver Operating Characteristic
RBM	Restricted Boltzmann Machine
SAE	Society of Automotive Engineers
SAVE-U	Sensors and System architecture for Vulnerable road Users Protection
SNR	Signal Noise Ratio
SRR	Short Range Radar
SVD	Singular Value Decomposition
SVM	Support Vector Machine
TMS	Traffic Management System
TN	True Negative
TP	True Positive
UWB	Ultra Wide-Band
VRUs	Vulnerable Road Users
WSD	Wavelet Shrinkage Denoising

CHAPTER 1

INTRODUCTION

Contents

In the first Chapter of this thesis book, we present the background and the problematic of this work. The discussion will be followed by the CYCLOPE project presentation from which the thesis work is one part. Then, followed by presenting general context of this research, the scope of the thesis presentation, the methodology used, thesis contributions, and the thesis organization.

1.1 Background

The increasing number of accidents and fatalities on the road has pushed the governments to make the road safety becomes one of the national priority concerns. Development of accident prevention systems is one of the focuses of this priority and it has also become one of the major concerns of automobile producers and researchers in the last decades in order to contribute to the improvement of the road safety.

Pioneered by the accident prevention systems, nowadays, the developments of the road safety systems have been shifted to the developments of Intelligence Transportation Systems (ITS) and Intelligent vehicles. With a huge improvement of the computer systems both hardware and software and the telecommunication technologies, it allows the vehicles to communicate among others and to the transportation infrastructures.

The integration of autonomous vehicles and ITS system, advanced mobility which a lot of safer, and environmental sustainability expected can be achieved soon or later. Related to the improvement of the road safety topic and toward the autonomous vehicles or drive-less cars system, the work of this thesis focus on the basic core of accident prevention system and it particularly concerns on developing a vulnerable road users identification that is robust to weather conditions.

Based on the Royal Society for the Prevention of Accidents, 95% of all accidents in the roads are caused by human error [1] and almost half of the road accident portions was the vulnerable road users including pedestrians, cyclists and motorcyclists [2]. The statistic shows that in 2015, 21% of the road victims were the pedestrians, 8% were the cyclists and 14% were the motorcyclists, and around 2000 people riding bicycle die every year in traffic accidents in EU countries [2][3].

One of the factors that contributes to the injuries and mortality of pedestrians and cyclists is mainly the lack of drivers visibility. As we know, there are many blind spots around the trucks and buses environments, that can make a fundamental crash accident between them and the uncovered road users [4]. Therefore, deploying an intelligent system of obstacles identification in automotive systems such as trucks and buses is crucial in order to have a better security in the road.

Over the years, many techniques have been proposed in order to reduce the injuries and mortality caused by road crash accidents. Crash warning and collision avoidance systems promise a great contribution in enhancing the protection of the pedestrians and the cyclists. This potential solutions include the computer vision for objects detection, identification as

suggested in [5] and the radar system. However, the performance of the camera-based system decreases when illumination is low, as an example in cloudy, smoggy, foggy and night conditions. On the other hand, the radar system is robust to such conditions, and it can provide accurately the distance both stationary and moving targets while the camera can't. By consequence, it detects and tracks their movement and can reduce this fatal error.

In order to achieve this objective, we propose to use Ultra-Wide Band (UWB) short-range radar to detect the movement of the vulnerable cyclists and pedestrians. The proposed radar uses UWB technology in order to reach high obstacles detection. This technology is characterized by a very low power density and a very small pulse width, from few picoseconds to few nanoseconds. The amplitude of the pulse should be normalized to comply with Federal Communications Commission (FCC) mask. The wide bandwidth used (several GHz) allows signals with a precise temporal resolution, a low probability of interception/detection and offers robustness against multipath fading [6]. Thus, the UWB radar offers a great interest in short range road safety applications [7].

In one side, since this technology is characterized by a very low power density, detection of radar targets that have low Radar Cross-Section (RCS) like pedestrians and cyclists is addressed for two main challenges. These challenges include in one hand, the ability to distinguish the target over the noises and static clutter and in the other hand, the ability to estimate the accurate detection targets position. Indeed, target detection for the low RCS is a complex process that includes such signal processing phases as raw radar data preprocessing, background subtraction and noise removal [8][9]. It is highly influenced by the capacity to distinguish successfully the useful radar information and the noise or clutter. This latter problematic becomes one of our focus studies in this thesis.

In another side, the use of UWB in the short-range radar system presents many advantages. Firstly, the brevity of UWB pulses with strong spectral contents makes it possible to obtain information on the target with a rich transitory response content. This allows the easy dissociation of various echoes at the reception stage. Then, the broad band spectrum authorizes to obtain results on the entire frequency band in a single measurement together with a strong capacity of detection. The pulse spectrum has abilities to penetrate through naturally screening materials [10]. Considering all these properties, UWB radar using very short pulses is of great interest for many applications such as obstacles detection and targets identification in short range[11].

Based on these advantages, in this thesis, we also proposed target recognition using machine learning dedicated to vulnerable road users especially for the pedestrians and the cyclists. Three methods have been investigated in this work i.e. Support vector Machines (SVM), Deep Belief Network (DBN) and Convolutional Neural Network (CNN).

The developed system was financed by the French National of Research Agency (ANR) through its project called CYCLOPE (Cyclo Protection Electronically) that will be presented briefly in the next section of this chapter. This project focuses on developing a technical solution to facilitate the integration of the cyclists into urban traffics by improving their safety and the comfort of the bus drivers. The UWB radar has been chosen for this system because it is a suitable candidate thanks to its immunity to various climatic and environmental con-

ditions and variations. This thesis was carried out in the IEMN (Institute of Electronics and Microelectronics and Nanotechnology) of Université Polytechnique Hauts-de-France (UPHF), Valenciennes and LEOST (Laboratory of Electronics, Wave and Signals for the Transports) of IFSTTAR villeneuve d'Ascq, Lille France.

1.2 Problematics

Since this technology is characterized by a very low power density, detection of radar targets that have low Radar Cross-Section (RCS) like pedestrians and cyclists is addressed for two main challenges. These first challenges include the ability to distinguish the target over the noises and static clutter and the second challenge consists in estimating the accurate detection targets position. Target detection for the low RCS is a complex process that includes signal processing phases such as raw radar data preprocessing, background subtraction and noise removal [8] [9], The target detection is highly influenced by the capacity to distinguish successfully the useful radar information over the noise or clutter. This problematic becomes one of our studies in this thesis beside the radar target recognition. Another challenge concerns the target recognition process. Indeed, in a real environment, recognizing the pedestrian and cyclist using UWB radar system is a complex problem and the solution to overcome this situation, requires a good machine learning system. Therefore in the current thesis work, we have evaluated, in the two different approaches, 1-D and 2-D radar data, the Support Vector Machine (SVM), Deep Belief Network and Convolution Neural Network (CNN) to recognize the radar target.

1.3 CYCLOPE Project

1.3.1 Brief Introduction

To accelerate energy transition, the promotion of soft mobility is becoming a major issue where cycling is one of the proposed solutions. However, in towns and cities, urban planning does not always favor cohabitation between transport modes such as cycling and buses. Our objective is to propose novel solution that facilitates the integration of the cyclist in urban traffic by providing the alerting systems to bus drivers. The project is organized in two parts, namely "Radio technology" and "Human and social sciences". The originality of the project relies on the strong interaction between the two parts to foresee the influence of the technical solutions chosen on the users behavior. Simulations should allow to test for the acceptability of the proposed solutions before experimental testing. The consortium relies on skills and expertise in Science and Information technology, Human and Social sciences, a technological SME-Small and Medium Enterprise- and public transport operator.

The objective of this national CYCLOPE project is to develop an innovative and low-cost technical solutions to facilitate the integration of cyclists in urban traffic by improv-

ing both their safety, the traffics flow and the comfort of bus drivers. We have a multi-disciplinary vision, STIC-Sciences and Technologies of Information and Communication- and HSS-Humanities and Social Sciences. Indeed, we are aware that technological innovation can not respond alone to the whole problem. The most targeted application in CYCLOPE is the collaborative sharing of bus lanes (or other public transport space). One means recommended in CYCLOPE is the exploitation of the radio signals emitted by the bus (communication signals to its terminal) using compact electronic devices embedded on the bicycle to alert cyclist of the bus presence in the lane adjacent to the bus lane so that Cyclist adopts the appropriate behaviors. The other means is to exploit signal radar to provide information to the bus driver. It is therefore not a matter of deploying a new radio link or modifying the current radar systems but of taking advantage of signals already present (which are diverted from their function initial or improved) by means of low cost devices integrated into the cyclist's equipment. More specifically, the CYCLOPE project consists of a "radio techniques" component and a "human and social sciences" component. One of the originality of CYCLOPE project consists in conducting the two parts in strong interaction in order to anticipate as soon as possible the impact of the technical choices made on the behavior of the users and to adapt the information delivered to favor the collaboration between them.

Another aim is to enhance the safety and sense of safety perceived by cyclists and bus drivers. Indeed, one of the brakes in the practice of advanced cycling by non-cyclists is the risk of a bicycle accident. It's clear that the incentive to cycle for citizens must first go through the demonstration of the safety of this means of transport. Safety is enhanced when bike lanes are offered or, at least, lanes sufficiently separated from those of cars. These other ways are naturally the clean or reserved ways of public transport (tram or bus). This security is also improved if cyclists get reliable information from their environment allowing them to anticipate the behavior of others and to adapt their own behavior. Thus, if the physical place is left to the cyclist to be cleared at the arrival of a bus, an element of security would be to inform the cyclist of the presence at the back of a bus, so that this cyclist can adapt his behavior. It also seems essential to study the behavior of motorists traveling in the lane adjacent to the bus when the cyclist riding in front of the bus is reintegrated into their lane. Indeed, this cyclist, informed of the arrival of a bus, could potentially arise unexpectedly while it can be partly or completely concealed by the bus. All these questions would find answers in the project CYCLOPE which proposes, through the approaches of technological studies and behaviors of the cyclists, to answer this societal problem, the cyclist's behavioral study approach is not limited to surveys, but is extended to the intensive use of a Bicycle Driving Simulator and Driving Simulator on various carefully selected topics and scenarios.

1.3.2 Relation between Thesis Work and CYCLOPE Project

This dissertation is one part of the national CYCLOPE project which aims to study and develop a radar system based on UWB technology in order to protect the vulnerable road users. The work is focused on detection and identification using Ultra-Wide Band (UWB) radar targets, particularly for pedestrian and cyclist.

1.4 The General Research Context

The Radar, stand for RAdio Detection And Ranging, works is based on radio-frequency waves that used to determine the presence of objects located in a close environment by analyzing parameters such as position, velocity, and direction of movement. Initially, the radar application was designed to military application, but since World War II, the radar technologies have been used in many industrial developments system. Now, many consumer applications have emerged in aeronautics, shipping, robotics, meteorology or even in driving assistant. In this thesis, we are particularly interested in UWB short range radars for automotive applications. The use of UWB in the short-range radar system presents many advantages. Firstly, the brevity of UWB pulses with strong spectral contents makes it possible to obtain information on the target with a rich transitory response content. This allows the easy dissociation of various echoes at the reception stage. Then, the broad band spectrum authorizes to obtain results on the entire frequency band in a single measurement together with a strong capacity of detection. The pulse spectrum has abilities to penetrate through naturally screening materials [10]. Considering all these properties, UWB radar using very short pulses is of great interest for many applications of obstacle detection and target identification in short range [11]. Based on these advantages, in this work, we have investigated the use of UWB short range radar for target detection and recognition of the vulnerable road users particularly the pedestrian and cyclist.

1.5 Thesis Scope

The scope of this thesis is to study and develop a recognition system based on UWB radar technology that is especially addressed to detect and identify the pedestrians and the cyclists. In this work, it is used only one UWB radar transceiver that allows detection of the distance and the determination of the nature of obstacle. For that, we used in this work, two different approaches of radar data, 1-D radar data (A-scan) and 2-D radar data (B-scan) that have been treated differently. In addition, in this work, we have analyzed possibilities of UWB radar to be used as a suited choice for developing a good protection system for the vulnerable road users.

1.6 Methodology

As mentioned in the previous section, the blind spots around the trucks and buses environments can be a fundamental cause of this crash accidents between them and these uncovered road users. This main problematic has motivated us to propose a better solution in order to facilitate the integration of the cyclists in urban traffics and to provide a better protection of the pedestrians as well. Our main objective is to encourage people to cycle safer by providing a better solution with the possibility of warning cyclists in case dangerous maneuvers of cars

and heavy goods vehicles (HGVs) occur[12].

This work is split in two stages i.e detection and recognition. In the first method of the detection stage, we have proposed and investigated a new robust UWB radar detector for 1-Dimensional raw radar data. It is a combination of HOS (Higher Order Statistics) and CA-CFAR (Cell-Averaging Constant False Alarm Rate) detector.

The first idea of using HOS is to have a good time delay estimation for the radar echo. As in literature [13], they used the HOS method to estimate time delay in unknown spatially correlated noise and the result is better than that of cross correlation method. Then compared to the second order statistics, HOS has better performance thanks to its ability to suppress the noise. Then, we compared the performance between automatic threshold CA-CFAR detector and HOS with fixed threshold value. From this comparison, we obtained that the performance of HOS detector with fixed threshold value is better than that of CA-CFAR. We noticed that the performances obtained using the CA-CFAR detector depend on the choice of the reference cell numbers. Indeed, if reference cell numbers is not properly set, CA-CFAR will result in a lot of false alarm or miss detection. Based on this fact, we proposed a robust UWB radar detector by combining the HOS and CA-CFAR. To do this combination, we firstly apply the HOS on the received signal. Indeed, not all the noise is eliminated, but most of them can be suppressed. As the noise has been eliminated, then we apply the CA-CFAR to have automatic threshold value. By doing this combination, we finally have an automatic threshold of UWB radar detector that is robust against the noise.

In the second part of detection stage, we proposed a novel noise removal method, that is a combination of Wavelet Shrinkage Denoising (WSD) and Higher Order Statistics (HOS). In this part, different types of noise removal techniques have been applied such as Principal Component Analysis (PCA), Singular Value Decomposition (SVD), WSD and HOS and the results have been analyzed. In addition, the signal-to-noise ratio (SNR) of the final results has been computed to compare the effectiveness of individual noise removal techniques.

For identification stage, different approaches have been investigated such as Support Vector Machines (SVM), Deep Belief Network (DBN) and Convolution Neural Network (CNN).

1.7 Thesis Contributions

This PhD work contributes to the area of the road safety by using UWB radar, particularly for detecting the presence or movement of the vulnerable road users that have a low RCS-Radar Cross Section- like pedestrians and cyclists. Specifically, based on the problematic described in the previous section. The detailed contributions of this work are:

1. The first, we have proposed, developed and tested a robust UWB radar detector by combining the HOS and CA-CFAR in the 1-D. This combination takes the advantages of the capability of HOS in suppressing the noise and of the ability of CA-CFAR in adapting with dynamic threshold values. As the result, this combination promises a good detector for 1-D UWB radar application and specifically, for cyclists and pedestrians detection.

2. The second, we have developed and tested the capability to recognize radar targets by using the normalized 1-D radar signature. In order to classify the radar target using this radar signature, we have evaluated two different types of machine learning algorithms, Support Vector Machine (SVM) and Deep Belief Networks (DBN). The input for both recognition methods is the normalized 1-D radar signature. The results show that both SVM and DBN can potentially be used efficiency to classify the UWB radar target. Moreover, DBN provides better accuracy than SVM.

3. The third, we have enhanced the capability of detection of this UWB radar by using 2-D approach. In this case, we have proposed a new method of noise removal. The combination of Wavelet Shrinkage Denoising (WSD) and the HOS have been applied and the signal-to-noise ratio (SNR) of the final results has been computed and compared with the other noise removal techniques. It is observed that a combination of WSD and HOS has better capability to remove the noise compared to the others; especially it is found that it allows to distinguish efficiency the pedestrian and cyclist over the noise and clutters whereas other noise removal techniques are not showing significant result.

4. The last but not the least, we have successfully applied the Convolutional Neural Network (CNN) to 2-D radar signature. After applying this latter, we use the sliding windows approach to search the candidates of the radar target. The 2-D radar signature can be obtained by implementing non-maximum suppression. This simple technique detects the energy surrounding of the radar target position and takes the one that is maximum, and eliminates the others. After transforming it into the power spectral, this signature can be used as an input of the CNN. Finally, we compared the CNN, DBN and SVM performance, and the results show that CNN gives better results in terms of accuracy compared to both DBN and SVM. And also, it allows efficient classification of UWB radar targets particularly the cyclist and pedestrian presenting low RCS.

As a final remark, it is important to acknowledge that most of the inspiration and motivation for this work remains a lot of challenges and we considered as the vision of the future perspectives. One of it is, to test this system in a real environment where the radar needs to be mounted on the vehicle such as bus or trucks. We have also combined this UWB radar system with the camera-based system in order to have a robust recognition system to protect the VRUs.

1.8 Thesis Organization

In Chapter 1, we presents the Introduction of Background and Problematic, national CYCLOPE Project, Thesis scope, Methodology, Thesis contributions and Thesis Organization. Chapter 2 makes a review on the State of the Art of Intelligent Transport System (ITS), System and Applications of Anti-Collision, System and Application of VRUs protection, and Ultra-wide Band Radar. Chapter 3 focuses on 1-D UWB Radar Targets Detection, the time delay estimation and detailed discussions on the proposed of robust UWB radar detector. Chapter 4 presents the Radar Targets Recognition on 1-D radar signature. In this chapter,

it will be discussed about the way the Support Vector Machine (SVM) recognizes the radar targets by using normalized 1-D radar signature as its input features, then followed by investigation of Deep Belief Network (DBN). In Chapter 5, we will focus on a new proposition of UWB radar target detection and identification using 2-D radar approach. In the first part of this section, the discussion will be focused on increasing radar Signal-to-Noise Ratio (SNR) by proposing a new noise removal method. We compared different methods such as Singular Value Decomposition (SVD), Principal Component Analysis (PCA), Wavelet Shrinkage Denoising (WSD), Higher Order Statistics (HOS) and our proposed method, a combination between HOS and WSD. After that, in the second part of this Chapter, a Convolution Neural network (CNN) is presented in order to enhance the recognition rate. Finally, the conclusion of this work and the perspectives are presented in the last Chapter 6.

1.9 List of Publications

Journals

- R. Sadli, C. Tatkeu, K. Hamidoun, Y. El Hillali and A. Rivenq, "UWB radar recognition system based on HOS and SVMs," in IET Radar, Sonar & Navigation, vol. 12, no. 10, pp. 1137-1145, 10 2018.

- K. Hamidoun, R. Sadli, Y. Elhillali, R. Elassali, A. Rivenq and K. Elbaamrani, 'Implementation of a New IR-UWB System Based On M-OAM Modulation On FPGA Component', International Journal of Wireless & Mobile Networks (IJWMN) Vol. 10, No. 2, April 2018.

- L. Maatougui, L. Sakkila, R. Sadli, A. Rivenq, "Computation of Correlation and Cumulant 4 under FPGA Architecture", International Journal of Engineering and Technical Research (IJETR), ISSN: 2321-0869 (O) 2454-4698 (P), Volume-5, Issue-3, July 2016.

Conferences

- R. Sadli, Y. El Hillali, L. Sakkila, A. Rivenq, C. Tatkeu, "The 4th Order Statistics on Ultra Wideband Short-Range Radar (UWB-SRR) for Cyclist Detection", 3rd Annual International Conference on Transportation 5-8 June 2017, Athens, Greece.

- R. Sadli, L. Maatougui, Y. El Hillali, C. Tatkeu, A. Rivenq, "Implementation of the 4th Order High Order Statistics (HOS) Based on FPGA for an UWB Short-Range Radar Receiver", Colloque International TELECOM'2017 & 10émes JFMMA 10, 11, & 12 Mai 2017, Rabat, Moroc.

Oral Presentations and Posters

- R. Sadli, Y. El Hillali, A. Rivenq, C. Tatkeu, "Implementation of the 4th Order High Order Statistics (HOS) Based on FPGA for an UWB Short-Range Radar Receiver", La journée Matinée De Chercheurs, 2017 à Mons, Belgique. (Poster).
- R. Sadli, Y. El Hillali, A. Rivenq, C. Tatkeu, "Study and Implementation of UWB radar in Combination with Vision-based System to Protect the Cyclists" La matinée des Doctorants IEMN-DOAE, Valenciennes, 20 october 2017. (Oral Presentation format Pecha Kucha).
- R. Sadli, Y. El Hillali, A. Rivenq, C. Tatkeu, "Development of Algorithms for Perception the Cyclists by the Bus", Les journées des doctorants 2017 du COSYS, Marne La Vallée les 22 et 23 juin 2016. (Oral Presentation)
- R. Sadli, Y. El Hillali, A. Rivenq, C. Tatkeu, "New UWB radar Recognition System Based on Higher Order Statistics and Support Vector Machine", Journée du club EEA 'véhicule autonome – transport intelligent', Valenciennes 9-10 novembre 2017. (Oral Presentation).
- L. Maatougui, R. Sadli, "Développement d'un système de détection et de suivi des piétons", Journée du club EEA 'véhicule autonome – transport intelligent', Valenciennes 9-10 novembre 2017. (Poster)

CHAPTER 2

STATE OF THE ART

Contents

This chapter will expose about the State of the Art relative to this research topic including the Intelligent Transport System (ITS), overview of Autonomous Vehicle, Environmental Perception Systems, Radar Systems, Radar Detector and Noise Removal Method.

2.1 Intelligent Transportation System (ITS)

Intelligent Transport Systems (ITS) are advanced technologies including Advanced Traffic Management (ATMS), Advanced Traveler Information Service (ATIS), and Advanced Vehicle Control and Safety (AVCS) [14] which aims to provide innovative service relative to different modes of transportation and management. ITS is a system that uses computers, controls, communications, and various automation technologies in order to improve the road safety, throughput, and efficiency of transportation. It is also targeted to reduce energy consumption and environmental impact [15].

By integrating communication and information technology applications into management and operation of the transportation system across all modes [16], These technologies enable users to be better informed and make safer, more coordinated, and smarter use of transport networks [17]. An advancing transportation safety, mobility, and environmental sustainability can be achieved by using these technologies.

The ITS has emerged in the year of 1991's since the electronic technologies have taken place an important role in optimizing surface transportation followed by dramatically improved the computer both hardware and software, communication and sensor technologies [18].

2.1.1 Applications of ITS

Nowadays, ITS has been integrated in several applications [19]:

Electronic toll collection (ETC); this technology equips vehicles with sophisticate communication system in order to collect the tolls electronically. The performance of toll service stations can be increased now by minimizing the delay.

Highway data collection; this system provides the traffic information to the traffic control center to analyze and identify areas of congestion on the road network. The collection of data are gathered by using Geo-referencing systems such as GPS, GSM, and GPRS.

Traffic management system (TMS); A system that plays an important role in management transportation that collects real-time information from different hardware components like cameras and speed sensors to improve the flow of vehicle traffic and safety.

Emergency vehicle pre-emption: This system makes it possible to provide these types of vehicles with solutions enabling them to circulate without problems in a road network even in situations of congestion without increasing congestion on the road network.

Road condition monitoring: [20], It proposes a road condition monitoring system using multiple sensors data for ensuring communication between vehicle and providing as a web application if some disaster occurs.

2.1.2 Key Underlying ITS Technologies

In this subsection, we present several key underlying technologies for ITS systems, such as Computer Technologies, Wireless Networks, Detection Technologies, Global Positioning System (GPS) and Anti-Collision Technologies.

Computer Technologies

Computer technologies play an important rule in developing ITS technologies. With the evolution of hardware and software dramatically, it allows us to integrate ITS technologies into transport infrastructures and management system. Computer technologies is a main key to solve problems in transport systems.

Wireles Network

The wireless networks are used to establish communication between vehicles and roadside infrastructures.

It is a standard communication commonly used for wireless internet access [21]. The technologies used are known: Wi-Fi, WiMax,DSRC, MBWA, UWB, etc [19].

Detection Technologies

Detection technologies including static and mobile detection is aimed to gather information related to the the road users activities. Static detections using sensors or detectors fixed along the road such as loops and magnetic sensors, cameras, acoustic RF sensors [19]. Mobile detection exploiting GPS on vehicles, or those installed on mobile phones and hybrid detection [19].

GPS technologies

GPS technologies have been integrated on the vehicles in order to localize their position. [21] GPS is a core technology behind the vehicle navigation and route guidance systems.

Anti-Collision and Recognition Technologies

Nowadays, with the increasing number of vehicles circulating in the roads, the technologies of anti-collision and obstacles recognition are necessary embedded on the vehicles. Radar, Lidar, Ultra-sonic, and Camera are the core of this technologies.

To summarize the discussion about the intelligence transportation system, the emerging of ITS is aimed to provide an innovative service that allows the road users to be better informed and make them safer, more coordinated and smarter use of transport networks. ITS technology is an integration of communication and information technology into different modes of transportation managements which brings a number of social, environmental, and economical benefits.

Relating to ITS technology and still in its framework, intelligent vehicles is important to be mentioned in this state of the art. Intelligence vehicles is a promising technology for future transportation systems.They share cooperatively essential information between vehicles and roadside units or infrastructures [22] and surely, they drives autonomously. The way it perceives the surrounding environment is a key of this technology. Therefore in the following section, we will discuss about this technology.

2.2 Overview of Autonomous Vehicles

Even though a statement that human beings are poor drivers as documented in popular culture is over-dramatized. The facts that human can sometimes be at times distracted, drowsy, drunk, drugged, and irrational decision maker are the proof evidences that can not be denied [23]. An extensive investigations of recent years showed, the large majority of road traffic accidents (50–70%) happened because of wrong actions of the drivers. The correctness of the decisions taken by the driver and driving skills are the basis for the majority of the reasons for the road traffic accidents [24]. These limitations led the researchers all over the world to develop a perception-control vehicle system called intelligent vehicle to overcome the human constrains.

These vehicles can communicate and share information with the ITS infrastructure and between self driving cars. However, designing and building the autonomous system that drives better than the average human drive is not easy. There is a complex problematic such as problems of localization, mapping, scene perception, vehicle control, trajectory optimization, and higher-level planning decisions associated with autonomous vehicle development still remain full of open challenges that have need to be fully solved [23] before this system is safely to be used. However, driver-less vehicle or autonomous vehicle is a promising technology for future transportation system. It potentially has the capacity to improve road safety and to have a better mobility. It also potentially will be used in order to overcome the human mobility constraints particularly for elderly and disabled people [25].

Nowadays, autonomous vehicle system has become one of the hottest topics in the artificial

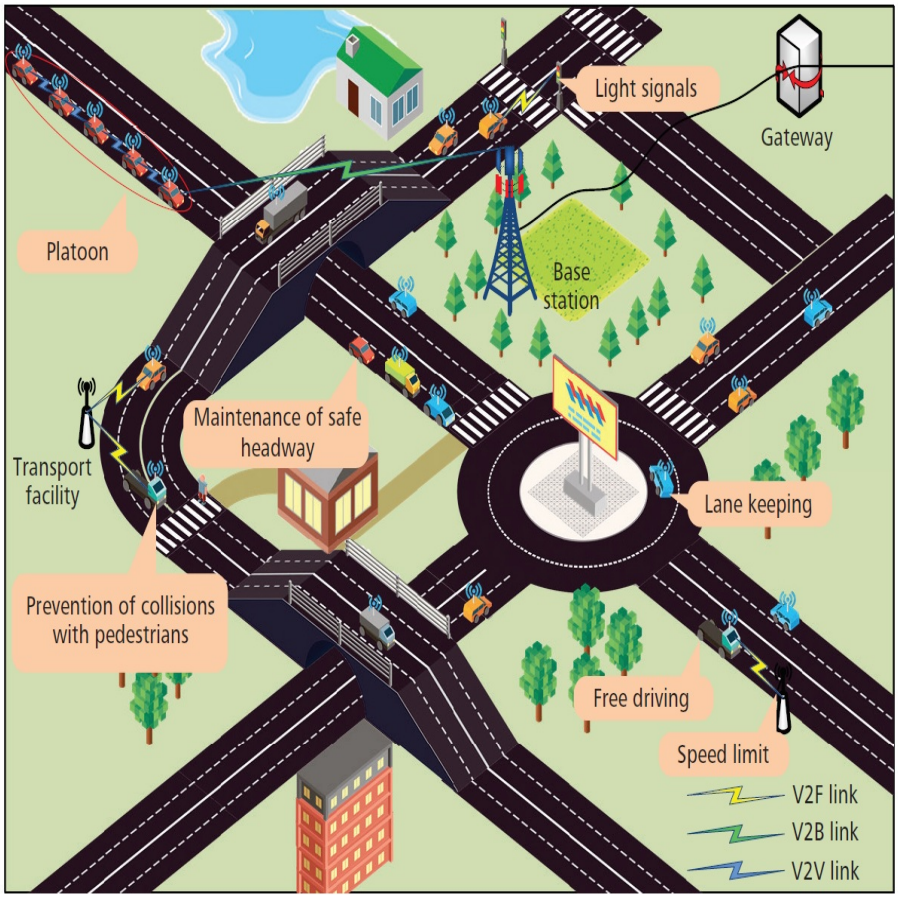


Figure (2.1) Illustration of connected vehicles and infrastructure [28]

intelligent research domain. In the last few decades, autonomous vehicles was considered just a science-fiction. Nowadays, we can find the technologies that allow a car to self-park [26] and to break automatically in emergency conditions. Soon, we will see cars will navigate without a driver under full conditions. Their presence will not only change the way of moving, but also will have an impact on the social evolution of the society, security and environment.

2.2.1 Autonomous Vehicles Operational Levels

There are five levels of autonomous driving vehicles defined by the Society of Automotive Engineers (SAE)[27]. The levels 1, 2 and 3 required driving license but levels 4 and 5 are driver-less. Tabel 2.2 [29] and figure 2.3 shows the summary of vehicles autonomous levels.

Level	Name	Narrative definition	DDT		DDT fallback	ODD
			Sustained lateral and longitudinal vehicle motion control	*OEDR*		
Driver* performs part or all of the *DDT						
0	**No Driving Automation**	The performance by the *driver* of the entire *DDT*, even when enhanced by *active safety systems*.	*Driver*	*Driver*	*Driver*	n/a
1	**Driver Assistance**	The *sustained* and *ODD*-specific execution by a *driving automation system* of either the *lateral* or the *longitudinal vehicle motion control* subtask of the DDT (but not both simultaneously) with the expectation that the *driver* performs the remainder of the *DDT*.	*Driver* and *System*	*Driver*	*Driver*	Limited
2	**Partial Driving Automation**	The *sustained* and *ODD*-specific execution by a *driving automation system* of both the *lateral* and *longitudinal vehicle motion control* subtasks of the *DDT* with the expectation that the *driver* completes the *OEDR* subtask and *supervises* the *driving automation system*.	***System***	*Driver*	*Driver*	Limited
***ADS* ("*System*") performs the entire *DDT* (while engaged)**			*System*	***System***	*Fallback-ready user (becomes the driver during fallback)*	Limited
3	**Conditional Driving Automation**	The *sustained* and *ODD*-specific performance by an *ADS* of the entire DDT with the expectation that the *DDT fallback-ready user* is *receptive* to *ADS*-issued *requests to intervene*, as well as to *DDT performance-relevant system failures* in other *vehicle* systems, and will respond appropriately.				
4	**High Driving Automation**	The *sustained* and *ODD*-specific performance by an *ADS* of the entire *DDT* and *DDT fallback* without any expectation that a *user* will respond to a *request to intervene*.	*System*	*System*	***System***	Limited
5	**Full Driving Automation**	The *sustained* and unconditional (i.e., not *ODD*-specific) performance by an *ADS* of the entire *DDT* and *DDT fallback* without any expectation that a *user* will respond to a *request to intervene*.	*System*	*System*	*System*	**Unlimited**

Figure (2.2) Roles of human driver and driving automation system by level of driving automation [29]

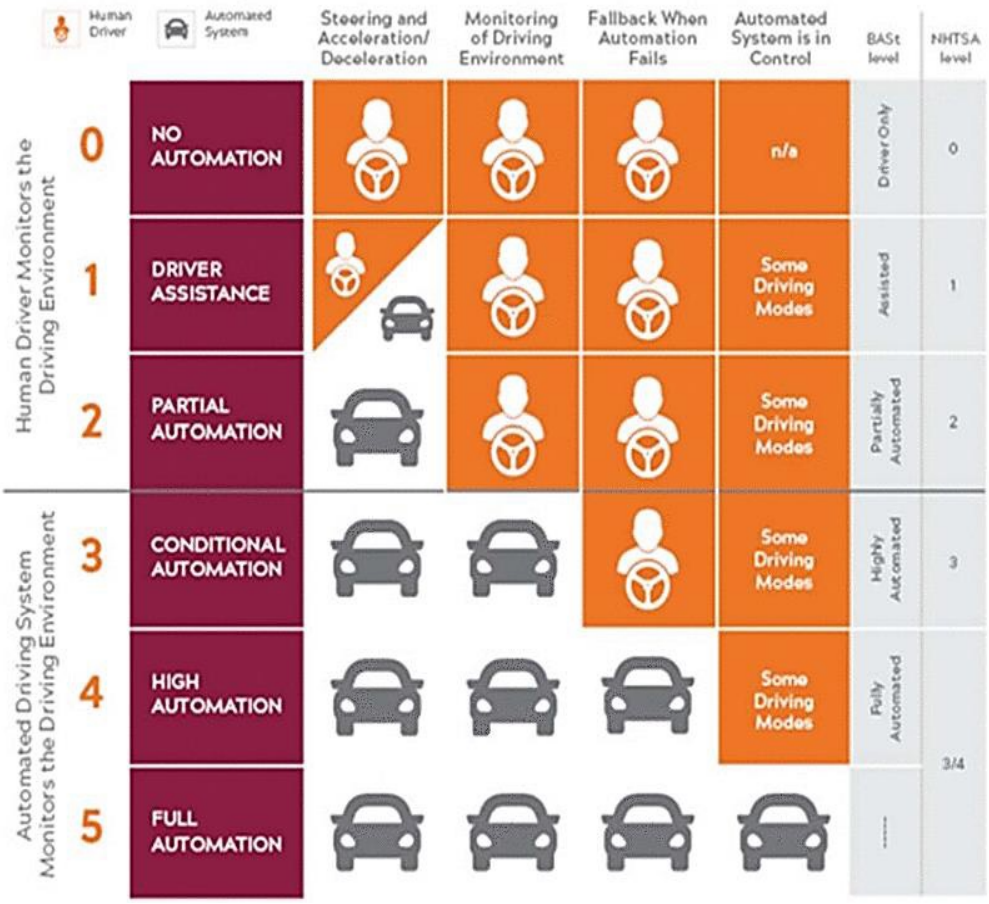


Figure (2.3) Autonomous Vehicles Levels [27]

2.2.2 Key Technologies in Autonomous Vehicle

This technology originated from multi-disciplinary scientific researches such as vehicle control systems, robotics, sensors, machine perception, innovation, mapping and machine learning [30]. The important keys in creation of this autonomous vehicle are telecommunication and navigation systems, sensors technologies and artificial intelligence systems. These core technologies can be derived into four parts based on the functionality of self driving car [31]: environment perception, car navigation, path planning, and the car control.

2.3 Environmental Perception Systems

Perceiving surrounding environment, so that the drivers can interact consequently, is the key success of reducing the road traffic accident. As earlier mentioned in the chapter 1, the focus of this thesis work is on developing a system that the bus can perceive the movement of the VRUs particularly in the blind spot areas, so that the system can give spatial alert to the bus driver to be aware of the presence of the VRUs around them. Therefore, related to this

topic, there are two main concerns of the discussion in this section: the anti-collision and the Vulnerable Road Users (VRUs) protection systems and applications.

2.3.1 Systems and Applications of Anti-Collision

In recent years, the obstacle and collision avoidance have attracted a lot of researcher's attention. With the aim of developing a technology that can facilitate to prevent or to reduce road accidents due to human error and to make drivers comfort, this technology has become very popular in these days. In order to achieve a good level precision, there are two difference technologies approach have been used, passive and active sensor system [32] [33] that it depends on the applications.

In passive technology, a sensor receives signals coming from environment. This kind of sensors includes in particular stereo-cameras or optical-flow cameras [33]. While, the active system works based on the response (information gathered) of transmitted signals from a sensor to analyze the actual situation.

There are many research projects and applications have been conducted and developed relating to Anti-collision systems and the some of them will be discussed in the following paragraphs.

In [34], they used a combination of the laser beams and camera to estimate position of the vehicle after 1 second, then project it on the road surface. The installed camera on the vehicle is used to capture the positions of each laser beams. The obstacles can be detected by measuring the difference between normal and abnormal road conditions.

In [35], an Anti-Collision System (ACS) that allows Automatic breaking system, has been developed by using a stereo multi-purpose camera and ultrasonic sensor. The camera has been used to get dimensional data of the vehicle and its environment and the obstacle is detected by ultrasonic sensor. These information are then sent to an automated emergency braking system that breaks the vehicle at an appropriate time and condition.

In [36], they developed a collision avoidance system in heavy traffic where the vehicles speed normally less than about 20 km/h. This system puts 8 ultra-sonic sensors to cover all around the car.

In [37], they reported that an autonomous cruise control (ACC) system was firstly introduced by Mercedes-Benz in 1999 for their S-class series car. This ACC is a 77 GHz radar-based. Since then, several type of radar system have emerged in developing the autonomous car system. They are based mainly on one of the following technologies: FMCW, FSK-CW (Frequency Shift-Keying-Continuous Wave) and MFSK-CW (Multiple Frequency Shift Keying-Continuous Wave). For the Short Range Radar (SRR) technologies mainly use UWB technologies: UWB pulse, spread spectrum techniques and radar "Stepped Frequency".

National Highway Traffic Safety Administration U.S. Department of Transportation entered a collaborative research project in Advanced Collision Avoidance System/Field Operational Test (ACAS/FOT) Program. The members are: Delphi-Delco Electronic Systems,

Delphi Chassis, HRL Laboratories, HE Microwave and UMTRI. In ACAS the FCW (Forward Collision Warning) and ACC (Adaptive Cruise Control) functions are implemented using a combination of a long-range forward radar-based sensor, a forward vision-based sensor and a Global Positioning Satellite (GPS). The objective of this research is to determine the practical suitability of the combined ACC/FCW function for widespread use by the driving public [38].

2.3.2 Systems and Applications of VRUs Protection

Many techniques have been proposed over the years in order to reduce the injuries and mortality caused by road crash accidents including the video and radar techniques. PROTECTOR, SAVE-U (Sensors and system architecture for Vulnerable road Users Protection) and PROSPECT (Proactive Safety for Pedestrians and Cyclists) were the EU projects concerned in protecting the uncovered road users. SAVE-U concerned in developing sensor-based driver assistance system by integrating three different technologies of sensor (radar, IR, camera) simultaneously using sensor fusion to optimize VRUs (vulnerable Road Users) detection [39]. In the PROTECTOR, they focused also on three sensors. Beside the camera and laser scanner, the microwave radar was used for obstacle detection, around the vehicle. Using this kind of radar, detection of pedestrian among the other objects is done by evaluating the reflected power, the power variant over tunnel, the dimension and prevailing dynamic of the obstacle [40]. PROSPECT is a collaborative research project funded by the European Commission, aimed at improving the protection of vulnerable road users (VRUs) with an emphasis on the two groups with the largest shares of fatalities: cyclists and pedestrians. The project started in May 2015 and involves many relevant partners from the automotive industry, academia and independent test labs [41].

The PROTECTOR and SAVE-U exploited the microwave 24 GHz narrow band radar and PROSPECT used 77 GHz narrow band high resolution radar. In one hand, using a narrow band radar requires a very complex process to extract information in order to classify correctly the type of radar target. On the other hand, using UWB radar is interesting to be investigated for this application because it has a rich transitory response thanks to its large bandwidth used. So, theoretically, it has a good capacity in target recognition.

After reviewing some of the anti-collision and VRUs protection systems above and compared to another system like cameras, lidar and ultra-sonic, the radar system has some additional important advantages of robustness in all weather conditions, good estimation of target distance, and faster target discrimination. The radar systems are suited for every weather condition like foggy, rainy and etc. Therefore, our next discussions will be focused on the radar systems.

2.4 Radar Systems

The Radar, RAdio Detection And Ranging, works by transmitting and receiving the electromagnetic waves. The analysis of their reflections is used to identify the presence and determine the parameters of objects located in the environment (position, speed, direction of movement, etc). Since the Second World War, radar technologies have been the subject of many industrial developments. While initially the applications were purely military, many consumer applications have emerged in aeronautics, shipping, robotics, meteorology or driving. Radar transmits and receives electromagnetic waves which aims to determine the range, velocity and altitude of the target objects.

Basically, the radar has two main different types, that are narrow band radars and wide-band radars. The narrow band radars are mostly used in long range applications and the wide-band radars are mostly used in short range applications. The VRUs protection systems normally need a short range radar with a high precision. In this case the UWB radar is suitable for this application.

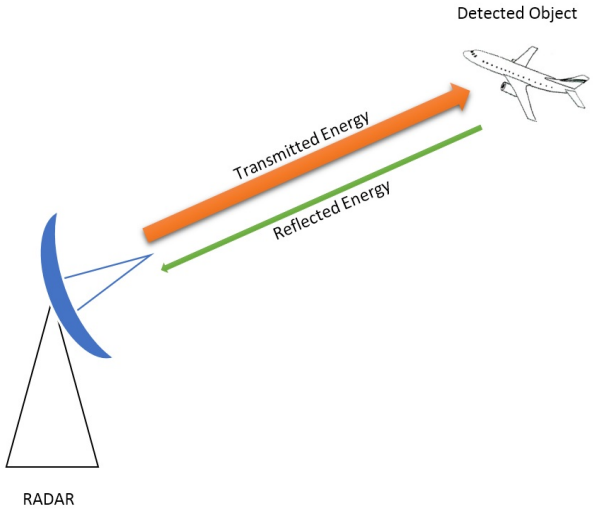


Figure (2.4) Basic Principle of Radar

So, that is way in this thesis work we are interested in developing an identification system for VRUs, especially for the pedestrian and cyclist, using UWB-SRR.

2.4.1 Radar Equation

When a target hit by the radar signal, it will reflect the radar energy back to the receiver. The amount of the radar receives this energy back depends on the the target size, orientation, shape and material that is proportional to a parameter called Radar Cross Section (RCS). It is the ratio of the power reflected by the target to the power density incident on it, and is

denoted by σ [42]. The power density (P_D) of the radar is calculated as:

$$P_D = \frac{P_t G}{4\pi R^2} \tag{2.1}$$

where P_t is the transmitted power, G is the gain of antenna, and R is the distance between measured position from a radar antenna.

Thus, the total power received by a radar receiver antenna is written as following:

$$P_r = \frac{P_t G \sigma A_e}{(4\pi R^2)^2} \tag{2.2}$$

where A_e is the effective antenna aperture that related to the antenna gain by:

$$A_e = \frac{G\lambda^2}{4\pi} \tag{2.3}$$

λ is the wavelength of transmitted signal

Substituting 2.3 into 2.3, then P_r can be written as:

$$P_r = \frac{P_t G^2 \lambda^2 \sigma}{(4\pi)^3 R^4} \tag{2.4}$$

2.4.2 Range Resolution

Range resolution is the ability of a radar to distinguish between two or more different targets at different ranges. The range resolution ΔR is given by:

$$\Delta R = \frac{c\tau_p}{2} = \frac{c}{2B} \tag{2.5}$$

where $B = \dfrac{1}{\tau_p}$ is the signal bandwidth.

2.4.3 Different type of Radar

Pulse Radar

These radars transmit a signal formed of pulses of radio waves periodically. The pulse consists of a frequency shifted carrier which is generated by mixing the local oscillator f_{LO} with an intermediate frequency (IF) signal at frequency f_{IF} in the transmitter. The working principle of this radar is that the pulse generator creates a pulse with duration τ and repeated in every T_r second ($T_r = 1/PRF$), PRF stands for Pulse Repetition Frequency [43].

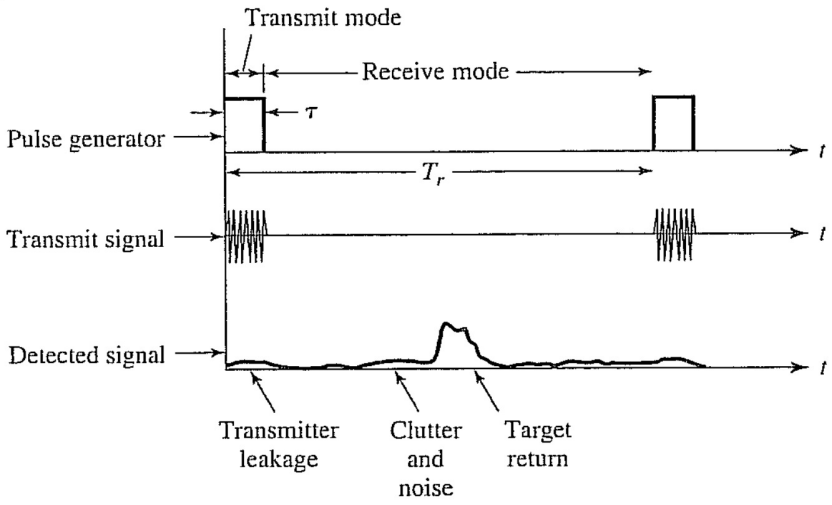


Figure (2.5) Description of pulse radar system (source:[43])

Radar FMCW

Frequency Modulated Continuous Wave (FMCW) radar is a radar transmitting a continuous carrier modulated by a periodic function. The signal is transmitted continuously by increasing and decreasing frequency ramps with a duration T_{CPI} and a frequency width f_{sweep}. The use of two frequency ramps makes it possible to obtain distance and speed information thanks to the high and low beats frequencies. The distance R and velocity v_r can be determined by [37]

$$\begin{cases} R = -\dfrac{(\Delta f_1 - \Delta f_2) c T_{CPI}}{4 f_{sweep}} \\ \\ V_r = -\dfrac{\Delta f_1 - \Delta f_2}{4} \lambda \end{cases} \tag{2.6}$$

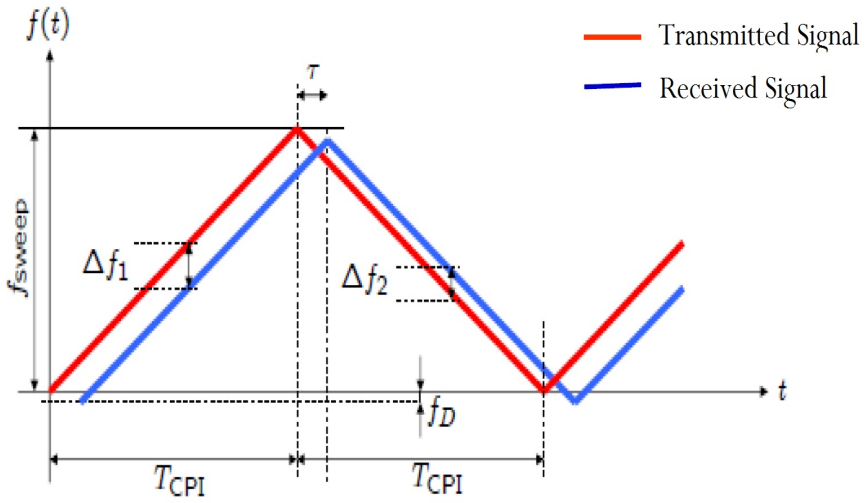


Figure (2.6) FMCW signal with two ramps, transmitted and received signal with delay τ and Doppler f_D (source:[37])

Doppler Radar

Doppler radar uses a fix transmitted frequency to detect a moving target and its velocity [44] where transmitter and/or receiver are in motion [43]. It works based on the well-known Doppler effect that is when a constant frequency source moves toward or away from an observer, it produces a perceived up-shift or down-shift in frequency [43].

If a radar transmitter sending a signal at frequency f_0 and hit the moving target, the returned signal has the frequency $f_0 + f_d$, where f_d is the Doppler frequency shift given by [44]

$$f_d = \frac{2V_r}{c} f_0 \tag{2.7}$$

where c is the speed of light and v_r is the relative velocity of the target perceived by the radar and it is given by:

$$v_r = v_a cos\theta \tag{2.8}$$

where v_a is the actual velocity of a target and θ is the angle between the target trajectory and the line-of-sight, as illustrated in the figure 2.7.

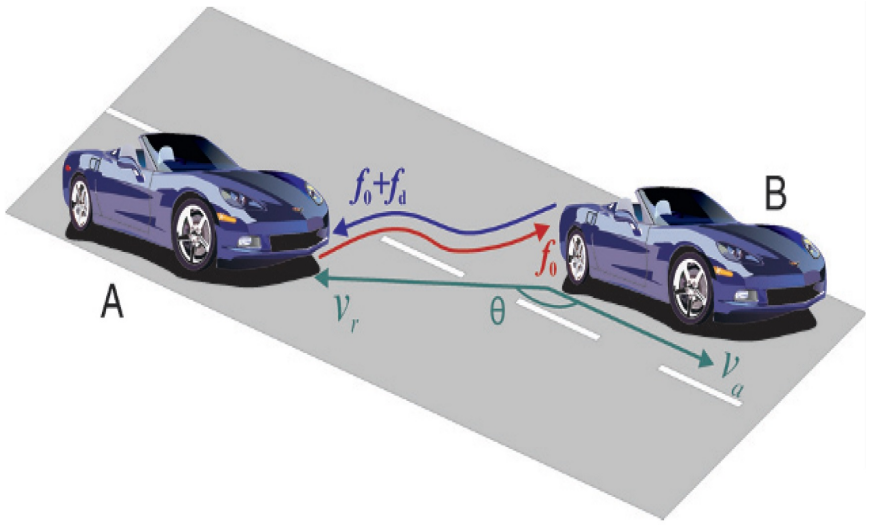


Figure (2.7) Illustration of Doppler Radar (source:[44])

2.5 Ultra-Wide Band Radar

In this thesis work, we are interested in investigating the capability of using UWB Radar in target identification particularly for the cyclist and pedestrian. To begin this discussion we are going firstly to look into UWB technology in general.

2.5.1 Ultra-Wide Band Technology

In February 2002, US Federal Communication Commission (FCC) permitted unlicensed UWB operation and commercial deployment of UWB devices as issued in the First Report and Order (R&O). Based on the R& O document, there are three classes of devices defined: (i) imaging systems (e.g., ground penetrating radar systems, wall imaging systems, through-wall imaging systems, surveillance systems, and medical systems), (ii) vehicular radar systems, and (iii) communications and measurement systems [45].

For wireless communications, the FCC regulated power levels are very low (i.e., –41.3 dBm/MHz in the frequency range of 3.1–10.6 GHz) [45][46]. This level is set to protect and avoid the interference with the existing communication systems. FCC has assigned two FCC masks for the indoor and outdoor UWB devices. For the indoor and outdoor UWB communications, the FCC radiation limits in the frequency range of 3.1–10.6 GHz are alike. While for the 1.61–3.1 GHz frequency range the outdoor radiation limit is 10 dB lower than the indoor mask [46].

In European, the organizations which involved in the regulation of UWB are ETSI (European Technical Standard Institute) and ECPT (European Conference of Postal and Telecommunications administration). These institutions have been working to conduct UWB compatibility and spectrum sharing studies and device regulatory mechanisms [46]. After completion

of spectrum compatibility studies by ECPT, ETSI establishes two TGs (ERM/TG31A and ERM/TG31B) to develop UWB regulation and standards for the European Union [46]. The ERM/TG31A is responsible for identifying a spectrum requirement and developing radio standards for short range UWB devices. The ERM/TG31B is responsible for developing standards and system reference for automotive UWB applications in higher frequency bands [47][48].

Within ECPT, SE24 is responsible for making regulation and spectrum management. One of the issues studied is spectrum sharing for frequency less than 6 GHz [45].

2.5.2 UWB Waveform

UWB impulses are typically narrow time pulses of sub-nanosecond or picoseconds order. The amplitude of the pulse should be normalized to comply with the FCC mask [45]. The short pulse UWB offers a great interest in short range radar application [49]. There are many types of waveforms that can be used to generate the UWB pulse. The Gaussian monocycles are commonly used as UWB impulses. A generic Gaussian pulse is given by [50]:

$$p_G(t) = \frac{A}{\sqrt{2\pi}\sigma} e^{(-\frac{t^2}{2\sigma^2})} \tag{2.9}$$

where A is the amplitude and σ is the spread of the Gaussian pulse. Gaussian monocycle can be defined as the first derivative of the Gaussian function as [50] [51]:

$$p_{G_1}(t) = -\frac{At}{\sqrt{2\pi}\sigma^3} e^{(-\frac{t^2}{2\sigma^2})} \tag{2.10}$$

Apart from the Gaussian function, the Gegenbauer polynomial also can be used to generate UWB pulse as introduced in [52]. The first four-order of Gegenbauer Polynomial is given by:

$$G_u(0,1,x) = 1 * (1-x^2)^{(1/4)} \tag{2.11}$$
$$G_u(1,1,x) = 2x * (1-x^2)^{(1/4)} \tag{2.12}$$
$$G_u(2,1,x) = (-1+4x^2) * (1-x^2)^{(1/4)} \tag{2.13}$$
$$G_u(3,1,x) = (-4x+8x^3) * (1-x^2)^{(1/4)} \tag{2.14}$$

Based on the published result on [49], the monocycle pulse gives the better performance for ultra wideband radar applications in terms of efficiency (ratio dynamics to peak width), implementation complexity and distance resolution reached, followed by the third order Gegenbauer. Figure 2.8 and 2.9 show respectively the plots of the Gaussian and its derivatives and the plots of the first four-order of Gigenbauer Polynomials.

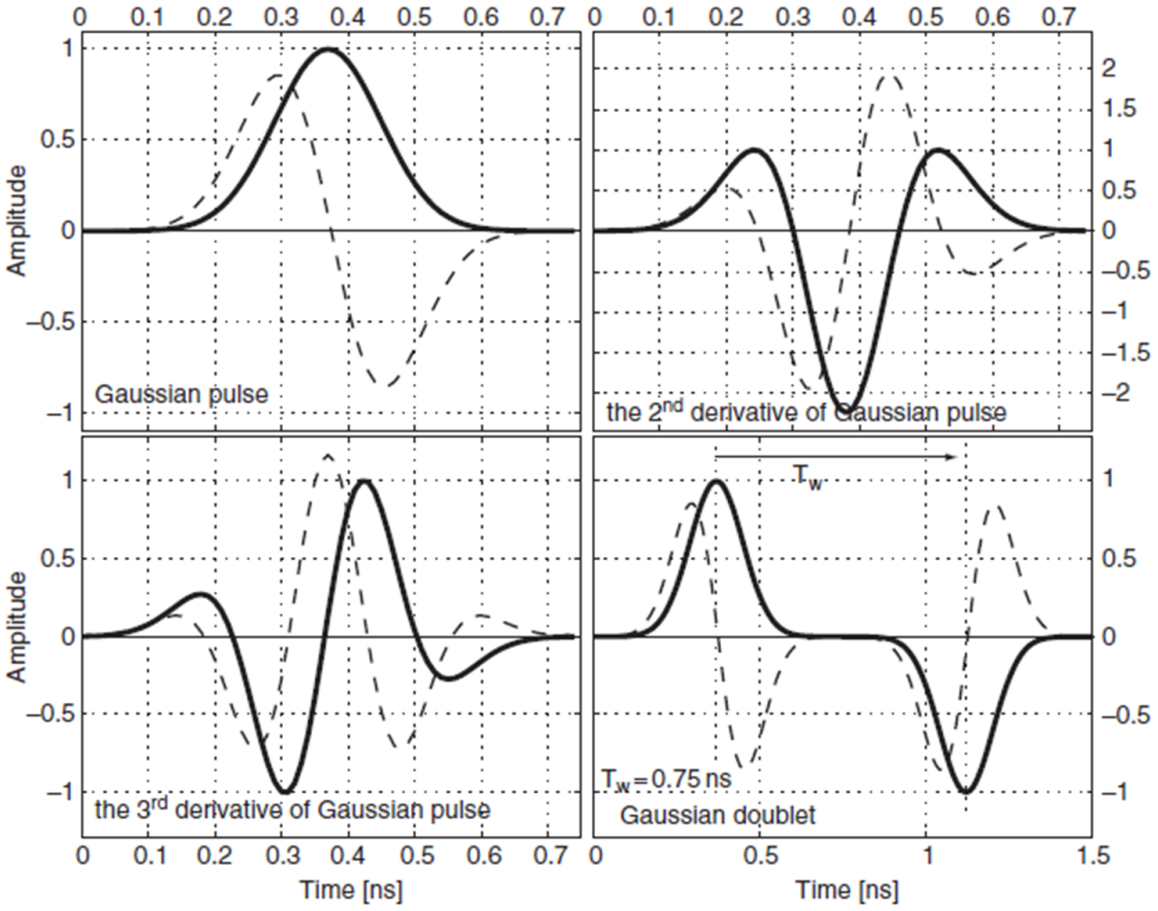


Figure (2.8) Gaussian-pulse-based waveforms [51]

2.5.3 Impulse UWB Radar

The emerge of Ultra Wide Band (UWB) technology for the radar application allows us to develop the low power and low cost radar sensors. Initially, intended for military purpose applications, UWB radar technology now has been developed to various applications. UWB radar has a good obstacle detection capability, it can be used to measure distances or positions with a better resolution than existing narrow band radar devices. It also can be used to locate the buried objects under the ground or placed behind the wall.

Ultra-Wide-Band (UWB) technology is based on the transmission of very short pulses with relatively low energy, which make it a promising option for short range radar and wireless communication systems [6].

The UWB Short-Range Radar (UWB-SRR) sends very short electromagnetic pulses. These pulses are typically narrow time pulses of sub-nanosecond or picoseconds order. The amplitude of the pulse should be normalized to comply with the FCC mask [53]. The short pulse UWB offers a great interest in short range radar application [7] because it has many advantages, such as:

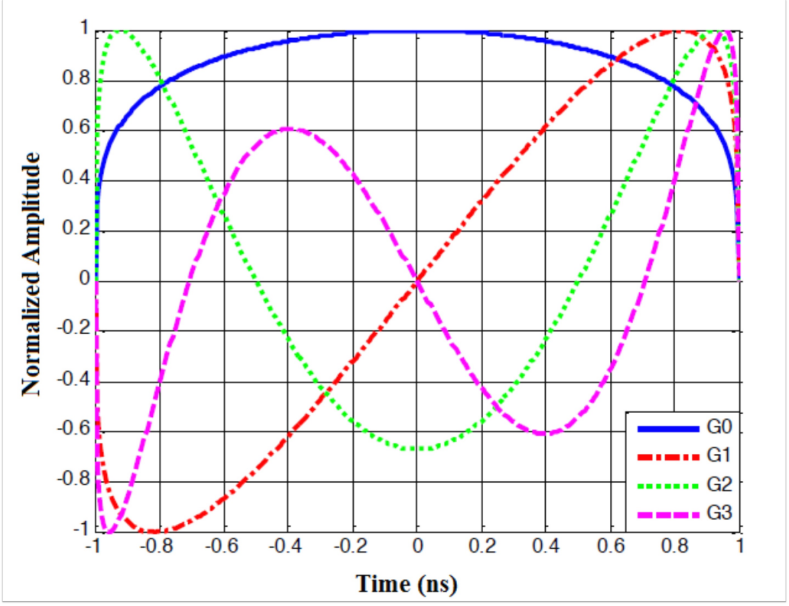


Figure (2.9) Plot of the first four-order of Gegenbauer Polynomial [52]

1. Its relative simplicity and likely lower cost to build than spread spectrum radios
2. Its substantially low consumed power, lower than existing conventional radios
3. Its implementation as a simple integrated circuit chipset with very few off-chip parts
4. Its high bandwidth capacity and multi-channel performance
5. Its high data rates for wireless communications

To measure the target distance, we measure the time delay $\triangle t$ between emission and reception. This distance d is given by the following equation:

$$d = (c.\triangle t)/2 \tag{2.15}$$

where c is the speed of light.

Using Ultra-Wide-Band (UWB) technology in the field of radar system gives many advantages such as: it allows detection of moving targets without using Doppler effect, its ability to measure both stationary and moving objects on and nearby the road, different materials and environments distort of pulses differently that allows this information to be used for better object identification [53]. Therefore, this system is very challenging for cyclist detection applications.

2.6 Radar Detector and Noise Removal Method

Constant False Alarm Rate (CFAR) detector is a very well-known radar detector. Its threshold is adaptively estimated based on local information on the background noise, from

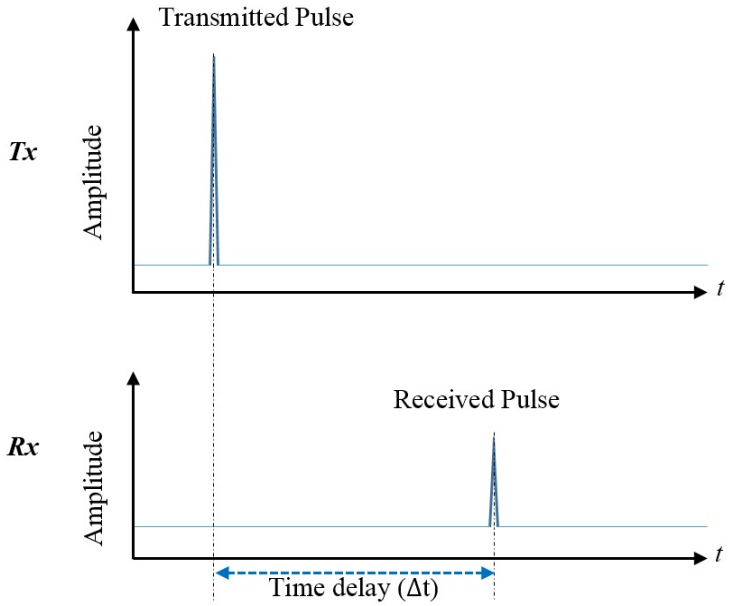


Figure (2.10) Basic UWB Radar Principle

both leading and lagging cells (called reference cells) surrounding the cell under test (CUT) [54][55][56]. The problem with this CFAR method is the difficulty to determine the number of the reference cells.

Since the technology of UWB radar is characterized by a very low power density, detection of radar targets that have low radar cross-section (RCS) like pedestrians and cyclists is addressed for two challenges. These challenges include the ability to distinguish the target over the noises and static clutter and estimate the accurate detection targets position. Target detection for the low RCS is a complex process that includes such signal processing phases as raw radar data preprocessing, background subtraction and noise removal [8][9], The target detection is highly influenced by the capacity of distinguishing the useful radar information over the noises or clutters.

There are many noises removal methods used in radar system like Principal Component Analysis (PCA) [57], Singular Value Decomposition (SVD) [58] and Wavelet-based [59][60]. Beside of these popular methods, higher order statistic (HOS) is also commonly used in noise suppression and time delay estimation for target detection. HOS consists of higher order moment spectra, which is defined for deterministic signals and Cumulant spectra for random process [61]. Indeed, many real world applications are truly non-Gaussian. Therefore, contrary to the second-order statistics, higher order statistics are applicable when we are dealing with non-Gaussian, because it can adapt to non-Gaussian noise or to non-linear channel characteristics and reveal the phase information [62] [63]. In [13], HOS has been used to estimate time delay in unknown spatially correlated noise and the result is better than that of used cross correlation method. Wavelet-denoising is also good in signal denoising, it has two terms in context of noise processing, smoothing and denoising. Smoothing removes

high frequencies and retains low frequencies whereas denoising attempts to remove whatever of the noise present and sustain whatever of the signal regardless of the frequency content of the signal [**?**]. Therefore, in order to obtain a robust radar detector, HOS is a good candidate to be combined either with CFAR to have robust automatic radar detector or with the Wavelet to have a robust noise removal method.

2.7 Conclusion of Chapter 2

This chapter has elaborated on the related State of the Art with the topic of this research including the Intelligent Transport Systems (ITS), Autonomous Vehicles, Anti-Collision Systems and Applications, Vulnerable Road Users (VRUs) Protections Systems, Ultra Wideband Radar System, Radar detector and Noise Removal Method. As it has been discussed in the Introduction Chapter, the main objective of this research is to develop a VRUs identification system using UWB radar. This system can be used to increase the awareness of bus/truck drivers if a VRU like pedestrian or cyclist is approaching, and this system can also be used as the additional sensors in the autonomous vehicles system in order to have a good environment perception particularly to cover the limitation of the camera based system at night or in the bad weather condition like foggy, rainy, and etc. In the next chapter 3, we will present the investigation of development of UWB radar detector.

CHAPTER 3

STUDY AND DEVELOPMENT OF UWB RADAR DETECTOR

Contents

In this chapter, we will discuss about how the UWB radar targets are localized by using a proposed method, that is a combination of Higher Order Statistics (HOS) and the well-known Cell-Averaging Constant False Alarm Rate (CA-CFAR) radar detector.

To begin the discussion, we firstly provide a basic information about the UWB radar module that has been used in this research, then it will be continued by presenting the basic of detection theory. After that, we will be focusing on the development of UWB radar detector.

3.1 UWB Radar Module

The UWB radar used in the following for experimentations, is a UWB radar Kit called HST-D3 developed by UMAIN corporation. The kit composes of the basic UWB radar module called HST-S1 and raspberry Pi 3. The received radar raw data is transmitted to the PC through a Raspberry Pi 3 via various communication including wired communication using LAN and wireless communication using wireless LAN or Bluetooth. Figure 3.1 shows the connection diagram between the radar antennas, radar module, raspberry pi 3, and the Laptop/PC. Figure 3.2 shows the upper view of HST-D3 kit.

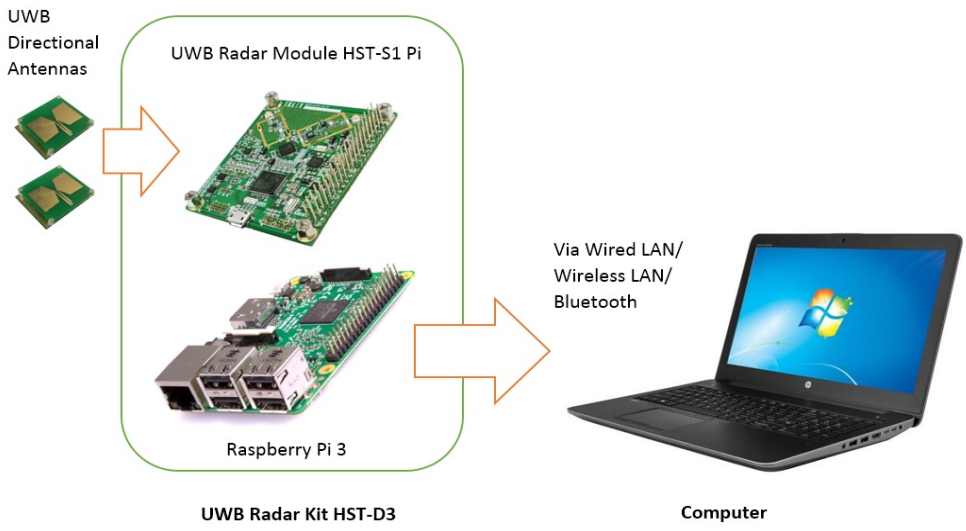


Figure (3.1) Diagram Connection of the HST-D3 with the PC

This is an impulse ultra-wide band radar operates at frequency range from 3 to 4 GHz with the output power is about -25dBm. The bandwith used is 450Mhz - 1GHz. Detail of the specification of this radar is shown in the table 3.1.

HST D3 UWB radar kit module renews the radar frame at every 22.5ms, that means the radar transmit pulses at average about 44.4 times per second. Figure 3.3 shows detail of timing diagram of HST D3 UWB radar kit.

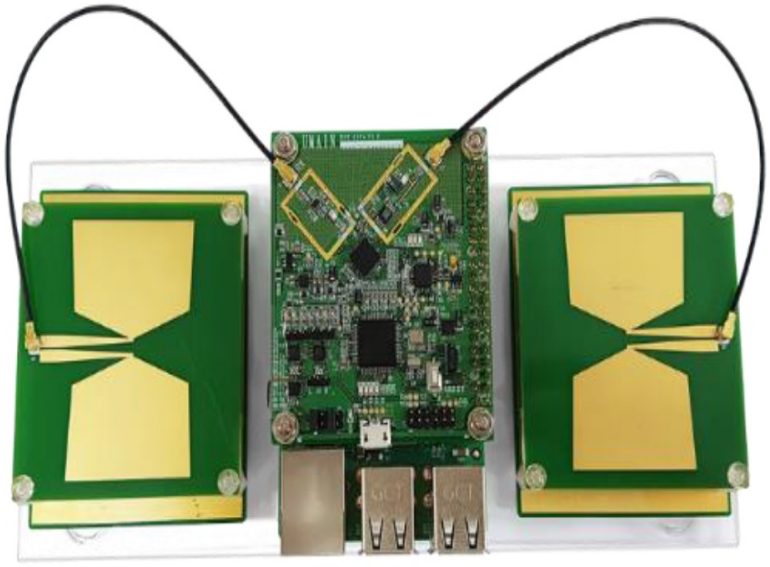


Figure (3.2) HST-D3 UWB radar kit

Table (3.1) UWB radar kit HST-D3 specification [65]

Parameter	Value
Detecting Range	3.0 ∼ 4.0 GHz
Bandwidth	0.45 ∼ 1.0 GHz
Output Power	Typ.-25 dBm
Distance Resolution	15 ∼ 33 cm

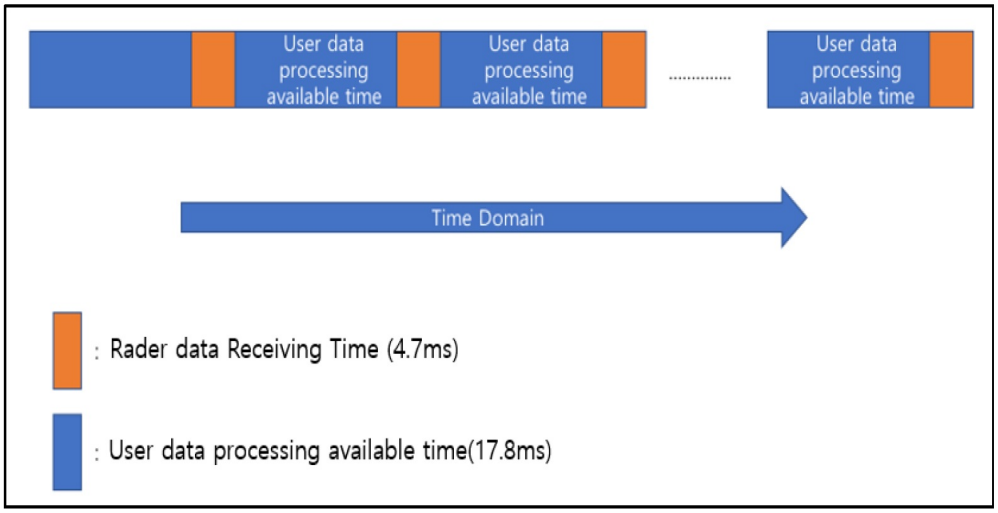


Figure (3.3) Timing diagram of HST D3 UWB radar kit [65]

Figure 3.4 and 3.5 show the example of typical radar data with the absence and presence of target respectively.

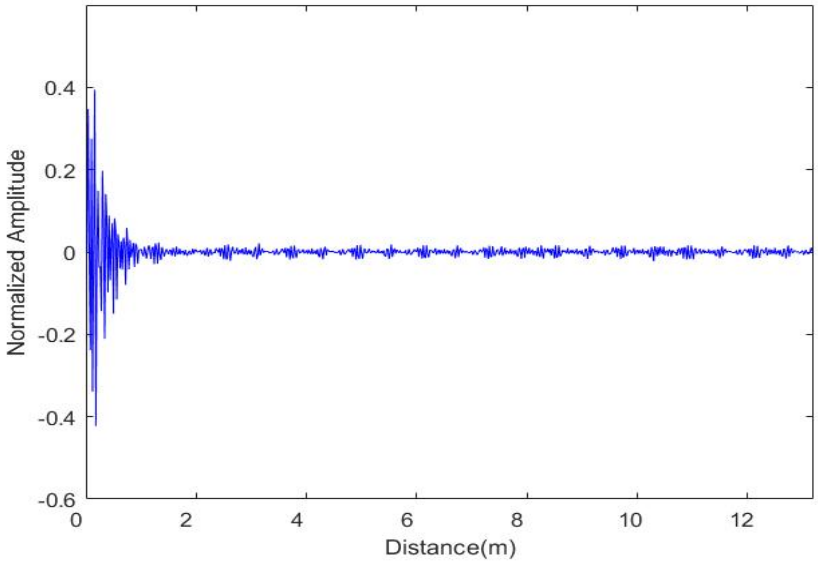


Figure (3.4) Example of radar signal in the absence of target

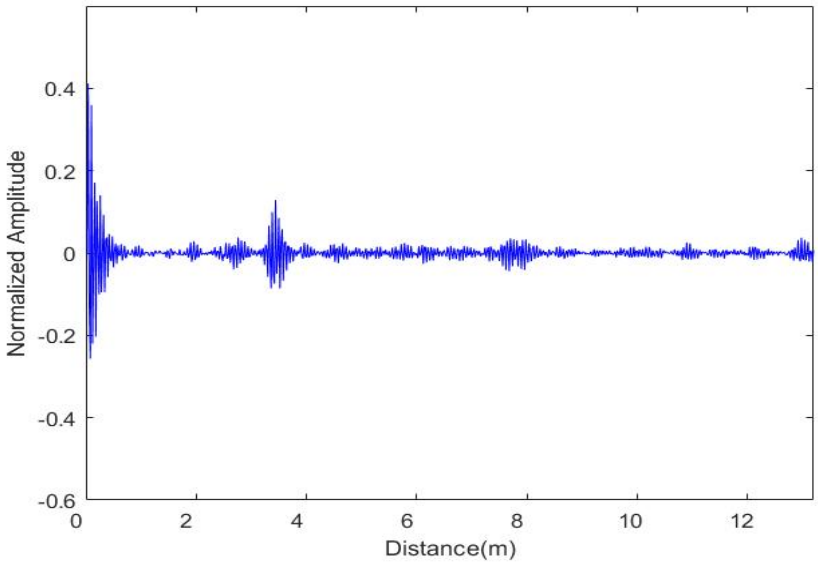


Figure (3.5) Example of radar signal in the present of target

3.2 Detection Theory

Before we are going to further discussion about the UWB radar detector design, we are now going to discuss firstly about the detection theory in general. The objective of detection in radar system is to decide on the presence or absence of the target. There are two terms in radar detection i.e Probability of Detection (P_D) and Probability of False Alarm (P_{FA}).

Detection of a radar target is based on the given threshold values at the output of the radar receiver.

3.2.1 Probability of Detection and Probability of False Alarm

We introduce two types of hypotheses, H_0 and H_1. H_0 is referred to as the *null hypothesis* and H_1 is referred to as *alternative hypothesis*. Since we have to chose between H_0 or H_1, this hypothesis is called *binary hypothesis*. The probability density functions of this hypothesis are shown in the figure 3.6.

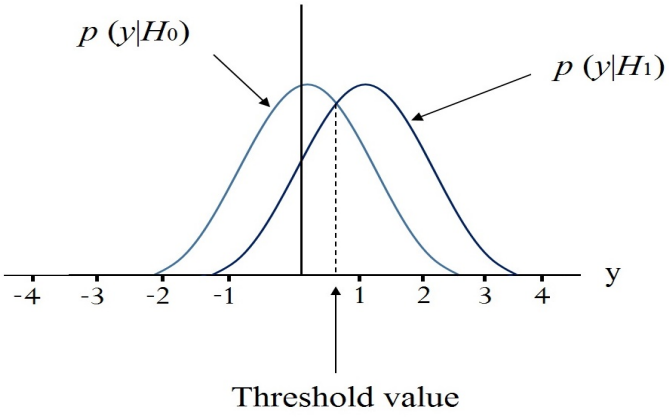


Figure (3.6) Probability densities for hypothesis testing problem

Let Y a random signal observed in observation zone Z which is divided into two spaces Z_0 and Z_1, disjoint with $Z_0 \cap Z_1 = \emptyset$ and $Z_0 \cup Z_1 = Z$. If $Y \in Z_0$ then must be decided H_0 and if $Y \in Z_1$ must be decided H_1. Then, we introduce two decisions events D_0 and D_1:

- D_0 : Decide H_0 is correct
- D_1 : Decide H_1 is correct

So, we have four possibilities of decisions as shown in the table 3.2.

Now, let the $p(Y|H_i)$ the conditional probability of Y given H_i (with $i = 0, 1$), then based on theory of probability:

Table (3.2) Detection possibilities

Decision	Signification	Probabilities
Decide D_1, given the correct H_1	Correct detection of target	P_D
Decide D_1, given the correct H_0	False Alarm	P_{FA}
Decide D_0, given the correct H_1	Miss Detection	P_M
Decide D_0, given the correct H_0	Correct rejection of noise	P_R

- The probability of detection P_D is the probability of correctly detecting a target when a target is actually present:

$$P_D = P(D_1|H_1) = \int_{Z_1} p(y|H_1)dy \tag{3.1}$$

- The probability of false alarm P_{FA} is the probability of deciding that a target is present when in fact there is only noise:

$$P_{FA} = P(D_1|H_0) = \int_{Z_1} p(y|H_0)dy \tag{3.2}$$

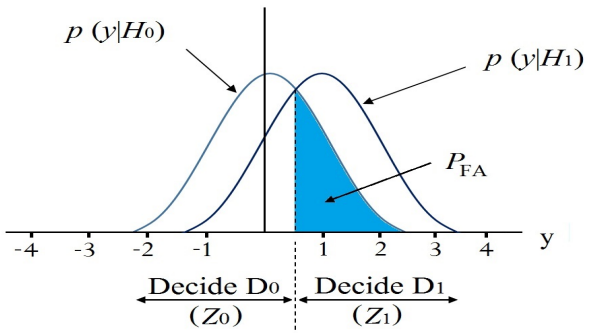


Figure (3.7) Probability densities for hypothesis testing problem of false alarm

- The probability of miss detection P_M is the probability of deciding that a target is not present when in fact a target is actually present:

$$P_M = P(D_0|H_1) = \int_{Z_0} p(y|H_1)dy \tag{3.3}$$

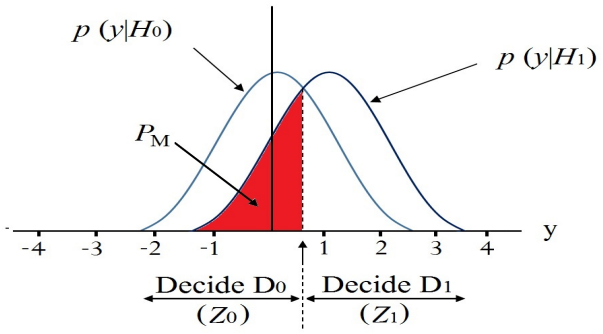


Figure (3.8) Probability densities for hypothesis testing problem of miss detection

- The probability of noise rejection P_R is the probability of correctly deciding that a target is not present when in fact there is only noise:

$$P_R = P(D_0|H_0) = \int_{Z_0} p(y|H_0)dy \tag{3.4}$$

Since $Z_0 \cap Z_1 = \emptyset$ and $Z_0 \cup Z_1 = Z$ we have:

$$\begin{aligned} \int_{Z_0} p(y|H_i)dy &= \int_{Z} p(y|H_i)dy - \int_{Z_1} p(y|H_i)dy \\ &= 1 - \int_{Z_1} p(y|H_i)dy \end{aligned} \tag{3.5}$$

with $i = 0, 1$

For $i = 0$, the probability of miss detection P_M can be derived as:

$$\begin{aligned} \int_{Z_0} p(y|H_1)dy &= 1 - \int_{Z_1} p(y|H_1)dy \\ P_M &= 1 - P_D \end{aligned} \tag{3.6}$$

For $i = 0$, the probability of noise detection P_{NS} can be derived as:

$$\begin{aligned} \int_{Z_0} p(y|H_0)dy &= 1 - \int_{Z_1} p(y|H_0)dy \\ P_R &= 1 - P_{FA} \end{aligned} \tag{3.7}$$

3.2.2 Neyman-Pearson Criteria

The principle of Neyman-Pearson criterion is to maximize the probability of detection P_D subject to the constraint of probability of false alarm $P_{FA} = \alpha$. We can constrain the P_{FA} by choosing the threshold γ.

To maximize the probability of detection P_D for a given probability of false alarm $P_{FA} = \alpha$, decide H_1 if:

$$\Lambda(y) = \frac{p(y \mid H_1)}{p(y \mid H_0)} > \gamma \tag{3.8}$$

The function of $\Lambda(y)$ is called *likelihood ratio* since it indicates for each value of y the likelihood of H_1 versus the likelihood of H_0. The entire test of eq.3.8 is called *likelihood ratio test*. By using this ratio, we can obtain a threshold γ_0 which corresponds to a probability of false alarm $P_{FA} = \alpha$.

$$P_{FA} = \int_{\gamma_0}^{\infty} p(\Lambda|H_0)d\Lambda = \alpha \tag{3.9}$$

Then, the probability of detection can be obtained by:

$$P_D = \int_{\gamma_0}^{\infty} p(\Lambda|H_1)d\Lambda \tag{3.10}$$

After presenting the basic concept of detection theory, in the next section, we will present the discussion about development of UWB radar detector.

3.3 Development of UWB Radar Detector

3.3.1 Moving Target Indication (MTI)

A moving target indication (MTI) helps for pre-processing radar signal before applying targets detection method because it can enhance the signal-to-clutter ratio (SCR) of the radar echo. One of the MTI methods is the so-called pulse canceller. The basic of the pulse canceller is called 2-pulse canceller as its diagram is shown in figure 3.9. The input signal R_i is delayed for one cycle z^{-1}, then it will be substracted by the new incoming signal. The result of absolute value is taken as output R_{out} of this pulse canceller. This simple pulse canceller is very effective to remove the static clutter and to detect the moving objects. To have a higher order of the pulse canceller, it can be done by cascading the 2-pulse canceller [66]. Transfer functions for several higher order of pulse cancellers are presented in the table 3.3.

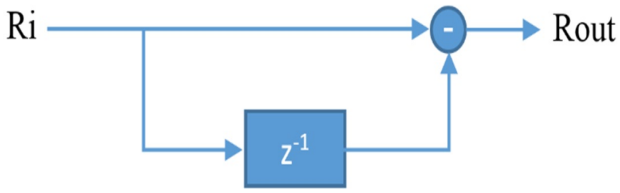


Figure (3.9) Basic 2-pulse canceller

Table (3.3) Transfer Functions For Several Higher Order Pulse Canceller [66]

Number of Pulses Processed	**Transfer Function**
2	$1 - z^{-1}$
3	$1 - 2z^{-1} + z^{-2}$
4	$1 - 3z^{-1} + 3z^{-2} - z^{-3}$
5	$1 - 4z^{-1} + 6z^{-2} - 4z^{-3} + z^{-4}$

Based on [67], the four pulse canceller performs effective reduction of static radar clutters. Therefore, in this experiment before implementing HOS and CFAR detector, we simply applied the four pulse canceller to reject the clutters and direct antenna coupling in our UWB radar. The four pulse canceller can be written as:

$$R_{out} = R_i - 3R_{i-1} + 3R_{i-2} - R_{i-3} \tag{3.11}$$

where R_{out} is the final output after pulse canceller, R_i is the input of original signal, and R_{i-n} is the nth original delayed signal.

After applying the pulse canceller on the received radar signal, the noises and clutters is not completely removed. Therefore, it is required to perform further advance processing steps to distinguish the targets and the noise or clutters. The accuracy of target detection depends on the success of separating the target over the noise or clutter[9].

3.3.2 Higher Order Statistics

Unlike second-order statistics, the Higher Order Statistics (HOS) algorithm is based on higher order moment spectra in order to interpret and analyze the characteristics of a random signal [68][69].

This technique offers many advantages. It reduces clearly the Gaussian noise and the secondary lobes, reconstructs the phase as well as magnitude response of signals or systems, detects and characterizes the non linearity in the data. In addition, the use of HOS allows detecting several obstacles at the same time and simplifying the automation of the process by applying a simple threshold.

HOS consists of higher order moment spectra, which is defined for deterministic signals and Cumulant spectra for random process [69]. Indeed, many real world applications are truly non-Gaussian. Therefore, contrary to the second-order statistics, higher order statistics are applicable when dealing with non-Gaussian, because it can adapt to non-Gaussian noise or to non-linear channel characteristics and reveal the phase information [68][70]. In [13], the HOS method has been used to estimate time delay in unknown spatially correlated noise and the result is better than that of used cross correlation method.

There are two HOS algorithms that have been investigated in this work i.e. 4^{th} *order Cumulant* based on *Tugnait4 algorithm* and 4^{th} *order Cross-Moment*. The basic concepts of them will be discussed in the following:

Tugnait4 Algorithm

Tugnait4 is an algorithm which uses 4^{th} order-Cumulant of the HOS to estimate time delay. It is developed by Jitendra K.Tugnait (1989). The advantage of Tugnait4 algorithm is its ability to suppress the Gaussian noise over the useful signal. The expression of the Tugnait4 algorithm is shown in the equation below [7][13]:

$$\boldsymbol{J}_4(i_0) = \frac{\boldsymbol{cum}_4(c(i-i_0), c(i-i_0), r(i), r(i))}{\sqrt{|\boldsymbol{cum}_4(c(i), c(i)||\boldsymbol{cum}_4(r(i), r(i)|}} \tag{3.12}$$

with:

c is the reference signal

r is the incoming signal

i_0 is the time index decision

where:

$$\begin{aligned} &\boldsymbol{cum}_4(c(i-i_0), c(i-i_0), r(i), r(i)) \\ &\quad = \frac{1}{N}\sum_{i=1}^{N-1} c^2(i-i_0) r^2(i) \\ &\quad - 2\left[\frac{1}{N}\sum_{i=0}^{N-1} c(i-i_0) r(i)\right]^2 \\ &\quad - \left[\frac{1}{N}\sum_{i=0}^{N-1} c^2(i-i_0)\right]\left[\frac{1}{N}\sum_{i=0}^{N-1} r^2(i)\right] \end{aligned} \tag{3.13}$$

4^{th} Order Cross-Moment

A novel design based on approach of a 4^{th} cross-moment is proposed with the proposed UWB radar. The 4^{th} order cross-moment m_4 for a stationary random process $x(n)$ with samples $x_0(n), x_1(n), x_2(n)$ and $x_3(n)$ is defined as [69]:

$$m_4(\tau_1, \tau_2, \tau_3) = E\{x_0(n)x_1(n+\tau_1) \\ x_2(n+\tau_2)x_3(n+\tau_3)\} \tag{3.14}$$

where $E\{*\}$ denotes statistical expectation.

For deterministic signal, it is replaced by a time summation over all time samples (for energy signals) or time averaging (for power signals). Under the assumption that $x(n)$ is of zero-mean, the fourth-order cross moment is calculated from the given data as [69]:

$$m_4(\tau_1, \tau_2, \tau_3) = \frac{1}{N}\sum_{n=0}^{N} x_0(n)x_1(n+\tau_1) \\ x_2(n+\tau_2)x_3(n+\tau_3) \tag{3.15}$$

where N is the length of the signal.

A better combination of the x_0, x_1,x_2, and x_3 for the proposed system is 'Reference', 'Signal', 'Reference', and 'Signal', respectively. Then "x_0, x_2" are replaced by the received echoes, and "x_1, x_3" are replaced by the references. The parameters of τ_1 is equal to zero, and $\tau_2 = \tau_3 = \tau$ so the equation becomes:

$$m_4(\tau_1, \tau_2, \tau_3) = \frac{1}{N}\sum_{n=0}^{N} \left[c(n)r(n+\tau)\right]^2 \tag{3.16}$$

with: c is the reference signal and r is the received signal.

3.3.3 Time Delay Estimation

Determination of radar targets requires the time delay estimation in order to localize the targets positions. Therefore, in this section we firstly present the investigation results of the performance comparison between two HOS methods i.e. 4^{th} *order Cumulant* (Tugnait4-based Algorithm) and 4^{th} *order Cross-Moment*. Then, we provided the results of comparison between the selected HOS and the ordinary second order statistics.

Figure 3.10 shows an example of the result performance of 4^{th} order Cumulant (green curve) and 4^{th} order Cross-Moment (red-dotted curve). Both methods were applied to the original echo of 1-D raw radar data. It can be seen that they have a very similar performance in terms of noise suppression. Therefore, since 4^{th} order Cross-Moment has less computation than that of 4^{th} order Cumulant, 4^{th} order Cross-Moment is more efficient to be use for time delay estimation in UWB radar system.

Then, 4^{th} order Cross-Moment is compared to the ordinary second order statistics as presented in figure 3.11 and 3.12. The figure 3.11 shows the comparison between 4^{th} order Cross-Moment and the ordinary second order statistics with the threshold value set to 0.1. The red curve and blue-dotted curve are the plots of HOS and the ordinary second order results, respectively. We can see that the first target has a lot of energy reflection which makes it easily to be detected with the given threshold value to 0.1 (black-dotted line on figure 3.11), while the second one has low energy reflection which makes it cannot be detected.

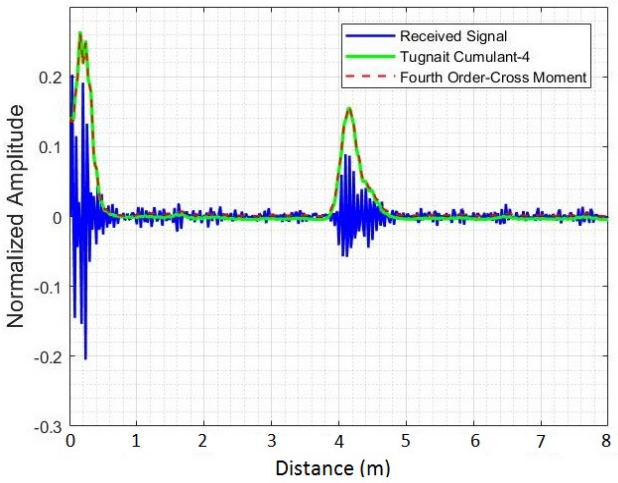


Figure (3.10) Performance comparison between 4^{th} order Cumulant (Tugnait4) and 4^{th} order cross-moment.

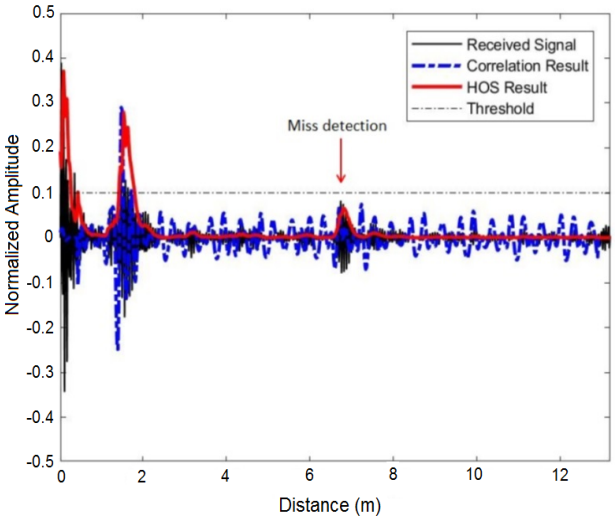


Figure (3.11) The comparison result between HOS and ordinary second order statistics in the presence of two radar targets with the threshold value set to 0.1

The figure 3.12 shows the comparison between 4^{th} order Cross-Moment and the ordinary second order statistics with the threshold value set to 0.05. We can see that with this threshold value, HOS positively detected the position of the second target thanks to its successful noise suppression, while the second order statistics produced a lot of false detection

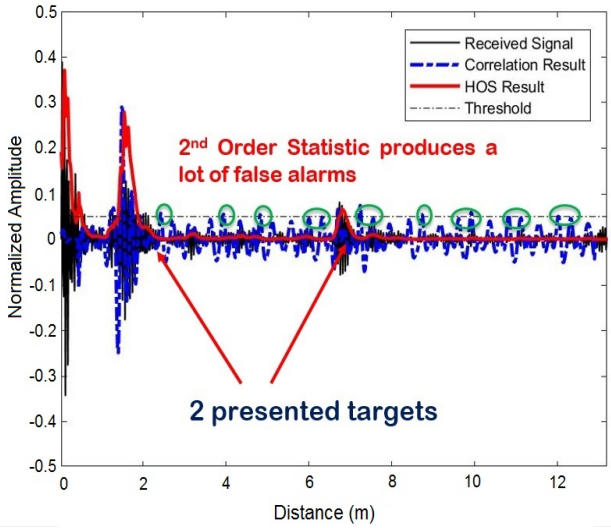


Figure (3.12) The comparison result between HOS and ordinary second order statistics in the presence of two radar targets with the threshold value set to 0.05

and it failed to detect the second target.

From this observation, we concluded that one of the advantages of using HOS in this UWB radar application is that it can enhance the radar sensitivity performance thanks to its ability to suppress the noise.

In the next subsection we will discuss about the investigation of the proposed UWB radar detector that is a combination of HOS and CA-CFAR (Cell Averaging-Constant False Alarm Rate) detector. This proposed detector is aimed to have an UWB radar detector which is robust against the noise and it works with adaptive threshold.

3.3.4 Proposed UWB Radar Detector

We begin this discussion by introducing the well-known CA-CFAR (Cell Averaging-Constant False Alarm Rate) detector, and then it will be continued by presenting the performance comparison results between the selected HOS type (4^{th} *order Cross-Moment*) and the CA-CFAR detector.

In a CA-CFAR detector, a threshold is adaptively estimated based on local information on the background noise, from both leading and lagging cells (called reference cells) surrounding the cell under test (CUT) [54][55][56]. The figure 3.13 describes the architecture of CA-CFAR.

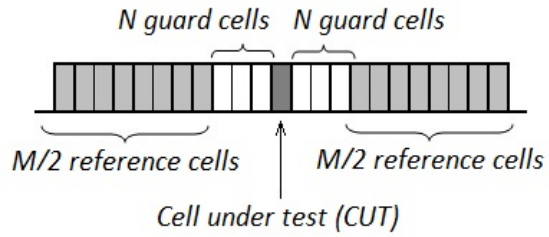


Figure (3.13) CA-CFAR algorithm architecture [54]

The noise estimate can be computed as [55][56]:

$$P_n = \frac{1}{M} \sum_{m=1}^{M} x_m \tag{3.17}$$

where M is the number of reference cells and x_m is the sample in each reference cell and P_n represents the estimated noise power. Then the detection threshold T is given by:

$$T = \alpha P_n \tag{3.18}$$

α is a scaling factor called the threshold factor and is calculated as:

$$\alpha = M(P_{fa}^{-1/M} - 1) \tag{3.19}$$

where P_{fa} is the desired false alarm rate.

As already mentioned above, we have investigated the performances comparison results between automatic threshold CA-CFAR which directly implemented in the radar signal without time delay estimation and the HOS with the fixed threshold value. It is important be noticed that both HOS and CA-CFAR has been implemented after 4-pulse canceller process. We obtained that the HOS with the fixed threshold value gives better performance than that of the CA-CFAR which directly applied to the signal without time delay estimation as we can see in table 3.20.

We found that the problem with CA-CFAR is the difficulty to determine the number of reference cells as shown in fig. 3.14 to fig. 3.16. Figure 3.14 shows the result of applying CA-CFAR directly to the radar signal without time delay estimation when the number of reference cells (M), the number of guard cells (N), and the false alarm rate (P_{fa}) were respectively set to, 80, 2, and 10^{-5}. With this configuration, the detector detected three targets, while in the real condition it was presented only one target at 5 meters away from the radar. This means that the false alarms have occurred. Figure 3.15 shows the result of the CA-CFAR performance with the same number of guard cells and false alarm rate as in the previous configuration but the number of reference cells was reduced to 20. With this configuration, the detector performed good. It can detect the right target position without false detection, but it had the problem for the another radar data frame as shown in fig. 3.16 that the detector has missed detection of the presented target. Therefore, if the number of reference cells is not well chosen, the detector will result in a lot of false alarms and most probably will have miss detection of the target.

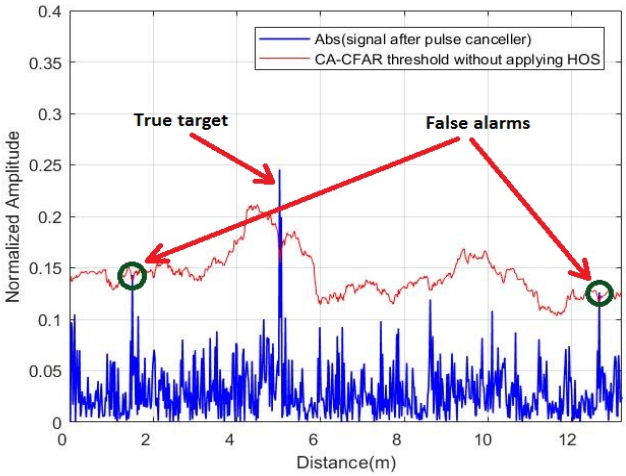


Figure (3.14) CA-CFAR detects false alarms, $P_{fa} = 10^{-5}$, $M = 80$, and $N = 2$.

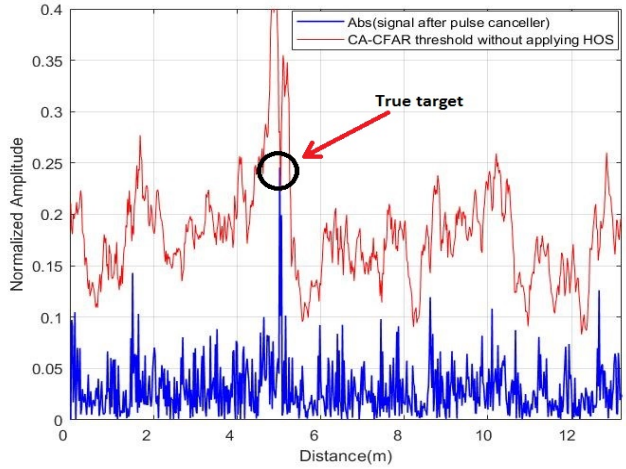


Figure (3.15) Correct detection of CA-CFAR, $P_{fa} = 10^{-5}$, $M = 20$, and $N = 2$.

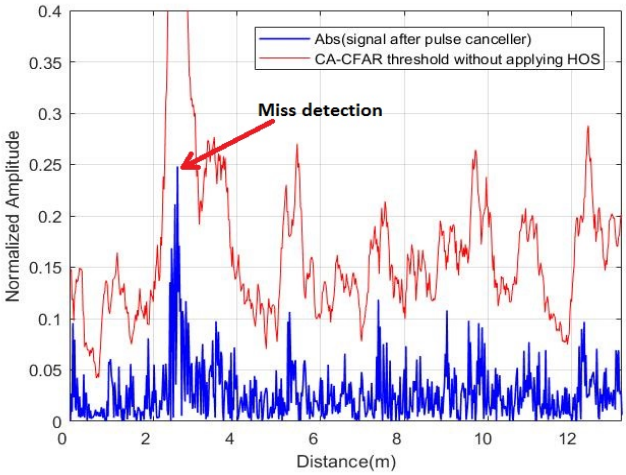


Figure (3.16) Miss detection of CA-CFAR, $P_{fa} = 10^{-5}$, $M = 20$, and $N = 2$.

In order to develop a robust radar detector, we proposed a solution for this problem, by firstly applying the HOS to estimate the time delay before running automatic threshold CA-CFAR detector. Figure 3.17, 3.18, and 3.19 show several examples of the comparison results between the CA-CFAR applied directly to the signal without time delay estimation and with time delay estimation (HOS). We can see clearly that, the implementation of the CA-CFAR detector after time delay estimation by HOS promises better performances compared to the CA-CFAR detector without applying time delay estimation.

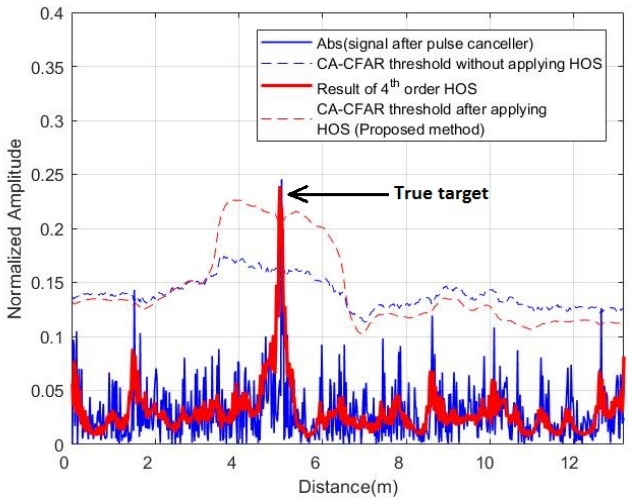


Figure (3.17) Comparison of CA-CFAR performance run directly on the signal without time delay estimation and with time delay estimation (HOS), $P_{fa} = 10^{-5}$, $M = 150$, and $N = 2$. A real target presents at 5 meters away from the radar.

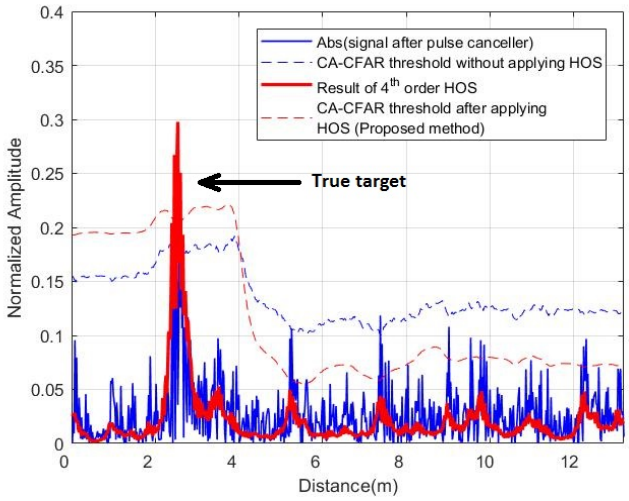


Figure (3.18) Comparison of CA-CFAR performance run directly on the signal without time delay estimation and with time delay estimation (HOS), $P_{fa} = 10^{-5}$, $M = 150$, and $N = 2$. A real target presents at 2.5 meters away from the radar.

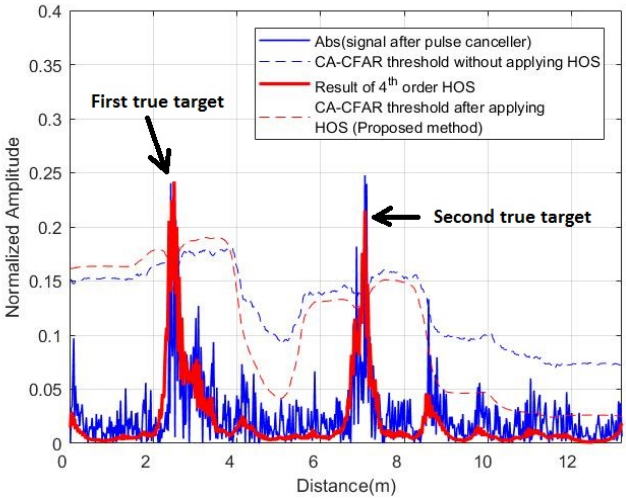


Figure (3.19) Comparison of CA-CFAR performance run directly on the signal without time delay estimation and with time delay estimation (HOS), $P_{fa} = 10^{-5}$, $M = 150$, and $N = 2$. Two real targets present respectively at 2.5 and 7 meters away from the radar.

In order to validate the proposed method, we have tested a number of 200 raw radar data where each of it contains a real single target. We consider two conditions including Positive Condition (PC) and Negative Condition (NC). The Positive Condition is when the radar gives the good decision and otherwise is for the Negative Condition. The following is the detail description for both decisions:

1. Positive Condition:
   (a) *when real condition presents one target and the observation only detect one target, that means there is no false detection.*
2. Negative Condition:
   (a) *when real condition presents one target, but the observation results more than one target, that means there are detection errors (false alarm).*
   (b) *when real condition presents one target, but the observation results no detected target.*

The performance of accuracy is calculated as:

$$Acc = \frac{Number\ of\ PC}{Number\ of\ PC + Number\ of\ NC}\mathbf{x}100\% \tag{3.20}$$

The following table 3.4 presents the performance evaluation results between HOS with the fixed threshold value (TH), CA-CFAR without time delay estimation, and CA-CFAR after

time delay estimation by HOS for different numbers of the reference cells M. The number of guard cells N and the false alarm rate P_{fa} are respectively set to 2 and 10^{-5}. From this table, we found that the performance of the HOS with the fixed threshold value is better than that of the CA-CFAR applied directly to the radar signal without HOS. But, the performance of CA-CFAR after applying HOS outperforms the others.

Table (3.4) Performance evaluation between HOS with the fixed threshold (TH), CA-CFAR without time delay estimation, and CA-CFAR after time delay estimation by HOS. (N = 2, $P_{fa} = 10^{-5}$)

Detector type	Number of True Decision	Number of False Decision	Decision accuracy
HOS (TH=0.05)	124	76	62.0%
HOS (TH=0.1)	**129**	71	**64.5%**
CA-CACFAR ($M = 100$)	102	98	51.0%
CA-CACFAR ($M = 150$)	115	85	57.5%
CA-CACFAR ($M = 200$)	**118**	**82**	**59.0%**
CA-CACFAR ($M = 250$)	106	94	53.0%
HOS - CACFAR ($M = 100$)	157	43	78.5%
HOS - CACFAR ($M = 150$)	164	36	82.0%
HOS - CACFAR ($M = 200$)	**175**	**25**	**87.5%**
HOS - CACFAR ($M = 250$)	172	28	86.0%

Figure 3.20 shows the performance of detection accuracy versus the number of reference cells between the CA-CFAR and the combination of CA-CFAR and HOS, $P_{FA} = 10^{-5}$ and $N = 2$. From this figure, we can see that both method achieve their best accuracy when the chosen of reference cell number is 200. In this figure also can be seen that the performance of the proposed method is better than that of the ordinary CA-CFAR, where it can achieve up to 87% of accuracy when $M = 200$.

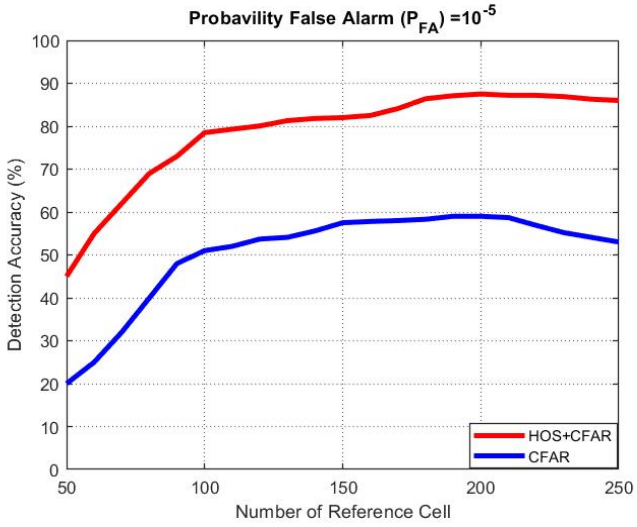


Figure (3.20) Comparison of CA-CFAR performance applied directly to the signal without time delay estimation and with time delay estimation (HOS), $P_{fa} = 10^{-5}$, and $N = 2$.

3.4 Conclusion of Chapter 3

In this chapter we have discussed about the development of UWB radar detector. There are several methods have been investigated including the ordinary second order statistics, higher order statistics (HOS), and the well-known cell-averaging false alarm rate (CA-CFAR) detector. After investigating the performance of these radar detectors, a combination of HOS and CA-CFAR detector has been proposed and investigated.

There are two types of HOS algorithms, 4^{th} order Cumulant (Tugnait4-based) and 4^{th} order Cross-Moment. Firstly, we investigated these both algorithms to obtain the optimal performance of the HOS before combining it with the CA-CFAR. Based on the investigation results, we found that both algorithms give a very similar performance in terms of noise suppression, but in terms of the complexity, 4^{th} order Cross-Moment has less calculation than that of 4^{th} order Cumulant. This means that 4^{th} order Cross-Moment is more efficient to be used in UWB radar application. After that, we compared the performance of time delay estimation between 4^{th} order Cross-Moment and the ordinary second order statistics. We noticed that the result of 4^{th} order Cross-Moment is much better than that of second order statistics. Therefore, 4^{th} order Cross-Moment has been considered to be use in our proposed UWB radar detector.

Finally, by combining 4^{th} order HOS and the CA-CFAR, an automatic UWB radar detector which is robust to the noise has been developed. To prove this idea, we have evaluated the performance between HOS with the fixed threshold value, CA-CFAR detector without time delay estimation, and a combination of HOS and CA-CFAR detector (proposed method). The results show that the combination of HOS and CA-CFAR promises a better performance for UWB radar detector compared to the other methods.

In the next Chapter, we will present the recognition radar target using the machine learning where the radar signature is exploited as input features. The machine learning methods that will be investigated for this purpose are Support Vector Machines (SVM) and Deep Belief Networks (DBN).

CHAPTER 4

STUDY AND DEVELOPMENT OF UWB RADAR TARGETS RECOGNITION

Contents

In the section 4.1 of this chapter, we will present the process of obtaining the radar signatures for different radar targets. This work relates to the previous chapter 3, as we concluded that the proposed of our UWB radar detector promises good performance for this UWB radar system, we then use it to obtain radar signatures. In section 4.2, we will discuss about the Support Vector Machine (SVM) as a supervised machine learning system used in this first proposition to identify the UWB radar target type which uses 1-D normalize radar signature as input features, followed by presenting its experimental results. The discussion of this chapter will be continued by presenting a Deep Learning approach for this UWB radar recognition system called Deep Belief Networks (DBN). DBN will be exposed in detail in the section 4.3 of this chapter as well as its performance results.

As it has been mentioned that one of the advantages of using UWB technology in the radar system is that the radar receiver will receive a pulse distortion differently for the different target nature that is called radar signature. Thus, we can take this advantage to develop a system recognition of radar target by using this signature. The radar signature must be normalized before feeding it into a machine learning system. The process of obtaining radar signature will be explained in detail on the following section 4.1.

4.1 Radar Signature

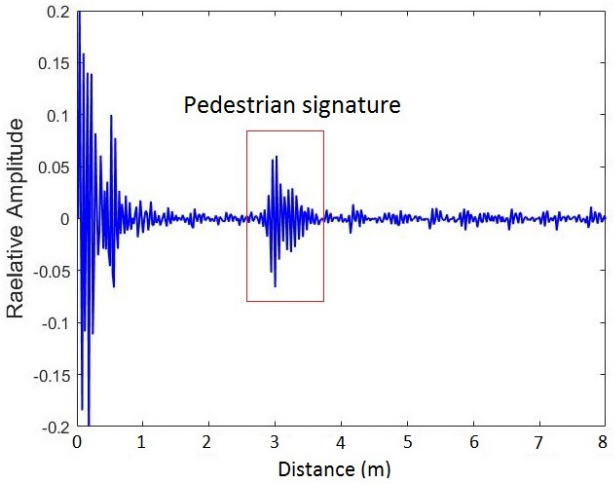
After investigating the performance of the radar detectors, we noticed that the combination of HOS and CA-CFAR (a proposed method) promises a good performance in UWB target detection. Then, we use this proposed method in our UWB radar system to detect the positions of the radar target and extract its signature.

4.1.1 Obtaining Radar Signature

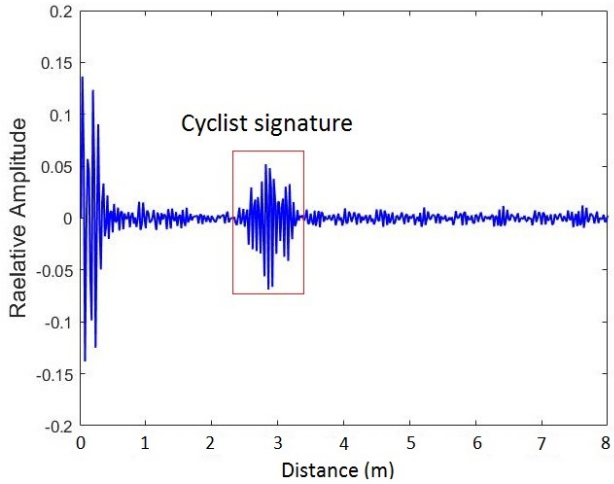
The HOS is firstly applied after pulse canceler, then the threshold value is estimated by using CA-CFAR on the result of the HOS. Once the position of the target is obtained, then we step back to the original received echo and perform the windowing around the known target position and keep it as a radar signature. Figure 4.1 describes the process of obtaining the radar signature.

Figure (4.1) Diagram of determination of radar signature

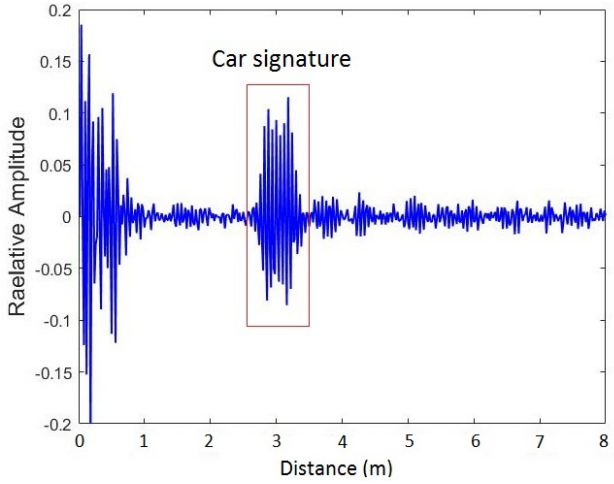
Figure 4.2 shows three different UWB pulses distortion by three different targets nature including: pedestrian, cyclist, and car. These differences in pulses distortion are called radar



(a)

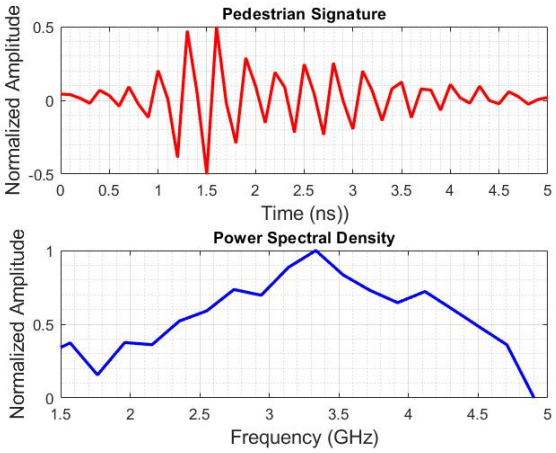


(b)

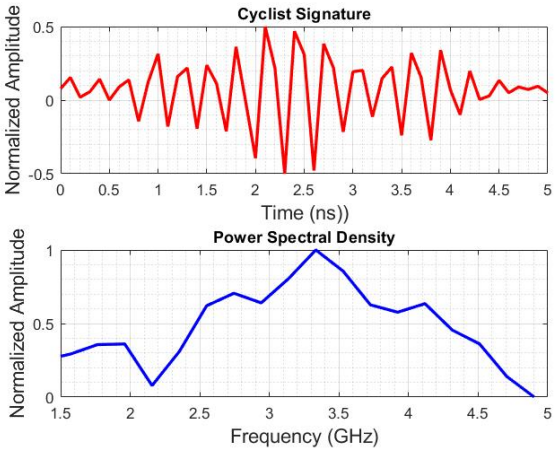


(c)

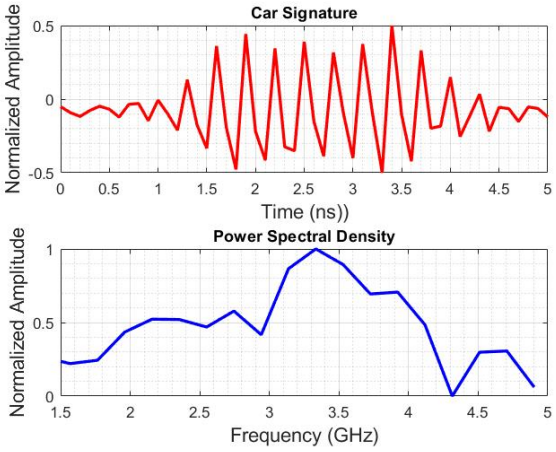
Figure (4.2) Signatures of the three different targets



(a) Pedestrian signature and its power spectral density



(b) Cyclist signature and its power spectral density



(c) Car signature and its power spectral density

Figure (4.3) Normalized signatures of the three different targets followed by their power spectral densities. As can be seen, the figure shows that the different nature of radar target yielded different radar signature

target signatures. Figure 4.3 shows their normalized radar signatures followed by their power spectral densities. We can see clearly the difference between them. The important things of these signatures are their normalized amplitudes that represent the features vector of each signature.

4.2 SVM-Based Approach

The recognition system has been developed by using Support Vector Machines (SVM) technique. SVM recognizes the target based on the result of training parameters.

In the second part of this chapter, we are going to discuss about how the radar data is required and the target position is obtained.

SVM is a supervised machine learning of binary classification technique for pattern recognition. This classifier creates a hyper-plane to separate the pattern of data into two classes (+1 or -1) with the maximum margin. The vectors that define the hyperplane are called *support vectors* [71].

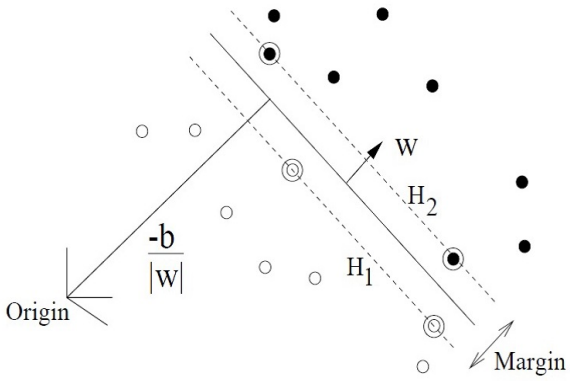


Figure (4.4) Illustration of Support Vectors, Hyperplane and Margin

Let us label the training data $x_i \in R^d$ with a label $y_i \in -1, 1$, for all the training data $i = 1, ..., n$, where n is the number of data, and d is the dimension of the problem. y_i is separated by a hyper plane with margin M, then for each training sample (x_i, y_i):

Negative class $w^T x_i + b \leq -1$, if $y_i = -1$

Positive class $w^T x_i + b \geq 1$, if $y_i = 1$

In order to maximize the margin, we need to minimize $||w||$. That is equivalent to minimize $1/2||w||^2$. Then, we can formulate a Quadratic Optimization (QP) problem and solve it for w and b.

By multiplying each class by its label, then we have the constrain of this QP optimization as:

$$y_i(w^T x_i + b) \geq 1 \tag{4.1}$$

Thus, the optimal margin can then be found by minimizing:

$$1/2||w||^2 \tag{4.2}$$

which subjects to:

$$y_i(w^T x_i + b) - 1 \geq 0, \forall i \tag{4.3}$$

4.2.1 Solving optimization problem

By using Lagrange multiplier method, we get the primal form for this minimization problem:

$$L = f(x) - \alpha.g(x) \tag{4.4}$$

where $g(x) \geq 0$.

Inserting $f(x) = 1/2||w||^2$ and $g(x) = y_i(w^T x_i + b) - 1$ into (9) we get primal form:

$$L_p = 1/2||w||^2 - \sum_{i=1}^{n} \alpha_i(y_i(w^T x_i + b) - 1) \tag{4.5}$$

By applying the KKT (Karush-Kuhn Tucker) conditions to the primal form, we get the primal-dual form:

$$L_{pd} = -1/2 \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j x_i x_j + \sum_{i=1}^{n} \alpha_i \tag{4.6}$$

As the vectors x_i, x_j and y_i are known and L_{pd} is optimal at $\frac{\partial L_{pd}}{\partial \alpha} = 0$, and at $\sum_{i=1}^{n} \alpha_i y_i = 0$, we can find α_i value.

The non-negative value of α_i will correspond to support vectors. Knowing the α_i, we can find w by:

$$w = \sum_{i=1}^{n} \alpha_i y_i x_i \tag{4.7}$$

The b parameter can then be determined by next equation:

$$\alpha_i[y_i(w^T x_i + b) - 1] = 0 \tag{4.8}$$

4.2.2 SVM kernel

In case of non-linear separable data, the transformations are performed by using variable kernel functions, such as sigmoid, polynomial, linear, Gaussian radial basis function (RBF),

etc. These kernel functions define an inner product in high dimensional space, given as below [72]

$$Linear \quad K(x_i, x_j) = {x_i}^T x_j \tag{4.9}$$

$$Polynomial \quad K(x_i, x_j) = (\gamma {x_i}^T x_j + r)^d, \ \gamma > 0 \tag{4.10}$$

$$RBF \quad K(x_i, x_j) = exp(-\gamma ||x_i - x_j||^2), \ \gamma > 0 \tag{4.11}$$

$$Sigmoid \quad K(x_i, x_j) = tanh(\gamma {x_i}^T x_j + r) \tag{4.12}$$

4.2.3 Experimental SVM Results

The SVM technique was realized by using the LIBSVM developed by [73]. LIBSVM is a library for SVM that has been widely used in many research areas. These four basic kernels are investigated and their performances are compared in order to have a better kernel that matches in our UWB-SRR system.

Dataset

In the case of UWB radar, we cannot find the dataset benchmark of radar signatures thus we performed our own dataset. For pedestrian and cyclist, the target signatures are collected from the moving, either pedestrian or cyclist randomly with different orientation (e.g: front, back, side, etc) in front of radar with different speed motion. For collecting car signatures, we simply moved the radar closer to and farther away the cars. Then, to extract the signature, we applied the method explained in subsection 4.1.1. In order to have a reliable dataset, we have also considered for short stop target.

In order to evaluate the SVM result performance, a total of 3000 radar signatures has been used as dataset which composes of three different target classes including car, cyclist, and pedestrian. Each target class has 1000 normalized radar signatures.

Training SVM Models

In order to measure the identification performance of the realized system, the testing and training data must be chosen from different data segments, which means that both training and testing data must be not the same. Therefore, after obtaining radar dataset, we randomized and divided them into two groups, 50% of them as a training set, and the remaining 50% as a testing set. The model data is trained based on data that has been grouped already into each class (labeled data).

To have a better performance of SVM classifier, before training and testing dataset, we need firstly to perform an investigation to have a better SVM kernel by using cross-validation technique [74]. The fourth basic SVM kernels including linear kernel, Polynomial kernel, Radial Base Function (RBF) kernel and Sigmoid kernel have been already investigated.

Basically, SVM kernel has two important parameters that are C and γ. Both parameters are used to control the over-fit weights and biases. Thus, before performing the process of training and testing data, it is important to have best parameters for C and γ, and then use them in training and testing process, so that the classifier can predict more accurately the unknown data. A common strategy to have these parameters is to separate the training data into $\boldsymbol{k}$ equal size bins and then one bin is used as the validation for testing the model and the remaining of $\boldsymbol{k}$-1 bins are used as the training data. The cross-validation process is then repeated $\boldsymbol{k}$ times (the folds), with each of the $\boldsymbol{k}$ bins used exactly once as the validation data. This technique is called $\boldsymbol{k}$-fold cross-validation. The $\boldsymbol{k}$ results from the folds can then be averaged to produce a single estimation. The advantage of this method is that all observations are used for both training and validation, and each observation is used for validation exactly once. The best performance of accuracy (99.49%) is achieved at C=4 and γ = 2. Table 4.1 shows the summary of best parameters for four used SVM kernels. Finally, we compared the performance of the four kernels by computing their cross-validation accuracy rates and we have chosen RBF kernel which has better performance for our model.

Table (4.1) Best Parameters for C and γ

SVM Kernels	C	gamma	Cross-validation
Linear	128	-	95.64%
Polynomial	3.051758e-05	64	98.72%
Radial Base Function (RBF)	4	2	99.49%
Sigmoid	1024	0.0625	96.41%

After evaluating the SVM kernels and obtaining the best parameters by using $k - fold$ cross-validation method, finally we trained the SVM using the RBF kernel with parameters C=4 and γ=2.

Testing SVM model

The remaining of 50% of the dataset were used to test the whole performance of SVM in recognition of UWB radar target. Testing dataset will be predicted as a specific label based on the parameters obtained from the result of SVM training. Table 4.2 shows the confusion matrix of the result performance of SVM in recognizing of UWB radar targets.

The table 4.2 above shows that SVM gives a good performance for our system where the recognition rate is up to 96.23%, 95.25% and 97.23% for cyclist, pedestrian, and car, respectively.

In the next section, we will present the investigation of Deep Learning approach (Deep Belief Network-DBN) for identifying of these uncovered vulnerable road users.

Table (4.2) Confusion matrix of using RBF kernel for three identification target

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	97.23%	1.31%	1.13%
	Cyclist	1.60%	96.23%	3.62%
	Pedestrian	1.17%	2.46%	95.25%

4.3 A Deep Belief Network Approach

Nowadays, Deep learning is becoming a very popular machine learning technique widely used in the field of artificial intelligent. The word "Deep" comes from the architecture of neural network that has more than 2 hidden layers in depth. Deep learning consists of multiple processing layers to learn representations of data with multiple levels of abstraction [75].

A very popular of first architecture of Deep Learning is LeNet, pionered by Yann LeCun. LeNet was one of the very first Deep Learning architecture which helped propel the field of Deep Learning. This pioneering work by Yann LeCun was named LeNet5 after many previous successful iterations since the year 1988. At that time, LeNet architecture was used mainly for character recognition tasks such as reading zip codes, digits, etc.

In 2006, Hinton et al. [77] have successfully trained a complicated NNs model to learn high-level structure of hand writing digits by stacking a few RBMs (Restricted Boltzmann Machines) on top of each other, thus leading to the so-called Deep Belief Networks (DBNs).

In this section, we will investigate the use of DBN in recognizing radar target and perform a comparative performance with the the result found by SVM. Before going to the further discussion about the DBN, we are going to firstly discuss about the basic principle of Neural Network processing.

4.3.1 Introduction of Artificial Neural Networks

Artificial neural Networks (ANNs), as it named, is an artificial neural system inspired and developed by the biological brain system. ANNs is based on a collection of units of neurons that are connected to one another that can process and send signals from one to the another.

The following discussion was inspired and adapted by Andrew Ng et al [78] from their tutorial about Deep Learning.

Simple Neural Network

A very basic or a simplest of NNs composes of only a single neuron, several input vector $\mathbf{x} = (x_1, x_2, .., x_n)$ and a bias b as illustrated in figure 4.5. All the inputs and the bias are connected to this neuron. These connections are called the synapses where every synapse

has the weight W. The hypothesis output of this simplest neural network is written as $h_{W,b}(\mathbf{x}) = f(\sum_{i=1}^{n} W_i x_i + b)$. Where the function of f is called the activation function. There are many kinds of activation function used in NNs implementation, the most commonly used are step function, sigmoid function, tanh and Rectifier Linear Unit (ReLu).

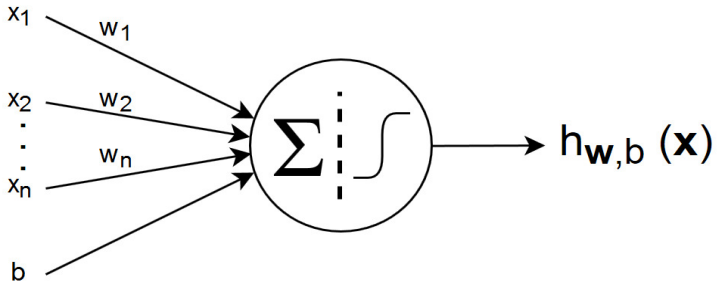


Figure (4.5) Simplest Neural Networks

Activation Function

The activation function is usually an abstraction representing the rate of action potential firing in the cell. In the above description of simplest neural network uses the sigmoid function as the activation function, it is written by:

$$f(z) = \frac{1}{1 + e^{-z}} \tag{4.13}$$

Yet, it is important to be noticed that there are another common choice of activation functions like hyperbolic tangent, or tanh and rectified linear unit (ReLU). The *tanh* function is written as:

$$f(z) = tanh(z)\frac{e^{z} - e^{-z}}{e^{z} + e^{-z}} \tag{4.14}$$

In practice for deep NNs rectified linear function often works better than sigmoid and tanh function. The rectified linear activation function is given by:

$$f(z) = max(0, z) \tag{4.15}$$

Figure 4.6 shows the plots of the sigmoid, tanh and rectified linear functions (ReLU).

Multi-Layer Neural Network

As the network model in the simplest neural network shown above is a very limited model to describe the input type, so the researcher has developed a complex neural network models. It is a neural network with multi-layer model as illustrated in the figure 4.7.

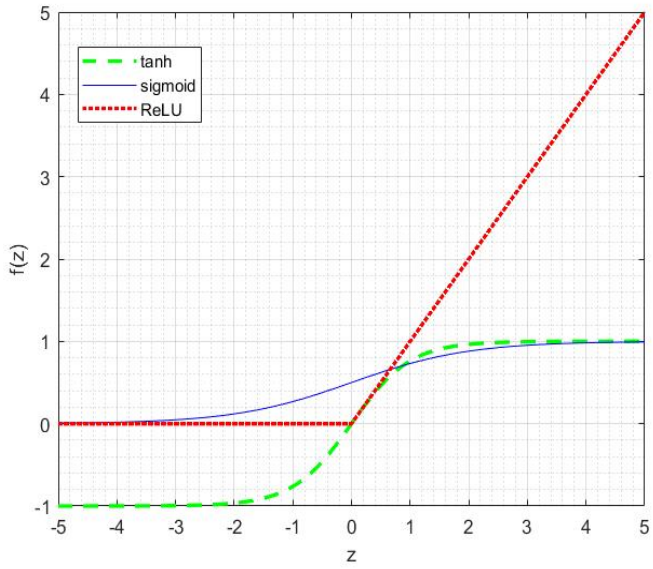


Figure (4.6) Plots of the sigmoid, tanh and rectified linear functions

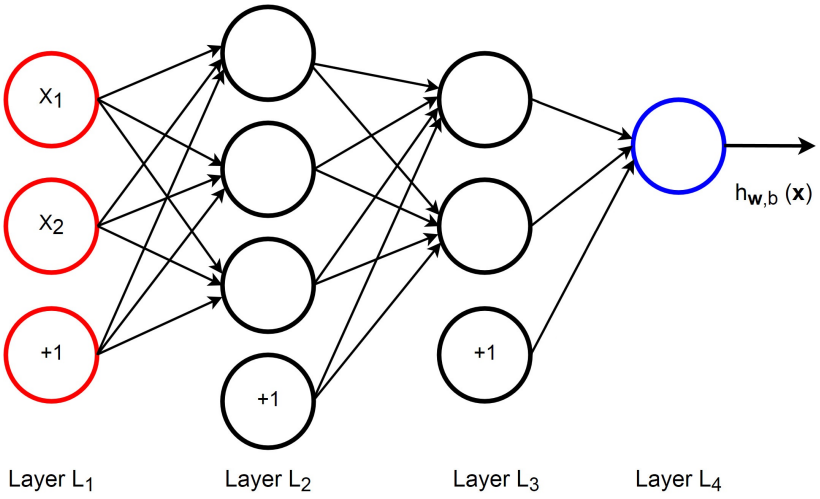


Figure (4.7) Network with two hidden layers

Figure 4.7 has 4 layers network with only two input parameters x_1 and x_2 in the input layer (the leftmost layer, layer L_1) and has one output in the output layer (the rightmost layer, layer L_4). This network has two hidden layers, layer L_2 and layer L_3. The circles labeled "+1" are called bias units, and correspond to the intercept term [78].

Feed Forward Propagation

Now, we are going to analyze the multi-layer neural network with two hidden layers presented in figure 4.8.

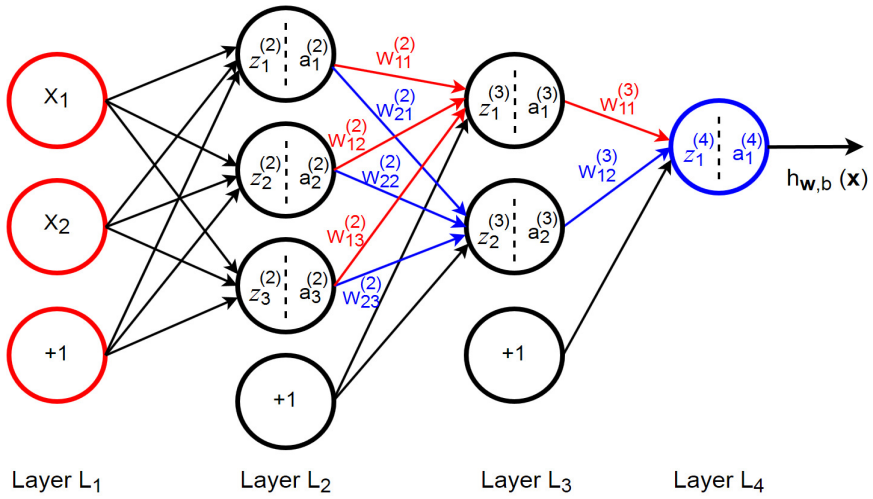


Figure (4.8) Detail explanation of the network with two hidden layers from figure 4.7

We write $W_{ij}^{(l)}$ to denote the parameter (or weight) associated with the connection between unit j in layer l, and unit i in layer $l+1$. The number of layers in the network is denoted by n_l, so in illustration of multi NNs model shown in figure 4.8, we have n_l=4.

The activation of unit i in layer l is denoted by a_i^l. For example $l = 1$, we also use $a_i^{(1)} = x_i$ to denote the $i-th$ input. Given a fixed setting of the parameters W and b, our neural network defines a hypothesis $h_{W,b}(x)$ that outputs a real number.

Now, we are going to look particularly at the last two layers, L_3 and L_4, where L_4 produces the output hypotheses and L_3 is the last hidden layer. Specifically, the computation of activation function in the layer L_3 can be derived as following:

$$\begin{aligned} a_1^{(3)} &= f(W_{11}^{(2)}a_1^{(2)} + W_{12}^{(2)}a_2^{(2)} + W_{13}^{(2)}a_3^{(2)} + b_1^{(2)}) \\ a_2^{(3)} &= f(W_{21}^{(2)}a_1^{(2)} + W_{22}^{(2)}a_2^{(2)} + W_{23}^{(2)}a_3^{(2)} + b_2^{(2)}) \end{aligned} \tag{4.16}$$

The hypothesis of the output of this neural network can be written as:

$$h_{wb}(x) = a_1^{(4)} = f(W_{11}^{(3)}a_1^{(3)} + W_{12}^{(3)}a_2^{(3)}) \tag{4.17}$$

If we let $z_i^{(l)}$ denoted as the total weighted sum of inputs to unit i in layer l, including the bias term, we will have:

$$z_i^{(l)} = \sum_{j=1}^{n} W_{ij}^{(l-1)} a_j^{(l-1)} + b_i^{(l-1)} \tag{4.18}$$

so that: $a_i^{(l)} = f(z_i^{(l)})$

or, for the computation of the next layer $l+1$ is:

$$z_i^{(l+1)} = \sum_{j=1}^{n} W_{ij}^{(l)} a_j^{(l)} + b_i^{(l)} \tag{4.19}$$

Using matrix-vectorial notation, the above equation 4.19 can be written as follow:

$$\mathbf{z}^{(l+1)} = \mathbf{W}^{(l)} \mathbf{a}^{(l)} + \mathbf{b}^{(l)} \tag{4.20}$$

so that:

$$\begin{aligned} \mathbf{a}^{(l+1)} &= f(\mathbf{z}^{(l+1)}) \\ &= f(\mathbf{W}^{(l)} \mathbf{a}^{(l)} + \mathbf{b}^{(l)}) \end{aligned} \tag{4.21}$$

The output of 4.17 can be written as:

$$\begin{aligned} h_{wb}(x) &= \mathbf{a}^{(4)} \\ &= f(\mathbf{z}^{(4)}) \\ &= f(\mathbf{W}^{(3)} \mathbf{a}^{(3)} + \mathbf{b}^{(3)}) \end{aligned} \tag{4.22}$$

Back Propagation

Suppose we have a fixed training set $\{(x^{(1)}, y^{(1)}), ..., (x^{(m)}, y^{(m)})\}$ of m training examples. We can train our neural network using batch gradient descent. For a single training example, the cost function with respect to single example is written as one-half square error as follow:

$$J(W, b; x, y) = \frac{1}{2} \| h_{W,b}(x) - y \|^2 \tag{4.23}$$

Given a training set of m examples, we then define the overall cost function to be:

$$\begin{aligned} J(W,b;x,y) &= [\frac{1}{m}\sum_{i=1}^{m} J(W,b;x^{(i)},y^{(i)})] + \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}(W_{ji}^{(l)})^2 \\ &= [\frac{1}{m}\sum_{i=1}^{m}(\frac{1}{2}\|h_{W,b}(x^{(i)}) - y^{(i)}\|^2)] + \frac{\lambda}{2}\sum_{l=1}^{n_l-1}\sum_{i=1}^{s_l}\sum_{j=1}^{s_{l+1}}(W_{ji}^{(l)})^2 \end{aligned} \tag{4.24}$$

The first term is an average sum-of-squares error term. The second term is a regularization term (also called a weight decay term) that tends to decrease the magnitude of the weights, and helps prevent over-fitting. The weight decay parameter λ controls the relative importance of the two terms.

Update Gradient Descent

To implement Gradient descent algorithm, the parameters of the gradient W, b must be updated as follows:

$$\begin{aligned} W_{ij}^{(l)} &= W_{ij}^{(l)} - \alpha\frac{\partial}{W_{ij}^{(l)}}J(W,b) \\ b_i^{(l)} &= b_i^{(l)} - \alpha\frac{\partial}{b_i^{(l)}}J(W,b) \end{aligned} \tag{4.25}$$

where α is the learning rate.

To do back propagation completely, the derivative of the overall cost function J(W,b) can be computed as:

$$\begin{aligned} \frac{\partial}{W_{ij}^{(l)}}J(W,b) &= \left[\frac{1}{m}\sum_{i=1}^{m}\frac{\partial}{W_{ij}^{(l)}}J(W,b;x^{(i)},y^{(i)})\right] + \lambda W_{ij}^{(l)} \\ \frac{\partial}{b_i^{(l)}}J(W,b) &= \left[\frac{1}{m}\sum_{i=1}^{m}\frac{\partial}{b_i^{(l)}}J(W,b;x^{(i)},y^{(i)})\right] \end{aligned} \tag{4.26}$$

Details of back propagation algorithm are:

1. Perform a feedforward pass, computing the activations for layers L_2, L_3, and so on up to the output layer L_{n_l}.
2. For each output unit i in layer n_l (the output layer), set:

$$\delta_i^{(n_l)} = \frac{\partial}{\partial z_i^{(n_l)}}\frac{1}{2}\|y - h_{W,b}(x)\|^2 = -(y_i - a_i^{(n_l)}.f'(z_i^{(n_l)}) \tag{4.27}$$

3. For $l = n_l - 1, n_l - 2, n_l - 3, ..., 2$
   For each node i in layer, set:

$$\delta_i^{(l)} = \left(\sum_{j=1}^{s_{l+1}} W_{ji}^{(l)} \delta_j^{(l+1)} \right) f'(z_i^{(l)}) \tag{4.28}$$

4. Compute the desired partial derivatives, which are given as:

$$\begin{aligned} \frac{\partial}{W_{ij}^{(l)}} J(W, b; x, y) &= a_j^{(l)} \delta_i^{(l+1)} \\ \frac{\partial}{b_i^{(l)}} J(W, b; x, y) &= \delta_i^{(l+1)} \end{aligned} \tag{4.29}$$

In the matrix-vectorial notation, the algorithm above can be re-written as:

1. Perform a feedforward pass, computing the activations for layers L_2, L_3, and so on up to the output layer L_{n_l}.
2. For output layer (layer n_l) , set:

$$\delta^{(n_l)} = -(y - a^{(n_l)} . f'(z^{(n_l)}) \tag{4.30}$$

3. For $l = n_l - 1, n_l - 2, n_l - 3, ..., 2$, set:

$$\delta^{(l)} = \left((W^{(l)})^T \delta^{(l+1)} . f'(z^{(l)}) \right) \tag{4.31}$$

4. Compute the desired partial derivatives, which are given as:

$$\begin{aligned} \nabla_{W^{(l)}} J(W, b; x, y) &= \delta^{(l+1)} (a^{(l)})^T \\ \nabla_{b^{(l)}} J(W, b; x, y) &= \delta^{(l+1)} \end{aligned} \tag{4.32}$$

Gradient Descent Algorithm

Now, we are ready to derive the full Gradient Descent Algorithm.

1. Set $\triangle W^{(l)} := \triangle b^{(l)} := 0$
2. For $i = 1$ to m,
   (a) Use back propagation to compute $\nabla_W^{(l)} J(W, b; x, y)$ and $\nabla_{b^{(l)}} J(W, b; x, y)$

(b) $\triangle W^{(l)} := \triangle W^{(l)} + \nabla_W^{(l)} J(W, b; x, y)$

(c) $\triangle b^{(l)} := \triangle b^{(l)} + \nabla_b^{(l)} J(W, b; x, y)$

3. Update the parameters:

$$\begin{aligned} W^{(l)} &= W^{(l)} - \alpha \left[\left(\frac{1}{m} \triangle W^{(l)} \right) + \lambda W^{(l)} \right] \\ b^{(l)} &= b^{(l)} - \alpha \left[\frac{1}{m} \triangle b^{(l)} \right] \end{aligned} \tag{4.33}$$

Finally, the NNs now can be trained by repeating the gradient descent steps to reduce the cost function $\boldsymbol{J}(\boldsymbol{W}, \boldsymbol{b})$.

After presenting the Introduction of Artificial Neural Networks, now we are going to present the Restricted Boltzmann Machines (RBM), which is the core stacks of Deep Belief Networks (DBN).

4.3.2 Restricted Boltzmann Machine (RBM)

Restricted Boltzmann Machines are some of the most common building blocks of deep probabilistic models. They are undirected probabilistic graphical models containing a layer of observable variables and single layer of latent variables. The discussion of this RBM concept based on the course note presented by Ali Ghodsi [79].

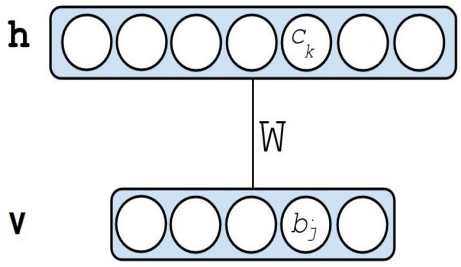


Figure (4.9) Illustration of basic RBM architecture

The RBM is the energy-based function which is defined as:

$$\begin{aligned} E(\mathbf{v}, \mathbf{h}) &= -\mathbf{b}^T \mathbf{v} - \mathbf{c}^T \mathbf{h} - \mathbf{v}^T W \mathbf{h} \\ &= -\sum_k b_k v_k - \sum_j c_j h_j - \sum_j \sum_k W_{jk} h_j v_k \end{aligned} \tag{4.34}$$

Conditional and Join Distribution

The join distribution can be written as:

$$p(\mathbf{v}, \mathbf{h}) = \frac{1}{Z} e^{-E(\mathbf{v},\mathbf{h})} \tag{4.35}$$

where Z is partition function:

$$Z = \sum_{v} \sum_{h} e^{-E(\mathbf{v},\mathbf{h})} \tag{4.36}$$

Then, conditional distribution can be derived as the following:

$$\begin{aligned} P(\mathbf{h}|\mathbf{v}) &= \frac{P(\mathbf{h}, \mathbf{v})}{P(\mathbf{v})} \\ &= \frac{P(\mathbf{h}, \mathbf{v})}{\sum_{\mathbf{h}} P(\mathbf{h}, \mathbf{v})} \\ &= \frac{\frac{1}{Z} e^{\mathbf{b}^T\mathbf{v} + \mathbf{c}^T\mathbf{h} + \mathbf{v}^T W \mathbf{h}}}{\sum_{\mathbf{h}} \frac{1}{Z} e^{\mathbf{b}^T\mathbf{v} + \mathbf{c}^T\mathbf{h} + \mathbf{v}^T W \mathbf{h}}} \\ &= \frac{\exp\left\{\mathbf{c}^T\mathbf{h} + \mathbf{v}^T W \mathbf{h}\right\}}{\sum_{\mathbf{h}} \exp\left\{\mathbf{c}^T\mathbf{h} + \mathbf{v}^T W \mathbf{h}\right\}} \end{aligned} \tag{4.37}$$

In the scalar form, the eq.4.37 can be written as:

$$P(\mathbf{h}|\mathbf{v}) = \frac{\exp\left\{\sum_{j=1}^{n} c_j h_j + \sum_{j=1}^{n} \mathbf{v}^T W_{:j} h_j\right\}}{\sum_{\mathbf{h}} \exp\left\{\sum_{j=1}^{n} c_j h_j + \sum_{j=1}^{n} \mathbf{v}^T W_{:j} h_j\right\}} \tag{4.38}$$

If we have exponential of sum, then it can be written as a product of exponential. Since the values of $\mathbf{h}$ are binary 0 and 1, the eq.4.38 then can be written as:

$$P(\mathbf{h}|\mathbf{v}) = \frac{\prod_{j=1}^{n} \exp\left\{c_j h_j + \mathbf{v}^T W_{:j} h_j\right\}}{\prod_{j=1}^{n} \sum_{h_j \,\epsilon\{0,1\}} \exp\left\{c_j h_j + \mathbf{v}^T W_{:j} h_j\right\}} \tag{4.39}$$

Then, we can compute for the $\sum_{h_j \,\epsilon\{0,1\}} \exp\left\{c_j h_j + \mathbf{v}^T W_{:j} h_j\right\}$ over two terms $h_j = 0$ and $h_j = 1$ that will give us:

$$P(\mathbf{h}|\mathbf{v}) = \prod_{j=1}^{n} \frac{\exp\left\{c_j h_j + \mathbf{v}^T W_{:j} h_j\right\}}{\left(1 + \exp\left\{c_j h_j + \mathbf{v}^T W_{:j} h_j\right\}\right)} \tag{4.40}$$

The inferring of the eq.4.40 indicates that the expression of $\frac{\exp\{c_j h_j + \mathbf{v}^T W_{:j} h_j\}}{(1+\exp\{c_j h_j + \mathbf{v}^T W_{:j} h_j\})}$ is known as the conditional probability for each h_j given the $\mathbf{v}$ values. Finally, we can write it as:

$$P(\mathbf{h}|\mathbf{v}) = \prod_{j=1}^{n} p(h_j|\mathbf{v}) = \prod_{j=1}^{n} sigm(c_j + \mathbf{v}^T W_{:j}) \tag{4.41}$$

and the probability for one individual unit h_j can be written as:

$$P(h_j = 1|\mathbf{v}) = sigm(c_j + \mathbf{v}^T W_{:j}) \tag{4.42}$$

The same thing can be derived for the conditional probability for each v_j given the $\mathbf{h}$ values, we have:

$$P(\mathbf{v}|\mathbf{h}) = \prod_{i=1}^{n} p(v_i|\mathbf{h}) = \prod_{i=1}^{n} sigm(b_i + W_{i:}\mathbf{h}) \tag{4.43}$$

and the probability for one individual unit v_i can be written as:

$$P(v_i = 1|\mathbf{h}) = sigm(b_i + W_{i:}\mathbf{h}) \tag{4.44}$$

The notations of the $W_{:j}$ and $W_{i:}$ are respectively all rows of column j of the matrix W and all columns of row i of matrix W.

Training RBM

To learn the parameters of the model, it is done by using the Gradient Descent method. In the case of energy function-based model, we need to minimize the energy with respect to the parameters ($W, \mathbf{b}$, and $\mathbf{c}$). The most obvious thing is to use the maximum log-likelihood. In this section, we will discuss how to optimize the parameters by using this maximum log-likelihood method.

Let's start by supposing we have some observations, and so we have bounce of $\mathbf{v}$ in our training dataset, $\mathbf{h}$ is not observed because it is just the hidden variable.

The log-likelihood l w.r.t parameters ($W, \mathbf{b}$, and $\mathbf{c}$) is given by:

$$\begin{aligned} l(W, \mathbf{b}, \mathbf{c}) &= \sum_{t=1}^{n} logP(\mathbf{v}^{(t)}) \\ &= \sum_{t=1}^{n} log \sum_{h} P(\mathbf{v}^{(t)}, \mathbf{h}) \end{aligned} \tag{4.45}$$

where: $P(\mathbf{v}^{(t)}, \mathbf{h}) = \frac{1}{Z} e^{-E(\mathbf{v}^{(t)}, \mathbf{h})}$, then we can write eq 4.45 as:

$$\begin{aligned} l(W, \mathbf{b}, \mathbf{c}) &= \sum_{t=1}^{n} log \sum_{h} \frac{1}{Z} e^{-E(\mathbf{v}^{(t)}, \mathbf{h})} \\ &= \sum_{t=1}^{n} \left[log \sum_{h} e^{-E(\mathbf{v}^{(t)}, \mathbf{h})} - log\, Z \right] \\ &= \sum_{t=1}^{n} log \sum_{h} e^{-E(\mathbf{v}^{(t)}, \mathbf{h})} - n\, log\, Z \\ &= \sum_{t=1}^{n} log \sum_{h} e^{-E(\mathbf{v}^{(t)}, \mathbf{h})} - n\, log \sum_{v,h} e^{-E(\mathbf{v}, \mathbf{h})} \end{aligned} \tag{4.46}$$

If we want to optimize the model, we need to maximize the log-likelihood. Let θ is the representation of the model parameters ($W, \mathbf{b}$, and $\mathbf{c}$), to optimize these parameters is as the following:

$$l(\theta) = \sum_{t=1}^{n} log \sum_{h} e^{-E(\mathbf{v}^{(t)}, \mathbf{h})} - n\, log \sum_{v,h} e^{-E(\mathbf{v}, \mathbf{h})} \tag{4.47}$$

To do this, we need to take derivative of eq.4.47 with respect to θ.

$$\begin{aligned} \nabla_\theta\, l(\theta) &= \nabla_\theta \sum_{t=1}^{n} log \sum_{h} e^{-E(\mathbf{v}^{(t)}, \mathbf{h})} - n\, \nabla_\theta log \sum_{v,h} e^{-E(\mathbf{v}, \mathbf{h})} \\ &= \sum_{t=1}^{n} \frac{\sum_h e^{-E(\mathbf{v}^{(t)}, \mathbf{h})} \nabla_\theta \{-E(\mathbf{v}^{(t)}, \mathbf{h})\}}{\sum_h e^{-E(\mathbf{v}^{(t)}, \mathbf{h})}} \\ &\quad - n\, \frac{\sum_{v,h} e^{-E(\mathbf{v}, \mathbf{h})} \nabla_\theta \{-E(\mathbf{v}, \mathbf{h})\}}{\sum_{v,h} e^{-E(\mathbf{v}, \mathbf{h}\}}} \end{aligned} \tag{4.48}$$

Finally we have two terms, the first term is the sum of the expectation of the derivative of energy function $\nabla_\theta \{-E(\mathbf{v}^{(t)}, \mathbf{h})\}$ with respect to the conditional distribution $P(\mathbf{h}|\mathbf{v}^{(t)})$ and the second term is expectation of the derivative of energy function $\nabla_\theta \{-E(\mathbf{v}, \mathbf{h})\}$ with respect to the join distribution $P(\mathbf{h}, \mathbf{v})$. Therefore, we can write the eq. 4.48 as follow:

$$\nabla_\theta\, l(\theta) = \sum_{t=1}^{n} \mathbb{E}_{P(\mathbf{h}|\mathbf{v}^{(t)})} \left[\nabla_\theta \{-E(\mathbf{v}^{(t)}, \mathbf{h})\} \right] - n\, \mathbb{E}_{P(\mathbf{h}, \mathbf{v})} \left[\nabla_\theta \{-E(\mathbf{v}, \mathbf{h})\} \right] \tag{4.49}$$

The first term is computable, but the second term is completely intractable because we need to sample from the unknown values both h and v. Fortunately, there is a method to

solve this problem, Hinton et al propose a method that so-called Contrastive Divergence. The idea is to replace the point estimate $\tilde{\mathbf{v}}$, then obtain the point $\tilde{\mathbf{v}}$ by using Gibb sampling. Do the sampling chain at $\mathbf{v}^{(t)}$.

Before going to continue discussing about the Contrastive Divergence, we firstly look at the derivative of the negative energy.

Derivation of Negative Energy

The gradient of the negative energy function with respect to $W, \mathbf{b}$, and $\mathbf{c}$, respectively are written as:

$$\nabla_W\{-E(\mathbf{v},\mathbf{h})\} = \frac{\partial}{\partial W}\left(\mathbf{b}^T\mathbf{v} + \mathbf{c}^T\mathbf{h} + \mathbf{v}^T W\mathbf{h}\right) = \mathbf{h}\mathbf{v}^T \tag{4.50}$$

$$\nabla_{\mathbf{b}}\{-E(\mathbf{v},\mathbf{h})\} = \frac{\partial}{\partial \mathbf{b}}\left(\mathbf{b}^T\mathbf{v} + \mathbf{c}^T\mathbf{h} + \mathbf{v}^T W\mathbf{h}\right) = \mathbf{v} \tag{4.51}$$

$$\nabla_{\mathbf{c}}\{-E(\mathbf{v},\mathbf{h})\} = \frac{\partial}{\partial \mathbf{c}}\left(\mathbf{b}^T\mathbf{v} + \mathbf{c}^T\mathbf{h} + \mathbf{v}^T W\mathbf{h}\right) = \mathbf{h} \tag{4.52}$$

So, the gradients of log-likelihood of 4.49 w.r.t $W, \mathbf{b}$, and $\mathbf{c}$, are respectively written as:

$$\nabla_W l\{\theta\} = \sum_{t=1}^{n} \hat{\mathbf{h}}^{(t)}\mathbf{v}^{(t)^T} - n\,\mathbb{E}_{P(\mathbf{h},\mathbf{v})}\left[\mathbf{h}\mathbf{v}^T\right] \tag{4.53}$$

$$\nabla_{\mathbf{b}} l\{\theta\} = \sum_{t=1}^{n} \mathbf{v}^{(t)^T} - n\,\mathbb{E}_{P(\mathbf{v},\mathbf{h})}\left[\mathbf{v}\right] \tag{4.54}$$

$$\nabla_{\mathbf{c}} l\{\theta\} = \sum_{t=1}^{n} \hat{\mathbf{h}}^{(t)} - n\,\mathbb{E}_{P(\mathbf{v},\mathbf{h})}\left[\mathbf{h}\right] \tag{4.55}$$

where: $\hat{\mathbf{h}}^{(t)} = sigm(\mathbf{c} + \mathbf{v}^{(t)}W)$.

Contrastive Divergence

The second term of the log-likelihood gradient in eq.4.49 is intractable. Fortunately, as proposed by Hinton et al.[77], this problematic can be solved by using Contrastive Divergence algorithm. Figure 4.10 illustrates the Constrastive Divergence process based on Markov Chain Monte Carlo (MCMC) using Gibbs sampling.

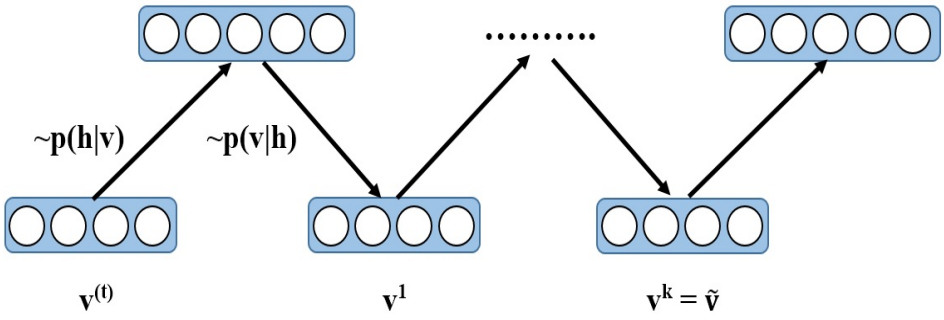


Figure (4.10) Constrastive divergence process based on MCMC using Gibbs sampling

The idea of the algorithm is as the following:

1. Replace the expectation by point estimate at $\tilde{\mathbf{v}}$
2. Obtain the point $\tilde{\mathbf{v}}$ by Gibbs sampling
3. Start sampling chain at $\mathbf{v}^{(t)}$

So we have an approximation of the derivative of energy function for the second term as follow:

$$\mathbb{E}_{P(\mathbf{h},\mathbf{v})}\left[\nabla_\theta\{-E(\mathbf{v},\mathbf{h})\}\right] \approx \nabla_\theta\{-E(\mathbf{v},\mathbf{h})\}|_{\mathbf{v}=\tilde{\mathbf{v}},\mathbf{h}=\tilde{\mathbf{h}}} \tag{4.56}$$

The pseudo code for this Contrastive Divergence algorithm can be written as the following:

1. For each training example $\mathbf{v}^{(t)}$
   (a) Generate a negative sample $\tilde{\mathbf{v}}$ using k step of Gibbs sampling, starting at $\mathbf{v}^{(t)}$
   (b) Update parameters:

$$\begin{aligned} W &\Longleftarrow W + \alpha\left(\mathbf{h}(\mathbf{v}^{(t)})\mathbf{v}^{(t)} - \mathbf{h}(\tilde{\mathbf{v}})\tilde{\mathbf{v}}^T\right) \\ \mathbf{b} &\Longleftarrow \mathbf{b} + \alpha\left(\mathbf{h}(\mathbf{v}^{(t)}) - \mathbf{h}(\tilde{\mathbf{v}})\right) \\ \mathbf{c} &\Longleftarrow \mathbf{c} + \alpha\left(\mathbf{v}^{(t)} - \tilde{\mathbf{v}}\right) \end{aligned} \tag{4.57}$$

2. Go back to 1 until stopping criteria

4.3.3 Training Deep Belief Network

Finally, we arrived at the training of the whole DBN network. To train the DBN, we have two important steps. The first step is called the pretraining. It is recursive unsupervised training of the RBMs stacks which is the step to learn the weights and biases in each RBM such that the input and the reconstruction are as close as possible. In this step, Contrastive Divergence proposed by Hinton et al.[77] is exploited. The second step is to add a class label layer and perform the supervised fine-tuning using Back Propagation and Gradient Descent algorithm to fit the parameters. Detail of the training DBN algorithm is written as the following and illustrated in figure 4.11:

1. Recursive training RBM stacks
   (a) Fit parameters for 1^{st} RBM
   (b) Freeze the parameters vector W^1 , and use sample W^1 as the data for training the 2^{nd} layer of RBM.
   (c) Freeze the parameters vector W^2 from the 2^{nd} layer RBM , and use sample W^2 as the data for training the 3^{rd} layer of RBM.
   (d) Do the same recursively for the next layers
2. Add output layer, run Back Propagation (BP) and Gradient Descent algorithms to fit the parameters

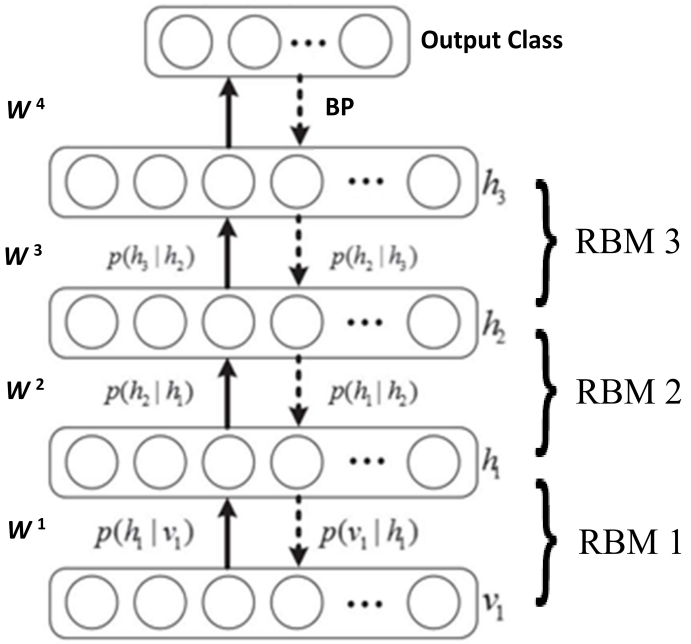


Figure (4.11) Ilustration of training DBN

4.3.4 Results and Discussion

To investigate the classification performance of the proposed system, in this section we will discuss about the DBN performances influenced by several factor including how many layers stacked in the network, number of training data and number of unit in every layer. To do this, we have analyzed different DBN architectures and settings. Different scenarios with different parameters and structures has been used to train the DBN using the radar signature dataset and the results of performances have been calculated and compared to each other.

Dataset

A total of 6000 data samples, 2000 dataset for each class including the pedestrian, cyclist and car, has been used as the training data and a total of 3000 data samples, 1000 samples per class, has been tested.

Experimental Observation for Impact of Many Layers

It is difficult to answer how many hidden layers must be stacked in DBN in order to have a good performance because there is no absolute answer for that. In one hand, stacking a few hidden layers for example, will be faster to train the network but result in poor performance in terms of accuracy because the network cannot extract sufficiently the features of the training dataset. In another hand, too many hidden layers stacked in the network may result in overfitting and slow learning time. Therefore, in order to evaluate the best architecture for our UWB radar system, there are three proposed DBN architectures with different number of hidden layers have been investigated in this work. The proposed DBN architectures for this UWB radar have been evaluated using MATLAB 2017 with DeepLearnToolbox [80]. The dataset has been trained and tested using a laptop with specification of Intel(R) Core(TM) i7-6700HQ CPU @2.6GHz with 16GB memory installed.

The results of the training errors have been calculated during the learning process and presented in the following figures 4.13, 4.15, and 4.17 respectively for the first, the second, and the third architecture. The training time consuming as the impact of the number of layer has been compared to each other as presented in the table 4.3.

We have set 3000 units in the layer connected to input, and 1000 units in the layer connected to output. The hidden layers inside have 750 units respectively. The detail architectures of the proposed DBN networks are set to 3000-2000, 3000-750-1000, and 3000-750-750-2000, respectively for the first, the second and the third.

The figures 4.12, 4.14, and 4.16 show the comparison of accuracy between the training and the validation data respectively for the first, the second, and third of the proposed DBN architectures. The figures 4.13, 4.15, and 4.17 show the loss comparison between the training and the validation data respectively for the first, the second, and third of the proposed DBN architectures. From these figures, we can see that the third DBN architecture with three

hidden layers performs better among the two others.

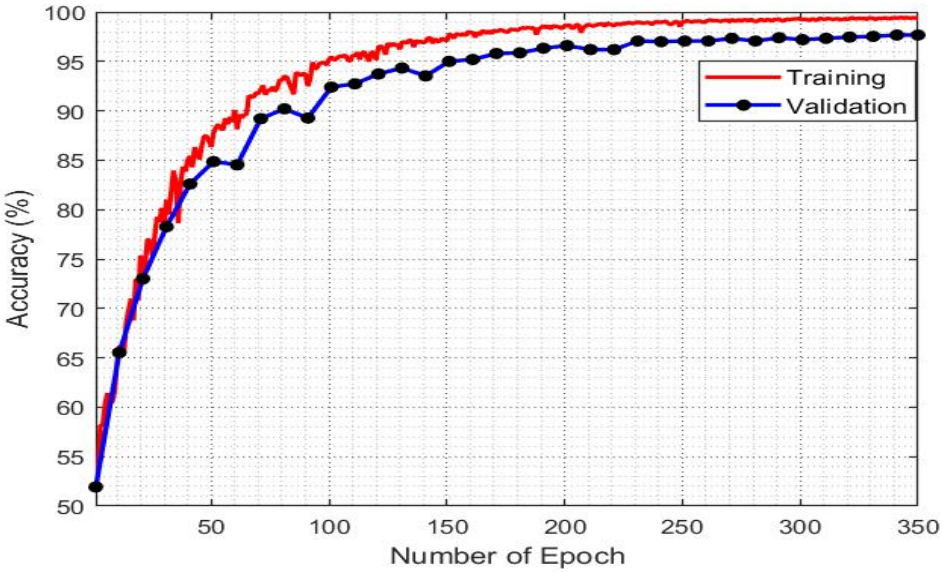


Figure (4.12) The accuracy comparison between the training and the validation data for the first architecture with two hidden layers: 3000-2000.

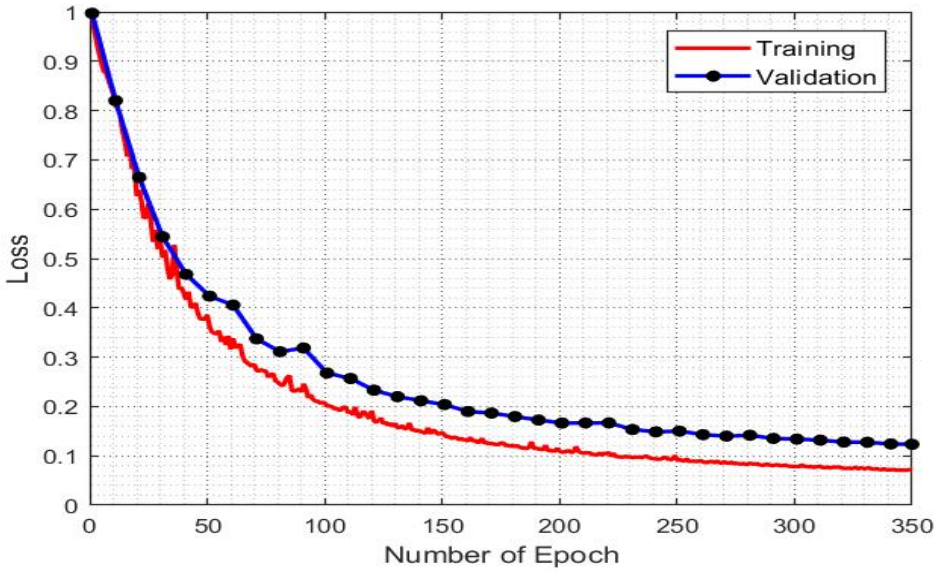


Figure (4.13) The loss comparison between the training and the validation data for the first architecture with two hidden layers: 3000-2000.

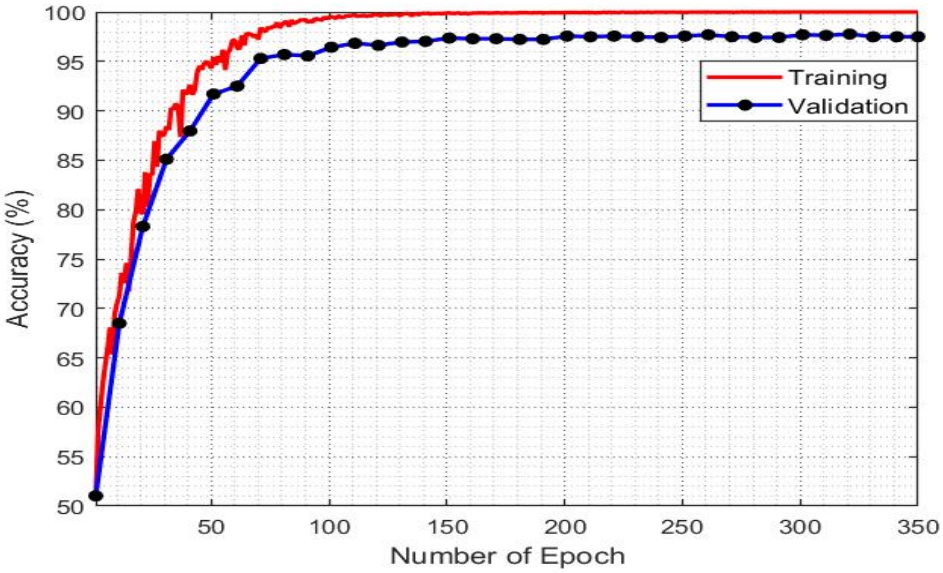


Figure (4.14) The accuracy comparison between the training and the validation data for the second architecture with three hidden layers: 3000-750-2000.

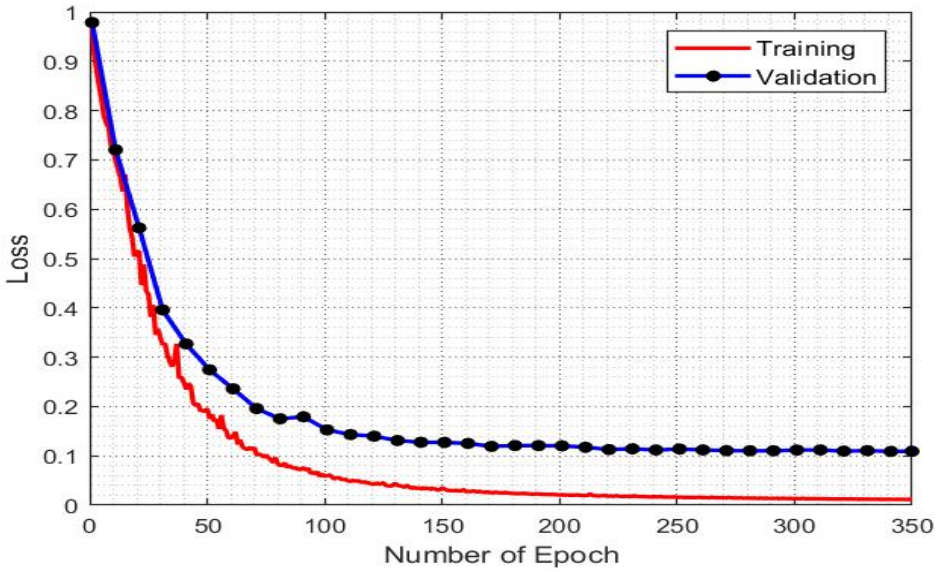


Figure (4.15) The loss comparison between the training and the validation data for the second architecture with three hidden layers: 3000-750-2000.

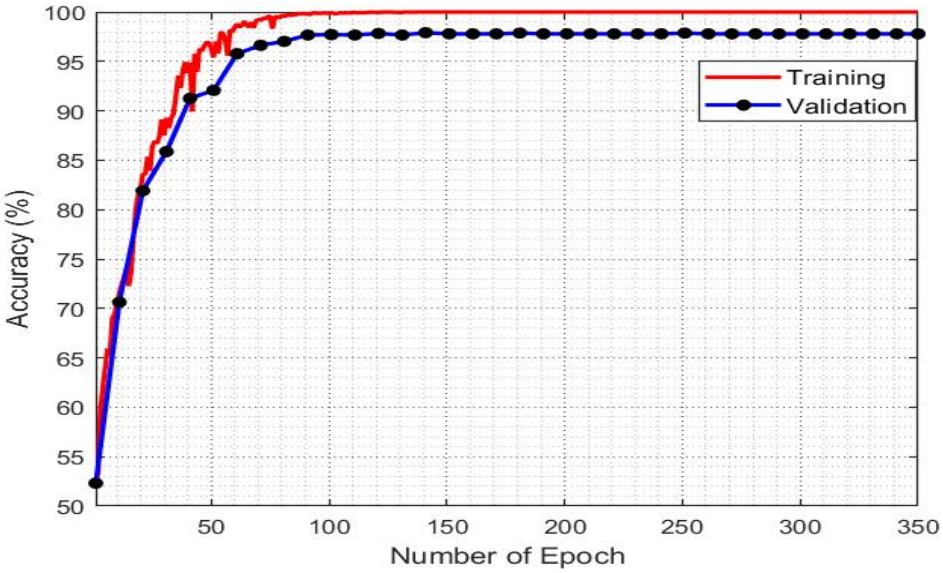


Figure (4.16) The accuracy comparison between the training and the validation data for the third architecture with four hidden layers: 3000-750-750-2000.

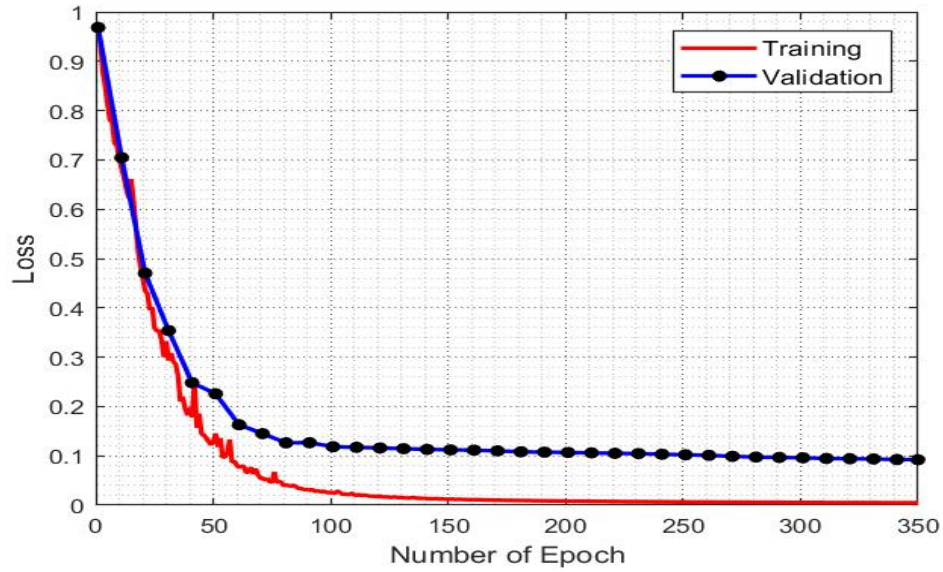


Figure (4.17) The loss comparison between the training and the validation data for the third architecture with four hidden layers: 3000-750-750-2000.

The table 4.3 shows that with the increase of number of layer, the training time operation also increases. We can see that the DBN architecture with two hidden layers consumes less time compared to that of DBN architecture with 3 and 4 hidden layers. In terms of time operation for training the DBN, the less hidden layers, the less time of training operation. But, in terms of error, it shows contrary that the less hidden layers, the more error.

Table (4.3) The result of DBNs with different numbers of layers

Number of Hidden Layers	Unit Layer	Training Time (m)	Error
2	3000-2000	30.46	0.124
3	3000-750-2000	50.25	0.109
4	3000-750-750-2000	100.17	0.092

The confusion matrix of the classification results for the three different DBN architectures are shown in table 4.4, 4.5, and 4.6, respectively. The table 4.7 shows the summarize of these confusion matrix in terms of total accuracy. In this table we also compare these DBN architectures performances with the previous results of SVM method. It can be seen that the third DBN architecture with three hidden layer (i.e 3000-750-750-2000) gives better results compared to SVM with RBF kernel and the other of two DBN architectures.

Table (4.4) Confusion matrix for the configuration of hidden layer architecture: 3000-2000

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	96.24%	1.04%	2.60%
	Cyclist	0.45%	97.92%	1.07%
	Pedestrian	3.31%	1.04%	96.33%

Table (4.5) Confusion matrix for the configuration of hidden layer architecture: 3000-750-2000

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	96.96%	1.50%	1.04%
	Cyclist	2.44%	97.15%	0.74%
	Pedestrian	0.60%	1.35%	98.22%

Table (4.6) Confusion matrix for the configuration of hidden layer architecture: 3000-750-750-2000

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	96.86%	0.62%	1.18%
	Cyclist	1.35%	98.01%	0.30%
	Pedestrian	1.79%	1.37%	98.52%

Table (4.7) Comparison results of the three DBN architectures and SVM with RBF kernel

Machine Learning Model	**Maximum Average Accuracy**
DBN: 3000-2000	96.83%
DBN: 3000-750-2000	97.44%
DBN: 3000-750-750-2000	97.80%
SVM: RBF Kernel	96.24%

4.3.5 Performance Comparison of DBN and SVM

Figure 4.18 shows the comparison result of DBN and SVM in terms of accuracy vs number of training samples. The figure shows that the performance of SVM achieved its highest average accuracy up to 96.24% when the use of number of training samples about 500 per target class. While DBN can achieved up to 97.80% of its average accuracy when the use of number of training samples around 2000 per target class. It can be concluded that the advantage of using DBN is that the more data training samples used, the better accuracy can be achieved.

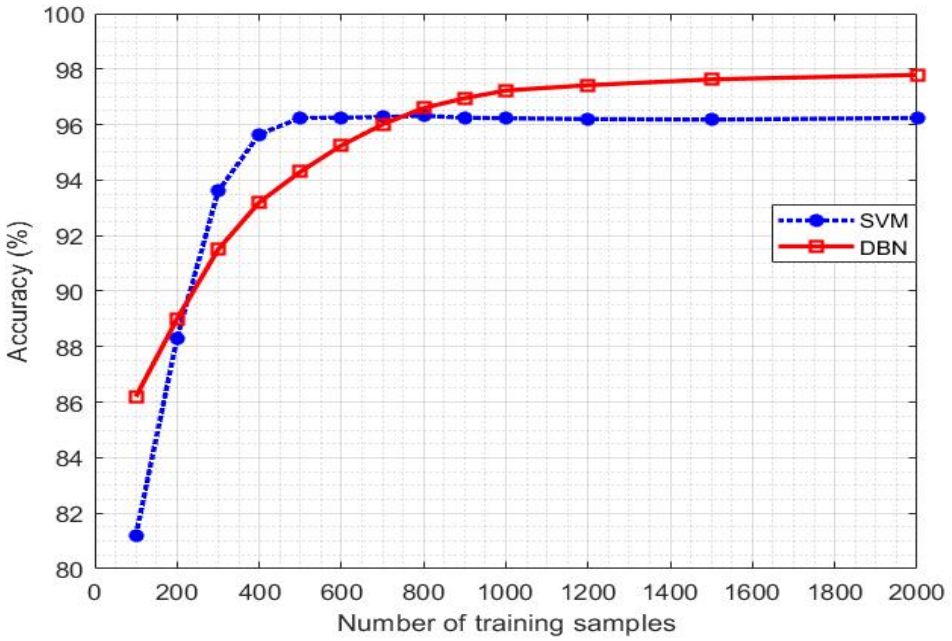


Figure (4.18) Comparison SVM and DBN performance for the use of different numbers of training samples

4.4 Conclusion of Chapter 4

As one of the advantages of using UWB technology in radar system is, that the UWB radar can distort pulse differently for different target nature. The difference of these pulses distortion are called radar signatures. These radar signatures can be used to classify the target type using a machine learning. In this chapter, we have discussed about two different types of machine learning algorithms in order to classify the UWB radar target types. Support Vector Machine (SVM) and Deep Belief Networks (DBN) have been investigated. The input feature of this recognition method is based on 1-D normalized radar signature.

SVM is a supervised machine learning algorithm of binary classification technique for pattern recognition. This classifier creates a hyper-plane to separate the pattern of data into two classes (+1 or -1) with the maximum margin. The vectors that define the hyperplane

are called *support vectors*. There are several SVM kernels have been analyzed in this radar identification including Linear, Sigmoid and Radial Based Function (RBF) kernel. Among these kernels, the best performance is achieved with the uses of RBF kernel.

DBN is a Deep Learning method, also has been investigated in this work. There are three types of DBN's architecture have been proposed. The first architecture is a DBN with two hidden layers: 3000-2000, the second is a DBN with 3 hidden layers: 3000-750-2000, and the third is a DBN with four hidden layers: 3000-750-750-2000. The performances of these architectures have been evaluated and compared to each other. We found that the DBN architecture with four hidden layers performs better than the others. We have also compared the performances of SVM and DBN, and the result is DBN performs better than SVM for this UWB radar.

After investigating of exploiting 1-D radar signature in identifying of radar targets, in the next chapter, it will be discussed about the evaluation of detecting and recognizing of UWB radar target using 2-D radar approach.

CHAPTER 5

ENHANCING RADAR TARGET DETECTION AND IDENTIFICATION

Contents

This chapter will be focused on a new proposition of UWB radar target detection and identification using 2-D (B-Scan) radar approach. After investigating the 1-D raw radar data, we realized that for the targets like pedestrian and cyclist, the radar receives a very low SNR signal, and it frequently losses the target position. Therefore, in order to enhance the target detection and identification rate, we have evaluated 2-D radar data approach.

In this work, we firstly focused on increasing radar Signal-to-Noise Ratio (SNR) by proposing a new noise removal method. After that, we concerned to enhance the radar target identification rate by investigating the use of Convolutional Neural Network (CNN).

5.1 Study of Noise Removal Techniques

The target detection is highly influenced by the capacity to distinguish the useful radar information and the noise successfully. In this section, a novel method of noise suppression of UWB short range radar dedicated to cyclists and pedestrians detection is presented. It is a combination of Wavelet Shrinkage Denoising (WSD) and Higher Order Statistics (HOS). To evaluate the performance of the proposed method, different types of noise removal techniques including Principal Component Analysis (PCA), Singular Value Decomposition (SVD), WSD and HOS have been applied, and the results have been analyzed. The SNRs of the final results have been computed to compare the effectiveness of individual noise removal techniques.

We begin this discussion by presenting the basic principles of different noise removal methods including the PCA, SVD, WSD, and our proposed method, that is a the combination of HOS and WSD. This section will be finished by presenting the results and discussion of the performance of these evaluated methods.

5.1.1 Principal Component Analysis (PCA)

PCA specifically is the technique of array processing and data analysis [82] and is also one of the most popular technique in signal processing. It works very well for image compression, finding patterns in high-dimensional data, relationship among variables, face recognition and in other fields. Basically, the PCA is a statistical technique used to reduce data dimensionality by finding the orthogonal linear transformation that maximizes the variance of the variables. PCA also has the ability to reduce clutters and noise of the radar data by obtaining the most principal eigenvalues which contain the useful information of radar impulse. To use the PCA in clutter and noise removal, there are several steps that need to be followed:

Finding the mean of data dimensions

The original 2-D radar data $\mathbf{X}$ where rows correspond to particular measurements and columns represent their data samples is firstly processed to find the mean value for each

measurement:

$$\mu_i = \frac{1}{N}\sum_{j=1}^{N} x_{ij} \tag{5.1}$$

with μ_i is the mean of the corresponding i-th measurement and x_{ij} denotes the j-th data sample of the i-th measurement, where $i = 1, 2, ..., M$ and $j = 1, 2, ..., N$. The M and N are the number of measurement and the number of their data samples.

Normalizing the data

After, finding the mean value for each measurement, the original data is then normalized by subtracting the mean value from each measurement. This process is also called adjusted mean where each measurement is adjusted to have the zero-mean.

$$x_{ij(Adj)} = x_{ij} - \mu_i \tag{5.2}$$

The new adjusted radar data $\mathbf{X}_0$ whose mean is zero takes the following form:

$$\mathbf{X}_0 = [\mathbf{x}_{1(Adj)}\ \mathbf{x}_{2(Adj)}\ ...\ \mathbf{x}_{i(Adj)}]^T \tag{5.3}$$

Calculating covariance C of the adjusted mean data matrix $\mathbf{X}_0$

The covariance matrix $\mathbf{C}$ can be then expressed as a dot product of the adjusted matrix $\mathbf{X}_0$ that is divided by the number of data samples N:

$$\mathbf{C} = conv(\mathbf{X}) = \frac{1}{N}\mathbf{X}_0\mathbf{X}_0^T \tag{5.4}$$

Calculating the eigenvalue λ_i and eigenvectors $\mathbf{e}_i$ of the co-variance matrix

From the covariance matrix $\mathbf{C}$, eigenvalues λ_i and eigenvectors $\mathbf{e}_i$ can be found by using the following relation:

$$\mathbf{C}\mathbf{e}_i = \lambda_i\mathbf{e}_i \tag{5.5}$$

Choosing the principal components

Then, eigenvectors are ordered by their eigenvalues from highest to lowest, which returns the components in order of significance. Eigenvector with the highest eigenvalue is the most

principle component. It is the most significant relationship between the data dimensions. We can select the k-first greater eigenvectors $\mathbf{Z}$ to reconstruct the free-noise radar data as shown in eq.5.6.

$$\mathbf{Z} = [\mathbf{e}_1\ \mathbf{e}_2\ \mathbf{e}_3\ \mathbf{e}_4\ ...\ \mathbf{e}_k] \tag{5.6}$$

Reconstructing radar data

Finally, the free-noise radar data $\hat{\mathbf{X}}$ can be reconstructed based on the chosen of k principal components using the relation of eq.5.6. The dimension of reconstructed data is the same as the original data.

$$\hat{\mathbf{X}} = \mathbf{X}\mathbf{Z}\mathbf{Z}^T \tag{5.7}$$

5.1.2 Singular Value Decomposition (SVD)

In clutter and noise removal, SVD can be used to separate the data matrix into complementary subspaces called signal and noise. SVD has been used as a noise reduction technique in many fields application such as speech processing, image processing, and radar application [83].

For noise removal using SVD in radar system, B-scan radar data is represented by a matrix $\mathbf{X}$, the same as in PCA method, where rows correspond to particular measurements and columns represent their data samples. The dimension of $\mathbf{X}$ is M x N, where M is the number of measurements and N is the number of their data samples for each measurement. Decomposition of data matrix $\mathbf{X}$ can be written as:

$$\mathbf{X} = \mathbf{U}\mathbf{S}\mathbf{V}^T \tag{5.8}$$

The matrix $\mathbf{U}$ called matrix of left singular vectors, whose dimension is M x M and $\mathbf{V}$ called matrix of right singular vectors, whose dimension is N x N. The matrix $\mathbf{S}$ is a diagonal matrix R x R, where the non-zero of its elements are called singular values. The entries of this singular values are positive and sorted in decreasing order ($\sigma_1 > \sigma_2 ... > 0$).

Singular values contain clutter, target, and noise information. If these values are known, clutter and noise can be suppressed by eliminating these singular values, then the data matrix can be reconstructed using inverse transformation of the SVD. SVD is written as:

$$\mathbf{X} = \sum_{i=1}^{N} \sigma_i u_i v_i^T \tag{5.9}$$

The $\mathbf{X}$ can be decomposed into two subspaces, signal and noise as follow [8]:

$$\mathbf{X} = \sum_{i=1}^{k} \sigma_i u_i v_i^T + \sum_{i=k+1}^{N} \sigma_i u_i v_i^T \tag{5.10}$$

The first term of the $\mathbf{X}$ in eq.5.10 is expected to be associated with the target signal, and the second term is supposed to be the clutter and the noises.

5.1.3 Wavelet Shrinkage Denoising (WSD)

There are two terms in context of wavelet noise processing, smoothing and denoising. Smoothing removes high frequencies and retains low frequencies whereas denoising attempts to remove whatever of the noise present and sustain whatever of the signal regardless of the frequency content of the signal [84]. Noise removal using Wavelet transforms domain involves three steps: a linear forward wavelet transform, nonlinear thresholding step and a linear inverse wavelet transform. The result of Wavelet noise removal preserve the signal characteristics, and regardless of its frequency contents. Let's say that the observed signal is:

$$X(t) = S(t) + N(t) \tag{5.11}$$

where S(t) is the uncorrupted signal with additive noise N(t). Let W(.) and W^{-1} (.) denote the forward and inverse wavelet transform operators, respectively. So the denoising procedure can be written as:

$$\begin{aligned} Y &= W(X) \\ Z &= D(Y, \lambda) \\ S &= W^{-1}(Z) \end{aligned} \tag{5.12}$$

Let $D(., \lambda)$ denotes the denoising operator with threshold λ. We intend to denoise $X(t)$ to recover $\hat{S}(t)$ as an estimate of $S(t)$ [85].

Threshold determination is an important question when denoising the data using WSD. A small threshold may yield a result close to the input, but the result may still be noisy. A large threshold on the other hand, produces a signal with a large number of zero coefficients. This leads to a smooth signal will destroy a lot of useful information. So we have chosen the universal threshold $\lambda_{UNIV} = \sqrt{2lnN}\sigma$, ($N$ being the signal length, σ being the noise standard deviation) [85].

5.1.4 Combination of HOS and WSD (Proposed Method)

The motivation of this proposition is the methods like SVD and PCA provide barely good separation between the radar signal and the noise. Consequently, when dealing with the low SNR radar signal, it is difficult to determine the radar target. In addition, we have evaluated the performance of HOS and WSD differently, both methods promise a good improvement of radar SNR compared to SVD and PCA. Starting from these points, a promising result can be achieved if we success to combine these two methods.

The idea is simple, the good capacity of noise suppression of the HOS and the strong signal separation technique of Wavelet Shrinkage Denoising are combined. The signal after applying the pulse canceller is firstly processed by the HOS, then the result of this first stage of proposed noise removal is then processed by WSD as the second stage of proposed noise removal. With this method, we have a better performance compared to the other clutter removal methods. The comparison results between the proposed and the other noise removal methods are presented in next subsection.

5.1.5 Results and Discussion

To measure the performance of the several noise removal methods, we used the Signal-to-Noise Ratio (SNR) comparison. The SNR is calculated by calculating the average energy of target region signal (S_r) divided by the average energy of the noise and clutter (N_r). To obtain the S_r and N_r, a B-scan data after noise removal is firstly normalized and applied a threshold value (T_r).

The separation of signal (S_r) and noise (N_r) after applying a threshold value is written as:

$$\boldsymbol{B\text{-}scan}_{afterTr} = \begin{cases} S_r & B\text{-}scan > T_r \\ N_r & otherwise \end{cases} \tag{5.13}$$

The SNR value is given by 5.14 where A_t is the size of target region (i.e. the number of pixels where the target is located), A_{nc} is supposed to be the region of the noise and clutter.

$$SNR = 10log\left(\frac{\frac{1}{A_t}\sum_{(i,j)\epsilon A_t}|S_r(i,j)|^2}{\frac{1}{A_{nc}}\sum_{(i,j)\epsilon A_{nc}}|N_r(i,j)|^2}\right) \tag{5.14}$$

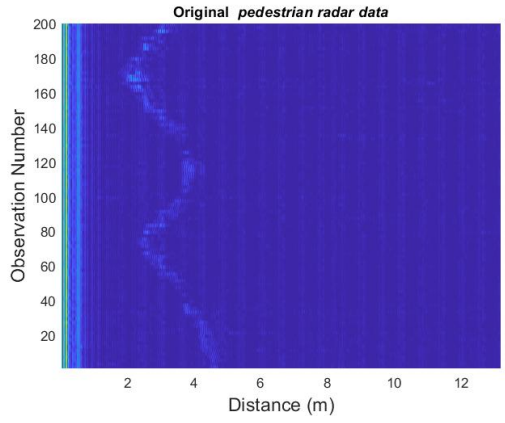
Table 5.1 shows the change in the SNR (dB) for all these techniques, as well as for the proposed technique (WSD+HOS).

Figure 5.1 shows the original radar data for three different targets including pedestrian, cyclist, and car. Figure 5.2, 5.3 and 5.4 present the results of implementation several noise removals on pedestrian, cyclist, and car radar data, respectively. If we compare these methods, we can see that the performance of combination WSD and HOS outperforms the other

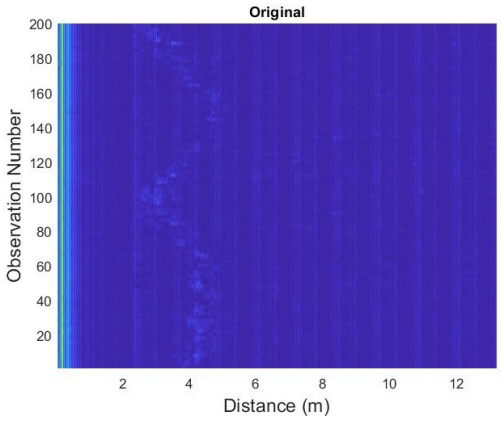
Table (5.1) Performance Comparison SNR(dB) For Different Noise Removal Methods

Target Types	***Clutter Removal Methods***					
	Pulse Canceller	**PCA**	**SVD**	**WSD**	**HOS**	**WSD+ HOS**
Pedestrian	16.34	19.12	20.53	25.18	25.18	**32.47**
Cyclist	15.68	18.26	19.48	23.22	26.02	**29.13**
Car	20.82	21.44	21.44	29.30	33.99	**38.61**

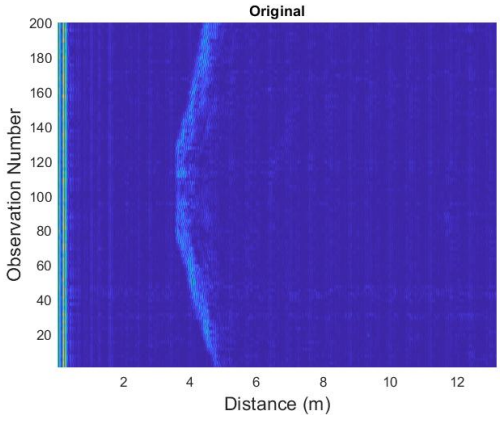
methods. The results also show that the combination of WSD and HOS promises a good SNR compared to the other methods as shown in Table 5.1. The SNR has increased up to 16.13dB, 13.45dB and 17.79dB measured, after applying the pulse canceller, for respectively the pedestrian, cyclist, and car B-scan data.



(a) Original pedestrian data



(b) Original cyclist data



(c) Original car data

Figure (5.1) Original B-scan radar data for three different targets

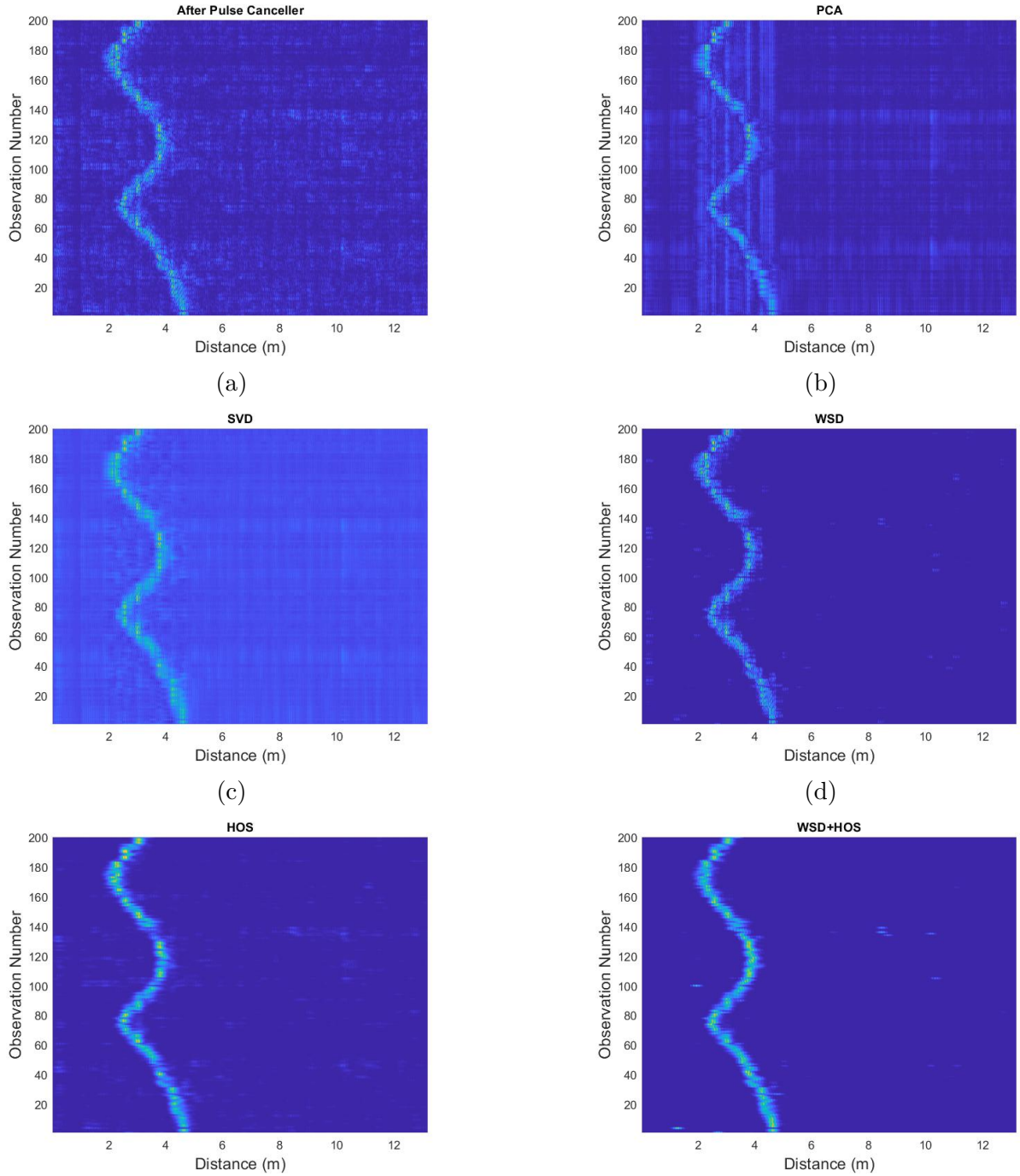


Figure (5.2) Results of applying noise removal methods for pedestrian data (a) B-scan after pulse canceller, direct coupling has been eliminated, (b) B-scan after PCA processing, (c) B-scan after SVD processing, (d) B-scan after WSD processing, (e) B-scan after HOS processing, (f) and B-scan after using combination WSD and HOS.

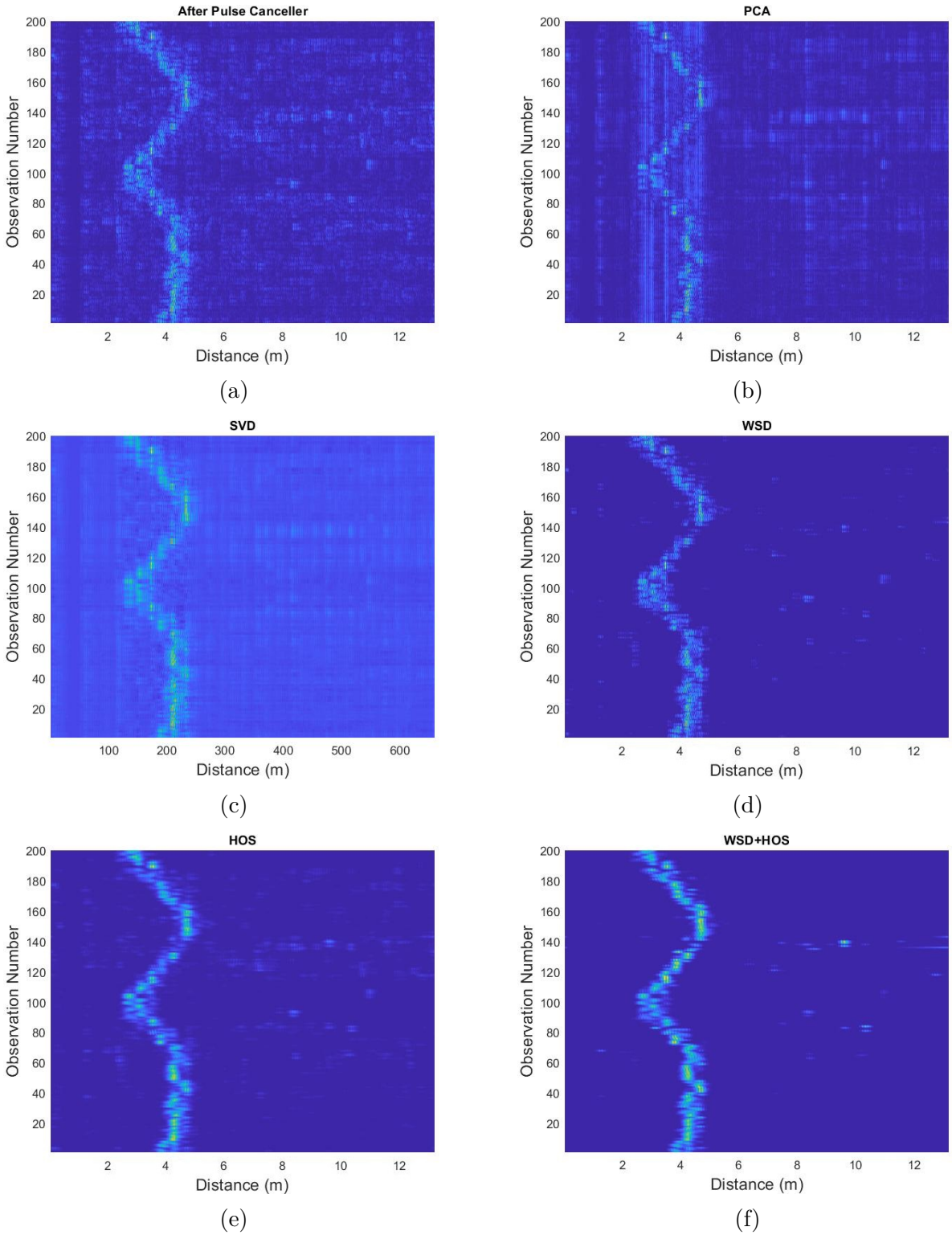


Figure (5.3) Results of applying noise removal methods for cyclist data (a) B-scan after pulse canceller, direct coupling has been eliminated, (b) B-scan after PCA processing, (c) B-scan after SVD processing, (d) B-scan after WSD processing, (e) B-scan after HOS processing, (f) and B-scan after using combination WSD and HOS.

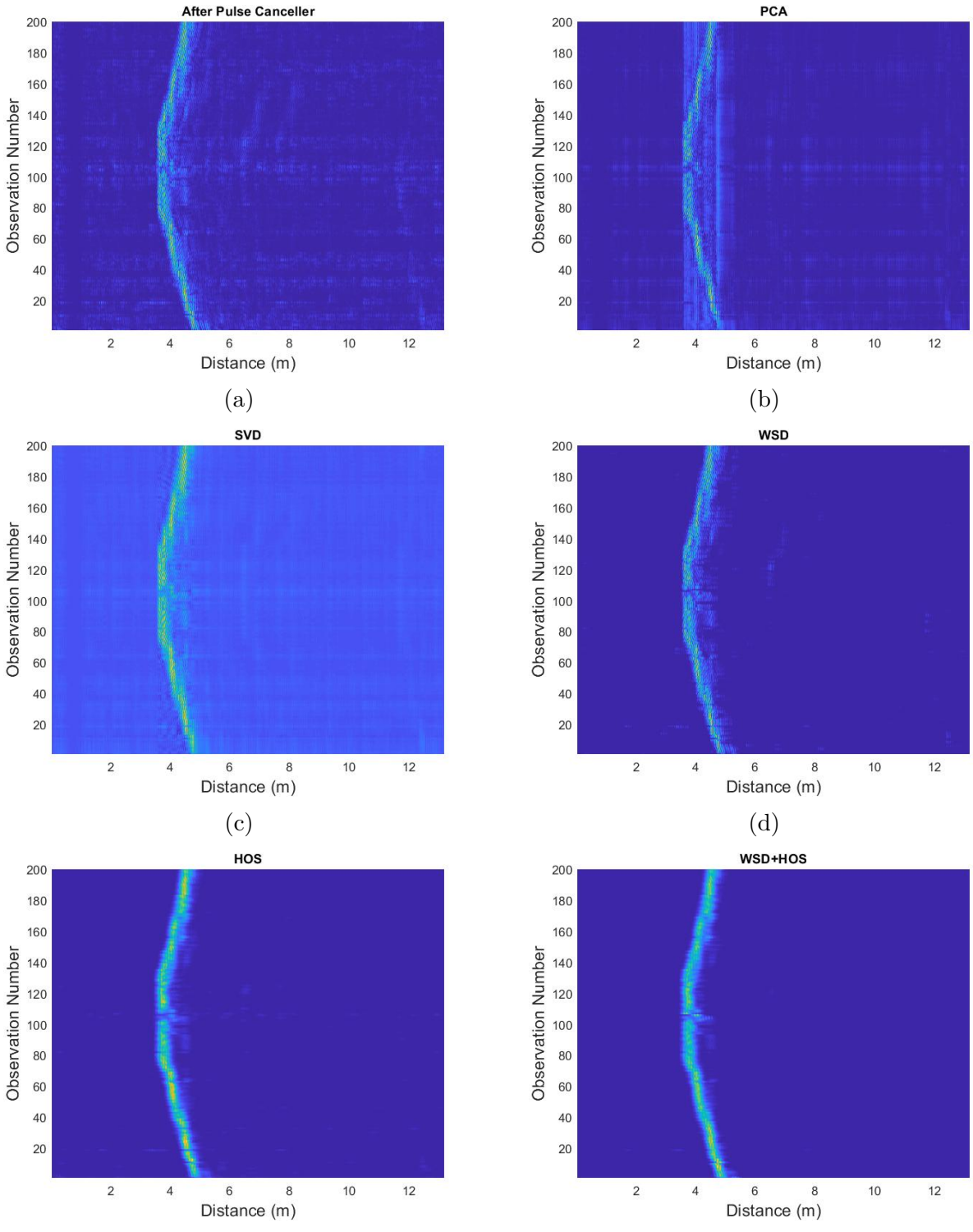


Figure (5.4) Results of applying noise removal methods for car data (a) B-scan after pulse canceller, direct coupling has been eliminated, (b) B-scan after PCA processing, (c) B-scan after SVD processing, (d) B-scan after WSD processing, (e) B-scan after HOS processing, (f) and B-scan after using combination WSD and HOS.

5.2 Convolutional Neural Network (CNN) Approach for Enhancing the Identification of UWB Radar Targets

In this section, we will discuss about the use of Convolution Neural Network (CNN) for enhancing UWB radar targets recognition. We address this issue particularly on identification of vulnerable road users (VRUs) like cyclists and pedestrians. We begin the discussion by presenting the preprocessing 2-D radar data, following by introducing a general idea of CNN, then providing the proposed CNN architectures. We will finish this section by presenting the experimentation results and discussions.

This work relates to the previous section 5.1, where the result of noise removal on 2-D radar data will be used as the input of CNN classifier.

5.2.1 Preprocessing 2-D Radar Data

Sliding Windows

We begin this discussion by presenting the preprocessing input data for CNN. Figure 5.5 shows a sliding window process on the result of radar data after applying the proposed noise removal on B-Scan radar data. The result is an image of size 200 x 660 pixels density. Target detection can be determined by applying sliding window over the newly incoming radar data. We limit the searching window for *height* and *width* is 20 and 50 pixels, respectively. This means that, we take 20 successive radar frames then we slide the window from the left to the right until the the last possible position in the searching space.

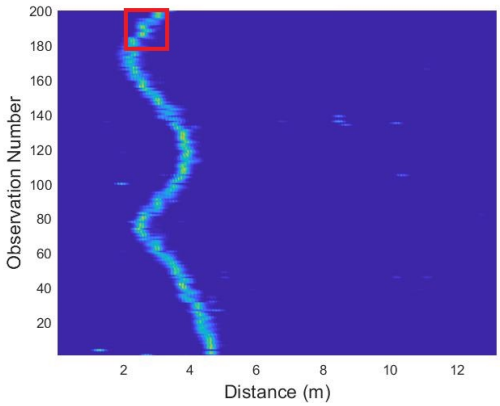


Figure (5.5) Sliding window process on the result of radar data after applying noise removal method

Non-Maximum Suppression

This method results the possibilities of the same target as several different targets. Therefore the non-maximum suppression (NMS) is required to be applied. We apply a very simple of NMS method that works effectively in our case. Figure 5.6 presents the result of applying NMS after the sliding window searched. The first step of NMS is to find the groups of detection range that distributes in a bin of range detection. Then, we apply the edge detection to find the cluster separation. Once the groups of the range candidate is determined then search for the maximum energy value and suppress any of non-maximum candidates for each group. We save this coordinate information to the next step.

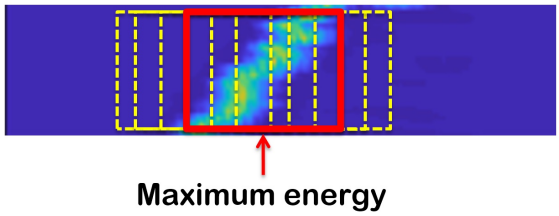


Figure (5.6) Description of applying Non-Maximum Suppression to obtain the precision position on the overlapped windows

After obtaining a good position of the radar target, then keep this area of this detected maximum energy as the radar target position area.

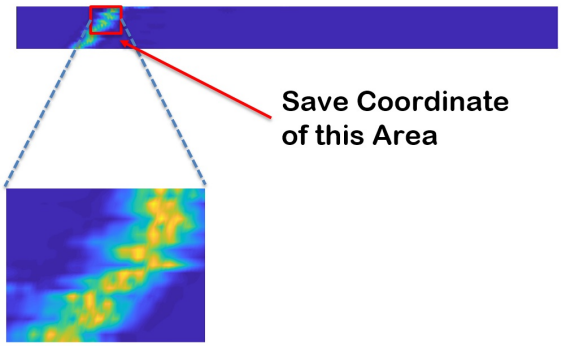


Figure (5.7) Saving co-ordinate of the detected maximum energy area

Centralizing of Energy Distribution

In order to have a zero means of energy distribution of detected target, we align the detected area as shown in figure 5.8

Transforming Data Into Power Spectral Density

After centralizing the energy distribution of the data, the detected area now is transformed into power spectral density. The aim of this process is to have a robust and strong char-

Figure (5.8) Centralizing of Energy Distribution

acteristics for different targets classes. The result examples of this transformation for three different target classes can be seen in figure 5.9.

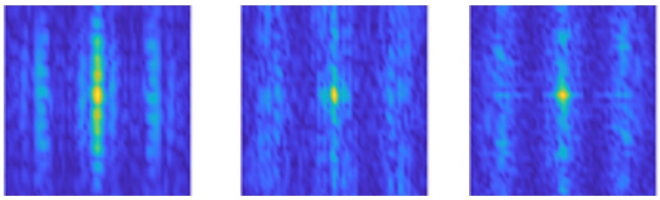


Figure (5.9) Result examples of transforming detected region into power spectral density

5.2.2 Convolution Neural Network

CNN is a deep neural network topology which consists of convolutional layers and activation layers, these layers transform the input into features and connect the features map to the output class probabilities. In this work, we have investigated the use of CNN in enhancing UWB radar target recognition rate.

Basic Concept of CNN

The input to a convolutional layer is a m x m x r image where m is the height and width of the image and r is the number of channels, e.g. an RGB image has $r = 3$. The convolutional layer will have k filters (or kernels) of size n x n x q where n is smaller than the dimension of the image and q can either be the same as the number of channels r or smaller and may vary for each kernel [86].

We present the basic principle of CNN for LeNet-5 architecture as shown in figure 5.10 (LeCunn, 1998) as a classical CNN architecture appeared in 1998. This network has been used to train MNIST digit dataset with 60K training examples. The architecture is illustrated as the following:

The architecture composes of convolutional layer (C1), pooling layer (S2), convolutional layer (C3), pooling layer (S4) and convolutional layer (C5). The convolutional layer C1 has 6 feature maps and each feature map contains 28x28 neurons. The pooling layer S2 contains 6

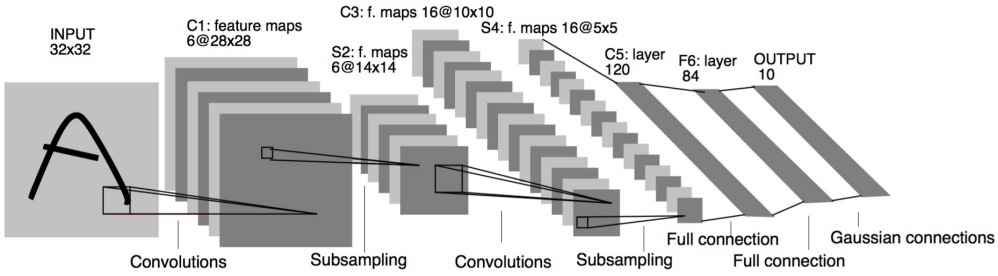


Figure (5.10) Architecture of LeNet-5 [88]

feature maps and 14x14 units neurons for each feature map. Layer C3 is a convolution layer with 16 feature maps and 10x10 neurons in each feature map. Layer S4 is a pooling layer with 16 feature maps and 5x5 units for each feature map. C5 and F6 are last convolution layer and fully connected layer, respectively. The last layer is the output layer contains 10 units and performs classification.

There are four main operations in the CNN architecture shown in Figure 5.10 above: Convolution, Non Linearity (ReLU), Pooling or Sub Sampling, and Classification (Fully Connected Layer).

Convolution layer

The main purpose of Convolution in a CNN is to extract features from the input image. Convolution preserves the spatial relationship between pixels by learning image features using small squares of input data. The input and output of each stage are sets of arrays called feature maps [75].

The basic principle of convolution between two function f and g is written as:

$$(f * g)(x) = \sum_{t} f(t)g(x+t) \tag{5.15}$$

Let K be the kernel of size $2h_1+1$ x $2h_2+1$, the convolution between image I and kernel K at the pixel (r, s) is performed as [76]

$$(I * K)(r, s) = \sum_{u=-h_1}^{h_1} \sum_{v=-h_2}^{h_2} K(u, v)I(r+u, s+v) \tag{5.16}$$

In CNN, a convolution kernel is called filter. 2-D convolution is done by spatially moving a filter over the image. If x and y are respectively denoted as input and output image, and y is the output then b is a bias, k_{ij} is the kernel that connects input feature map x_i to output feature map y_j, then the output feature map y_j is computed as [75]:

$$y_j = b_j + \sum_{i} k_{ij} * x_i \tag{5.17}$$

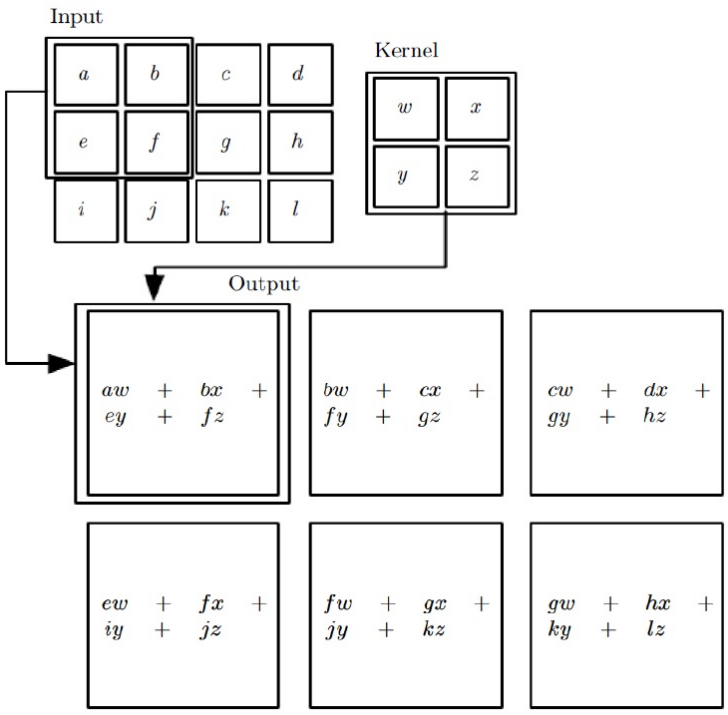


Figure (5.11) Illustration of 2-D convolutional process (source: [87]

where $*$ is the 2D discrete convolution operator and b_j is a trainable bias parameter. Figure 5.11 illustrates the process of 2-D convolution.

Non-linearity layer

The purpose of this layer is to introduce non-linearity to a system that basically has just been computing linear operations during the convolution layers. *Tanh* function was used in traditional convolution network as an activation function, but recently, researchers use more sophisticated non-linearity function like rectified linear unit (ReLU) because the network is able to train a lot faster [75]. The mathematical structure of the ReLU function is a piecewise nonlinear operator with a max output indicative function. The output of a ReLU is a rectified feature map, given by [88]:

$$y_r = ReLU(y_i) = max(0, y_i) \tag{5.18}$$

This function produces zero for negative input and conveys a non-negative value.

Pooling layer

The pooling layer will then simply perform down-sampling along the spatial dimensionality of the given input, further reducing the number of parameters within that activation. There

are two widely used pooling techniques, max pooling and mean pooling [89]. In this work, max pooling is used. Figure 5.12 describes an example of max pooling with 2x2 filters and strides by 2.

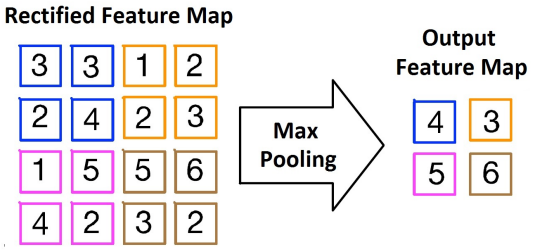


Figure (5.12) Description of applying Max Pooling

Fully-connected layer

The Fully Connected layer is a traditional Multi Layer Perception that uses a softmax activation function in the output layer (other classifiers like SVM can also be used, but will stick to softmax in this post). The term "Fully Connected" implies that every neuron in the previous layer is connected to every neuron on the next layer [90]. The purpose of the Fully Connected layer is to use these features for classifying the input image into various classes based on the training dataset.

The fully-connected layers will then perform the same duties found in standard ANNs and attempt to produce class scores from the activations, to be used for classification. It is also suggested that ReLU may be used between these layers, as to improve performance.

Softmax classifier

In this work, softmax classifier has been used to classify the characteristics of radar target. Softmax is developed from logistic regression in order to solve multi-class problems. The function of softmax can be expressed as:

Given a test input x we want our hypothesis to estimate the probability that p(y = j | x) for each value of $j = 1, \ldots, k$. I.e., we want to estimate the probability of the class label taking on each of the k different possible values. Thus, our hypothesis $h_\theta(x)$ will output a k dimensional vector (whose elements sum to 1) giving us our k estimated probabilities [91].

$$h_\theta(x) = \begin{bmatrix} p(y^{(i)} = 1|x^{(i)}; \theta) \\ p(y^{(i)} = 2|x^{(i)}; \theta) \\ \cdots \\ \cdots \\ \cdots \\ p(y^{(i)} = k|x^{(i)}; \theta) \end{bmatrix}$$

The cost function of the softmax classifier is written as:

$$J_{(\theta)} = -\frac{1}{m}\left[\sum_{i=1}^{m}\sum_{i=1}^{k} 1\left\{y^{(i)} = j\right\} log\frac{e^{\theta_j^T x^{(i)}}}{\sum_{l=1}^{k} e^{\theta_l^T x^{(i)}}}\right] \tag{5.19}$$

Proposed CNN Architectures for UWB Radar

There are three architectures of the CNN model that have been evaluated in this work. The main goal of the study is to propose an efficient CNN architecture which is suitable for our UWB radar targets identification system. Detail of these propositions will be discussed in the following subsection.

Architecture 1

The first proposition of CNN architecture in this study is the modified of Alexnet architecture. We used the term of 'modified' because in principle, this architecture is the same as Alexnet but we modified it to meet the need to our system. In Alexnet, the input of image size and the output classes are 224x224x3 and 1000 respectively, while in our system we have the size of input image and the output classes are 100x100x3 and 3, respectively. To get the idea of this architecture, we will present firstly the original of Alexnet architecture model as following. Alexnet is a model of CNN designed by Alex Krizhevsky with his colleague Ilya Sutskever and Geoffrey Hiton, Krizhevsky's PhD advisor. This model becomes one of the popular models after winning the ImageNet challenge in 2012, Large Scale Visual recognition Challenge (LSVRC). The structure of the AlexNet is given in Figure 5.13 [92]. It consists of eight learned layers with the compositions as follow: five convolutional and three fully connected layers.

The first convolutional layer uses 96 kernels of size 11 x 11 x 3 to filter the 224 x 224 x 3 input image with a stride of 4 pixels. The second convolutional layer has 256 kernels of size 5 x 5 x 48 to filter the output of the first convolutional layer (response-normalized and pooled). The third, fourth, and fifth convolutional layers are connected to one another without any intervening pooling or normalization layers. The third convolutional layer uses 384 kernels of size 3 x 3 x 256 connected to the (normalized, pooled) outputs of the second convolutional layer. The fourth convolutional layer has 384 kernels of size 3 x 3 x 192, and the fifth convolutional layer has 256 kernels of size 3 x 3 x 192. The fully-connected layers have 4096 neurons each [92].

From this original AlexNet structure, we modified its architecture to be adapted to our system, where we have the image size of 100 x 100 x 3 instead of 224 x 224 x 3 as Alexnet input image size and we only have three output classes. Therefore, we modified the architecture of Alexnet to have the input size of 100 x 100 x 3 and to have 3 output classes. Detail of this architecture can be seen in the table 5.2.

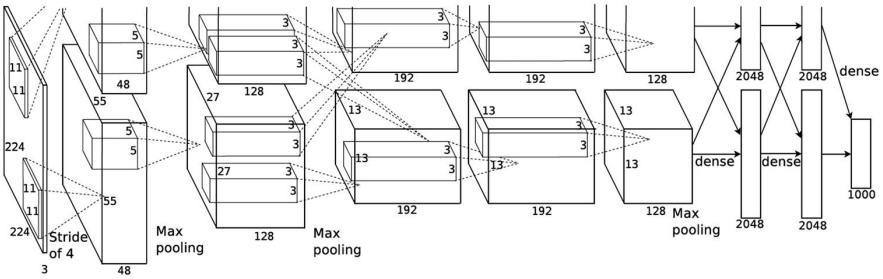


Figure (5.13) AlexNet Architecture source: [92]

Architecture 2

The second proposition is an architecture composes of tree convolutional layers, and one fully connected layer. We use the batch normalization, ReLU and max pooling layers successively after each convolutional layer. The batch normalization is also applied after fully connected layer. This architecture ends up with a softmax classifier and output layer. In the first convolutional layer, 32 kernels of size 3x3 are used to filter the input image of size 100x100x3. After batch normalization and ReLU, the max pooling with a kernel of size 2x2, padding and stride respectively are 2 and 0 is used to perform sub-sampling of the feature maps. Then, we double the number of kernel filters to the next stage layers, 64, and 128 on the convolutional layer 2 and 3, respectively by keeping the same configuration as the first convolutional layer. A fully connected layer with the size of 3 takes as input from the output of the last max pooling layer. We have only one fully connected layer, and ends up with softmax classifier.

Architecture 3

The third proposition is an architecture composes of five convolutional layers, and two fully connected layers. The same as the second proposition, it has the batch normalization, ReLU and max pooling layers successively after each convolutional layer. The differences are the numbers of convolutional layers and the number of fully connected layers. From the first convolutional layer to the third, the configurations are the same as the second architecture. In this proposition we enlarge the second proposed architecture to have more convolutional layers by adding two more covolutional layers, covolutional layer 4 and 5. As in the second proposition, the number of kernel filters is multiplied by 2 from the number of kernel filters of the previous layer. Therefore, the convolutional layer 4 and 5 in this architecture have respectively 256 and 512 kernels. The configurations of these two layers are keeping the same as those of the three-lower convolutional layers. That means, the layer 4 and 5 use the same padding and stride as the convolutional layer 1, 2 and 3. In this architecture, two fully connected layers with the size of 100 and 3 are respectively used. This architecture also ends up with softmax classifier.

Table (5.2) Proposed CNN architectures for UWB radar identification system

Architecture 1 (Modified AlexNet)	**Architecture 2**	**Architecture 3**
Input: 100x100x3	Input: 100x100x3	Input: 100x100x3
Conv1: 96, 11, 1,1	Conv1: 32, 3,1,1	Conv1: 32, 3,1,1
ReLU	BatchNormalization	BatchNormalization
CrossChannelNormalization	ReLU	ReLU
Max Pool: 3,2,0	Max Pool: 2x2,2,0	Max Pool: 2x2,2,0
Conv2: 256, 5,1,2	Conv2: 64, 3,1,1	Conv2: 64, 3,1,1
Relu	BatchNormalization	BatchNormalization
CrossChannelNormalization	ReLU	ReLU
Max Pool: 3,2,0	Max Pool: 2x2,2,0	Max Pool: 2x2,2,0
Conv3: 384, 3,1,1	Conv3: 128, 3,1,1	Conv3: 128, 3,1,1
ReLU	BatchNormalization	BatchNormalization
Conv4: 384, 3,1,1	ReLU	ReLU
ReLU	Max Pool: 2x2,2,0	Max Pool: 2x2,2,0
Conv5: 384, 3,1,1	FC1:3	Conv4: 256, 3,1,1
ReLU	Softmax	BatchNormalization
Max Pool: 3,2,0		ReLU
FC1: 4096		Max Pool: 2x2,2,0
ReLU		Conv5: 512, 3,1,1
Dropout 50%		BatchNormalization
FC1: 4096		ReLU
ReLU		FC1:100
Dropout 50%		FC2:3
FC3: 3		Softmax
Softmax		

5.2.3 Results and Discussions

In this section, we present the results of the implementation of Convolutional Neural Network on identifying the UWB radar targets. The experimentation has been performed on the 2-D images dataset created based on the method explained in the previous subsection 5.2.1. There are three different proposed CNN architectures as explained above have been investigated in order to obtain a suitable architecture for this system.

The beginning of this discussion is commenced by showing some random features map resulted by certain convolutional layers.

Feature Maps Visualisation

Features in a convolutional network are simply numbers that represent how present a certain pattern is. We have visualized several feature maps in this discussion as we can see them in the following figures 5.14, 5.15, and 5.16. The figure 5.14 and 5.15 respectively show the 30 random feature maps on the convolution layer 3 and 5. The figure 5.16 presents the feature maps in the last fully connected layer 2, the last layer before softmax classification layer. Obviously, the final feature maps look different both the structures and colors for three different classes, that will be easy to classify by softmax layer.

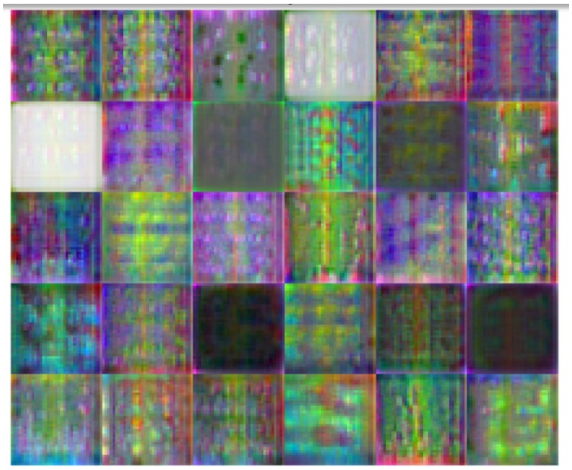


Figure (5.14) Visualisation of 30 random feature maps on convolution layer 3

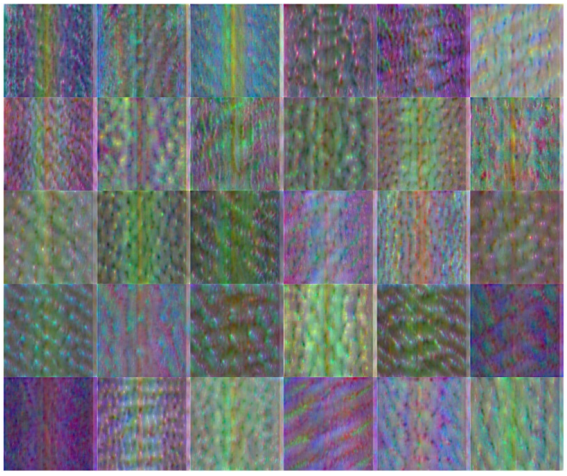


Figure (5.15) Visualisation of 30 random feature maps on convolution layer 5

Performance Evaluation

The proposed CNN architectures for this UWB radar have been used and tested using MATLAB 2017 with Neural Network Toolbox. The dataset has been trained and tested using a laptop with specification of Intel(R) Core(TM) i7-6700HQ CPU @2.6GHz with 16GB memory installed and a GPU NVIDIA Quadro M2000M with 4GB memory. The tests have been

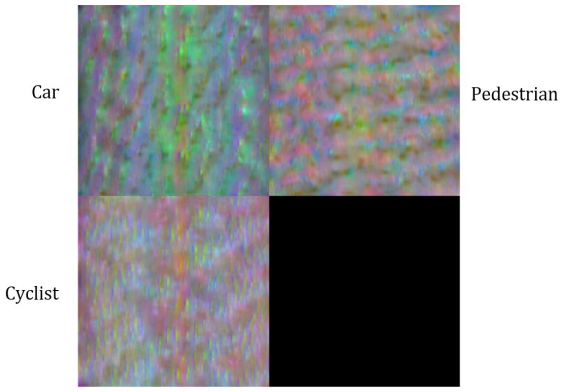


Figure (5.16) Visualisation of feature maps on fully connected layer 2

performed for three different types of the radar targets dataset, including pedestrian, cyclist and car. Table 5.3 shows distribution of dataset.

Table (5.3) Distribution of Dataset

Target Types	Number of Dataset
Car	6870
Cyclist	7058
Pedestrian	7103

Results Comparison Between The Proposed CNN Architectures

The investigation results of the proposed CNN architectures are presented in table 5.8. Since the training and testing data were chosen randomly and the weight and bias values were generated randomly as well, tests were carried out five times to reduce the effects of randomness. The table 5.8 represents the average performances of the three proposed CNN architectures. The performances have been measured using a total of 12000 samples of training data that has been used to train the CNN networks where each class consists of 4000 samples. To validate the performance, a total of 7500 samples of testing data has been tested where each target class consists of 2500 samples.

The configuration of the learning parameters used for training these architectures is presented in the table 5.4.

Table (5.4) Learning Parameters

Parameters	**Values**
Learning method	Stochastic Gradient Descent
Initial learning rate	0.01
Learning rate drop factor	0.02
Minibatch Size	128

The figures 5.17, 5.18, and 5.19 show the performance of training vs validation for the

proposed of the three CNN architectures, respectively for architecture 1, 2, and 3. The proposed CNN architecture 1 took 571s to complete 15 epochs of training process where the training performance converged after epoch 11 as we can see it in the figure 5.17 and table 5.8. As presented in the table 5.8, this proposed architecture achieved up to 99.27% of total accuracy. The detail of performance for every target class of this proposed architecture can be seen from the confusion matrix presented in table 5.5.

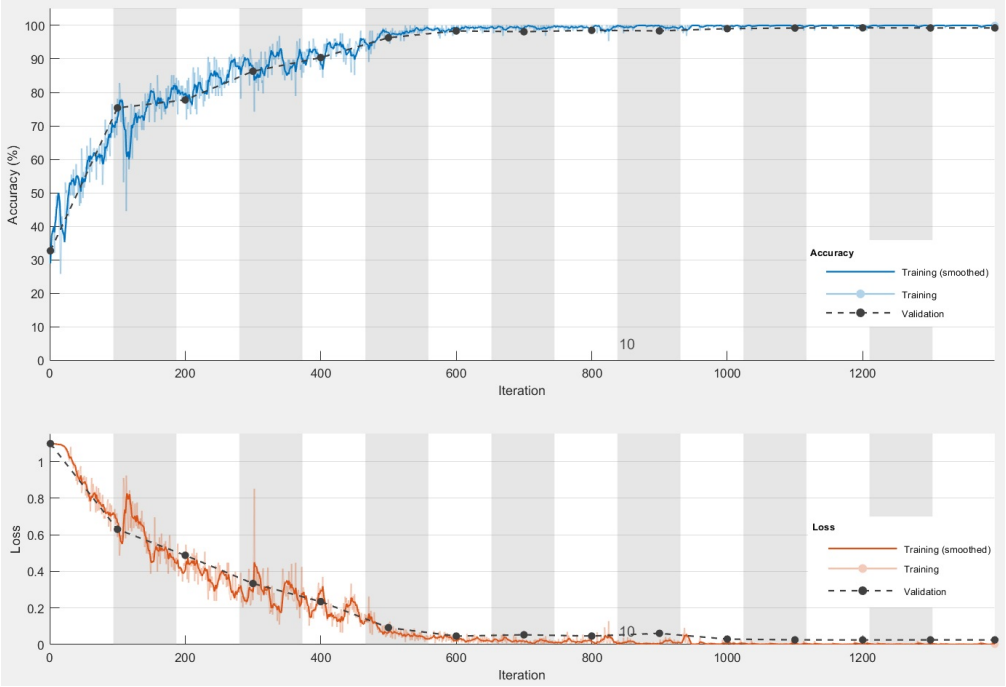


Figure (5.17) Training vs validation performance of the CNN architecture 1

Table (5.5) Confusion matrix of the first proposed CNN architecture based on investigating dataset

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	99.60%	0.20%	0.00%
	Cyclist	0.40%	98.88%	0.68%
	Pedestrian	0.00%	0.92%	99.32%

The proposed CNN architecture 2 took 709s to complete 15 epochs as we can see it in the figure 5.18 and table 5.8. Even though this architecture has less layer than that of the architecture 1, it took more time to finish 15 epochs of training process compared to the architecture 1. Our analysis is because it is difficult to find the global optimum when training the CNN networks, so it needs a lot of iteration for every epoch. The training vs validation performance of this proposed CNN architecture is presented in the figure 5.18. This architecture achieved up to 96.16% of total accuracy as presented in the table 5.8. The detail of performance for every target class of this proposed architecture can be seen from the confusion matrix presented in table 5.6.

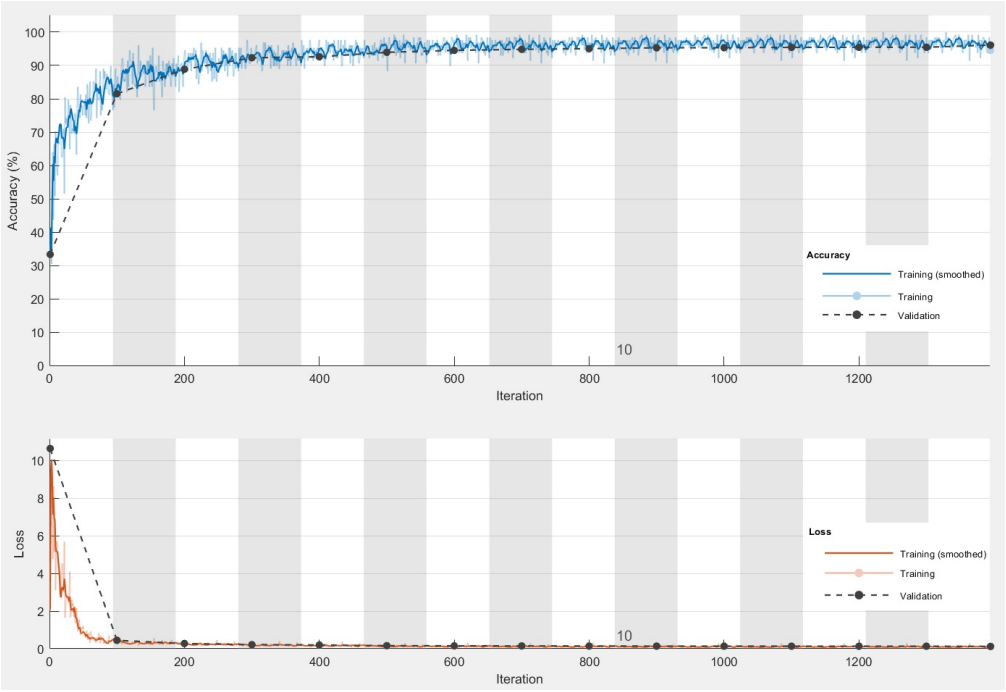


Figure (5.18) Training vs validation performance of the CNN architecture 2

Table (5.6) Confusion matrix of the second proposed CNN architecture based on investigating dataset

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	97.92%	2.18%	0.00%
	Cyclist	1.96%	93.04%	2.47 %
	Pedestrian	0.12%	4.78%	97.53%

The proposed CNN architecture 3 took 312s to complete 15 epochs of training process as we can see it in the figure 5.19 and table 5.8. The training vs validation performance of this proposed CNN architecture is presented in the figure 5.19. It can be seen that the training performance converged after epoch 6. This architecture performs very well compared to that of the others. This architecture achieved up to 99.59% of total accuracy. The detail of performance for every target class of this proposed architecture can be seen from the confusion matrix presented in table 5.7.

From the tables 5.5 to 5.7 and table 5.8, we can conclude that the proposed architecture 3 provides better accuracy and more efficient compared to the others.

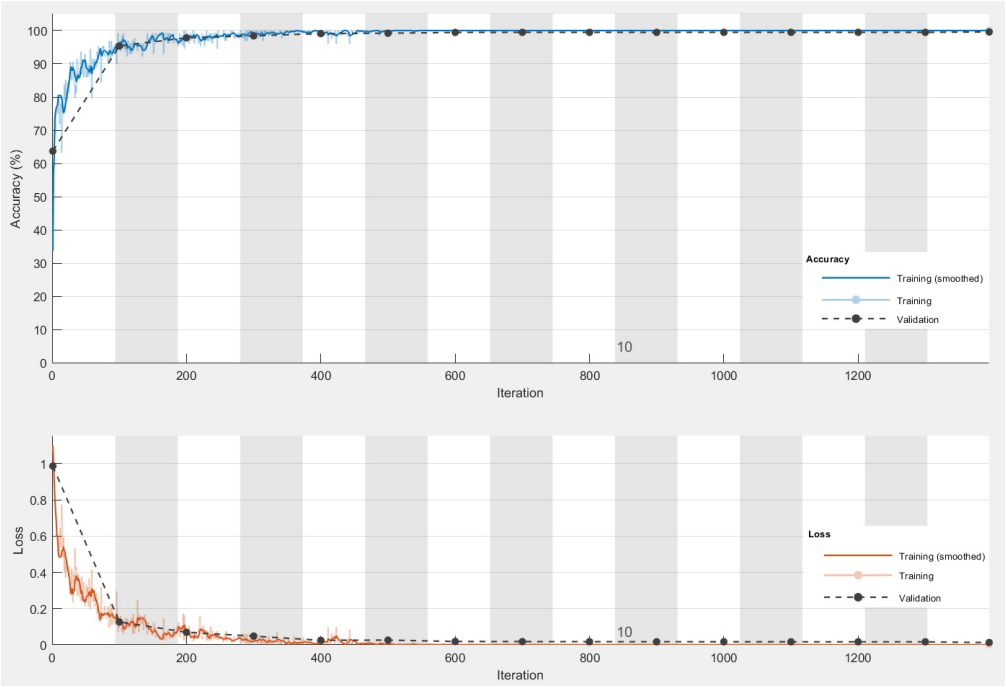


Figure (5.19) Training vs validation performance of the CNN architecture 3

Table (5.7) Confusion matrix of the third proposed CNN architecture based on investigating dataset

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	99.88%	0.12%	0.00%
	Cyclist	0.12%	99.36%	0.48 %
	Pedestrian	0.00%	0.52%	99.52%

Table (5.8) Comparison results of the three proposed CNN architectures

Proposed CNN architecture	**Training Time**	**Epoach to reach maximum accuracy**	**Maximum accuracy**
Architecture 1	571s	12	99.27%
Architecture 2	709s	10	96.16%
Architecture 3	312s	7	99.59%

In the next subsection we will present the the results of impact of using different numbers of training samples. In this study we focused only on the chosen architecture 3.

Impact of Different Number Training Samples

The number of training samples used when training the CNN network will impact to the performance of accuracy and to the training time process. Logically, the more samples data used in the training process, the more accurate the system can be achieved, but in

consequence, it will increase the training time operation. Therefore, in order to evaluate the performance of the proposed CNN architecture on identifying of UWB radar targets, we have investigated the impact of five different numbers of using the training samples. As it has been mentioned in the previous section, this study has been performed on the chosen of architecture 3.

In this study, the numbers of training samples will be varied as the following: 500, 1000, 2000, 3000, and 4000 per target class, and the results are presented in the table 5.9 to 5.13. A total of 2500 testing samples per class has been used for every investigation.

Table (5.9) Confusion matrix of the CNN performance for the usage of 500 training data per class

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	97.25%	4.72%	0.40%
	Cyclist	2.50 %	90.71%	5.31%
	Pedestrian	0.25 %	4.57 %	94.29 %

Table (5.10) Confusion matrix of the CNN performance for the usage of 1000 training data per class

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	98.91%	1.86%	0.12 %
	Cyclist	1.05%	95.02%	2.81 %
	Pedestrian	0.04%	3.12%	97.07 %

Table (5.11) Confusion matrix of the CNN performance for the usage of 2000 training data per class

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	99.20%	0.44%	0.00%
	Cyclist	0.76%	98.16%	1.12%
	Pedestrian	0.04%	1.40%	98.88%

Table (5.12) Confusion matrix of the CNN performance for the usage of 3000 training data per class

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	99.76%	0.32%	0.04%
	Cyclist	0.24%	98.95%	1.32%
	Pedestrian	0.00%	0.73%	98.64%

Table (5.13) Confusion matrix of the CNN performance for the usage of 4000 training data per class

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	99.88%	0.12%	0.00%
	Cyclist	0.12%	99.36%	0.48%
	Pedestrian	0.00%	0.52%	99.52%

Table 5.14 presents the summarize of the performances comparison for the use of the different numbers of training dataset. From this table, we can see that the maximum accuracy of 99.59% has been achieved when the use of 4000 training samples per class. We can conclude that the more samples data used in the training process, the more accurate the system can be achieved, but in consequence, it will increase the training time operation.

Table (5.14) Comparison results of the use of different numbers of training examples

Number of training examples per class	**Training Time**	**Epoach to reach maximum accuracy**	**Maximum accuracy**
500	28s	13	94.08%
1000	57s	10	97.00%
2000	111s	8	98.75%
3000	167s	8	99.12%
4000	312s	7	99.59%

5.2.4 Validation of the Developed System

After investigating the proposed noise removal method and CNN architectures, we tested and analyzed the overall developed algorithm in order to detect and to recognize radar targets directly from the radar data stream. In this experimentation, we have performed the observation for every radar target nature independently, means that we have investigated their performances sequentially one after another.

All of the experiments were conducted yet in the parking area of our laboratory. The radar was attached on a tripod and placed on a stroller. The height of the radar from the ground is about 1.8 meters. Figure 5.20 illustrates more details about this radar setup. The distance of the targets is within the range of 0.5 to 8 meters far from the radar.

Figure (5.20) Radar setup and an example of experiment where a cyclist moved in front of the radar

To do this observation, we have recorded for 5 minutes radar data stream for three different target natures (*pedestrian*, *cylist*, and *car*) based on the following scenarios:

1. Identification of the pedestrian and the cyclist: both pedestrian and cyclist can move randomly in front of the stationary radar within the range mentioned above (from 0.5 to 8 meters far from the radar). The pedestrian speed varies from 3 to 5 km/h while the cyclist speed varies from 6 to 10 km/h.

2. Identification of car: we performed the experiments on the stationary of three different cars while the stroller with the radar is moving toward and backward to the cars with the speed varying from 4 to 5 km/h. For this experimentation, the velocity of the cars is not the main purpose, but we focus on the nature of the target.

In order to validate the detector performance for this recorded radar data, we considered two conditions including Positive Condition (PC) and Negative Condition (NC). The Positive Condition is when the radar gives the good decision and otherwise is for the Negative Condition. The following is the detailed description for both decisions:

1. Positive Condition:

   (a) *when real condition presents one target and the observation result gives only one target detected, that means there is no false detection*

2. Negative Condition:

(a) *when real condition presents one target, but the observation results give more than one target, that means there are detection errors (false alarm).*

(b) *when real condition presents one target, but the observation result gives no detected target.*

The performance of detection accuracy is calculated as:

$$Acc = \frac{Number\ of\ PC}{Number\ of\ PC + Number\ of\ NC}\mathbf{x}100\% \tag{5.20}$$

Finally, we investigated the recognition accuracy of these recorded radar data. We firstly investigated the recognition of the moving pedestrian. The system has successfully identified a pedestrian as a pedestrian with accuracy of about 93.65%. The second observation was conducted for identifying a cyclist where the system has successfully recognized a cyclist as a cyclist with accuracy of about 92.63%. The final test was performed to identify a car, and the system has successfully identified a car as a real car with accuracy of about 94.30%. The confusion matrix for these observations results is summarized in the table 5.15 and the average accuracy of detection and identification are summarized in the table 5.16.

Table (5.15) Confusion matrix of the CNN performance based on recorded data stream

		Prediction		
		Car	Cyclist	Pedestrian
Actual	Car	94.30%	2.11%	2.12%
	Cyclist	3.11%	92.63%	4.23%
	Pedestrian	2.59%	5.26%	93.65%

Table (5.16) Validation of system on the recorded radar data stream using the proposed noise removal and CNN classification

Phase	Average Accuracy
Detection	95.33%
Recognition	93.53%

5.3 Conclusion of Chapter 5

In this chapter, we have discussed about 2-D radar approach for detecting and recognizing of the VRUs, i.e. cyclist and pedestrian. The task of detecting and identifying of the cyclist and pedestrian using UWB radar requires a good separation of two subspaces data, the signal and the noise or clutter. Therefore, in the first part of this chapter, we have presented the proposed of new noise removal method for this UWB radar system. A combination of Wavelet Shrinkage Denoising (WSD) and Higher Order Statistics (HOS) has been proposed in order

to enhance the UWB radar detection. The algorithm performs well in simple scenario where one target is presented for each time. The results prove that, with this algorithm, we can accurately perform the detection of a moving pedestrian and a cyclist. The results have been compared to the another methods including the Principal Component Analysis (PCA), Singular Value Decomposition (SVD), WSD and HOS. With this method, we have a better performance compared to the others. The SNR has increased up to 16.13dB, 13.45dB and 17.79dB for detecting of pedestrian, cyclist, and car, respectively.

In the second part, we have presented a discussion about the investigation of the Convolutional Neural Network (CNN) method in order to have a better result of UWB radar target classification. The 2-D radar signature has been used in this CNN approach. It is obtained from the B-scan radar data after involving the proposed of noise removal method. Using the sliding windows approach, the 2-D radar signature can be found by applying Non-Maximum Suppression method. This simple technique calculates the energies surrounding of the position of target candidate and takes the one that is maximum and eliminates the others. After transforming it into the power spectral, this signature can be used as an input of the CNN classifier.

Three CNN architectures have been investigated. The first architecture is a modified of Alexnet model, the second is an architecture with three convolutional layers and one fully connected layer, and the third is an architecture with five convolutional layers and two fully connected layers. The performances of these proposed architectures have been evaluated and compared to each other. We found that the third architecture has a good performance, and its accuracy achieves up to 99.59% with the use of 4000 number of training samples.

Finally, we concluded that the CNN is suitable to be used in VRUs identification using UWB radar system.

CHAPTER 6

CONCLUSION AND PERSPECTIVES

This dissertation is one part of the french national project called CYCLOPE which aims to study and develop a radar system based on UWB technology in order to detect and protect the vulnerable road users. This work is composed of two stages, detection and recognition. The objective of the first stage is to evaluate the performance of radar detectors in order to obtain a robust detection system. The objective of the second stage is to investigate the performance of the recognition system using Support Vector Machine (SVM) and Deep Belief Network (DBN) for input dataset of 1-D radar signature and Convolutional neural Network (CNN) for input dataset of 2-D radar signature.

In the first part of detection stage, we proposed to combine the HOS and the well-known automatic CA-CFAR detector. As there are two types of HOS algorithms, 4^{th} order Cumulant (Tugnait4-based) and 4^{th} order Cross-Moment, we investigated firstly these both algorithms to obtain the optimal performance of the HOS before combining with the CA-CFAR. Based on the investigation results, we found that both algorithms give a very similar performance in terms of noise suppression, but in terms of the complexity, 4^{th} order Cross-Moment has less calculation than 4^{th} order Cumulant. This means that 4^{th} order Cross-Moment is more efficient to be used in UWB radar. Then, we compared the performance of time delay estimation between 4^{th} order Cross-Moment and the ordinary second order statistics. We noticed that the performance of time delay estimation of 4^{th} order Cross-Moment is much better than that of second order statistics. Therefore, 4^{th} order Cross-Moment has been considered to be used in our proposed UWB radar detector. Finally, by combining 4^{th} order HOS and the CA-CFAR, UWB radar detector which is robust to noise and works with adaptive threshold has been developed. To prove this idea, we have evaluated the performance between the HOS with the fixed threshold value, CA-CFAR detector without HOS, and a combination of HOS and CA-CFAR detector (proposed method). The result is, the combination of HOS and CA-CFAR, gives better performance compared to the others and promises a good performance for UWB radar detector.

After investigating the performance of the proposed detector, we applied this detector to detect the positions of the radar targets. Once the position of the target is obtained, then we step back to the original of radar data and perform windowing around the known target

position and take it as a 1-D radar signature.

In order to classify the radar target using this radar signature, we have evaluated two different types of machine learning algorithms i.e. the Support Vector Machine (SVM) and Deep Belief Networks (DBN). The input feature of this recognition method is based on 1-D normalized radar signature. SVM is a supervised machine learning algorithm for pattern recognition that has been widely used in many research areas. There are four SVM kernels that have been analyzed in this first recognition approach including Linear, Polynomial, Sigmoid and Radial Based Function (RBF) kernel. Among of these kernels, the best performance is achieved with the RBF kernel. The results show that the SVM gives a good performance for the proposed system where the recognition rate could reach respectively up to 96.23%, 95.25% and 97.23% for the cyclist, pedestrian, and car.

After evaluating the SVM performance, we evaluated the use of DBN for this radar recognition system. There are three types of DBN architectures have been proposed. The first architecture is a DBN with two hidden layers: 3000-2000, the second is a DBN with 3 hidden layers: 3000-750-2000, and the third is a DBN with four hidden layers: 3000-750-750-2000. The performances of these architectures have been evaluated and compared each other. We found that the DBN architecture with four hidden layers performs better than the others. The results show that the performance of its accuracy achieves up to 97.80%. This result also proves that the performance of DBN better than that of SVM (96.24%) for this UWB radar target recognition system.

In the second part of detection stage, different types of noise removal techniques including Principal Component Analysis (PCA), Singular Value Decomposition (SVD), Wavelet Shrinkage Denoising (WSD) and Higher Order Statistics (HOS) have been applied, and the results have been analyzed. The signal-to-noise ratio (SNR) of the final results has been computed to compare the effectiveness of individual clutter removal techniques. It is observed that a combination of WSD and HOS has better capability to remove the noise compared to other applied techniques. Especially, it is found that it has potentially the capability to distinguish the pedestrian or the cyclist over the noise and clutters whereas other noise removal techniques are not showing significant result. The result of this work is then applied to target recognition that uses Convolutional Neural Network (CNN).

The aim of using Convolutional Neural Network (CNN) is to enhance the UWB radar target identification system. The 2-D radar signature has been used with this CNN approach. It is obtained from the B-scan radar data after involving the proposed noise removal method detailed in chapter 5. Using the sliding windows approach, the 2-D radar signature can be found by implementing non-maximum suppression. This simple technique detects the energy surrounding of the radar target position and takes the one of maximum, and eliminates the others. After transforming it into the power spectral, this signature can be used as an input of the CNN.

Three CNN architectures have been investigated. The first architecture is the modified Alexnet model, the second is an architecture with three convolutional layers and one fully connected layer, and the third is an architecture with five convolutional layers and two fully connected layers. The performance of these proposed architectures have been evaluated and

compared. We obtained that the architecture 3 has a good performance and it could achieve up to 99.59% of accuracy.

For the last part of this work, we compared the CNN, DBN and SVM performances, and the results show that CNN has a better result in terms of accuracy compared to that of DBN and SVM. It can potentially classify correctly the UWB radar targets like cyclist and pedestrian.

Finally, we have several perspectives for our future works, one of it is to test this system in the real environment where the radar is must be mounted to the vehicle (bus or truck). We have also considered to combine this UWB radar system with the camera-based system in order to have a robust recognition system to protect the VRUs.

Bibliography

[1] RSPA, 'Road Safety Factsheet: Road Crashes Overview', November, 2017. 2

[2] European Commission, 'Road Safety In The European Union: Trends, statistics and main challenges', 2016. 2

[3] ERSO, Traffic safety basic facts 2015: cyclists, *Tech. Rep, European Road Safety Observatory*, 2015. Cyclists. Mobile Information Systems. 2017. 1-13. 10.1155/2017/8149348. 2

[4] Niewoehner, W., Alexander Berg, F.,Endangerment Of Pedestrians And Bicyclists At Intersections By Right Turning Trucks, *DEKRA Automobil GmbH*, 2018 2

[5] Farooq, J., "Object detection and identification using SURF and BoW model", In *Computing, Electronic and Electrical Engineering (ICE Cube)*, International Conference, April 2016, pp. 318-323 3

[6] Hamidoun, K., et al., A New Multi-user Ultra Wide Band System Based on Modified Gegenbauer Functions and M-OAM Modulation for Communication of Intelligent Transportation Systems, *Wireless Personal Communications*, 2015, 82.4, pp. 2115-2134 3, 26

[7] Sakkila, L., Elhillali, Y., Zaidouni, J., Rivenq, A., Tatkeu, C., & Rouvaen, J. M., "High order statistic receiver applied to UWB radar", In *Communications, Computers and Signal Processing*, August 2009. PacRim 2009. IEEE Pacific Rim Conference on pp. 643-647 3, 26, 40

[8] Abujarad, F., Ground penetrating radar signal processing for landmine detection, *Disertation*, genehmigt durch die Fakultät für Elektrotechnik und Informationstechnik der Otto-von-Guericke-Universität Magdeburg, 2007 3, 4, 28

[9] D Sabushimike, SY Na, JY Kim, NN Bui, KS Seo, GG Kim, Low-Rank Matrix Recovery Approach for Clutter Rejection in Real-Time IR-UWB Radar-Based Moving Target Detection, *Sensors* 16 (9), 1409 3, 4, 28

[10] Damien Scherrer, 'OFCOM Infomailing No.8', Federal Office of communication, pp.11-15, September 2007 3, 6

[11] Sakkila, L., Rivenq, A., Boukour, F., et al., "Collision avoidance radar system using UWB waveforms signature for road applications", In : *Intelligent Transport Systems Telecommunications*,(ITST), 2009 9th International Conference on. IEEE, 2009. pp. 223-226 3, 6

[12] Dardari, D., Decarli, N., et al., High-Accuracy Tracking Using Ultrawideband Signals for Enhanced Safety of Cyclists, *Mobile Information Systems*. 2017. 1-13 7

[13] Tugnait,J. K., Time delay estimation with unknown spatialy correlated gaussian noise, *IEEE, Transaction on signal processing*, February 1993, vol. 42, no.2, pp. 549-558 7, 28, 40

[14] Y. Lin, P. Wang & M. Ma, "Intelligent Transportation System(ITS): Concept, Challenge and Opportunity", 2017 *IEEE 3rd international conference* on *big data security on cloud (bigdatasecurity)*, *IEEE international conference* on *high performance and smart computing (hpsc)*, and *IEEE international conference* on *intelligent data and security (ids)*, Beijing, 2017, pp. 167-172. doi: 10.1109/BigDataSecurity.2017.50 12

[15] A. Eskandarian (ed.), *Handbook of Intelligent Vehicles*, section 10 "Fully Autonomous Driving", edited by C. Laugier, Springer, New York, March 2011, ISBN: 978-0-85729-084-7 (Print) 978-0-85729-085-4 (Online) 12

[16] Toshiyuki Yokota, Richard J. Weiland, ITS System Architectures For Developing Countries, ITS Technical Note For Developing Countries, July 22, 2004 12

[17] Wikipedia, Intelligent transportation system, [https://en.wikipedia.org/wiki/Intelligent_ transportation_system], Accessed June, 2018 12

[18] Shaheen, Susan & Finson, Rachel. (2013). Intelligent Transportation Systems. *Encyclopedia of Energy*. 10.1016/B978-0-12-409548-9.01108-8 12

[19] Martin Luther Mfenjoua, Ado Adamou Abba Aria, Wahabou Abdouc, François Spies, Kolyang, 'Sustainable Computing: Informatics and Systems', *Journal of Sustainable Computing: Informatics and Systems* 19 (2018) 96–111 12, 13

[20] K. Ito, G. Hirakawa, Y. Arai, Y. Shibata, A road condition monitoring system using various sensor data in vehicle-to-vehicle communication environment, *Int. J. Space-Based Situat. Comput.* 6 (1) (2016) 21–30 13

[21] Sheng-hai An, Byung-Hyug Lee, Dong-Ryeol Shin, "A Survey of Intelligent Transportation Systems", *Third International Conference* on *Computational Intelligence, Communication Systems and Networks*, 2011 13

[22] L. Han and K. Wu, "Radar and radio data fusion platform for future intelligent transportation system," *The 7th European Radar Conference*, Paris, 2010, pp. 65-68 14

[23] Lex Fridman, Daniel E. Brown, Michael Glazer, et al,'MIT Autonomous Vehicle Technology Study: Large-Scale Deep Learning Based Analysis of Driver Behavior and Interaction with Automation, *Jounal CoRR*, Volume: abs/1711.06976, 2017, http://arxiv.org/abs/1711.06976 14

[24] Sergei Korjagin, Pavel Klachek, Innovative Development of Intelligent Transport Systems Based on Biocybernetical Vehicle Control Systems, *Transportation Research Procedia*, Volume 20,2017,Pages 326-333, ISSN 2352-1465 14

[25] N. Wu, W. Huang, Z. Song, X. Wu, Q. Zhang and S. Yao, "Adaptive dynamic preview control for autonomous vehicle trajectory following with DDP based path planner", 2015 *IEEE Intelligent Vehicles Symposium* (IV), Seoul, 2015, pp. 1012-1017. doi: 10.1109/IVS.2015.7225817 14

[26] Ricardo A. Daziano, Mauricio Sarrias, Benjamin Leard, Are consumers willing to pay to let cars drive for them? Analyzing response to autonomous vehicles, *Transportation Research Part C: Emerging Technologies*, Volume 78, 2017, Pages 150-164 15

[27] Litman, T, Autonomous Vehicle Implementation Predictions: Implications for Transport Planning, *Transportation Research Board 94th Annual Meeting*, Washington, DC United States 20001, January 2015 viii, 15, 17

[28] Kan Z, Qiang Z, Haojun Y, Long Z, Lu H, Chatzimisios P, Reliable and efficient autonomous driving: the need for heterogeneous vehicular networks, *Commun Mag IEEE* 53:72–79, 2015 viii, 15

[29] SAE, J3016 (2014). Taxonomy and Definitions for Terms Related to On-Road Motor Vehicle Automated Driving Systems. viii, 15, 16

[30] C. B. S. T. Molina, J. R. d. Almeida, L. F. Vismari, R. I. R. González, J. K. Naufal and J. B. Camargo, "Assuring Fully Autonomous Vehicles Safety by Design: The Autonomous Vehicle Control (AVC) Module Strategy," 2017 *47th Annual IEEE/IFIP International Conference* on *Dependable Systems and Networks Workshops* (DSN-W), Denver, CO, 2017, pp. 16-21 17

[31] Jianfeng Zhao, Bodong Liang, Qiuxia Chen, The key technology toward the self-driving car, *International Journal of Intelligent Unmanned Systems*, 2018, Vol. 6 Issue: 1, pp.2-20 17

[32] TAHRI, Tarik « Systèmes radars coopératifs multimodes pour la détection, l'identification des obstacles sur les voies, la localisation et la transmission de données trainsinfrastructures », Thèse de doctorat de l'Université de VALENCIENNES ET DU HAINAUT-CAMBRESIS, September 2014 18

[33] C Gianny, M Balsi, S Esposito, P Fallavolita, "Obstacle Detection System Involving Fusion of Multiple Sensor Technologies", *International Conference* on *Unmanned Aerial Vehicles in Geometics*, 4-7 September 2017, Bonn, Germany 18

[34] Nishi, T., Yamazaki, N., Koike, S ., Kuno, T ., Umezaki, T., "Collision Avoidance System Using Laser Beams", *Image and Vision Computing - IVC*, 2005 18

[35] Niveditha.P.R & Shanmugam, Gowri. Collision Warning System using Ultrasonic Sensors and Automatic Brake System, *Proc. of Int. Conf.* on *Recent Trends in Information, Telecommunication and Computing,ITC*, 2014 18

[36] B. Varghese, J. R. Thomas, F. Kamar, and S. R. Ali, Collision avoidance system in heavy traffic and blind spot assist using ultrasonic sensor, *International Journal of Computer Science and Engineering Communications*, vol. 2, 2014, pp. 93–96 18

[37] Nadjah Touati. Optimisation des formes d'ondes d'un radar d'aide à la conduite automobile, robustes vis-à-vis d'environnements électromagnétiques dégradés. Traitement du signal et de l'image. Thèse, Université de Valenciennes et du Hainaut-Cambresis, 2015 viii, 18, 22, 23

[38] U.S Department of Transportation, Automotive Collision Avoidance System Field Operational Test, *Phase I Interim Report*, May 2002 19

[39] Marchal, P & M Gavrila, D & Letellier, Laurent & Meinecke, Marc-Michael & Morris, R & Töns, M & Marchal Faurecia, Philippe & Sys, Rd. (2018). SAVE-U: An innovative sensor platform for vulnerable road user protection. 19

[40] D. M. Gavrila, M. Kunert and U. Lages, "A multi-sensor approach for the protection of vulnerable traffic participants the PROTECTOR project," IMTC 2001. *Proceedings of the 18th IEEE Instrumentation and Measurement Technology Conference. Rediscovering Measurement in the Age of Informatics* (Cat. No.01CH 37188), Budapest, 2001, pp. 2044-2048 vol.3 19

[41] Andres Aparicio, Ilona Cieslik, Johann Stoll, et al, "Advancing active safety towards the protection of Vulnerable Road Users by evolution of ADAS solutions that meet real-world deployment challenges: The project PROSPECT", *Proceedings of 7th Transport Research Arena TRA* 2018, April 16-19, 2018, Vienna, Austria 19

[42] V. Jain and P. Heydari, *Automotive Radar Sensors in Silicon Technologies*, Springer 2013 21

[43] Prof. Sean Victor Hum, S 2018, *Course Note: ECE422 Radio and Microwave Wireless Systems*, University of Toronto, 2018. viii, 21, 22, 23

[44] V. Issakov, *Microwave Circuits for 24 GHz Automotive Radar in Silicon-base Technologies*, Springer-Verlag Berlin Heidelberg 2010 viii, 23, 24

[45] Chia-C Chong, F Watanabe, and H Inamura,"Potential of UWB Technology for the Next Generation Wireless Communications", *IEEE Ninth International SympS.*on *Spread Spectrum Techniques and Applications*, 2006 24, 25

[46] Nikookar H, Prasad R, *Signal and Communication Technology: Introduction to Ultra Wideband for Wireless Communications*, Springer, 2009 24, 25

[47] Roy S, Foerster J R, SOMAYAZULU V.S, And Leeper D G, "Ultra wideband Radio Design: The Promise of High-Speed, Short-Range Wireless Connectivity", *Proc.IEEE*, VOL. 92, NO. 2, Feb.2004 25

[48] Matti H, H Veikko, L A Matti, "Survey to Ultra-Wideband Systems", *Euro-Cost*, July 8-9,1999. 25

[49] L.Sakkila, Y. Elhillali, A. Rivenq, C. Tatkeu, J-M. Rouvaen, "Short range automotive radar based on UWB pseudo-random coding", *International Conference* on *ITS Telecommunications*, Pages: 1 - 6, E-ISBN :1-4244-1178-5 , Sophia Antipolis, 2007. 25

[50] Adrian POPA, "An Optimization of Gaussian UWB Pulses",*10th International Conference* on *DEVELOPMENT AND APPLICATION SYSTEMS*, Suceava, Romania, May 27-29, 2010 25

[51] I.Opperman, H. Matti,J. Linatti, *UWB Theory and Aplications*, England: John Wiley & Sons, Ltd, 2004. viii, 25, 26

[52] F. Elbahhar, A. Rivenq-Menhaj, J.M. Rouvaen, Multi-user Ultra Wide Band communication system based on modified Gegenbauer and Hermite functions, *Wireless Personal communications*, Volume 34, Issue 3, August 2005. viii, 25, 27

[53] Dr. J. Reed, Dr. R. M. Buehrer, Dong S. Ha. Introduction to UWB: Impulse Radio for Radar and Wireless Communications. [http://sss-mag.com/pdf/uwbcars.pdf], Accessed April, 2017 26, 27

[54] Galushko,V. G., "Analysis of the CA CFAR algorithm as applied to detection of stationary Gaussian signals against a normal noise background", *9th International Kharkiv Symposium* on *Physics and Engineering of Microwaves, Millimeter and Submillimeter Waves (MSMW)*, Kharkiv, 2016, pp. 1-3 viii, 28, 43, 44

[55] Richards, M. A.,*Fundamentals of Radar Signal Processing*, New York: McGraw-Hill, 2005 28, 43, 44

[56] Mathworks, 'Constant False-Alarm Rate (CFAR) Detectors', [https://fr.mathworks.com/help/phased/examples/constant-false-alarmrate-cfar-detection.html], Accessed June, 2018 28, 43, 44

[57] Vaclav KABOUREK, Petr ČERNÝ, Miloš MAZÁNEK, Clutter Reduction Based on Principal Component Analysis Technique for Hidden Objects Detection, *RADIOENGINEERING*, VOL. 21, NO. 1, APRIL 2012 28

[58] Z. Li, X. Jing, W. Li and J. Wang, "A wavelet-based strong clutter removal technique for UWB life detection," *14th International Conference* on *Ground Penetrating Radar (GPR)*, Shanghai, 2012, pp. 957-960

[59] M. Garcia-Fernandez et al., "SVD-Based clutter removal technique for gpr," *IEEE International Symposium* on *Antennas and Propagation & USNC/URSI National Radio Science Meeting*, San Diego, CA, 2017, pp. 2369-2370 28

[60] S Allabakash, P Yasoda, P Srinivasulu, S Venkatramana Reddy, "Wavelet transform based methods for removal of ground clutter from the radar wind profiler data", *9th International Radar Symposium India - 2013 (IRSI - 13)* 28

[61] Qasim, Syed Manzoor, Khan, Ateeq Ahmad, Alshebeili, Saleh, et al, "FPGA based architecture for the computation of fourth-order cross moments", *In : Intelligent and Advanced Systems*, 2007. ICIAS 2007. *International Conference* on. *IEEE*, 2007. p. 1400-1403 28

[62] J. M. Mendel, 'Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications', Proc. of IEEE, vol. 79, No.3, 1991, pp.278, 305 28

[63] M. Sanullah, A Review of Higher Order Statistics and Spectra in Communication Systems, *Global Journal of Science Frontier Research Physics and Space Science*, vol. 13, Issue 4, 2013 28

[64] J. K. Tugnait. Time delay estimation with unknow spatialy correlated gaussian noise. *IEEE, Transaction on signal processing*, Febrary 1993, vol. 42, no.2, pages 549-558 28

[65] UMAIN.Inc, *User's Manual HST-D3 Evaluation Kit with Raspberry Pi 3 Development Version*, Doc No. HST-D3_Manual_V1.9.3_ENG, 2017 Manual available at:http://www.umain.co.kr, Accessed November 2017 viii, xi, 33

[66] Ewell,G.W., Design Of Digital Moving Target Indication Radar Processors, *PhD Thesis*, Georgia Institute of Technology April, 1974 xi, 38, 39

[67] Meikle, H., *Modern Radar Systems*, Artech House, 2008 39

[68] Mendel, J. M., "Tutorial on higher-order statistics (spectra) in signal processing and system theory: Theoretical results and some applications", *Proc. of IEEE*, vol. 79, No.3, 1991, pp.278, 305 39, 40

[69] Qasim, S. M., Khan, A. A., Alshebeili, S., and Abbasi,S. A.: "FPGA based architecture for the computation of fourth-order cross moments", *International Conference* on *Intelligent and Advanced Systems*, Kuala Lumpur, 2007, pp. 1400-1403 39, 40, 41

[70] Sanullah, M., A Review of Higher Order Statistics and Spectra in Communication Systems, *Global Journal of Science Frontier Research Physics and Space Science*, vol. 13, Issue 4, 2013 40

[71] Byun, H., and Seong-Whan Lee, S-W., Applications of Support Vector Machines for Pattern Recognition: A Survey, SVM 2002, *LNCS* 2388, pp. 213-236, 2002 55

[72] Hsu, C.-W., Chang,C.-C., Lin,C.-J., A practical guide to support vector classification, *Department of Computer Science National Taiwan University*, Taipei 106, Taiwan 57

[73] Chih-Chung Chang and Chih-Jen Lin, LIBSVM: A library for support vector machines, *ACM Transactions* on *Intelligent Systems and Technology*, 2:27:1–27:27, 2011
Software available at: http://www.csie.ntu.edu.tw/ cjlin/libsvm 57

[74] Liu, J., Fang, N., Xie, Y. J., and Wang, B. F., Radar target classification using support vector machine and subspace methods, in *IET Radar, Sonar & Navigation*, vol. 9, no. 6, pp. 632-640, 7 2015 58

[75] Y. LeCun, K. Kavukcuoglu and C. Farabet, "Convolutional networks and applications in vision," *Proceedings of 2010 IEEE International Symposium* on *Circuits and Systems*, Paris, 2010, pp. 253-256 59, 94, 95

[76] David Stutz, Understanding Convolutional Neural Networks, *Fakultät für Mathematik, Informatik und Naturwissenschaften Lehr- und Forschungsgebiet Informatik* VIII, August 30, 2014 94

[77] Hinton, G.E., Osindero, S., Teh, Y.W., A fast learning algorithm for deep belief nets, *Neural Computation* 18(7), 1527–1554 (2006) 59, 71, 72

[78] Andrew Ng et.al, 'Welcome to the Deep Learning Tutorial!: Multi-Layer Neural Network', [http://deeplearning.stanford.edu/tutorial/] Accessed: August 2, 2018 59, 62

[79] Ali Ghodsi. Deep Learning: Restricted Boltzmann Machines (RBM) [PDF slides]. [https://uwaterloo.ca/data-analytics/sites/ca.data-analytics/files/uploads/files/dbn2.pdf], Accessed: August 2, 2017 66

[80] https://github.com/rasmusbergpalm/DeepLearnToolbox, 2012 73

[81] R. A. Maxion and and R. R. Roberts, Proper Use of ROC Curves in Intrusion/Anomaly Detection, *Technical Report Series CS-TR-871*, School of Computing Science, University of Newcastle upon Tyne, November 2004

[82] E. Tebchrany, F. Sagnard, V. Baltazart, J. P. Tarel and X. Dérobert, "Assessment of statistical-based clutter reduction techniques on ground-coupled GPR data for the detection of buried objects in soils", *Proceedings of the 15th International Conference* on *Ground Penetrating Radar*, Brussels, 2014, pp. 604-609. 82

[83] Mohamed Hussein Emam Mabrouk, Signal Processing Of Uwb Radar Signals For Human Detection Behind Walls, *Disertation*, Ottawa-Carleton Institute for Electrical and Computer Engineering School of Electrical Engineering and Computer Science University of Ottawa 84

[84] C. Taswell, The what, how, and why of wavelet shrinkage denoising, in *Computing in Science & Engineering*, vol. 2, no. 3, pp. 12-19, May/Jun 2000.doi: 10.1109/5992.841791 85

[85] Raghuram Rangarajan, Ramji Venkataramanan and Siddharth Shah, Image Denoising Using Wavelets, December 16, 2002 85

[86] Standford,'Convolutional NeuralNetwork', `http://ufldl.stanford.edu/tutorial/supervised/ConvolutionalNeuralNetwork/`, Accessed: August 2, 2018 93

[87] S. Lee, T. Chen, L. Yu and C. Lai, Image Classification Based on the Boost Convolutional Neural Network, in *IEEE Access*, vol. 6, pp. 12755-12768, 2018. x, 95

[88] Lecun,Y., Bottou, L., Bengio, Y., & Haffner, P., Gradient-based learning applied to document recognition, in *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov. 1998 x, 94, 95

[89] Boureau Y-L., Ponce, J., & LeCun, Y., "A theoretical analysis of feature pooling in visual recognition", In *International Conference* on *Machine Learning*, 2010 96

[90] ujjwalkarn,'An Intuitive Explanation of Convolutional Neural Networks', `[https://ujjwalkarn.me/2016/08/11/intuitive-explanation-convnets/]`, Accessed: August 10, 2018 96

[91] Standford,'Softmax Regression', `[http://ufldl.stanford.edu/tutorial/supervised/SoftmaxRegression/]`, Accessed: August 2, 2018 96

[92] Krizhevsky, A., Sutskever, I., & Hinton, G. (2012). ImageNet classification with deep convolutional neural networks. *NIPS'12 Proceedings of the 25th international conference* on *neural information processing systems*, 1, 1097–1105 x, 97, 98
h