

**SİLAH SESLERİNİN KONVOLÜSYONEL
YAPAY SİNİR AĞLARI İLE SINIFLANDIRILMASI
VE ATICI KONUM TESPİTİNİN YAPILMASI**

**2018
DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ**

İlyas ÖZER

**SİLAH SESLERİNİN KONVOLÜSYONEL YAPAY SINIR AĞLARI İLE
SINIFLANDIRILMASI VE ATICI KONUM TESPİTİNİN YAPILMASI**

İlyas ÖZER

Karabük Üniversitesi

Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalında

Doktora Tezi

Olarak Hazırlanmıştır

KARABÜK

Ekim 2018

İlyas ÖZER tarafından hazırlanan “SİLAH SESLERİNİN KONVOLÜSYONEL YAPAY SİNİR AĞLARI İLE SINIFLANDIRILMASI VE ATICI KONUM TESPİTİNİN YAPILMASI” başlıklı bu tezin Doktora Tezi olarak uygun olduğunu onaylarım.

Doç. Dr. Oğuz FINDIK
Tez Danışmanı, Bilgisayar Bilimleri Anabilim Dalı



Bu çalışma, jürimiz tarafından oy birliği ile Bilgisayar Mühendisliği Anabilim Dalında Doktora tezi olarak kabul edilmiştir. 01/10/2018

Ünvanı, Adı SOYADI (Kurumu)

İmzası

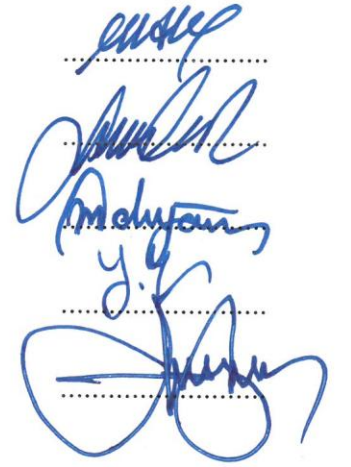
Başkan : Prof. Dr. Mehmet AKBABA (KBÜ)

Üye : Doç. Dr. Oğuz FINDIK (KBÜ)

Üye : Doç. Dr. Muharrem DÜĞENCİ (KBÜ)

Üye : Dr.Öğr.Üyesi Yasemin GÜLTEPE (KÜ)

Üye : Dr.Öğr.Üyesi Turgut ÖZSEVEN (GOP)



...../...../2018

KBÜ Fen Bilimleri Enstitüsü Yönetim Kurulu, bu tez ile, Doktora derecesini onamıştır.

Prof. Dr. Filiz ERSÖZ
Fen Bilimleri Enstitüsü Müdürü





“Bu tezdeki tüm bilgilerin akademik kurallara ve etik ilkelere uygun olarak elde edildiğini ve sunulduğunu; ayrıca bu kuralların ve ilkelerin gerektirdiği şekilde, bu çalışmadan kaynaklanmayan bütün atıfları yaptığımı beyan ederim.”

İlyas ÖZER

ÖZET

Doktora Tezi

SİLAH SESLERİNİN KONVOLÜSYONEL YAPAY SİNİR AĞLARI İLE SINIFLANDIRILMASI VE ATICI KONUM TESPİTİNİN YAPILMASI

İlyas ÖZER

**Karabük Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı**

Tez Danışmanı:

Doç. Dr. Oğuz FINDIK

Ekim 2018, 115 sayfa

Günümüz savaş alanında ve yasadışı silah kullanımının yaygın olduğu yerleşim yerlerinde atıcı konumunu tespit edebilen sistemler giderek kritik hale gelmektedir. Bir silah ateşlendiğinde konumu hakkında bilgi veren parlama, şok dalgası ve namlu ucundan yayılan ses dalgaları gibi birçok fiziksel imza oluşur. Bu çalışmada silah sesine ait akustik olaylar kullanılarak, örüntü tanıma yaklaşımı ile silah türünü belirleyen ve atıcı konumunu tespit edebilen bir sistem gerçekleştirilmiştir.

Atıcı konumunun tespit edilmesinde kalibre ve mermi boyu gibi parametrelere ihtiyaç duyulmaktadır. Fakat sadece genlik ve dalga boyu ölçülebilmektedir. Pek çok durumda eksik parametreler bazı ön kabuller ile şok dalgası modeli üzerinden belirlenmektedir. Bu tez çalışmasında ön kabullere bağlı kalmadan mermi boyu ve çapına ait parametreler tespit edilmektedir. Ayrıca bu sayede ortam gürültüsü, ekolar,

çoklu atış durumları ve aynı kalibredeki silahların belirlenmesinde karşılaşılan hataların etkisi azaltılmaktadır.

Bu kapsamda yapılan çalışmaları üç ana başlık altında toplamak mümkündür. İlk aşama yüksek derecede örtüşen spektrogram resim özelliklerini ve konvolusyonel sinir ağlarını bir arada kullanarak yüksek gürültü altında namlu ucu patlaması, şok dalgası (ŞD) ve yansıma olmak üzere silah sesine ait üç farklı akustik olayı tespit edebilen bir sistemin geliştirilmesidir. Bu yaklaşım ile özellikle uzun mesafeden yapılan atışlarda sinyal gürültü oranının düşmesine bağlı olarak yaşanan hatalı tespit problemleri büyük ölçüde azaltılmıştır. Ayrıca bu aşamada sesin hangi silaha ait olduğu tespit edilerek, ateşlenen silaha ilişkin parametreler tespit edilmektedir.

Bu doğrultuda öncelikli olarak darbesel ses olaylarını tespit edebilen bir otomatik ses tanıma yaklaşımı önerilmiştir. Önerilen model standart bir ses olayı kıyaslama görevi üzerinde test edilerek başarımlı literatürdeki güncel teknikler ile kıyaslanmıştır. Önerilen yaklaşım ile uyuşmayan durumlarda yapılan testlerde sınıflandırma başarımlı %63,4'lük bir bağıl hata azaltma oranına eşdeğer %4,5'lik bir artış sağlanmıştır. Çoklu durum testlerinde ise %98,63'lük bir sınıflandırma başarımlı ile çok daha iyi bir skor elde edilmiştir.

Atıcı konumunun belirlenmesi üzerine literatürden yapılan birçok çalışmada silah seslerine ait namlu patlamalarının (NP) tutarlı bir şekilde tespit edilemediği görülmektedir. Bu doğrultuda önerilen darbesel ses olayı tanıma yaklaşımı silah sesine ait akustik olayların tespit edilmesi görevine uyarlanmış ve gerçek silah atışlarına ait kayıtlarla hazırlanan veri seti üzerinde test edilmiştir. Önerilen yaklaşım tüm gürültü ve mesafe koşullarında şok dalgasını %100 başarı ile tespit etmiş ve %99,4'lük bir sınıflandırma başarımlı ile mermiye ilişkin parametreleri tespit etmiştir. Ayrıca çok daha zorlu bir görev olan NP tespitinde %98,5'lik tespit oranı ve %96'lık bir sınıflandırma başarımlı ile çok iyi performans sergilemiştir.

İkinci aşamada ise çapraz ilinti tabanlı algoritmalar kullanılarak sinyal varış zaman farkı hesabına dayalı olarak atıcı konumu hesabı yapılmaktadır. Bu kapsamda varış

zaman farkı kullanılarak şok dalgasına ve namlu patlamasına ait varış doğrultuları tespit edilerek, tek bir sensör dizisi ile atıcı konumu hesaplanmaktadır.

Son olarak sistemin donanım tasarımı, yüksek miktarda paralel işlem yapabilme kabiliyetleri nedeniyle alanda programlanabilir kapı dizileri (FPGA) kullanılarak çok yüksek hızlı entegre devre tanımlama dili (VHDL) ile gerçekleştirilmiştir.

Önerilen sistem iki farklı silah ile 40 farklı atış yapılarak test edilmiş ve algılayıcı dizin üzerindeki tüm mikrofonlar ses olaylarını doğru bir şekilde tespit edip sınıflandırmıştır. Sistemin yanca konum tespit hatası $2,1^o$ ve mesafe hatası %7,9 olarak gerçekleşmiştir. Ayrıca literatürdeki birçok çalışmada 100 m ve üzerindeki mesafelerde NP'lerin tespit edilmesinde problem yaşanması sebebiyle 200 m mesafeden 2 farklı silah ile 5'er atış yapılarak NP'lerin sistem tarafından tespit edilebilirliği kontrol edilmiştir. Tüm atışlarda algılayıcı dizin üzerindeki tüm mikrofonlar NP'yi doğru bir biçimde tespit edip sınıflandırmıştır.

Anahtar Sözcükler : Silah türünün tespiti, atıcı konum tespiti, spektrogram resim özellikleri, konvolusyonel sinir ağları.

Bilim Kodu : 924. 1.014

ABSTRACT

Ph. D. Thesis

CLASSIFICATION OF WEAPON SOUNDS WITH CONVOLUTIONAL NEURAL NETWORKS AND DETECTION OF SHOOTER LOCATION

İlyas ÖZER

Karabük University

Graduate School of Natural and Applied Sciences

Department of Computer Engineering

Thesis Advisor:

Assoc. Prof. Dr. Oğuz FINDIK

October 2018, 115 pages

Systems are now increasingly critical in today's battlefield and in places where illegal weapons use is widespread. Many physical signatures, such as flare, shockwave and sound waves emitted from the barrel tip, provide information about the position when a gun is fired. In this study, a system which can determine the weapon type and determine the position of the shooter by using the acoustic phenomena of the gun sound, has been realized.

In determining the position of the shooter, parameters such as caliber and bullet length are needed. But only amplitude and wavelength can be measured. In many cases, the missing parameters are determined by some preliminary assumptions and the shock wave model. In this thesis, the parameters related to the length and diameter of the bullets are determined without being dependent on the preliminary assumptions. In addition, the effect of errors in the determination of ambient noise, echoes, multiple

firing situations and weapons of the same caliber is reduced.

In this context, it is possible to group the studies under three main headings. The first stage is the development of a system that can detect three different acoustic events of gun sound, including muzzle blast, shock wave (SW) and reflection under high noise, using a combination of highly overlapping spectrogram picture features and convective neural networks. With this approach, faulty detection problems experienced due to the decrease of signal noise ratio in long shot shots have been greatly reduced. Also, the parameters of the firing weapon are determined by determining which weapon belongs to the sound at this stage.

In this respect, an automatic sound recognition approach has been proposed which can primarily detect impulsive sound events. The proposed model was tested on a standard sound event evaluation task and its performance was compared with the state-of-the-art techniques in the literature. In the tests performed in cases which mismatch with the proposed approach, an increase of 4.5% was achieved in the classification performance, which is equivalent to a relative error reduction rate of 63.4%. In multi-condition tests, a better score was obtained with a classification performance of 98.63%.

In the second stage, the position of the shooter is calculated based on the signal arrival time difference by using cross-correlation based algorithms. In this context, by using the time difference of arrival, the direction of arrival of the shock wave and the barrel explosion is determined and a single sensor sequence and the shooter position are calculated.

Finally, the system design of the system was carried out with a very high speed integrated circuit identification language (VHDL) using field programmable gate arrays (FPGAs) due to the high capacity of parallel processing.

The proposed system was tested with 40 different shots with two different weapons, and all microphones on the sensor range accurately identified and classified sound events. Azimuth angle of the system was $2,1^{\circ}$ and the distance error was 7,9%. In

addition, due to problems in detecting MBs at distances of 100 m and over in many studies in the literature, the detectability of MBs by the system was checked by taking 5 shots with 2 different weapons at a distance of 200 m. All the microphones on the sensor array accurately detect and classify the MB.

Key Word : Weapon type detection, shooter position detection, spectrogram image features, convolutional neural networks.

Science Code : 924. 1.014



TEŐEKKÜR

Bu tez alıŐmasının planlanmasında, araŐtırılmasında, yürütülmesinde ve oluşumunda ilgi ve desteęini esirgemeyen, engin bilgi ve tecrübelerinden yararlandıęım, yönlendirme ve bilgilendirmeleriyle alıŐmamı bilimsel temeller ışığında Őekillendiren sayın hocam Do. Dr. Oęuz FINDIK'a;

Tez izleme komitesinde yer alan ve deęerli katkılarıyla alıŐmama destek veren Sayın Prof. Dr. Mehmet AKBABA ve Sayın Do. Dr. Muharrem DÜŐENCİ'ye;

Saha testlerinde sağladıkları destekten dolayı Polis Özel Harekat TeŐkilatına, Karabük Polis Meslek Eęitim Merkezi'nin tüm alıŐanlarına, 1. Sınıf Emniyet Müdürü Sayın Mahmut KARABULUT'a, atıŐ eęitmeni amcam Sayın Osman ÖZER'e;

Birlikte yapmıŐ olduęumuz alıŐmalardan dolayı eŐim Sayın Zeynep ÖZER'e

Her konuda desteklerini hiçbir zaman esirgemeyen sevgili aileme tüm kalbimle teŐekkür ederim.

İÇİNDEKİLER

	<u>Sayfa</u>
KABUL.....	ii
ÖZET.....	iv
ABSTRACT.....	vii
TEŞEKKÜR.....	x
İÇİNDEKİLER	xi
ŞEKİLLER DİZİNİ.....	xiv
ÇİZELGELER DİZİNİ	xvi
SİMGELER VE KISALTMALAR DİZİNİ	xvii
BÖLÜM 1	1
GİRİŞ	1
1.1. ÇALIŞMANIN TEMEL AMACI VE LİTERATÜRE KATKILARI.....	10
1.2. TEZİN ORGANİZASYONU	12
BÖLÜM 2	15
SİLAH SESİNE AİT AKUSTİK ÖZELLİKLER.....	15
2.1. NAMLU PATLAMASI	16
2.2. ŞOK DALGASI	18
BÖLÜM 3	21
DARBESEL SES TESPİTİ VE SİLAHA AİT AKUSTİK OLAYLARIN SINIFLANDIRILMASI.....	21
3.1. OTOMATİK SES OLAYI TANIMA	22
3.2. DBN SINIFLANDIRICI	25
3.3. DBN İÇİN SIF ÖZELLİKLERİ.....	27
3.4. DBN İÇİN DENEYSEL KURULUM	28
3.5. DBN İÇİN SONUÇLAR VE TARTIŞMA	29
3.6. CNN İÇİN SPEKTROGRAM RESİM ÖZELLİKLERİ.....	31
3.7. RESİM YENİDEN BOYUTLANDIRMA YÖNTEMLERİ.....	32

	<u>Sayfa</u>
3.8. KONVOLUSYONAL SİNİR AĞI	33
3.9. ÖNERİLEN DARBESEL SES TANIMA YAKLAŞIMI	34
3.10. DENEYSEL DEĞERLENDİRME	37
3.10.1. Ses olayı ve Gürültü Veri Tabanı.....	37
3.10.2. Deneysel Kurulum	37
3.10.3. CNN Mimarisi	38
3.11. SONUÇLAR VE TARTIŞMA.....	39
3.11.1. Diğer Sistemler İle Karşılaştırma	39
3.11.2. Yeniden Boyutlandırma Metodun Etkisi	40
3.11.3. Yeniden Boyutlandırma Faktörü ve Renk Haritası Etkisi	41
3.11.4. Çoklu Durum Testleri	43
BÖLÜM 4	45
SİLAH SESİNE AİT AKUSTİK BİLEŞENLERİN TESPİT EDİLMESİ	45
4.1. VERİ SETİ	47
4.2. DENEYSEL KURULUM	50
4.3. CNN MİMARİSİ.....	53
4.4. SONUÇLAR VE TARTIŞMA.....	54
BÖLÜM 5	60
ATICI KONUMUNUN BULUNMASI.....	60
5.1. GENELLEŞTİRİLMİŞ ÇAPRAZ KORELASYON.....	60
5.2. VARIŞ DOĞRULTUSU HESAPLANMASI	62
BÖLÜM 6	69
SİSTEM DİZAYNI.....	69
BÖLÜM 7	81
FPGA	81
7.1. FPGA TABANLI YAZILIMLAR	83
7.1.1. Spektrum Analizörü.....	84
7.1.1.1. Hamming Penceresi	84

	<u>Sayfa</u>
7.1.1.2. Spektrogram Resim Elde Edilmesi	86
7.1.1.3. Ses Arabirimi	87
7.1.1.4. Bellek Kontrolü.....	89
7.1.1.5. Renk Haritası	89
7.1.2. CNN FPGA.....	90
7.1.3. Konum Hesaplama.....	91
BÖLÜM 8	93
ATIŞ TESTLERİ	93
BÖLÜM 9	97
SONUÇ VE ÖNERİLER	97
KAYNAKLAR	100
EK AÇIKLAMALAR A. AKIŞ DİYAGRAMI	110
EK AÇIKLAMALAR B. KÜRESEL KOORDINAT SİSTEMİ.....	112
ÖZGEÇMİŞ	115

ŞEKİLLER DİZİNİ

Sayfa

Şekil 1.1. Soldan sağa doğru sırasıyla piyasada ticari ürün olarak bulunan Boomerang ve PILAR atış konum tespit sistemleri.	9
Şekil 2.1. Silahın ateşlenmesi sonucu oluşan Şok Dalgası ve Namlu Patlaması (Muzzle Blast) akustik olaylarının yayılım şekilleri.	16
Şekil 2.2. M4 namlu patlaması ve AK47 namlu patlaması.....	17
Şekil 2.4. İki farklı mikrofon tarafından kaydedilen Şok Dalgası, Şok Dalgasına ait yankı ve Namlu patlamalarının genlik zaman grafiği.	18
Şekil 2.3. Şok dalgası genlik zaman grafiği.	19
Şekil 3.1. Soldan sağa doğru sırasıyla “summer”, “hot”, “hsv” ve “jet” renk haritaları kullanılarak elde edilmiş Korna sesine ait lineer nicelenmiş spektrogram resimler ve yukarıdan aşağıya doğru gürültüsüz ve 0 dB gürültü altındaki görüntüleri.	34
Şekil 3.2. CNN tabanlı otomatik ses olayı tanıma sistemine ait blok diyagram.	36
Şekil 3.3. “summer” renk haritası ile k yeniden boyutlandırma faktörüne bağlı hata oranlarının değişimi.	42
Şekil 4.1. Kaydırıcı üzerindeki iki adet Vesper VM1000 mems mikrofon.	47
Şekil 4.2. Zoom H6 ses kayıt cihazının ayarlanması.	48
Şekil 4.3. Kayıt ekipmanının farklı açılardan görünümü.	48
Şekil 4.4. M4’e ait ses kaydının genlik-zaman grafiği.	50
Şekil 4.5. AK47’ye ait ses kaydının genlik-zaman grafiği.	51
Şekil 4.6. AK47’ye ait ses kaydına 0 dB SNR değerinde insan konuşmalarına ait gürültü eklenmiş halinin genlik zaman grafiği.	51
Şekil 4.7. AK47’ye ait silah sesinin gürültüsüz halinin spektrogram görüntüsü.	52
Şekil 4.8. AK47’ye ait ses kaydına 0 dB SNR değerinde insan konuşmalarına ait gürültü eklenmiş haline ait spektrogram görüntüsü.	52
Şekil 5.1. İki farklı mikrofon tarafından kaydedilen silah sesine ait akustik olaylar arasındaki zaman farkı.	63
Şekil 5.2. Önerilen modelin akış diyagramı.	65
Şekil 5.3. Silah sesine ait akustik olayların yatay düzlemdeki geometrisi.	66
Şekil 6.1. Tez çalışmasında kullanılan Vesper VM1000 mikro elektro mekanik mikrofon.	69
Şekil 6.2. WM8731 ses kodeğine ait blok diyagram.	70

Sayfa

Şekil 6.3. Genesys2 geliştirme kartı üzerindeki PMOD konnektörlerinin pin bağlantısı.	72
Şekil 6.4. I2C seri veri yoluna ait örnek bir bağlantı.	73
Şekil 6.5. WM8731 ses kodeğinin monte edildiği ses kartı.	75
Şekil 6.6. Harici anten ve GPS alıcısı.	76
Şekil 6.7. BMP280 dijital basınç sensörü.	77
Şekil 6.8. Genesys2 kartı.	78
Şekil 6.9. Algılayıcı dizin.	79
Şekil 6.10. Sistemde kullanılan tüm donanım ekipmanları.	79
Şekil 6.11. Tez çalışmasında geliştirilen sistemin görüntüsü.	80
Şekil 7.1. FPGA genel yapısı.	82
Şekil 7.2. FPGA ortamında hazırlanan yazılımların genel yapısı.	85
Şekil 7.3. Test aşamasında kullanılan spektrogram resim oluşturma yönteminin blok diyagramı.	87
Şekil 7.4. Silah sesine ait akustik olayların sınıflandırılmasında kullanılan CNN mimarisinin şematik görünümü.	91
Şekil 7.5. Kartezyen koordinat sistemindeki konum bilgisinin küresel koordinat sistemine dönüştüren sistemin şematik görünümü.	92
Şekil Ek A.1. Ses ara yüzüne ait akış diyagramı.	111
Şekil Ek B.1. Koordinat Sistemi.	113

ÇİZELGELER DİZİNİ

	<u>Sayfa</u>
Çizelge 3.1. Sistem parametreleri	29
Çizelge 3.2. Sınıflandırma Başarımı	29
Çizelge 3.3. En son güncel yöntemler için sınıflandırma doğruluğu	30
Çizelge 3.4. Farklı gizli düğümler için %5 yeniden boyutlandırma faktöründe sınıflandırma doğruluğu	30
Çizelge 3.5. En son güncel yöntemler ve Lanczos2-CNN modelinin sınıflandırma doğruluğu	40
Çizelge 3.6. Yeniden boyutlandırma faktörünün etkisi	40
Çizelge 3.7. Farklı renk haritalarına sahip Lanczos2-CNN modeli kullanılarak $k = 6$ ve $513 \times 155 \times 6$ giriş büyüklüğü ile elde edilen sınıflandırma performansı	42
Çizelge 3.8. Farklı yeniden boyutlandırma yöntemleri kullanılarak yapılan çoklu durum testleri	43
Çizelge 4.1. Farklı SNR değerleri altında silah sesine ait akustik olayların ve diğer ses olaylarının sınıflandırma başarımı.	55
Çizelge 4.2. Silah sesine ait akustik olayların tespit edilebilme başarımı.	56
Çizelge 4.3. Çoklu durum testlerinde farklı SNR değerleri altında silah sesine ait akustik olayların ve diğer ses olaylarının sınıflandırma başarımı	57
Çizelge 4.4. Çoklu durum testlerinde silah sesine ait akustik olayların tespit edilebilme başarımı	57
Çizelge 5.1. Algılayıcı dizin sayısına bağlı olarak konum, atış doğrultusu ve mermi yörüngesinin hesaplanabilirliği	64
Çizelge 6.1. Ses kodeğine ait bağlantıların listesi	74
Çizelge 7.1. Ses arabiriminde kullanılan varlık tanımlamasına ait port bilgileri ve işlevleri	88
Çizelge 7.2. WM8731 kullanılan varlık tanımlamasına ait port bilgileri ve işlevleri	89
Çizelge 7.3. Renk haritası	90

SİMGELER VE KISALTMALAR DİZİNİ

SİMGELER

- M : Mach Sayısı
C : Ses hızı
 θ_M : Mach Açısı
V : Mermi Hızı
 C_0 : Sıfır Celcius Derecedeki Ses Hızı
A : Şok Dalgası Genliği (Amplitude)
L : Şok Dalgası Dalga Boyu
 Φ : Mermi Kalibresi
l : Mermi Boyu
d : En Yakın Yaklaşma Mesafesi
f : Spektral Vektör
k : Resim Yeniden Boyutlandırma Faktörü
 x_p : Kartezyen koordinat sistemi üzerinde tanımlı bir noktanın x ekseninde konumu
 y_p : Kartezyen koordinat sistemi üzerinde tanımlı bir noktanın y ekseninde konumu
 z_p : Kartezyen koordinat sistemi üzerinde tanımlı bir noktanın z ekseninde konumu
 φ_p : Küresel koordinat sisteminde tanımlı bir noktanın azimut açısı
 θ_p : Küresel koordinat sisteminde tanımlı bir noktanın kutupsal açısı
R : Küresel koordinat sisteminde tanımlı bir noktanın radyal mesafesi

KISALTMALAR

- ADC : Analog Digital Converter (Analog dijital dönüştürücü)
- ASIC : Application Specific Integrated Circuit (Uygulamaya Özgü Bütünleşik Devre)
- CNN : Convolutional Neural Network (Konvolusyonel Sinir Ağlarını)
- DAC : Digital Analog Converter (Dijital Analog Dönüştürücü)
- DBN : Deep Belief Network (Derin İnanç Ağı)
- DFT : Discrete Fourier Transform (Ayrık Fourier Dönüşümü)
- DNN : Deep Neural Network (Derin Yapay Sinir Ağı)
- DoA : Direction of Arrival (Varış Doğrultusu)
- FFT : Fast Fourier Transform (Hızlı Fourier Dönüşümü)
- FPGA : Field Programmable Gate Array (Alanda Programlanabilir Kapı Dizileri)
- GCC : Generalized Cross Correlation (Genelleştirilmiş Çapraz Korelasyon)
- GMM : Gaussian Mixture Models (Gauss Karışımı Modelleri)
- GPS : Global Positioning System (Küresel Konumlama Sistemi)
- GPU : Graphic Processing Unit (Grafik İşlemci Birimi)
- HDL : Hardware Description Language (Donanım Tanımlama Dili)
- HMM : Hidden Markov Model (Gizli Markov Modeli)
- IP : Intelligent Property
- I2C : Inter-Integrated Circuit
- MEMS : Microelectromechanical systems (Mikro-Elektro-Mekanik Sistemler)
- MFCC : Mel-Frequency Cepstral Coefficients (Mel Frekanslı Cepstral Katsayıları)
- MLP : Multi-Layer Perceptrons
- NP : Namlu Patlaması
- OST : Otomatik Ses Tanıma
- Phat : Phase Transform (Faz Dönüşümü)
- PMOD : Peripheral Module
- PLP : Perceptual Linear Prediction
- RBM : Restricted Boltzman Machine (Kısıtlanmış Boltzman Makinesi)
- ROM : Read Only Memory (Sadece Okunabilir Bellek)
- RGB : Red Green Blue (Kırmızı Yeşil Mavi)
- RWCP : Real World Computing Partnership

SAI : Stabilized Auditory Image
SCL : Serial Clock Line (Seri Saat Hattı)
SDA : Serial Data Line (Seri Veri Hattı)
SIF : Spectrogram Image Features (Spektrogram Resim Özellikleri)
SNR : Sinyal Noise Ratio (Sinyal Gürültü Oranı)
SVM : Support Vector Machine (Karar Destek Makinesi)
ŞD : Şok Dalgası
TDOA : Time Difference of Arrival (Varış Zaman Farkı)
USB : Universal Serial Bus
VHDL : VHSIC Donanım Tanımlama Dili
VHSIC : Very High Speed Integrated Circuit (Çok Hızlı Bütünleşik Devre)

BÖLÜM 1

GİRİŞ

Atış tespit ve konum tayin etme sistemleri günümüz kara harbinde özellikle keskin nişancı konumunun tespitinde ve yasadışı silah kullanımının yaygın olduğu yerleşim birimlerinde güvenlik güçlerinin beka kabiliyetinin artırılması ve karşı reaksiyon sürelerinin azaltılması bakımından son derece kritik bir öneme sahiptir. Ayrıca askeri üs bölgelerine hafif silahlarla yapılan taciz atışları ve saldırıların önüne geçilebilmesi ve stratejik tesislerin korunması açısından da oldukça önemlidir. Dünya üzerinde bu sistemlerin geliştirilmesine yönelik çalışmalar doksanların ortasından itibaren Bosna savaşı ile birlikte giderek hız kazanmaya başlamıştır. Öte yandan ülkemizde bu sistemler üzerine yapılan çalışmalar son derece limitlidir.

Bir silah ateşlendiği esnada konumu hakkında bilgi sağlayan akustik ve optik olmak üzere iki ana grup altında toplanabilecek birçok fiziksel imza oluşur. Silahın içerisindeki kimyasal reaksiyon sonucu oluşan yüksek basınç ve sıcaklıkta ki gazlar namlunun açık ucundan dışarı doğru çıkarken namlu ucu patlamasını ve parlamasını meydana getirirler. Öte yandan ateşlenen merminin ses üstü hızlara çıkması durumunda sıkışan hava sonik patlamaya neden olur ve merminin hareket yörüngesinden dışarı doğru konik bir biçimde yayılan ŞD'yi meydana getirir. Buna ilave olarak ortam sıcaklığından daha yüksek sıcaklığa sahip olan mermi bir termal imza oluşturur. Literatürdeki atıcı konum belirleme sistemleri silahın ve hareket halindeki merminin oluşturduğu bu fiziksel imzaları kullanarak atışın nereden yapıldığını belirlerler.

Atış konum tespit sistemleri üzerine yapılan ilk çalışmalar yüksek kalibreli mermiye sahip topçu atışlarının konumunun belirlenmesi üzerine yapılmıştır [1]. Çalışmada üç çift mikrofon kullanılarak merminin çıkardığı sesin her bir mikrofona geliş zamanı arasındaki farklara göre atışın yapıldığı konum hesaplanmıştır. Benzer şekilde bir

başka çalışmada da topçu atışının konumunun tespit edilmesi hedeflenmiştir [2]. Bununla birlikte diğer çalışmadan farklı olarak birbirine yakın dizilime sahip dört mikrofondan oluşan iki farklı algılayıcı grup kullanılarak atışın yapıldığı konum tespit edilmiştir.

Literatürdeki ilk çalışmalarda akustik tabanlı atıcı konum tespit sistemleri ile yüksek kalibreli top mühimmatlarıyla yapılan atışların konumu belirlenmeye çalışılırken ilerleyen yıllardaki çalışmalarda daha küçük kalibreli silahların konumunun belirlenmesine odaklanılmıştır. Bu doğrultuda yerleşim birimlerindeki düşük kalibreli silahlara ait seslerin konumunun belirlenmesi için çalışmalar yapılmıştır [3]. Bunun için bir parkta 5 farklı noktaya yerleştirilen mikrofonlar kullanılarak silah seslerinin konumunun belirlenmesi hedeflenmiştir [3]. Önerilen sistem tespit edilen seslerin konumunu bulmada başarılı olurken silah seslerini çevredeki diğer darbesel seslerden ayırmada başarısız olmuştur.

Atıcı konum tespit sistemleri üzerine yapılan çalışmalar 90'ların ortasından itibaren literatüre "Sniper Alley" olarak geçen ve Sırp keskin nişancıların Saraybosna Kent'inin etrafındaki tepelerden yaptıkları atışlarla resmi rakamlara göre binin üzerinde sivilin yaralanması ve aralarında çocuklarında bulunduğu yüzlerce kişinin öldürülmesine karşılık olarak Birleşmiş Milletler Koruma Gücü tarafından keskin nişancı konum tespit çalışmalarının başlatılmasıyla hızlanmıştır [4]. Bu doğrultuda konum belirleme çalışmalarında silaha ait akustik özelliklerin yanı sıra namlu parlaması [5,6] ve hareket eden merminin oluşturduğu termal iz [7] gibi diğer fiziksel özelliklerde kullanılmıştır.

Namlu parlamalarından atıcı konumunun tespit edilmesi üzerine yapılan çalışmalarda 3-5 μm bandında yoğunlaşan parlamaları tespit edebilmek için genellikle IR kameralar kullanılmaktadır [5,6]. Bununla birlikte güneş ışınları gibi çevresel faktörler sistemlerde yanlış algılama problemlerine neden olabilmektedir. Bu problemin üstesinden gelebilmek adına IR kameraların yanı sıra görünür ışık kamerası [5] veya akustik sensörler [6] kullanılarak hatalı alarm durumlarının azaltılması hedeflenmiştir. Akustik sensör kullanılan sistemde namlu parlaması IR kamera ile algılanırken, silah atışından kaynaklanan NP doğrulama amaçlı olarak kullanılmıştır.

Atışın yapıldığı konumun belirlenmesinde termal izler ve namlu parlamaları gibi fiziksel özellikler de kullanılmakla birlikte, akustik tabanlı atıcı konum tespit çalışmaları diğer fiziksel özelliklere göre literatürde daha fazla çalışılan bir konudur ve iki temel aşamadan oluşmaktadır. Bunlardan birincisi mikrofon dizileri ile elde edilen sesin NP veya ŞD olup olmadığının tespiti, ikincisi ise tespit edilen seslerinin mikrofonlara geliş zaman farklarını bularak atıcı konumunun hesaplanmasıdır.

Atışın yapıldığı konumun belirlenebilmesi için silah sesine ait akustik olayların doğru bir biçimde tespit edilebilmesi kritik öneme sahip konuların başında gelmektedir. Akustik olayların belirlenmesi aşamasında yapılan yanlışlıklar konum belirlenmesi işleminde çok yüksek miktarda hata oluşmasına neden olmaktadır. Bu aşamada akustik tabanlı atış tespit sistemlerinde silaha ait iki temel akustik özellik olan NP ve ŞD ayrı ayrı veya birlikte kullanılarak atışın yapıldığı konum tespit edilebilmektedir. Bununla birlikte her iki akustik olayda farklı özelliklere sahiptir ve tespit edilmesi bakımından birtakım zorluklar barındırırlar. ŞD'ler genlik-zaman alanındaki özellikleri bakımından çok hızlı yükselen ve düşen kenarlara sahiptir ve büyük "N" harfine benzeyen kendine has bir şekilleri vardır. Bu durum NP'lere göre tespit edilebilirliklerini daha kolay hale getirirken hareket halinde ki merminin mikrofon dizinin uzağından geçmesi durumunda bu karakteristik kenarların algılanabilirliği ciddi anlamda düşmektedir. Öte yandan NP'lerin algılanması ŞD'lere göre çok daha zorlu bir görevdir. Özellikle uzun mesafeli atışlarda NP'lerin ses seviyesi ŞD'lerin çok çok altında kalmaktadır ve buna bağlı olarak mikrofonlar tarafından algılanmaları ciddi anlamda zorlaşmaktadır. Ayrıca ŞD'ler gibi karakteristik bir şekle sahip olmamaları ve çoğu durumda ŞD yankılarının NP'den önce mikrofonlara ulaşması hatalı algılama problemlerini meydana getirmektedir.

Silaha ait NP'ler ve ŞD'lerin son derece geçici olmaları ve mikrofon üzerinde çok yoğun bir etki bırakmaları nedeniyle sıradan ve düzenli sesler için geliştirilmiş mikrofonlar ile algılanmalarında problem yaşanmaktadır. Mikrofon membranlarının mekanik olarak sönümlenmesi bu problemi azaltsa da sorunu tam olarak ortadan kaldırmamaktadır. Bu problemin üstesinden gelebilmek için geliştirilen bir yaklaşımda, yüksek örnekleme hızlarında çalışabilecek ve birbirini takip eden NP ve ŞD'leri özellikle seri atış durumlarında ayırt edebilecek durum makinası mantığı

kullanan basit bir algoritma önerilmiştir [8]. Bu yaklaşımda silah sesine ait akustik olayların tespit edilmesi aşamasında sensörler tarafından elde edilen ses olaylarının genlik ve sürelerinin önceden belirlenmiş örüntüler ile uyuşup uyuşmadığını kontrol edilmektedir. ŞD'nin N harfine benzeyen spesifik bir şekli olması nedeniyle önerdikleri algoritmada öncelikli olarak hızlı bir şekilde yükselen ve nispeten kısa süren bir sinyal aranmaktadır. NP ise ŞD'ye göre çok daha karmaşık bir sinyaldir ve sadece genlik ve süre bilgileri ile tespit edilmesi oldukça zordur. Bu nedenle önerdikleri yöntemde sinyalin sıfır geçiş değerlerini kullanarak yine durum makinesi ile NP olup olmadığını karar vermişlerdir. Kullandıkları yöntem NP'den silah modelinin tespit edilmesinde herhangi bir karakteristik imza tespit edememiştir. Öte yandan ŞD'ler aracılığıyla tespit ettikleri mermi hızları ve kalibreler arasındaki bağıntıyı kullanarak 50 kalibrelik M107'yi ve 7,62 mm çapındaki AK47 ve M240 silahlarını başarılı bir şekilde ayırt edebilirken 5,56 mm çapındaki M4 ve M249 silahlarını ayırt etmede başarısız olmuştur. Aynı zamanda 5,56 mm çapındaki M16 piyade tüfeğini 50, 100 ve 200 metrede %100 başarı ile sınıflandırırken 300 metrede sınıflandırma başarısını %67'ye düşmüştür. Benzer şekilde M240 silahı da 50, 100 ve 200 metrede %100 başarı ile sınıflandırırken 300 metrede sınıflandırma başarısını %83'e düşmüştür. Ayrıca çalışmanın neticesinde önerdikleri yöntem ile 150 m ve üzerindeki mesafelerde NP'lerin 3 mikrofon tarafından tutarlı bir biçimde algılanamadığı ortaya koymuşlardır. Yapılan çalışmada veri füzyonuna dayanan ve askerler tarafından giyilebilen ağ tabanlı bir donanım mimarisi kullanılmıştır. Dolayısıyla elde edilen sınıflandırma sonuçları veri füzyonu ile birden fazla algılayıcı dizin kullanılarak elde edilen sonuçlardır. Bununla birlikte aynı çalışmada tek algılayıcı dizin kullanılarak atıcı konumunun tespit edilmesi üzerine bir değerlendirmede ortaya konulmuştur. 150 m ve üzerindeki mesafelerde NP'leri tutarlı bir biçimde algılayamamaları nedeniyle tek algılayıcı dizin kullanılarak yapılan çalışmada atışlar 50-100 m aralığında gerçekleştirilmiştir. Aynı çalışmanın devamı niteliğinde bir başka çalışmada da aynı ses olayı algılama yaklaşımı kullanılmıştır [9]. Farklı ortamda yapılan test çalışmalarında bir önceki çalışmayla benzer sonuçlar elde edilerek namlu patlamalarından silah sınıflandırılması konusunda başarısız olunmuştur. ŞD'ler ile sınıflandırma yapılırken de aynı şekilde merminin gözlemlenen hızı bir parametre olarak kullanılarak birden fazla algılayıcı dizin ile sınıflandırma yapılmıştır. Bu yöntem aynı kalibre ve farklı mermi hızlarına sahip silahların

sınıflandırılmasında başarılı olurken birbirine yakın mermi hızlarına sahip silahlar için başarısız olmuştur. Benzer şekilde 150 m ve üzerindeki mesafelerde NP'ler tutarlı bir biçimde algılanamamıştır.

Sıfır geçişli kodlama ses tanıma uygulamalarında yaygın olarak kullanılan yöntemlerden bir tanesidir. Silah sesine ait akustik olayların belirlenmesi aşamasında sadece NP'lerin tespit edilmesinde [8,9] kullanılabildiği gibi her iki akustik olayın tespit edilmesi aşamasında da kullanılabilmektedir [10,11]. Ayrıca NP'lerin ŞD'lere göre daha düşük frekanslı sinyaller olması nedeniyle sıfır geçişli kodlama işleminin yanı sıra kesme noktası 500Hz olan bir alçak geçiren filtre kullanılarak NP içerme olasılığına sahip olan sıfır geçişli kodlu parçaların içerisindeki yüksek frekans bileşenlerinin saf dışı edilmesi üzerine de çalışmalar sunulmuştur [10,11]. NP'lerin tespiti işlemi esnasında alçak geçiren filtre kullanılan çalışmaların her ikisinde de silah sesi sınıflandırması üzerine bir değerlendirme sunulmamıştır. Atıcı konumunun belirlenmesinde ise 150 m mesafeden G3 ve AK47 silahları ile yapılan atışlarda $\pm 4^\circ$ yanca hata ve konum belirleme işleminde ise %15'lik bir hata değeri elde edilmiştir. Çalışmalarda konum belirleme işlemleri iki adet algılayıcı dizin kullanılarak gerçekleştirilmiştir. Akustik olayların belirlenmesi aşamasında sıfır geçişli kodlama, genlik ve süre gibi özelliklerin yanı sıra sinyallerin baskın frekans özellikleri de kullanılabilmektedir [12]. Çalışmada silah sesine ait akustik olayların tespit edilmesi için iki farklı ŞD dedektörü kullanılmıştır. Bunlardan birincisi elde edilen sinyallerin frekans içeriğini inceleyerek 1,6 KHz üzerinde yüksek frekanslı ses sinyalleri olup olmadığını kontrol eder. Eğer sinyaller yüksek frekans içeriği tarafından domine edilmişse ŞD olarak sınıflandırılır. İkinci dedektör ise sinyalleri sürelerine göre değerlendirerek ŞD veya NP olup olmadığına karar verir. Gerçek atışlarla yapılan testlerde önerilen yöntemin akustik olayları tespit etme aşamasındaki başarımı hakkında bir değerlendirme verilmemiştir. Sadece konum belirleme ve kalibre tahmini gibi parametrelerin belirlenmesi hakkında değerlendirmeler yapılmıştır bu aşamada akustik olay tespitinden kaynaklı hatalar ihmal edilerek bir değerlendirme sunulmuştur. Benzer şekilde frekans alanı özellikleri kullanılan bir başka çalışmada [13] ise frekans analizi ve genlik-zaman alanı özellikleri birlikte kullanılarak, ilk aşamada frekans alanındaki enerji dağılımları yardımıyla darbesel olaylar tespit edilmektedir. Daha sonra genlik-zaman alanı özellikleri kullanılarak sinyalin başlangıç

anı belirlenir ve elde edilen sinyalin NP veya ŞD olup olmadığına dair bir güven değerini içeren kalite tanımlayıcısı ile birlikte merkezi füzyon algoritmasına gönderilerek atıcı konumu burada tespit edilir. Çalışmada silah ve mühimmat türüne bağlı olarak NP algılama aralığı 30-150 m aralığında değişmiştir. Öte yandan ŞD algılama aralığı ise 30-50 m aralığında değişmiştir. Bununla birlikte yağış altında yapılan testlerde doğrudan mikrofonlar üzerine düşen yağmur damlaları ilgili mikrofonun hatalı alarm oluşturmasına sebep olmuştur. Fakat ağ merkezli dağıtık bir mimari kullanılması nedeniyle bu durum düşük ve orta şiddetli yağış koşullarında problem oluşturmamıştır. Kurulan ağın ortasından 12 atış yapılarak gerçekleştirilen testlerde atıcı, sensör alanının kenarından yaklaşık 100 metre uzaktayken ortalama yanca hata $0,66^{\circ}$, ortalama yükselme hatası $0,61^{\circ}$ ve ortalama mesafe hatası 2,56 metre olarak elde edilmiştir. Ağın yan tarafından yine aynı mesafeden 11 atış yapılarak testler tekrarlanmış ve sensör dizilerinin bir kısmının kapsama alanının dışında olması sebebiyle hata oranları yükselerek yanca $1,41^{\circ}$, yükselme sırasında $1,11^{\circ}$ ve mesafede 6,04 metre olarak gerçekleşmiştir. Bununla birlikte kontrol edilmek istenen alan arttıkça daha fazla düğüme ihtiyaç duyulmakta bu da bant genişliği ve gecikme problemlerini önemli ölçüde artırmaktadır.

Özellikle yerleşim yerlerinde kullanılan atıcı konum tespit sistemlerinde silah sesine ait akustik olayların yanı sıra çığlık ve benzeri diğer ses anomalilerinde tespit edilmesi istenebilmektedir [14]. Bu doğrultuda kullanılan ses özellikleri sadece silah sesine ait akustik olayların algılanması için kullanılanlara göre farklılık göstermektedir. [14]'deki çalışmada MFCC (Mel-frequency cepstral coefficients) ve sıfır geçişli kodlama gibi 49 farklı özellik GMM (Gaussian Mixtures Model) ile sınıflandırıldıktan sonra konum tespiti yapılmıştır. Sınıflandırma performansını gözlemlemek için yapılan testlerde eğitim ve test verilerinin SNR değerleri 0-20 dB aralığında değişecek şekilde 5'er dB'lik aralıklarla ayarlanarak yanlış tespit oranları gözlemlenmiştir. Çığlık ve gürültüyü ayırt etme performansı SNR değerlerinin düşmesiyle birlikte dramatik bir şekilde düşüş kaydetmiştir. Özellikle eğitim setinin SNR değerlerinin düşmesi yanlış tespit oranlarını sistematik bir şekilde artırmıştır. Öte yandan eğitim setinin SNR değeri sabit bir değere ayarlandıktan sonra test setinin SNR değeri düşürülmeye devam edilmesi durumunda hatalı reddedilme durumu açısından daha kötü bir performans ortaya koymaktadır. Bu davranışı hesaba katarak eğitim veri

setinde neredeyse gürültüsüz seslerle sistemin eğitilmesi gerektiğini ortaya koymuşlardır. Aksi durumda sistemin silah veya çığlık sesini algılamak üzerine eğitilmek yerine tam tersi şekilde gürültüyü algılamak üzere eğitildiği sonucuna ulaşmışlardır. Bununla birlikte sistemin yüksek SNR oranları ile eğitilmesi durumunda yüksek gürültülü bir ortamda test edilirse gerçek olayların yalnızca çok küçük bir kısmını algılayabileceğini ortaya koymuşlardır. Dolayısıyla önerilen model gürültüye karşı son derece dirençsizdir.

Yukarıda da belirtildiği gibi atış konum tespit sistemlerinde silah sesine ait akustik olaylar olan NP ve ŞD bir arada [8,9-12,15] kullanılarak atışın yapıldığı konum hesaplanabildiği gibi ŞD [16,17] veya NP'ler [18-19] ayrı ayrı kullanılarak da konum tahmini yapılabilmektedir. Bununla birlikte sadece ŞD'lerden yararlanarak gerçekleştirilen sistemlerde atıcıya ait mesafe bilgisini tam olarak hesaplamak imkansızdır [16]. Merminin yavaşlama hızına bağlı olarak mermiye ait yörünge eğriliğini tahmin etmek suretiyle mesafe bilgisi hakkında bir değerlendirme yapılması mümkün olmakla birlikte çok geniş bir sensör ağına ihtiyaç duyulması nedeniyle pratikte çok kullanışlı bir yöntem değildir [16]. Bu nedenle çalışmada tam olarak atışın yapıldığı konumu hesaplamak yerine atışın yapıldığı doğrultunun belirlenmesine odaklanılmıştır [16]. Ayrıca ŞD'ler ile atışın yapıldığı doğrultunun belirlenebilmesi için en az iki farklı sensör dizisine ihtiyaç duyulmaktadır, tek bir sensör dizini ile atışın yapıldığı doğrultu belirlenmemektedir. Yalnızca ŞD'lerden yararlanarak atış yapılan konumun belirlenmesinde birtakım kısıtlamalar bulunmasına rağmen çalışmada sadece ŞD'leri kullanmalarının temel sebebi olarak, NP'lerin karakteristik olmamasını ve 100 m ve üzerindeki mesafelerde tutarlı bir şekilde algılanamamalarını göstermişlerdir.

Süpersonik hızdaki küçük kalibreli bir merminin oluşturduğu balistik ŞD, atış konum tespit uygulamalarında oldukça önemlidir ayrıca kendilerine özgü bir şekle sahip olmaları ve kolay tespit edilebilmeleri nedeniyle atış konum tespit sistemleri üzerine yapılan çalışmalarda yaygın olarak kullanılmaktadırlar. ŞD'lerin balistik analizi, bir silah sesi olayının temel balistik özelliklerinin tahminini mümkün kılmaktadır. ŞD'lerden atıcı konumunun belirlenmesi işleminde ŞD'nin tepe basınç değeri ve sinyalin süresi merminin algılayıcı dizine en yakın olarak geçtiği mesafe hakkında bir

bilgi sunmaktadır. Bölüm 2.2’de de belirtildiği gibi ŞD’nin yükseliş süresi genellikle ihmal edilerek hesaplama yapılmaktadır. ŞD’nin yükseliş süresini etkileyen temel faktör atmosferik soğurma etkisidir [20]. Yükseliş süresinin dikkate alındığı yaklaşımlar merminin algılayıcı dizine en yakın geçtiği mesafeyi bulma noktasında daha doğru bir sonuç vermekle birlikte bu sürenin genellikle 2-3 μ s arasında değişen çok kısa bir zaman olması nedeniyle çok yüksek örnekleme hızı gerektirmektedir. Bu nedenden dolayı bu tez çalışmasında yükseliş süresi ihmal edilerek hesaplama yapılmaktadır.

ŞD silahın balistik karakteristikleri hakkında bilgi sağlamakla birlikte Eşitlik 2.5’te de görüldüğü gibi sadece genlik ve dalga boyu ölçülebilirken mermi hızı, merminin kalibresi ve boyu gibi bilgilere ihtiyaç duyulmaktadır. Literatürdeki birçok yaklaşımda eksik parametreler bazı ön kabuller ile çok değişkenli balistik model kullanarak tamamlanmaktadır [9,21-24]. Bununla birlikte bir diğer yaklaşımda da dalgacık özellikleri kullanılarak namlu patlamaları, şok dalgaları ve yankılanmalar sınıflandırılarak akustik olaylar tespit edilmeye çalışılmış ve böylelikle mermi boyu ve kalibre parametreleri ön kabuller olmadan tespit edilmiştir [25]. Bununla birlikte çalışmada sadece akustik olayların tespit edilmesi ele alınmış konum tespiti üzerine bir çalışma yapılmamıştır.

Silah atışına ait akustik olayların tespit edilmesinde karşılaşılan en büyük zorluklardan bir tanesi yankılanmalar ve hedef vurma sesi gibi benzer darbesel seslerin varlığı nedeniyle ŞD’den sonra gelen NP’lerin net bir şekilde seçilememesidir [26-27]. Bir diğer önemli zorlukta özellikle uzun mesafeden yapılan atışlarda sinyal gürültü oranının düşmesine bağlı olarak NP’lerin algılanamamasıdır [9]. Bunlara ilave olarak atıcı mesafesinin çok yakın olması durumunda ŞD’ler ve NP’ler arasındaki sürenin çok kısa olmasına bağlı olarak her iki sinyalin tek bir akustik olay olarak algılanmasıdır [28]. Literatürdeki çalışmalardan da anlaşıldığı gibi silah sesine ait akustik olayların tespiti aşamasında NP’lerin tutarlı bir biçimde algılanabilmesi oldukça zorlu bir görevdir. Öte yandan sadece ŞD’den yararlanılarak atıcı konumu tespit edilmek istediğinde de mesafe bilgisi hesaplanamadığı gibi tek bir algılayıcı dizin ile doğrultu tahmini yapılamamaktadır. Bunun aksine tek bir algılayıcı dizin ile sadece NP’lerin algılanması durumunda atışın yapıldığı doğrultu hesaplanabildiği gibi

yine tek bir algılayıcı dizin ile ŞD ve NP birlikte tespit edilebilirse konum hesabı da yapılabilmektedir. Bu nedenle NP'lerin tutarlı bir biçimde tespit edilmesi oldukça önemli bir görev olarak öne çıkmaktadır.

Ayrıca atıcı konum tespit sistemlerini yerel durum farkındalığını artırmak amacıyla genellikle tek bir sensör düğümünden oluşan sistemler ve geniş alan kapsamı için çoklu sensör düğümlerinden topladıkları bilgileri veri füzyonuyla birleştiren sistemler olarak da gruplandırmak mümkündür. Yerel durum farkındalığını artırmak için kullanılan sistemlere verilebilecek örneklerden bir tanesi araç üzerine monte edilebilen ve aynı zamanda ticari bir ürün olan Boomerang sistemidir [23]. Benzer şekilde [16,29,30] de yapılan çalışmalarda yerel durum farkındalığını artırmayı amaçlayarak yapılmıştır. Öte yandan bir grup çalışmada da çok sayıda dağıtık sensör düğümü kullanılarak geniş alan kapsamının sağlanması hedeflenmiştir [9,26,31]. Ayrıca ticari bir ürün olan PILAR sistemi ilk versiyonlarında yerel durum farkındalığını artırmayı hedeflemiş olmakla birlikte geliştirilmiş versiyonunda veri füzyonu ile geniş alan savunması kabiliyeti de kazandırılmıştır [32]. Şekil 1.1'de ticari bir ürün olan Boomerang ve PILAR sistemleri görülmektedir. Bunlara ilave olarak geniş alan savunması için kullanılan dağıtık algılayıcı dizin temelli mimarilerde akustik ağın optimizasyonu da önemli bir gereklilik olarak ortaya çıkmaktadır [11,33].



Şekil 1.1. Soldan sağa doğru sırasıyla piyasada ticari ürün olarak bulunan Boomerang ve PILAR atış konum tespit sistemleri.

Silah sesine ait akustik olayların net bir şekilde tanımlanabilmesi konum belirleme doğruluğu açısından kritik bir öneme sahiptir. Literatürde ki birçok çalışmada kendine has bir şekli olması nedeniyle şok dalgalarını kolaylıkla ayırt edebilirken çoğu durumda daha karmaşık sinyaller olan namlu patlamaları yankılar ve hedef vurma sesleri gibi diğer darbesel sesler ile karıştırılmaktadır [9,27-29]. Bazı durumlarda da artan ortam gürültüsüne veya uzun mesafelerden yapılan atışlarda ses sinyallerinin zayıflamasına bağlı olarak NP'ler hiç algılanamamaktadır. Bu durumlarda ya konum bilgisinin hiç tespit edilememesine veya konum belirleme hatalarının artmasına neden olmaktadır.

Akustik olay tespitinden kaynaklı hataların azaltılabilmesi için bu tez çalışmasında öncelikli olarak SIF ve konvolüsyonel sinir ağlarının bir arada kullanıldığı yeni bir darbesel ses sınıflandırma yaklaşımı sunulmuştur. Ayrıca önerilen yaklaşım ile silah sınıflandırılması yapılarak şok dalgası modeli üzerinde gerekli olan kalibre ve mermi boyu gibi parametrelerinin ön kabullere bağlı kalmadan yüksek doğruluk oranlarında tespit edilmiştir. Sonrasında ise çapraz ilinti tabanlı algoritmalar kullanılarak sinyal varış zaman farkı hesabına dayalı olarak atıcı konumu hesabı yapılmıştır. Son olarak özellikle CNN'lerin yüksek hesaplama maliyetlerinin üstesinden gelebilmek ve konum belirleme sürelerini kısaltabilmek amacıyla yüksek paralelleştirme kabiliyetlerine sahip olan FPGA kullanılarak VHDL ile donanım tasarımı gerçekleştirilmiştir. Önerilen yöntem özellikle yüksek ortam gürültüsü altında akustik olayları başarılı bir şekilde tespit etmiştir.

1.1. ÇALIŞMANIN TEMEL AMACI VE LİTERATÜRE KATKILARI

Atıcı konum tespit sistemleri özellik 90'ların ortasından itibaren çok yaygınlaşmaya başlamıştır [4]. Son yıllarda çalışmalar akustik tabanlı atıcı konum tespit sistemleri üzerinde yoğunlaşmıştır. Akustik tabanlı sistemlerin çalışması iki temel aşama altında gruplandırılabilir. Birinci aşamada silah sesine ait akustik bileşenler tespit edilir. İkinci aşamada ise bu akustik olayların her bir mikrofona ulaştıkları zaman esas alınarak atışın yapıldığı konum hesaplanır.

Bu tez çalışmasının temel amacı yüksek gürültü altında silah sesine ait akustik bileşenleri doğru bir biçimde tespit edebilen bir yöntem önermektir. Bu doğrultuda öncelikli olarak karakteristik bir şekle sahip olmaması nedeniyle literatürdeki birçok çalışmada tespit edilmesinde güçlük yaşanan [8,9,26,27] NP'lerin tutarlı bir biçimde algılanabilirliklerini sağlamaktır. İkincil bir amaç ise silah sesine ait akustik bileşenleri örüntü tanıma yaklaşımı ile sınıflandırarak ŞD modeli üzerinde ihtiyaç duyulan parametrelerin bulunmasını sağlamaktır. Bir diğer amaçta NP'lerin tespit oranlarını artırmak suretiyle birden fazla algılayıcı dizine ihtiyaç duymadan, tek bir algılayıcı dizin ile bile konum ve atış doğrultusunu hesaplayarak yerel durum farkındalığının artırmasını sağlamaktır.

Bu tez çalışmasında akustik bileşenlerin sınıflandırılabilmesi için SIF'i ve görüntü işlemede yaygın olarak kullanılan CNN'leri bir araya getiren yeni bir yaklaşım önerilmiştir. Literatürdeki çalışmalarda SIF'ler yüksek gürültü altında oldukça iyi bir başarımla sergilerken, gürültüsüz ortamda güncel tekniklerin bir miktar altında bir sınıflandırma performansı göstermektedir. Önerilen yöntem ile SIF'lerin gürültüsüz ortamdaki ses olaylarını sınıflandırma başarımları güncel teknikleri yakalarken, yüksek gürültü altında standart bir kıyaslama görevinde çok daha iyi bir performans göstermiştir.

Bu tez çalışmasında SIF'lerin kullanılmasının temel amacı tespit edilmek istenen seslerin son derece darbesel sesler olması [4] ve öte yandan çevresel gürültülerin genellikle daha dağınık ve uzun sinyaller olmalarına bağlı olarak darbesel seslerin spektrum üzerinde belirgin işaretler oluşturmalarıdır. Ayrıca gürültü spektrum üzerinde genellikle belirli bir bant bölgesinin bozulmasına neden olurken diğer bölgelerde herhangi bir bozulma meydana gelmemektedir. Bu durumda spektrogram resimlerin gürültüye karşı olan dirençlerini artırmaktadır. Spektrogramların sınıflandırılması için ise bu çalışmada CNN'ler kullanılmıştır. Bunun temel sebebi CNN'lerin zayıf mutlak yerelliğe sahip olmalarıdır. Yani herhangi bir ses olayı spektrogram üzerinde farklı bir konumda bulunabilir veya frekans aralığı hafif bir şekilde değişebilir. Bu durumda CNN'leri spektrogramların sınıflandırılması için iyi bir araç haline getirmektedir. Bununla birlikte ses olayları doğası gereği farklı uzunlukta olabilmektedirler. Bunun neticesinde elde edilen spektrogram resimler

farklı boyutlarda olmaktadır. Fakat CNN'ler giriş katmanında sabit uzunlukta görüntülere ihtiyaç duymaktalar. Bu tez çalışmasında önerilen yaklaşım ile Lanczos kernel tabanlı resim yeniden boyutlandırma yöntemi ile farklı uzunluktaki darbesel ses olayları sınıflandırılmış ve güncel tekniklere göre daha iyi bir performans elde edilmiştir.

Ayrıca literatürdeki silah sesine ait akustik bileşenlerin tespit edilmesi üzerine sunulan yaklaşımlar NP'leri yüksek gürültü altında ayırt etmek noktasında genellikle problem yaşamaktadır [8,9]. Önerilen yaklaşım ile NP'lerin yüksek gürültü altında tespit edilebilirliği test edilmiş ve oldukça iyi bir performans sergilemiştir.

Bunlara ilave olarak akustik olayların sınıflandırmasına dayanan bir yaklaşımın atıcı konum tespit sistemlerinde kullanılabilirliği test edilmiş ve önerilen yöntemin atıcı konum tespit sistemlerinde kullanılabilirliği teyit edilmiştir.

Bununla birlikte önerilen yöntem yüksek hesaplama maliyetleri gerektirmektedir. Bu problem üstesinden gelebilmek adına FPGA tabanlı bir donanım kullanarak yüksek derece paralel çalışabilen bir mimari ortaya konulmuştur.

1.2. TEZİN ORGANİZASYONU

Silah sesine ait akustik bileşenleri sınıflandırma yaklaşımı ile tespit ederek, silah seslerinin sınıflandırılmasına ve atıcı konumun tespit edilmesine odaklanan çalışmanın bölümleri aşağıdaki gibi organize edilmiştir.

Birinci bölümde atıcı konum tespit sistemleri hakkında kısa bir bilgilendirme yapılmıştır. Ayrıca silah sesine ait akustik bileşenlerin neler olduğu hakkında bilgi verilmiştir. Sonrasında literatürdeki çalışmalar, tezin amacı ve literatüre katkıları anlatılmıştır.

İkinci bölümde silah sesine ait akustik özellikler, bu özelliklerin yayılma karakteristikleri ve nasıl oluştukları hakkında bilgiler verilmiştir. Balistik ŞD modeli anlatılmış ve model için gerekli olan parametreler hakkında bilgi verilmiştir. Ayrıca

akustik bileşenlerin tespit edilmesinde karşılaşılan temel zorluklar hakkında bilgilendirme yapılmıştır.

Üçüncü bölümde öncelikle darbesel ses tespiti üzerine iki farklı yaklaşım ortaya konulmuş ve önerilen yaklaşımlarda kullanılan araçlar hakkında bilgi verilmiştir. Önerilen yaklaşımlar standart bir ses olayı sınıflandırma görevi üzerinde test edilmiş ve elde edilen sonuçlar literatürdeki çalışmalar ile kıyaslanmıştır.

Dördüncü önerilen yaklaşım silah sesine ait akustik olayların sınıflandırılması görevine uyarlanmıştır. Bu bölümde ayrıca hazırlanan silah sesi veri seti hakkında bilgi verilerek önerilen yaklaşım bu veri seti üzerinde test edilerek sonuçları değerlendirilmiştir.

Beşinci bölümde mikrofonlar arasındaki varış zaman farklarının nasıl tespit edildiği anlatılmıştır. Ayrıca genelleştirilmiş çapraz korelasyon işlemi hakkında bilgi verilmiştir. NP ve ŞD arasındaki yayılma gecikmelerinin nedenleri değerlendirilerek bu gecikmelerin formülasyonu hakkında bilgi verilmiştir. Buna ilave olarak örnek bir atış geometrisi üzerinde sinyallerin algılayıcı dizinlere nasıl ulaştığı anlatılmıştır. Sonrasında silah sesine ait akustik bileşenler kullanılarak tek bir algılayıcı dizin ile atıcı konumun nasıl hesaplandığı hakkında bilgiler verilmiştir.

Altıncı bölümde geliştirilen atıcı konum tespit sisteminin dizaynı hakkında bilgiler verilmiştir. Ayrıca kullanılan donanım ekipmanları ve özellikleri tanıtılmıştır.

Yedinci bölümde FPGA üzerinde geliştirilen gömülü yazılım hakkında bilgi verilmiştir. Tasarlanan yazılımın bölümleri tanıtılmış ve ses olaylarının toplanması, spektrogram resim oluşturulması, CNN tabanlı sınıflandırılma işleminin yapılması ve atıcı konum hesabının yapıldığı program bölümleri hakkında bilgi verilmiştir.

Sekizinci bölümde tek sensör dizisi kullanılarak yapılan saha testleri hakkında bilgiler verilmiş ve elde edilen sonuçlar değerlendirilmiştir.

Dokuzuncu bölümde tez çalışmasında yapılan çalışmaların sonuçları değerlendirilmiş ve gelecek çalışmalar hakkında bilgi verilmiştir.

Kaynaklar bölümünde tez çalışmasının gerçekleştirilmesi esnasında yararlanılan kaynaklar verilmiştir.



BÖLÜM 2

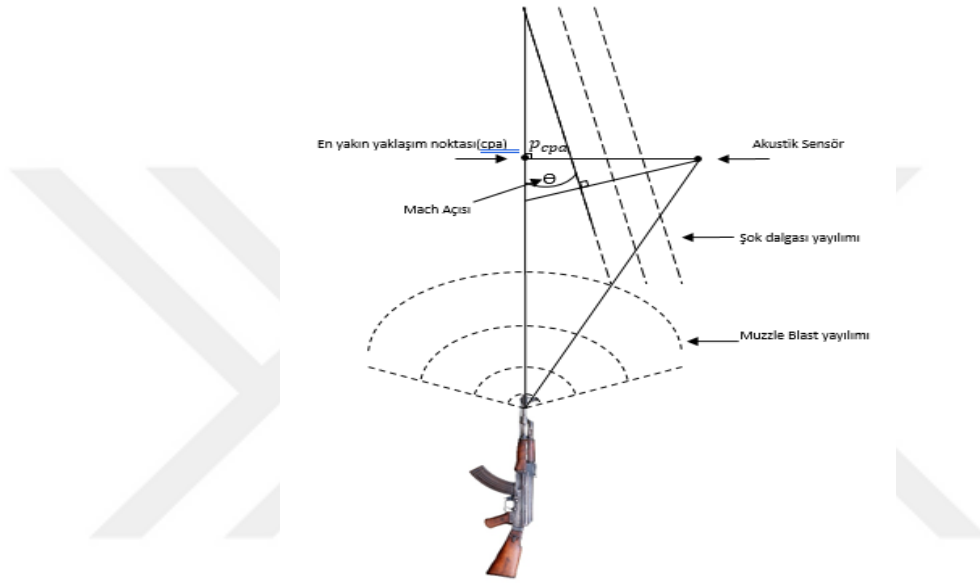
SİLAH SESİNE AİT AKUSTİK ÖZELLİKLER

Silah sesinden atıcı konumunun tespit edilmesinde kullanılan iki temel akustik özellik bulunmaktadır [4,8,9,18]. Bunlardan birincisi silahın içerisinde ki kimyasal reaksiyon sonucu oluşan ve namlunun açık ucundan ses hızıyla havaya yayılan küresel biçimli ses dalgasıdır [25]. Bir diğer akustik özellik ise ses üstü hızlara çıkan merminin havayı sıkıştırmasına bağlı olarak meydana getirdiği sonik patlamayla oluşan ŞD'dir [34]. NP tüm silahlarda görülürken ŞD sadece merminin ses üstü hızlara çıktığı silahlarda görülmektedir. Silah sesinden atıcı konumunun tespitinde bu özelliklerden biri veya her ikisi birlikte kullanılarak geometrisi bilinen mikrofon dizinleri üzerinde konum tespiti için gerekli olan varış zaman farkı (TDOA-Time Difference of Arrival) gibi parametreler hesaplanmaktadır [18].

Şekil 2.1'de de görüldüğü gibi NP silahın ateşlenmesiyle birlikte havada küresel biçimde ses hızıyla yayılırken ŞD için bu durum biraz farklıdır. Mermi yörüngesi boyunca ses üstü hızla hareketini sürdürürken, ŞD mermi yörüngesi üzerinde bulunan bir noktadan sensör arasındaki mesafe boyunca ses hızında yayılır [26]. Bu durumun bir neticesi olarak ŞD'ye ait sinyaller mikrofonu NP'den önce ulaşırlar [26]. Şekil 1.3'te farklı konumlarda bulunan iki ayrı mikrofonun algıladığı ŞD, NP ve ŞD yankıları görülmektedir. Ayrıca aynı şekil üzerinde iki mikrofon üzerindeki varış zaman farkları da görülebilmektedir.

NP sürekli olarak ses hızında hareket ettiği için varış zamanı (ToA-Time of Arrival) doğrudan atışın yapıldığı konum ve sensör arasındaki mesafe ile orantılıdır [26]. Öte yandan ŞD içinse merminin yörüngesi üzerinden sensöre doğru yayılmaya başladığı nokta ve bu noktadan sensöre doğru ses hızında yayıldığı süreyle ilişkilidir [26]. Yörünge üzerinden ayrılmaya başladığı noktaya kadar olan hız ses hızının üzerindeyken bu noktadan sensöre kadar olan hız ise ses hızına eşittir. Dolayısıyla atıcı

konumunun başarılı bir şekilde hesaplanabilmesi için silah sesine ait akustik olayların doğru bir şekilde tespit edilebilmesi kritik öneme sahip bir konudur. Bu noktada ŞD herhangi bir doğal fenomen tarafından doğrudan taklit edilemeyen ayırt edici özelliklere sahip olduğu için tespit edilmesi çok daha kolaydır [35]. Öte yandan NP ise yer yüzü şekilleri, engeller ve mesafe gibi faktörlerden çok daha fazla etkilenir ve tespit edilmesi zordur.

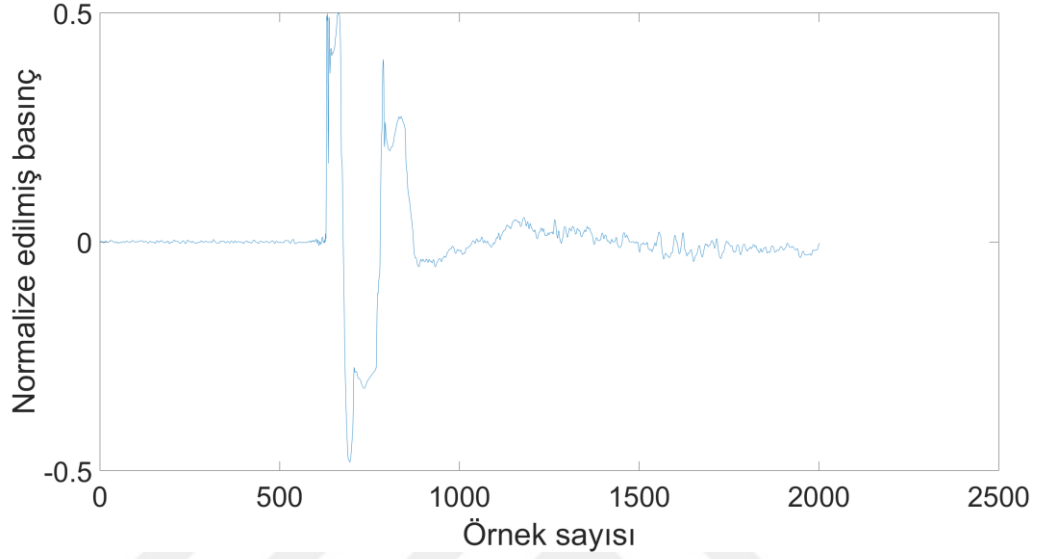


Şekil 2.1. Silahın ateşlenmesi sonucu oluşan Şok Dalgası ve Namlu Patlaması (Muzzle Blast) akustik olaylarının yayılım şekilleri.

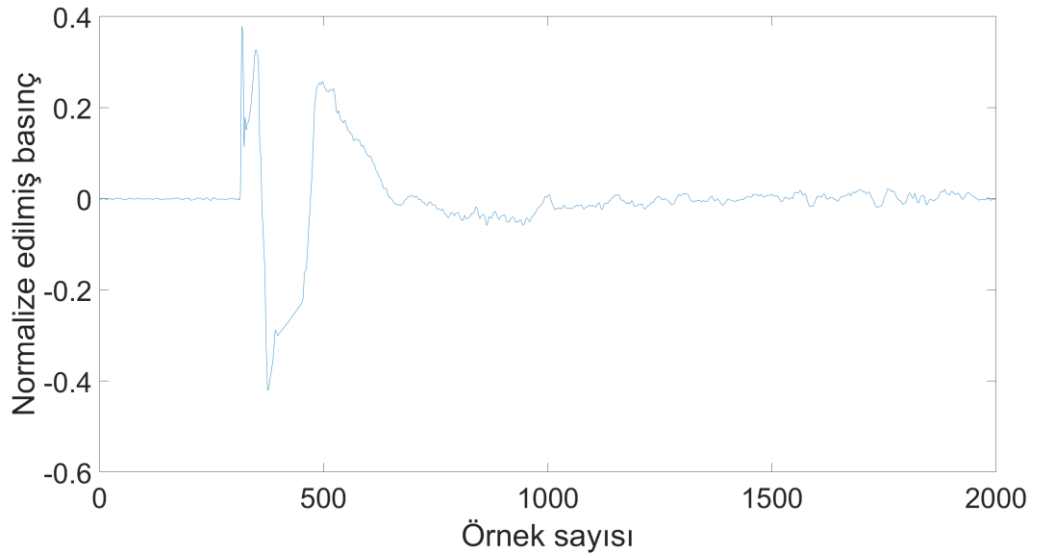
2.1. NAMLU PATLAMASI

Yukarıdaki bölümlerde de belirtildiği gibi NP silahın içerisindeki kimyasal reaksiyon sonucu oluşan ve silahın açık ucundan havaya yayılan küresel biçimli ses dalgasıdır [36]. Bu ses dalgası her yönde yayılmakla birlikte atış doğrultusu istikametinde diğer yönlere göre daha güçlüdür [37-38]. Ayrıca silaha olan mesafe arttıkça ses basınç seviyesi (SPL) azalır. Tipik olarak 500 Hz'in altında [11] ve 1-5 ms uzunluğunda sinyallerdir [35]. Bunlara ilave olarak yer yüzü şekilleri, engeller, sıcaklık, nem ve rüzgar yönleri gibi faktörler ile etkileşim içerisinde. Eğer atış çok yakın bir mesafeden yapılmış ise birincil akustik sinyal olarak NP'yi gözlemlemek mümkündür. Öte yandan çok uzak mesafeden yapılan atışlarda doğrudan ses yolu engeller nedeniyle gizlenebilir ve bu durumda gözlemlenen sinyalde çok yönlü yansımalar ve yankılar

görülebilmektedir. Böyle bir durumda geliş doğrultusu hesabının bu sinyallere göre yapılması yanlış varış doğrultusu (DoA-Direction of Arrival) hesaplanmasına neden olur. Şekil 2.2’de 2 ayrı silaha ait NP’ler görülmektedir.



(a)

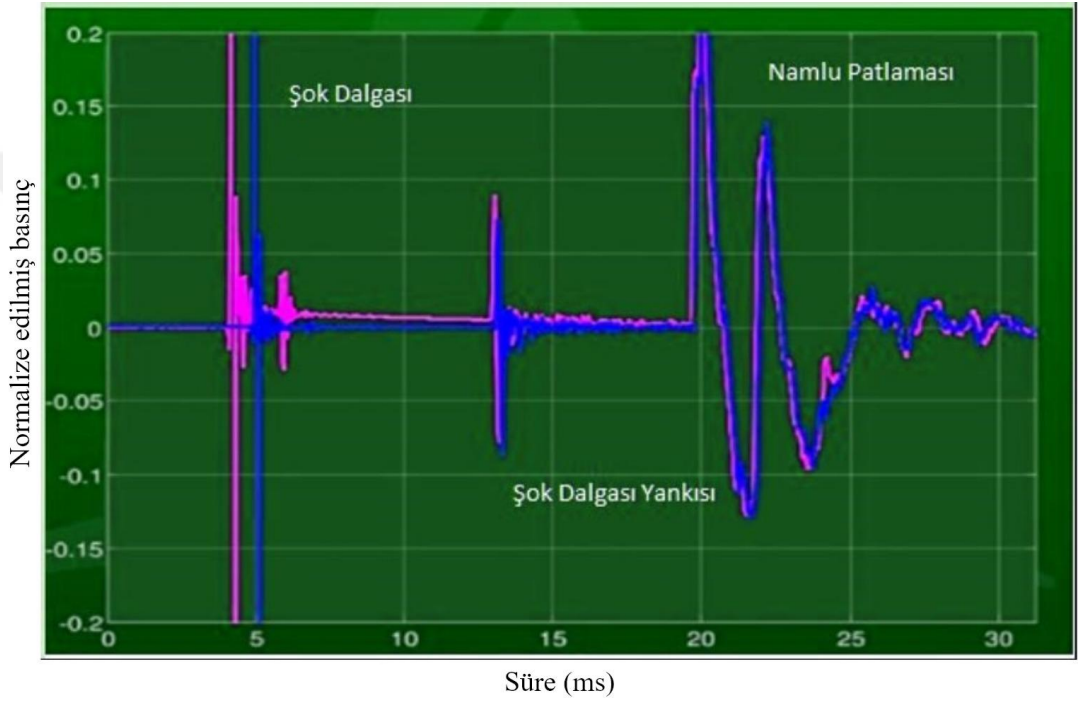


(b)

Şekil 2.2. a) M4 namlu patlaması, b) AK47 namlu patlaması.

2.2. ŞOK DALGASI

ŞD havada hareket eden merminin ses üstü hızlara çıkması sonucunda meydana gelir [39-43]. Bunun sonucu olarak merminin hareket yörüngesinden dışarı doğru konik bir biçimde yayılan şok dalgası meydana gelir. Şok dalgası tipik olarak 200-300 mikro saniye uzunluğunda çok hızlı şekilde yükselen ve düşen kenarlara sahip 1-4 kHz bandında Şekil 2.3'te de görüldüğü gibi büyük "N" harfine benzeyen bir sinyaldir.



Şekil 2.4. İki farklı mikrofondan kaydedilen Şok Dalgası, Şok Dalgasına ait yankı ve Namlu patlamalarının genlik zaman grafiği.

Ayrıca Şekil 2.4'te iki farklı mikrofondan kaydedilen bir silaha ait akustik olaylar görülmektedir. Şekilden de anlaşıldığı gibi ŞD her iki mikrofonu da NP'den önce ulaşmaktadır. Bunun ardından ŞD'lerin yankıları mikrofonlara ulaşırken en son olarak NP'ler mikrofonlara ulaşmaktadır. Sinyallerin tepe genlik değerleri atışın yapıldığı doğrultuya, merminin algılayıcı dizine göre ne kadar uzaktan geçtiğine, çevresel engellere ve atışın uzaklığına göre değişkenlik gösterebilmektedir.

V merminin hızı ve C sıcaklığa bağlı ses hız olmak üzere M şu şekilde ifade edilebilir:

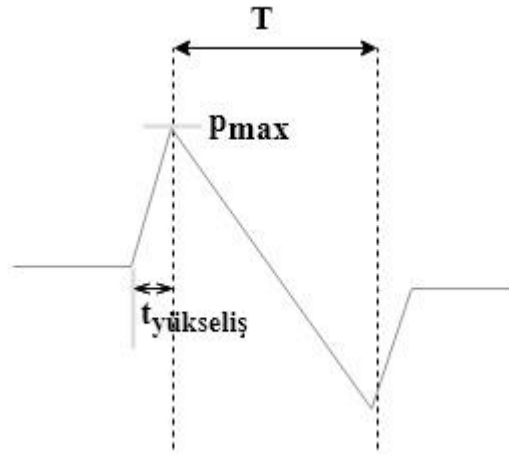
$$M = V/C \quad (2.1)$$

Mach açısı olarak ta bilinen ŞD konisinin iç açısı θ_M şu şekilde ifade edilebilir:

$$\theta_M = \arcsin(1/M) \quad (2.2)$$

Ses hızı hava sıcaklığına bağlı bir değişkendir T Celcius cinsinden hava sıcaklığı ve $C_0 = 331,3 \text{ m/s}$ sıfır Celcius derecedeki ses hızı olmak üzere C şu şekilde ifade edilebilir:

$$C = C_0 \sqrt{1 + \frac{T}{273}} \quad (2.3)$$



Şekil 2.3. Şok dalgası genlik zaman grafiği.

Ayrıca ŞD'nin Şekil 2.3'te görülen yükseliş zamanı genellikle ihmal edilerek [44-45] matematiksel olarak şu şekilde ifade edilir:

$$\left\{ \begin{array}{l} A \left(1 - \frac{2}{L} (t - \tau) \right) \quad \tau \leq t \leq \tau + L \\ 0 \quad \quad \quad \text{diğer durumlarda} \end{array} \right\} \quad (2.4)$$

Φ mermi kalibresi, l mermi boyu (yaygın olarak $4,35\Phi$), P_0 atmosferik basınç, M mach sayısı, d Şekil 2.1’de görülen en yakın yaklaşma mesafesi olmak üzere en yüksek genlik değeri A , ve dalga boyu L şu şekilde tanımlanır:

$$A = \frac{0.53P_0(M^2-1)^{1/8}\Phi}{d^{3/4}l^{1/4}} \quad (2.5a)$$

$$L = \frac{1.82Md^{1/4}\Phi}{C(M^2-1)^{3/8}l^{1/4}} \quad (2.5b)$$

Eşitlik 2.5’te de görüldüğü gibi atıcı konum hesabı için kalibre, mermi boyu ve mermi hızı gibi parametrelere ihtiyaç duyulur fakat sadece genlik ve dalga boyu ölçülebilmektedir. Birçok yaklaşımda eksik parametreler Whitham’ın şok dalgası modeli ile tahmin edilmektedir. Öte yandan bu çalışmada önerilen yaklaşım ile silaha ait kalibre ve mermi boyu parametrelerinin tahmin edilmesine gerek kalmamaktadır.

BÖLÜM 3

DARBESEL SES TESPİTİ VE SİLAHA AİT AKUSTİK OLAYLARIN SINIFLANDIRILMASI

Tez çalışmasının bu bölümünde öncelikli olarak bir darbesel ses olayı tanıma yaklaşımı sunulmaktadır. Sonrasında önerilen yaklaşım silah sesine ait akustik olayların sınıflandırılması görevine uyarlanmaktadır. Böyle bir yol izlenmesinin temel amacı önerilen darbesel ses olayı tanıma yaklaşımını standart bir kıyaslama görevi üzerinde test ederek literatürdeki güncel tekniklere göre başarımlar düzeyini tespit edebilmektir. Yazarın bilgisine göre silah seslerine ait akustik olayların tespit edilmesi ve sınıflandırılması için halihazırda literatürde kullanılan standart bir görev bulunmamaktadır. Bu nedenle çalışma başka darbesel seslerden oluşan bir veri tabanı üzerinde test edilmiştir. Standart kıyaslama görevi üzerinde tatmin edici bir başarıya ulaşan model silah sesine ait akustik olayların tespit edilmesi ve sınıflandırılması görevine uyarlanmıştır.

Darbesel ses tespiti işitme engelli insanların araba kornası ve alarm gibi seslere karşı uyarılması, güvenlik sistemlerinde çığlık ve cam kırılmasına benzer ses anomalilerinin tespiti ve atıcı konum tespit sistemlerinde silah sesine ait akustik olayların belirlenmesi gibi birçok alanda kullanılabilir. Bu ses olaylarının otomatik olarak tanınması son yıllarda dikkat çekici bir araştırma alanıdır. Karmaşık bir ses ortamında ses olaylarını otomatik olarak tanıyabilme özelliği, makine işitmesi, akustik gözetleme ve multimedya erişim uygulamaları için çok kullanışlıdır. Öte yandan birçok geleneksel metot artan ortam gürültüsü altında darbesel seslerin tespiti konusunda oldukça zayıf bir performans sergilemektedir. Bu durum ses olaylarının otomatik olarak tanınmasını oldukça zorlu bir görev haline getirmektedir. Son çalışmalar, SIF'lerin gürültü altında yüksek performansa sahip olduğunu, temiz koşullardaki başarı oranlarının ise en son teknolojiye göre bir miktar daha düşük olduğunu göstermiştir. Bu tez çalışmasında yüksek derecede örtüşen spektrogramlar lineer nicelenmiş görüntülere

dönüştürüldükten sonra çeşitli resim yeniden boyutlandırma metotları uygulanarak boyutları azaltılıp, görüntü sınıflandırmada oldukça yüksek performans sergileyen CNN'ler ile özellik çıkarımı ve sınıflandırılması yapılmaktadır. Ayrıca derin inanç ağlarını (DBN-Deep Belief Net) kullanan bir başka yaklaşım daha önerilmiştir. Her iki önerilen metodun başarımını diğer çalışmalarla kıyaslayabilmek adına öncelikle literatürde yaygın olarak kullanılan RWCP veri setindeki 50 farklı ses sınıfında testler gerçekleştirilmektedir. Sonrasında önerilen yöntem silah sesine ait akustik özelliklerin tespit edilmesine ve sınıflandırılmasına uyarlanarak sahada yapılan gerçek atış testleri ile yapılan ses kayıtları üzerinde test edilmektedir.

3.1. OTOMATİK SES OLAYI TANIMA

Literatürdeki ilk çalışmalarda sesler, konuşma ve konuşma olmayan olmak üzere iki gruba ayrılmış olmakla birlikte birbiriyle yakından ilgili bu iki alandan konuşma tanıma üzerinde göreceli olarak çok daha fazla çalışma yapılmıştır. Konuşma haricindeki seslerin otomatik olarak tanınmasına yönelik araştırmalar ise son yıllarda ivme kazanmaya başlamış ve multimedya [46], biyoakustik izleme [47], yaban hayatı bölgelerine izinsiz giriş tespiti [48], akustik gözetim ve güvenlik [49] ve çevresel seslerin algılanması [50], gibi çok sayıda uygulama alanına yayılarak çeşitlilik kazanmıştır.

Bir otomatik ses tanıma (OST) sisteminde ses olayları, sinyal işleme ve makine öğrenmesi teknikleri kullanılarak tanımlanmaktadır. Ses tanıma aslında bir örüntü tanıma problemidir ve sinyal ön işleme, özellik çıkarma, sınıflandırma olmak üzere 3 temel aşamadan oluşur. Sinyal ön işleme aşamasında genellikle giriş sinyali daha küçük parçalara bölünür ve Hamming penceresi gibi bir pencere fonksiyonu uygulanarak özellik çıkarmaya hazır hale getirilir. Özellik çıkarma karmaşık olan verinin boyutunu azaltarak bir özellik vektörü olarak temsil edilmesidir. OST üzerine yapılan ilk çalışmalar konuşma tanıma uygulamalarına dayanmaktadır [51]. Konuşma tanıma uygulamalarında kullanılan sıfır geçiş oranı, çerçeve güç enerjisi, ses perdesi ve buna benzer özellikler k-NN ve karar ağaçları gibi basit sınıflandırıcılarla sınıflandırılmıştır. Fakat son yıllarda veri miktarının çok büyük oranda artması ve daha fazla sınıfın tanımlanmasına duyulan ihtiyaç OST sistemlerine özel yeni tekniklerin

geliştirilmesini zorunlu hale getirmiştir. SAI (Stabilized auditory image)) [52], SIF [53], MFCC (Mel-frequency cepstral coefficients) [54], PLP (perceptual linear prediction) [55] son yıllarda yaygın olarak kullanılmaya başlanılan özelliklerden birkaçıdır. Bunlara ilave olarak SVM (Support Vector Machine) [55], GMM (Gaussian mixture models) [56], MLP (multi-layer perceptrons) [57] ve göreceli olarak daha yeni olan DNN (Deep Neural Network) [58] OST sistemlerinde sınıflandırıcı olarak kullanılmasıyla sistem performanslarında artış gözlenmiştir.

Diğer taraftan OST sistemlerinde kullanılan birçok özellik gürültüsüz ortamda oldukça başarılı sınıflandırma performansı sergilemesine rağmen, ortam gürültüsünün artmasıyla birlikte sınıflandırma performansı da dikkat çekici bir şekilde azalmaktadır. Ancak son dönemde zaman-frekans alanı SIF özelliklerini kullanan yeni bir yaklaşım gürültülü koşullar altında göreceli olarak daha başarılı sonuçlar vermektedir.

SIF üzerine [59]'da yapılan çalışmada spektrogram öncelikli olarak sabit aralıkta gri tonlamalı olarak normalize edilmiş, sonrasında gri tonlamalı resim kırmızı, yeşil, mavi (RGB) komponentlerine nicelendirilerek monokrom resim elde edilmiştir. En son aşamada monokrom resimler parçalara bölünüp her birinin istatistiksel dağılımı elde edilerek özellik vektörü çıkarılmıştır. 50 farklı ses sınıfında 0, 10, 20 dB gürültü altında ve gürültüsüz koşullarda yapılan çalışmada ortalama %88,5'lik bir sınıflandırma başarımı elde edilmiştir. [60]'da yapılan çalışmada ise sınıflandırma performansından ödün vermeden blokların satır ve sütun boyunca ortalama ve standart sapma değerleri kullanarak özellik vektörünün boyutu azaltılmıştır. [58]'de yapılan çalışmada enerji tabanlı bir model kullanılarak spektrogramlar bloklara bölünmeden ve merkezi momentleri kullanılmadan doğrudan aşağı örnekleme yapılarak özellik çıkarımı yapılmış [59]'da belirtilen test koşullarında %92,3'lük bir sınıflandırma başarımı elde edilmiştir. [61]'de yapılan çalışmada ise ön ve arka plan ses modellerinin havuzlanmış spektral varyansının en üst düzeye çıkarılmasına dayanan bir metot önerilmiş, [59] ve [60]'da belirtilen test koşullarında %92,9'lük bir sınıflandırma başarımı elde edilmiştir. [59]'daki çalışma ile gürültü altındaki sınıflandırma performansı en son güncel tekniklere göre aynı kıyaslama görevinde belirgin bir şekilde artmasına rağmen gürültüsüz koşullardaki başarımda da en son güncel tekniklere göre %8'in üzerinde ciddi bir düşüş göstermiştir. Öte yandan [58] ve [61]'de yapılan çalışmalarda

gürültüsüz koşullarda [59]'daki çalışmaya göre artış sağlamış olmakla birlikte çıkarılan özellik boyutu çok ciddi bir şekilde artış göstermiştir.

SIF'in gürültüye karşı sağlamlığı darbesel ses sinyalinin spektrogramının seyrek olmasına karşılık gürültünün daha fazla yayılmış bir spektrograma sahip olması ve genellikle gürültü spektrogramının dinamik aralığının belirli bölgelerinde bozulmaya neden olurken diğer bölgelerin sağlam kalması ile açıklanabilir [53]. Spektrogram görüntülerin oluşturulması kadar önemli bir diğer konuda bu görüntülerin doğru bir sınıflandırıcı ile bir arada kullanılabilmesidir. Ayrıca elde edilen spektrogram resimlerin boyutları pencere uzunluğuna, örtüşme oranına ve ses kaydının uzunluğuna bağlı olarak değişmektedir ve genellikle boyutları çok yüksektir. Literatürde elde edilen resmin sınıflandırıcıya nasıl uygulanacağı konusunda da standartlaşmış bir yöntem yoktur. Bu tez çalışmasında hem spektrogram resmin oluşturulmasında kullanılan renk haritaları, lineer ve logaritmik ölçekleme yöntemleri ve resim yeniden boyutlandırma gibi faktörler değerlendirilmekte hem de farklı tipteki derin yapay sinir ağı modelleri kullanılarak nihai bir yöntem belirlenmektedir.

Spektrogram görüntü bir sinyalin frekans spektrumunun görsel temsili olması nedeniyle görüntü ile ilgili uygulamalarda kullanılan yöntemler spektrogram görüntülerden özellik çıkarımında ve sınıflandırılmasında kullanılabilir. Bu doğrultuda testler literatürde ki çalışmalarda benzer görevlerde kullanılan BBN ve CNN'ler ile test edilmiştir. CNN'ler de son yıllarda bilgisayar görmesi ve görüntü ile ilgili uygulamalarda çok ciddi bir popülerite kazanmış ve dijital mikrobiyoloji görüntüleme [62], sosyal sinyal işleme [63], görsel takip [64] ve çoklu hata tespiti [65] gibi birbirinden farklı birçok alanda başarı ile uygulanmıştır.

Spektrogram üzerinde zayıf mutlak yerelliğe sahip olan darbesel ses sinyalleri genellikle seyrekler ve ayırt edilebilen örüntülere sahiptirler. Zayıf mutlak yerellik, fark edilebilir bir ses olayının farklı zamanlarda ve biraz farklı frekans aralıklarında görülebileceği anlamına gelir. Öte yandan CNN'lerin örüntünün resim üzerindeki konumuna karşı duyarsız olması [66] bunları SIF'lerin sınıflandırılması için uygun bir araç yapmaktadır. Bununla birlikte CNN'lerin giriş katmanına SIF'lerin nasıl uygulanabileceği konusu bugüne kadar tam olarak ortaya çıkartılmamıştır.

Bu tez çalışmasında önerilen yöntemlerde spektrogramların kullanımı derin yapay sinir ağının türüne göre farklılık göstermektedir. DBN için logaritmik olarak ölçeklenmiş gri tonlamalı resimler kullanılırken CNN'ler için lineer nicelenmiş RGB resimlere dönüştürüldükten sonra, farklı tiplerde yeniden boyutlandırma metotları ile tüm resimlerin boyutları sabit bir değere ayarlandıktan sonra, DBN ve CNN ile standart bir gürültü sağlamlık görevi üzerinde sınıflandırma performansı kıyaslanmaktadır. Önerilen CNN tabanlı metot tüm koşullarda güncel metotlarla kıyaslandığında mükemmel performans sergilemektedir.

3.2. DBN SINIFLANDIRICI

DBN önceden eğitilmiş Kısıtlanmış Boltzmann Makinelerinin (RBM) bir araya getirilmesi ile oluşturulan çok katmanlı bir yapıdır. RBM generatif stokastik sinir ağıdır ve [67]'de önerilmiştir. Görünür (giriş) ve gizli olmak üzere iki katmandan oluşan yapının her bir katmanın kendi içerisindeki bağlantıları kısıtlanmıştır. RBM enerji tabanlı bir denetimsiz öğrenme modeli olduğundan modeldeki parametrelerin dağılımı bir skaler enerji olarak ilişkilendirilir.

Bu makalede Gaussian-Bernoulli ve Bernoulli-Bernoulli olmak üzere iki farklı RBM yapısı kullanılmıştır. DNN'in ilk katmanı Gaussian-Bernoulli RBM'den oluşurken ara ve son katman Bernoulli-Bernoulli RBM'den oluşmaktadır. Görünür node sayısı i , gizli birim sayısı j , b_i^v ve b_j^h gerçek değerli bias, w_{ji} i 'inci görünür birimle j 'inci gizli birim arasındaki ağırlık katsayısı ve nodelerin binary ($v_{bb} \in \{0,1\}^V$ ve $h_{bb} \in \{0,1\}^H$) olduğu kabul edilirse Bernoulli-Bernoulli RBM için enerji fonksiyonu $E_{bb}(v, h)$ şu şekilde tanımlanabilir:

$$E_{bb}(v, h) = - \sum_{i=1}^V \sum_{j=1}^H v_i h_j w_{ji} - \sum_{i=1}^V v_i b_i^v - \sum_{j=1}^H h_j b_j^h \quad (3.1)$$

Z bölümlenme fonksiyonu olmak üzere bütün görünür ve gizli vektörlerin kombinasyonlarının enerjileri toplamı şu şekilde ifade edilebilir:

$$Z = \sum_{i=1}^V \sum_{j=1}^H e^{-E(v, h)} \quad (3.2)$$

Bu durumda RBM ağındaki görünür ve gizli vektörler arasındaki birleşik olasılık şu şekilde tanımlanabilir:

$$p = (v, h) = \frac{1}{Z} e^{\{-E(v,h)\}} \quad (3.3)$$

Gaussian-Bernoulli RBM'sin de görünür birimler gerçek değerli ($v_{gb} = \in R^v$) ve gizli birimler binary değerli $h_{gb} \in \{0,1\}^H$ olduğundan (3.1)'de ki eşitlik şu şekilde gelir:

$$E_{gb}(v, h) = - \sum_{i=1}^V \sum_{j=1}^H \frac{v_i}{\sigma_i} h_j w_{ji} - \sum_{i=1}^V \frac{(v_i - b_i^v)^2}{2\sigma_i^2} - \sum_{j=1}^H h_j b_j^h \quad (3.4)$$

Ön eğitim aşamasında RBM maksimum olabilirlik öğrenme algoritması olan karşılaştırmalı ıraksamayı (CD) kullanır. CD Markov zincirini kullanarak verilen görünür birim v 'ye göre gizli birim h 'nin değerini günceller, bunun için Markov zincirinin yakınsamasını beklemez sınırlı sayıda Gibbs örnekleme ile değeri günceller. Elde edilen yeni h değeri ile de v değeri güncellenir. İlk katmanın eğitimi tamamlandıktan sonra bu katmana ait gizli birim h bir sonraki katmanın görünür verisi haline gelir. Tüm ağın eğitimi tamamlanıncaya kadar bu işlem tekrarlanır.

DBN üzerinde denetimsiz eğitim işlemi tamamlandıktan sonra ince ayar işlemi için C boyutlu bir softmax sınıflandırıcı ilave edilir. Denklemi şu şekildedir:

$$P(C_i | v, W, b) = \text{softmax}_i(W_v + b) = \frac{e^{W_i v + b_i}}{\sum_j e^{W_j v + b_j}} \quad (3.5)$$

Softmax aktivasyon ile çıkışta beklenen sınıf sayısı boyutunda bir katman RBM yığınının üzerine eklenmiş olur. Bu fonksiyon çıkış katmanında toplam sonuç 1 olacak şekilde çıkışları normalize ederek her bir sınıfa ait çıkış değerinin o sınıfa ait olma olasılığını vermesini sağlar. Daha sonra tahmin edilen sınıf bilgisi ile gerçek sınıf bilgisi arasındaki çapraz entropi hatası minimize edilecek şekilde tüm ağa geri yayılma işlemi uygulanır.

Bu çalışmada DBN tüm durumlarda gürültüsüz seslerle eğitilerek farklı gürültü seviyeleri altındaki sınıflandırma performansı değerlendirildi. Her eğitimde dropout değeri 0,1 mini-yığın boyut değeri 100 ve devir sayısı 1000 olarak ayarlandı.

3.3. DBN İÇİN SIF ÖZELLİKLERİ

Spektrogram resim bir sinyalin frekans spektrumunun görsel temsilidir ve hızlı Fourier dönüşümü (FFT) ile elde edilen magnitude spektrumlarının oluşturduğu bir yığından meydana gelir. Verilen ses vektörü, yüksek oranda örtüşen w_s uzunluklu parçalara ayrılarak FFT yapılır ve spectral vektör f_F elde edilir. Bu durumda mevcut pencere F 'nin spektral magnitude vektörü f_F şu şekilde elde edilir:

$$s_f(n) = s(F\delta + n)w(n) \text{ for } n = 0 \dots (w_s - 1) \quad (3.6)$$

$$f_F(k) = \sum_{n=0}^{w_s-1} s_f(n)e^{-j2\pi nk/w_s} \text{ for } k = 1 \dots (w_s/2 - 1) \quad (3.7)$$

Burada $w(n)$ hamming çerçeve fonksiyonunu, δ ise ardışıl iki pencere arasındaki ilerlemeyi ifade eder. Lineer ve log değerleri şu şekilde elde edilir:

$$S_{Linear}(k, F) = |f_F(k)| \quad (3.8)$$

$$S_{Log}(k, F) = \log(S_{Linear}(k, F)) \quad (3.9)$$

Log değeri elde edildikten sonra Lanczos kernel kullanılarak şu şekilde yeniden boyutlandırılır:

$$L(x) = \begin{cases} \text{sinc}(x)\text{sinc}(x/a) & \text{if } -a < x < a \\ 0 & \text{diğer durumlar} \end{cases} \quad (3.10a)$$

$$L(x, y) = L(x) \cdot L(y) \quad (3.10b)$$

$$S(x, y) = \sum_{k=|x|-a+1}^{|x|+a} \sum_{F=|y|-a+1}^{|y|+a} S_{Log}(k, F)L(x - k)L(y - F) \quad (3.11)$$

R yeniden boyutlandırma faktörü, B ve D orijinal matris boyutları olmak üzere yeniden boyutlandırma işleminden sonra yeni matris boyutları $x = [BR]$ ve $y = [DR]$ olur ve [58] belirtilen şekilde skaler enerji metriği hesaplanır.

Daha sonra özellik vektörüne şu şekilde tanımlanan $v(x) = \sum_{y=0}^{y-1} S(x, y)$ skaler enerji metriği eklenerek özellik vektörünün boyutu $y+1$ olarak elde edilir. Her bir özellik $S'(x, y) = S(x, y)/v(x)$ denklemi ile elde edilir. Özellik vektörünün son halini elde etmek için tüm enerji şu şekilde hesaplanır:

$$enerji = \sum_{x=0}^{x-1} \sum_{y=0}^y S'(x, y) \quad (3.12)$$

Özelliklerin son hali $S'(x, y)/enerji$ denklemi ile elde edilir. Skaler enerji metriğinin kullanılması gürültülü koşullar altında sistem performansın %10-20 arasında artış kaydetmektedir.

Bunlara ilave olarak gürültü altında sınıflandırma performansını artırmak için [58] te belirtilen şekilde basit bir gürültü giderme işlemi uygulanır. Şu şekilde hesaplanır:

$$S_{an}(x, y) = S(x, y) - \min(S(x, y)) \quad (3.13)$$

3.4. DBN İÇİN DENEYSEL KURULUM

Tüm durumlarda özellikler 2048 nokta ile hamming çerçevesi kullanılarak, 2032 nokta ile yüksek derecede örtüşecek şekilde logaritmik ölçekli olarak elde edilmiştir.

Yeniden boyutlandırma işlemi için a değeri 2 olmak üzere Lanczos kernel kullanılmıştır. Yeniden boyutlandırma işlemi değişik yeniden boyutlandırma faktör değerlerinde yapılarak Çizelge 3.1'de belirtilen farklı uzunluktaki özellik vektörleri elde edilmiştir.

Her durumda 2 katmanlı DNN kullanılarak sınıflandırma performansı test edilmiştir. Ancak yeniden boyutlandırma faktörüne bağlı olarak özellik vektörlerinin boyutunun değişmesi nedeni ile her bir katmanda kullanılan düğüm sayısı değişiklik

göstermektedir. Aynı özellik setinde farklı düğüm sayıları kullanılarak sınıflandırma performansları kıyaslanmıştır. Çizelge 3.1’de kullanılan sistem parametreleri görülmektedir.

Çizelge 3.1. Sistem parametreleri.

Yeniden Boyutlandırma Faktörü	Özellik Boyutu	Çerçeve (ms)	Gizli Katman	Gizli Düğüm
0,05	53	128	2	75
0,1	104	128	2	100
0,2	206	128	2	210

3.5. DBN İÇİN SONUÇLAR VE TARTIŞMA

3 farklı yeniden boyutlandırma faktörü ile elde edilen sınıflandırma performansları Çizelge 3.2’de görülmektedir. Buna göre her 3 yeniden boyutlandırma faktörü ile gürültüsüz koşullar altında elde edilen sınıflandırma performansı birbirine çok yakındır. %10 yeniden boyutlandırma faktöründe 104 özellik ile %97,13’lük bir başarımla görülürken, sadece 53 özellik kullanılan %5 yeniden boyutlandırma oranında %96,4’lük bir sınıflandırma başarımla gözlemlenmiştir. Toplam sistem performansı açısından en iyi sonuç %20 yeniden boyutlandırma oranında 206 özellik ile %64,35’tir.

Çizelge 3.2. Sınıflandırma başarımları.

Yeniden Boyutlandırma (%)	Gürültüsüz	20 dB	10 dB	0 dB	Ortalama
20	96,73	82,00	52,60	26,07	64,35
10	97,13	73,00	41,47	22,73	58,58
5	96,40	74,47	42,53	24,27	59,42

Standart kıyaslama görevi kullanılması sistem performansını literatürdeki diğer çalışmalarla kıyaslanmasına olanak sağlamaktadır. Çizelge 3.3’te sınıflandırma performanslarına ait bir liste verilmektedir. Listeye göre burada önerilen yöntemin sınıflandırma performansının SIF ve ETSI-AFE haricindeki sınıflandırmalardan daha

iyi bir sonuç sağladığı görülmektedir. Öte yandan Dennis SIF, SIF-DNN-DN-e ve NLM-SIF-v yöntemlerinde sırasıyla 486,721 ve 901 özellik kullanılırken bizim önerdiğimiz yöntemde %20 yeniden boyutlandırma faktörü için 206 özellik kullanılmaktadır.

Çizelge 3.3. En son güncel yöntemler için sınıflandırma doğruluğu ([54,56]).

Sistem	Gürültüsüz	20 dB	10 dB	0 dB	Ortalama
MFCC-HMM	99,4	71,9	42,3	15,7	57,4
MFCC-SVM	98,5	28,1	7,0	2,7	34,1
ETSI-AFE	99,1	89,4	71,7	35,4	73,9
MPEG-7	97,9	25,4	8,5	2,8	33,6
Gabor	99,8	41,9	10,8	3,5	39,0
GTCC	99,5	46,6	13,4	3,8	40,8
MP+MFCC	99,4	78,4	45,4	10,5	58,4
Dennis SIF	91,1	91,1	90,7	80,0	88,5
SIF-DNN-DN-e	96,0	94,4	93,5	85,1	92,3
NLM-SIF-v	98,0	96,3	93,3	84,0	92,9

Bunlara ilave olarak sistem performansı üzerine gizli düğüm sayısının etkisini görmek için %5 yeniden boyutlandırma faktöründe 45, 60, 75 ve 90 düğüm sayılarında sınıflandırma yapılmıştır. Çizelge 3.4’te de görüldüğü gibi en yüksek başarımlar %59,42 ile 75 gizli düğüm sayısında elde edilirken düğüm sayısı artırıldığında veya azaltıldığında sınıflandırma performansının düştüğü görülmüştür.

Çizelge 3.4. Farklı gizli düğümler için %5 yeniden boyutlandırma faktöründe sınıflandırma doğruluğu.

Gizli Düğüm Sayısı	Gürültüsüz	20 dB	10 dB	0 dB	Ortalama
45	94,27	61,67	36,27	16,73	52,23
60	96,33	71,47	43,67	17,93	57,35
75	96,40	74,47	42,53	24,27	59,42
90	96,07	72,67	36,67	17,13	55,63

Tez çalışmasının bu bölümünde Lanczos kernel kullanılarak SIF özelliklerinin azaltılması ve DBN ile sınıflandırılması üzerine bir yöntem önerildi. Tüm yeniden

boyutlandırma faktörlerinde gürültüsüz koşullar altında literatürdeki çalışmalara yakın sonuçlar elde edildi. Diğer taraftan gürültü altında MFCC, MPEG7 ve Gabor gibi özelliklerden daha iyi bir sonuç elde edilirken mevcut SIF tekniklerinin bir miktar altında sonuçlar elde edilmiştir.

Tez çalışmasının bu bölümünde yalnızca Lanczos kernel kullanılarak değişik yeniden boyutlandırma faktörleri ile işlem yapılmıştır. Elde edilen sonuçlara göre daha az özellik kullanımı ile gürültüsüz koşullarda literatürdeki çalışmalara yakın sonuçlar elde edilirken önerilen yöntemin gürültü altındaki performansı güncel tekniklerin altında kalmıştır. Bu nedenle silah sesine ait akustik bileşenlerin sınıflandırılması için uygun bulunmamıştır.

3.6. CNN İÇİN SPEKTROGRAM RESİM ÖZELLİKLERİ

DBN için kullanılan spektrogramlar logaritmik ölçekli olarak kullanılmış ve RGB komponentlerine nicelenmemiştir. CNN için kullanılan spektrogram resim özellikleri DBN için kullanılanlardan farklıdır.

Her iki yöntemde FFT yapılarak bir spektral yığın elde edilmektedir. Sonrasında ilk aşamada lineer ölçeklenmiş resim elde edilmektedir. Bu aşamadan sonra DBN için elde edilen resimler logaritmik ölçeklenmiş şekilde kullanılmaktadır. Ayrıca resimler Lanczos kernel kullanılarak yeniden boyutlandırılmaktadır ve son olarak basit bir gürültü giderme işlemi uygulanmaktadır. Gürültü giderme işlem sınıflandırma performansını %10-20 arasında artırmaktadır.

CNN ile birlikte kullanılan spektrogram resimlerde ise Eşitlik 3.8'den sonra elde edilen zaman-frekans matrisi [0, 1] aralığında normalize edilerek gri tonlamalı resim elde edilmektedir. Gri tonlamalı spektrogram resim elde etme işlemini aşağıda gösterilen şekilde tanımlamak mümkündür:

$$G(k, F) = \frac{S(k, F) - \min(S)}{\max(S) - \min(S)} \quad (3.14)$$

Bu aşamadan sonra gri tonlamalı resim kırmızı, yeşil, mavi (RGB) kompanantlerine nicelenir. m_c monokrom resim, f doğrusal olmayan fonksiyon ve c niceleme bölgeleri olmak üzere haritalama şu şekilde ifade edilebilir:

$$m_c(k, F) = f_c(G(k, F)) \quad \forall c \in (c_1, c_2, \dots, c_N) \quad (3.15)$$

3.7. RESİM YENİDEN BOYUTLANDIRMA YÖNTEMLERİ

Bu tez çalışmasında kullanılan ses veri tabanı farklı uzunluklardaki ses sinyallerini içerdiğinden, elde edilen spektrogram görüntülerinin boyutları birbirine eşit değildir. CNN'nin birinci tabakası, sabit boyutlu görüntüleri girdi olarak gerektirdiğinden, elde edilen spektral görüntülerin boyutları yeniden boyutlandırma yoluyla ayarladıktan sonra kullanılmaktadır. Bu tez çalışmasında, sınıflandırma performansı üzerine etkilerini görmek için en yakın komşu, bilinear, bikübik ve Lanczos ($a=2$ ve 3 ile) olmak üzere 5 farklı yeniden boyutlandırma yöntemi kullanılmıştır.

En yakın komşuluk en basit enterpolasyon yöntemidir ve değeri bulunmak istenen nokta en yakın komşusunun değerini alır. Bu yöntem, diğer enterpolasyon yöntemlerinden çok daha hızlı çalışır. Bilinear enterpolasyon yöntemi, yeni bir piksel değerini tahmin etmek için en yakın 2×2 komşusunun ağırlıklı ortalama değerlerini dikkate alır. Bikübik enterpolasyon, bilinear enterpolasyonun aksine komşu olan en yakın 16 pikselin değerini dikkate alır. Bikübik enterpolasyon, en yakın komşu ve doğrusal olmayan enterpolasyonlardan daha yavaş olmakla birlikte, enterpolasyondan sonra elde edilen yüzeyler daha pürüzsüzdür ve daha az enterpolasyon yapay dokularına sahiptir. Lanczos çekirdek yöntemi, orijinal görüntüyü daha düzgün görünmesini sağlamak için bir konvolüsyon çekirdeği kullanır. Bir sinc fonksiyonunun sonsuz desteği vardır, yani sınırsız süreye sahiptir ve bant sınırlı sinyal için en uygun yeniden yapılandırma veya rejenerasyon filtresidir. Lanczos filtrasyonu, sonlu destekle sinc filtrasyon tekniğine etkin bir şekilde yakınsamak için kullanılabilir. Her enterpolasyon değeri, $2a$ sıralı giriş örneklerinin ağırlıklı toplamından hesaplanır. Daha düzgün bir enterpolasyon veya keskin bir geçiş arasındaki bir seçim, uygun bir parametre ile $2a$ değeri değiştirilerek yapılabilir.

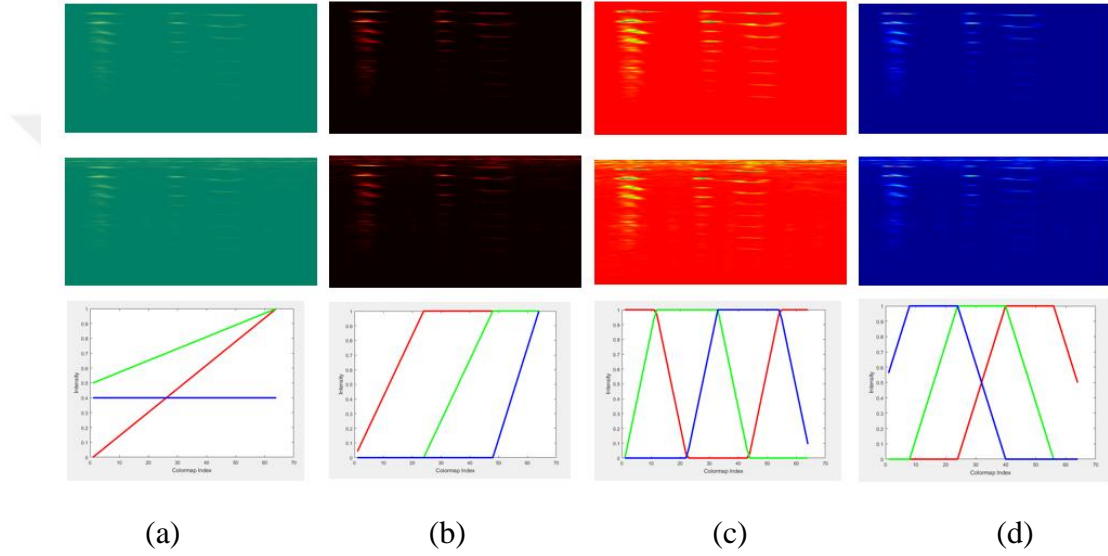
3.8. KONVOLUSYONAL SİNİR AĞI

CNN ilk olarak bilgisayar görmesi problemi için önerilmiş [67] tipik bir çok katmanlı yapay sinir ağı (multi-layer neural network) [68,69] modelidir ve makine öğrenmesinde resim ile ilgili uygulamalar için oldukça uygundur [70,71]. Genellikle CNN konvolusyon, havuzlama ve tam bağlı katmanlarından oluşmasının yanı sıra sıradan neural networklerden farklı olarak katmanları 3 boyutlu nöronlara sahiptir.

Konvolusyon katmanı filtre boyutu ve üretilen haritaların sayısı ile tanımlanır ve CNN'yi oluşturan en önemli birimdir. Bir nesnenin görüntüsünün resmin neresinde bulunduğundan bağımsız olduğu fikrinden hareketle her bir nöron girişin sadece küçük bir bölümüne bağlanmakta ve girişin tüm derinliği boyunca da uzanmaktadır. İleri yayılım aşamasında giriş ve filtreler arasında nokta çarpımı yapılarak 2 boyutlu bir aktivasyon haritası oluşturulmaktadır. CNN'in öğrenmesini istediğimiz verilerin büyük kısmı doğrusal olmayan veriler olmasına rağmen konvolusyon, matris çarpımı ve toplaması gibi işlemleri içeren lineer bir yöntemdir. Bu nedenle ReLU (Rectified Linear Units) ile doyma olmayan bir aktivasyon fonksiyonu $f(x)=\max(0,x)$ uygulanarak [72] CNN'in doğrusal olmama özelliği kazandırılır. Bir diğer önemli CNN katmanı doğrusal olmayan aşağı örnekleme işlemi yerine getiren havuzlama katmanıdır. Çeşitli tiplerde havuzlama katmanları mevcuttur fakat en yaygın olanları maksimum- havuzlama ve ortalama-havuzlamadır [73]. Bu çalışmada max-pooling yöntemi tercih edilmiştir. Maksimum-havuzlama yönteminde giriş örtüşmeyen dikdörtgenlere bölünür ve her bir alt parça arasından en büyük olanı alınır. Tamamen bağlı katmandaki nöronlar geleneksel sinir ağlarında olduğu gibi kendisinden önceki ve sonraki katmandaki her bir nörona bağlıdır ve çıkış katmanında softmax veya SVM gibi bir sınıflandırıcı kullanılır. Bu şekilde oluşturulan çok katmanlı yapı ağın daha soyut özellikleri bulmasını, daha karmaşık yapıları keşfetmesini sağlayarak, CNN'in resim sınıflandırma [74-77], yüz ifadesi tanıma [78] ve sahne etiketleme [79] gibi birçok uygulamada başarılı olmasını sağlamaktadır. Benzer şekilde [80] ve [81]'de konuşma tanıma ve [82]'de konuşma olmayan sesleri tanıma alanına uygulanmış ve başarılı sonuçlar elde edilmiştir.

3.9. ÖNERİLEN DARBESEL SES TANIMA YAKLAŞIMI

MFCC ve HMM gibi birçok geleneksel ses olayı tanıma sistemleri gürültüsüz koşullarda oldukça yüksek bir başarımla sergilerken ortam gürültüsünün artmasıyla birlikte sınıflandırma performansları da önemli ölçüde düşmektedir. Oysaki günlük hayatta ortamda gürültü olması kaçınılmazdır. Bu çalışmada SIF ve CNN ler bir arada kullanılarak tüm koşullarda yüksek sınıflandırma performansı sağlanmaktadır.



Şekil 3.1. Soldan sağa doğru sırasıyla “summer”, “hot”, “hsv” ve “jet” renk haritaları kullanılarak elde edilmiş Korna sesine ait lineer nicelenmiş spektrogram resimler ve yukarıdan aşağıya doğru gürültüsüz ve 0 dB gürültü altındaki görüntüleri. Son satırda ilgili renk haritasına ait index.

Gürültünün darbesel ses sinyaline göre daha yayılmış bir şekilde sahip olması ve Şekil 3.1’de de görüldüğü gibi gürültü etkisinin spektrogram üzerinde belirli bölgelerde bozulmalara neden olurken diğer bölgelerin bu bozulmadan etkilenmemesi, SIF’leri gürültüye karşı dayanıklı hale getirmektedir. Gürültüden farklı olarak, ses olayları genellikle kısa zaman periyotlarına ve daha belirgin zaman-frekans temsillerine sahiptir. Bu nedenle, seslerin spektrogramlarından elde edilen görsel bilgiler, ses sınıflandırması için iyi bir özellik olabilir. Spektrogramda ayırt edilebilen paternler zayıf mutlak yerelliğe sahiptir, bu da fark edilebilir bir ses olayının farklı zamanlarda ve biraz farklı frekans aralıklarında görülebileceği anlamına gelir. Buna ek olarak,

CNN, konvolüsyon ve alt örnekleme işleminden dolayı görüntü üzerindeki örüntünün pozisyonuna duyarsızdır, bu nedenle SIF'leri sınıflandırmak için uygundur.

Öte yandan, CNN ve SIF'nin birlikte kullanılmasında iki temel sorun vardır. İlk olarak, örtüşme oranına, ses vektörünün ve çerçevenin uzunluğuna bağlı olarak SIF'in boyutları, MFCC gibi diğer özelliklere kıyasla çok daha büyüktür. İkinci olarak, CNN'ler bir dizi zorlu koşullar altında çok iyi performans gösterirken, sabit büyüklükteki girdilere duyulan ihtiyaç nedeniyle ses olaylarının doğasına aykırı önemli bir eksikliğe sahiptirler. Bu tez çalışmasında, yüksek gürültü altında ses olaylarının sınıflandırılması CNN ve SIF'yi bir arada kullanan bir metot önerilmiştir. Çeşitli sınıflandırıcılar SIF özelliklerinin sınıflandırılmasında iyi bir performans göstermesine rağmen, SIF oluşumu ve CNN giriş katmanı arasındaki özellik çıkarma yöntemi bugüne kadar tam olarak ortaya konmamıştır.

Önerilen yöntemde [59]'da olduğu gibi, spektrogram ilk olarak sabit artışlarla gri tonlamada normalleştirilmiştir, daha sonra gri tonlamalı görüntü RGB bileşenine nicelenir ve tek renkli bir görüntü elde edilir. Spektrogram görüntüsü temel olarak FFT büyüklük spektrumlarının her bir çerçevenin yan yana istiflenmesiyle oluşan bir veri matrisidir. N ses vektörünün uzunluğu, w_s çerçeveler olmak üzere zaman-frekans matrisinin boyutları aşağıdaki gibi elde edilir:

$$F = \lfloor w_s/2 + 1 \rfloor \quad (3.16)$$

$$L = \lfloor ((N - (w_s - \delta)))/\delta \rfloor \quad (3.17)$$

F , zaman-frekans matrisinin frekans boyutunu belirtir ve L zaman boyutunu gösterir. Bu nedenle, her bir gri ölçekli spektrogram görüntüsü için bir $F \times L$ büyüklüğünde matrisi elde edilmesinin yanı sıra nicelenmiş görüntüler için $F \times L \times 3$ boyutu elde edilir. Çerçeve uzunluğu her bir ses vektörü için sabit olduğundan, zaman-frekans matrislerinin frekans boyutu her spektrogram görüntüsü için aynıdır. Diğer taraftan, her ses vektörünün uzunlukları eşit değildir, bu nedenle elde edilen zaman-frekans matrislerinin zaman boyutu her bir görüntü için farklıdır. Önerilen yöntemde CNN giriş katmanına uygulanacak görüntü boyutunu belirlemek için, n eğitim ve test

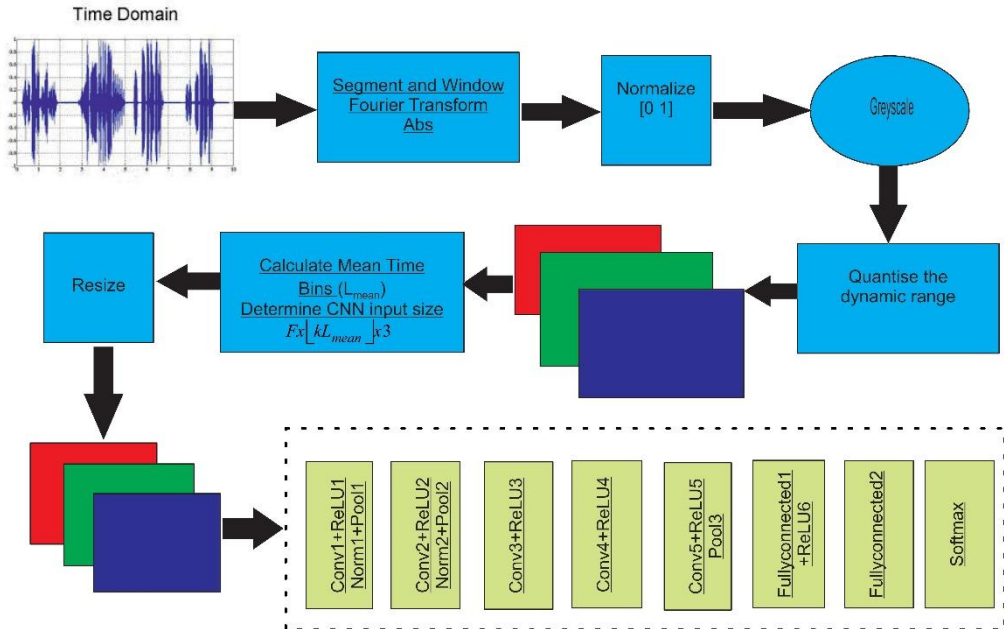
verisindeki ses vektörlerinin toplam sayısı olmak üzere tüm spektrogram görüntülerinin ortalama zaman boyutu aşağıdaki gibi belirlenir:

$$L_{mean} = \frac{1}{n} \sum_{i=1}^n \left\lfloor \frac{N_i - (w_s - \delta)}{\delta} \right\rfloor \quad (3.18)$$

Spektrogram görüntülerinin boyutunun k ile azaltılması için kullanılan yeniden boyutlandırma faktörüdür ve F frekans boyutu tüm görüntülerde sabittir, CNN ağına uygulanacak SIF boyutları aşağıdaki gibi elde edilir:

$$F \times \left\lfloor \frac{1}{k} L_{mean} \right\rfloor \times 3 \quad (3.19)$$

Tüm spektrogram görüntüleri, görüntü yeniden boyutlandırma yöntemleri ile Denklem 9'daki boyutlara sabitlenerek CNN'ye uygulanır. CNN parametrelerini değiştirmeden, sınıflandırma performansları en yakın komşuluk, bilinear, bikubik interpolasyon ve Lanczos kerneli (a=2 ve 3 değerleri kullanılarak) olmak üzere 5 farklı deneme metodu kullanılarak karşılaştırılmış önerilen metod Şekil 3.2'de tasvir edilmiştir.



Şekil 3.2. CNN tabanlı otomatik ses olayı tanıma sistemine ait blok diyagram.

3.10. DENEYSEL DEĞERLENDİRME

Bu bölümde öncelikli olarak bu tez çalışmasında kullanılan ses veri tabanı ve gürültü koşulları açıklanıyor. Daha sonra test işlemlerinde kullanılan sistem yapısı anlatılıyor.

3.10.1. Ses olayı ve Gürültü Veri Tabanı

Real Word Computing Partnership (RWCP) Sound Scene Database in Real Acoustical Environments [83] veri tabanından [59]'da tarif edilen şekilde toplam 50 sınıf seçilmiştir. Standart değerlendirme yöntemi kullanılması, elde edilen sonuçların daha önce yapılan çalışmalarla kıyaslanarak sistemin sınıflandırma performansının değerlendirilmesine olanak sağlamaktadır. Her bir sınıf yüksek sinyal gürültü oranı (SNR) değerlerinde kaydedilmiş 80 kayıt içeriyor ve her bir kaydın başlangıç ve bitiş bölümlerinde sessiz bölümler bulunuyor. Renk haritası etkisi dışındaki tüm testlerde beş katlı çapraz doğrulama kullanılmıştır. 80 ses dosyasından rasgele seçilen 50 tanesi eğitim seti için kalan 30 tanesi de test seti için kullanılmaktadır. Toplamda eğitim için 2500 ve test için 1500 dosya bulunmaktadır. Çoklu durum testleri haricinde eğitim işlemi sadece temiz ses dosyaları ile herhangi bir ön işlem yapılmadan yapılmaktadır. Test aşamasında ses dosyalarına NOISEX-92 veri tabanından [78] seçilen “Destroyer Control Room”, “Speech Babble”, “Factory Floor 1” ve “Jet Cockpit 1” isimli 4 farklı gürültüden bir tanesi 20, 10 ve 0 dB SNR seviyelerinde eklenmektedir. Eklenecek gürültü ve gürültü içerisindeki başlangıç noktası rasgele seçilmektedir. [59]'da seçilen bu gürültüler, durağan olmayan ve baskın olarak düşük frekanslı bölgeleri olan çevresel gürültülerin tipik örnekleridir.

3.10.2. Deneysel Kurulum

Tüm durumlarda spektrogramlar 16 kHz örnekleme hızında 1024 nokta Hamming çerçevesi kullanılarak, 1012 nokta ile yüksek derecede örtüşecek şekilde elde edilmiş ve [59]'daki çalışmada tarif edilen şekilde lineer ölçekli spektrogram resimlere dönüştürülmüştür. Bu aşamada “hot”, “summer”, “hsv” ve “jet” renk haritaları kullanılarak farklı lineer nicelenmiş resimler elde edilmiştir.

Yeniden boyutlandırma işlemi için en yakın komşu, bilinear, bikübik enterpolasyon ve a değeri 2 ve 3 olmak üzere Lanczos çekirdek metotları kullanılmıştır. Yeniden boyutlandırma faktörü $k=6$ seçilerek tüm resimler aynı boyutlara getirilip ağa uygulanmıştır. Sınıflandırma performansı üzerine etkilerini görmek adına aynı yeniden boyutlandırma metodu ile farklı k parametreleri ile yeniden boyutlandırma işlemi yapılmıştır.

3.10.3. CNN Mimarisi

Genellikle CNN'lerin eğitimi için çok fazla sayıda resim [66,85] kullanılmaktadır fakat standart bir kıyaslama görevi kullanılması nedeniyle eğitim için kullanabileceğimiz sınırlı sayıda veri bulunmakta bu nedenle ImageNet veri tabanında ki [86] 1,3 milyon resim kullanılarak 101 iterasyon ile önceden eğitilmiş (pre-trained) bir ağ kullanarak daha güçlü bir resim özelliği sunumu sağlamayı hedefledik. [87]'de yapılan çalışma CNN'de rasgele başlangıç parametreleri seçmek yerine önceden eğitilmiş ağ parametreleri kullanmanın olumlu etkilerini ortaya koymuştur.

Bu tez çalışmasında kullanılan CNN mimarisi Şekil 3.2'de görülmektedir. [66]'da kullanılan ağa benzer bir ağ mimarisi kullanılmıştır. Fakat [66]'dan farklı olarak 8 eğitilmiş layer yerine 5'i konvoluyonal ve 2'si tamamen bağlı katman olmak üzere 7 eğitilebilir katmandan oluşmaktadır. 2. Tamamen bağlı katmanın çıkışı 50-yollu softmax katmanına giriş olarak uygulanarak etiket bilgilerinin dağılımı elde edilmektedir. $513 \times 155 \times 3$ boyutlarında lineer nicelenmiş spektrogram resimler giriş katmanına girdi olarak uygulanmaktadır. İlk iki konvoluyonal katman 4 aşamadan oluşuyor bunlar sırasıyla konvolusyon, ReLU, normalizasyon ve havuzlama. 3. ve 4. konvoluyonal katman ise konvolusyon ve ReLU olmak üzere 2 aşamadan oluşuyor. 5. konvoluyonal katmanda 3. ve 4. katmandakilere ilave olarak havuzlama aşaması bulunmaktadır, bu katmanı iç çarpım ve ReLU olmak üzere 2 aşamadan oluşan 500 düğümlü tamamen bağlı katman takip etmektedir. Ayrıca 1'den 5'e kadar konvolusyon katmanlar sırasıyla , 96 özellik haritası ile birlikte filtre boyutu 11×11 adım 4 dolgu 0, 256 özellik haritası ile birlikte filtre boyutu 5×5 adım 1 dolgu 2, 384 özellik haritası ile birlikte filtre boyutu 3×3 adım 1 dolgu 1, 384 özellik haritası ile birlikte filtre boyutu 3×3 adım 1 dolgu 1 ve 256 özellik haritası ile birlikte filtre

boyutu 3x3 adım 1 dolgu 1 den oluşmaktadır. Birinci tamamen bağlı katman 500 düğümünden oluşurken ikinci tamamen bağlı katman 50 düğümünden oluşmaktadır ve son olarak bu katmanı softmax sınıflandırma katmanı takip etmektedir.

Her durumda 7 eğitilebilir katmanlı CNN kullanılarak sınıflandırma performansı test edilmiştir. Önerilen ağ stokastik gradyanlı alçalma kullanılarak, parti boyutu 25 örnek, momentum 0,9, ağırlık kaybı 0,0005, öğrenme oranı 1×10^{-4} olarak eğitilmiş ve eğitim 1000 iterasyondan sonra durdurulmuştur.

3.11. SONUÇLAR VE TARTIŞMA

Bu bölümde öncelikli olarak CNN tabanlı ses olayı sınıflandırma sisteminin sınıflandırma performansı rapor edilen diğer sistemlerle kıyaslanarak metodun etkinliği gözlemlenecektir, sonrasında farklı parametreler altındaki sınıflandırma performansı değerlendirilecektir.

3.11.1. Diğer Sistemler İle Karşılaştırma

Standart bir kıyaslama görevi kullanmanın sağladığı en büyük avantaj sistemin başarısını literatürde rapor edilen diğer çalışmalarla karşılaştırılmasına olanak sağlamasıdır. Çizelge 3.5'te aynı kıyaslama görevi kullanılarak yapılmış bazı çalışmaların sonuçları ve bu tez çalışmasında önerilen Lanczos çekirdek tabanlı CNN ("summer" renk haritası ile, $a=2$, $k=6$ ve CNN giriş boyutu $513 \times 155 \times 3$ olmak üzere) modelin sonuçları görülmektedir. Daha önceki çalışmalarda ortalamada en iyi sınıflandırma performansına sahip olan NLM-SIF-v ile kıyaslandığında gürültüsüz, 20, 10, ve 0 dB koşullarında sırasıyla %1,3, %2,9, %5,5, ve %8,5 daha iyi bir sınıflandırma performansı sergilediği görülmektedir. Ortalamada ise %63,4'lük bağıl hata azaltma oranına eşdeğer %4,5'lik doğruluk artışı sağlamıştır. Aynı zamanda Çizelge 3.5'te koyu renkli olarak gösterilen değerler ilgili koşul altındaki en iyi sonuçları göstermektedir. Bu tez çalışmasında önerilen Lanczos2-CNN metodu gürültüsüz koşullar haricindeki tüm durumlarda en yüksek başarıyı gösterirken, gürültüsüz koşullarda da en iyi değeri sergileyen Gabor'un yalnızca %0,5 altında %99,3'lük bir başarıyı sergilemiştir.

Çizelge 3.5. En son güncel yöntemler ve Lanczos2-CNN modelinin sınıflandırma doğruluğu ([58], [59], [61]) (Lanczos2-CNN modelinin sınıflandırma doğruluğu standart sapma ile birlikte verilmiştir).

Sistem	Gürültüsüz	20 dB	10 dB	0 dB	Ortalama
MFCC-HMM	99,4	71,9	42,3	15,7	57,4
MFCC-SVM	98,5	28,1	7,0	2,7	34,1
ETSI-AFE	99,1	89,4	71,7	35,4	73,9
MPEG-7	97,9	25,4	8,5	2,8	33,6
Gabor	99,8	41,9	10,8	3,5	39,0
GTCC	99,5	46,6	13,4	3,8	40,8
MP+MFCC	99,4	78,4	45,4	10,5	58,4
Dennis SIF	91,1	91,1	90,7	80,0	88,5
SIF-DNN-DN-e	96,0	94,4	93,5	85,1	92,3
NLM-SIF-v	98,0	96,3	93,3	84,0	92,9
Lanczos2-CNN	99,3 (0,1)	99,2 (0,2)	98,8 (0,2)	92,5 (0,6)	97,4 (0,2)

3.11.2. Yeniden Boyutlandırma Metodun Etkisi

Bu tez çalışmasında kullanılan ses veri tabanının farklı uzunlukta ses sinyalleri içermesi nedeniyle elde edilen spektrogram resimlerin boyutları birbiriyle aynı değildir. Fakat CNN nin ilk katmanına sabit boyutlardaki görüntünün giriş olarak uygulanması gerektiğinden elde edilen spektral görüntülerin boyutlarının tüm resimler için aynı olması gerekmektedir. Bu aşamada sınıflandırma performansı üzerine etkilerini görmek için en yakın komşu, bilinear, bikübik enterpolasyon ve a değeri 2 ve 3 olmak üzere Lanczos kernel metotları kullanılmıştır. Çizelge 3.6’da farklı yeniden boyutları metotları kullanarak “summer” renk haritası i ile $k = 6$ ve giriş boyutu = $513 \times 155 \times 3$ olmak üzere elde edilen sonuçlar görülmektedir.

Çizelge 3.6. Yeniden boyutlandırma faktörünün etkisi (standart sapma ile birlikte).

Sistem	Gürültüsüz	20 dB	10 dB	0 dB	Ortalama
Nearest-CNN	99,2(0,2)	99,2(0,2)	98,8(0,4)	91,7(1)	97,2(0,3)
Bilinear-CNN	99,3(0,2)	99,2(0,2)	98,08(0,2)	90,9(0,9)	97,0(0,2)
Bicubic-CNN	99,2(0,2)	99,2(0,2)	98,7(0,2)	91,2(0,9)	97,0(0,3)
Lanczos2-CNN	99,3(0,1)	99,2(0,2)	98,8(0,2)	92,5(0,6)	97,4(0,2)
Lanczos3-CNN	99,2(0,2)	99,1(0,3)	98,6(0,4)	91,5(1,6)	97,1(0,4)

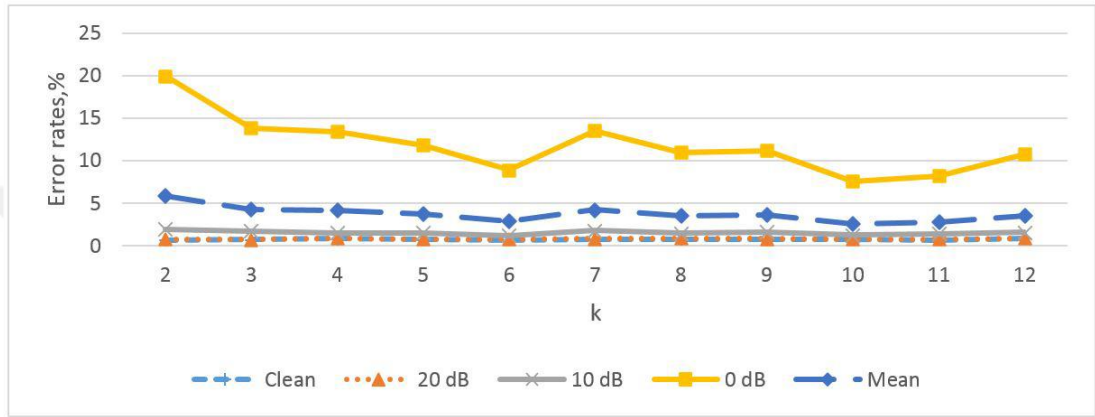
Gürültüsüz ve 20 dB gürültü koşulları altında tüm yöntemler benzer sonuçları verirken, 10 dB gürültü koşullarında Lanczos2-CNN ve en yakın-CNN yöntemleri %98,8 en iyi başarı oranını, ayrıca 0 dB koşullarında Lanczos2-CNN %92,5 ile en iyi sınıflandırma sonucu verir. Bu durum, ortam gürültü seviyeleri arttıkça yeniden boyutlandırma yönteminin doğru seçilmesinin önemini artırır.

Bilineer ve Bikübik yeniden boyutlandırma metotları, bilinmeyen pikselin değerini hesaplarırken bu pikseli çevreleyen ve değerleri bilinen sırasıyla en yakın 2x2 ve 4x4 komşusunu dikkate alır. Bu, aşağı örneklemede keskin özelliklerin bulanıklaşmasına neden olur. Bulanıklığın kapsamı, çekirdeğin genişliğine bağlıdır. Mevcut veri setiyle elde ettiğimiz spektrogram görüntüleri çoğunlukla aniden değişen bölgeleri içerdiğinden, bu durum sınıflandırma performansını etkilemektedir. Bununla birlikte, spektrogram görüntülerinin işlenmesi, basamaklanma yapay dokuları ve keskin kenarların korunması arasında bir dengeyi gerektirir. Lanczos filtresi ($a=2$ için) diğer enterpolasyon metotları ile karşılaştırıldığında örnek seyreltme (decimation) ve enterpolasyon için basamaklanma azaltılması, keskinlik ve en düşük parazit salınım açısından daha iyi bir uyum sağlar [89-91]. Bu nedenle, Lanczos2 daha iyi bir sınıflandırma performansı sağlar.

3.11.3. Yeniden Boyutlandırma Faktörü ve Renk Haritası Etkisi

Önerilen modelde öncelikli olarak eğitim ve test için kullanılmış olan tüm spektrogram resimlerindeki ortalama spektral çizgi sayısı (veri setinde ortalama zaman uzunluğu 933 olarak elde edildi) bulunduktan sonra Eşitlik 3.19’da belirtilen şekilde CNN giriş katmanına uygulanacak resim boyutları tespit edilmektedir. Şekil 3.3’te “summer” renk haritası ve Lanczos2-CNN modeli kullanılarak k ’nın 3’ten 10’a kadar olan değerleri için elde edilen hata oranları görülmektedir. Bu aşamada CNN’nin diğer parametrelerinde değişiklik yapılmadan giriş katmanına k ’nın 2 ve 12 değerleri için sırasıyla 513x466x3 ve 513x77x3 boyutlarında resimler girdi olarak uygulanmıştır. Şekil 3.3’te de görüldüğü gibi 0 dB gürültü oranında $k=10$ değeri için en küçük hata oranları elde edilirken k ’nın 6’dan küçük değerlerinde hata oranında doğrusala yakın bir artış göstermiştir. Öte yandan k ’nın 6’dan büyük değerlerinde hata oranı daha

yatay bir seyir izlemektedir. 10 dB gürültü altında da k parametresine bağlı değişim 0 dB’dekine benzer olmakla birlikte değişim 0 dB’ye oranla daha yumuşaktır. Grafikten çıkarılacak bir diğer önemli sonuçta gürültünün olmadığı veya çok az olduğu durumlarda k parametresinin çok önemli olmadığıdır. Bununla birlikte artan spektrogram resim boyutuyla birlikte yüksek gürültü altında ki performansta ciddi anlamda düşüş görülmektedir.



Şekil 3.3. “summer” renk haritası ile k yeniden boyutlandırma faktörüne bağlı hata oranlarının değişimi.

Bu tez çalışmasında önerilen modelin en fazla duyarlı olduğu noktalarda biri, Çizelge 3.7’de de görüldüğü gibi, doğru renk haritasının seçilmesidir. 10 ve 0 dB’de artan gürültü seviyeleriyle, renk haritasının etkisi açıkça görülebilir. “Hot” renk haritası, “summer” renk haritasına yakın bir başarı oranını %96,4 sağlamasına rağmen, başarı oranları “jet” ve “hsv” için sırasıyla %93 ve %90,1 ile önemli ölçüde düşmüştür.

Çizelge 3.7. Farklı renk haritalarına sahip Lanczos2-CNN modeli kullanılarak $k = 6$ ve $513 \times 155 \times 6$ giriş büyüklüğü ile elde edilen sınıflandırma performansı.

Sistem	Gürültüsüz	20 dB	10 dB	0 dB	Ortalama
Hot	99,2	99,1	98,5	88,6	96,4
Jet	98,9	98,7	97,5	77,0	93,0
Hsv	99,1	98,7	94,2	68,5	90,1

Her bir renk haritasının renk indeksine karşılık gelen yoğunluk değerleri incelendiğinde “hot” ve “summer” renk haritalarında her bir monokrom renk değerinin

düzgün bir şekilde artış gösterdiği veya sabit kaldığı görülürken “hsv” ve “jet” renk haritalarında harita indeksine bağlı olarak yoğunluk değerlerinin her bir renk değeri için artış ve azalış gösterdiği görülmektedir. “Hsv” ve “jet” renk haritalarındaki bu değişim gürültü altındaki sınıflandırma performansını olumsuz etkilemektedir.

3.11.4. Çoklu Durum Testleri

Çoklu durum testlerinde her bir ses sınıfında eğitim için kullanılan 50 ses dosyasından rastgele seçilen 10 tanesinin her birine “Destroyer Control Room”, “Speech Babble”, “Factory Floor 1” ve “Jet Cockpit 1” gürültülerinden bir tanesi 0 dB SNR oranında eklenirken kalan 40 tane ses dosyasına herhangi bir işlem yapılmadan gürültüsüz olarak kullanılmıştır. Çizelge 3.8’de de görüldüğü gibi 0 dB SNR değerinde sadece gürültüsüz sesler ile eğitilenlere oranla belirgin bir artış yaşanmıştır, buna ilave olarak 10 dB gürültü altındaki değerlerde de ortalama %0,45 lik bir artış olmuştur. Öte yandan Bilineer-CNN, Bikübik-CNN ve Lanczos2-CNN modelleri için gürültüsüz ve 20 dB koşullarında çok küçük miktarda düşüş yaşanırken Nearest-CNN ve Lanczos3-CNN için gürültüsüz ve 20 dB de herhangi bir düşüş yaşanmamıştır.

Ayrıca uyuşmayan koşullarda ki testlerde ortalamada elde edilen en iyi ve en düşük performanslar arasındaki %0,75’lik fark çoklu durum testlerinde yarı yarıyadan biraz daha fazla azalarak %0,37’ye inmiştir. Çoklu durum testlerinde Lanczos3-CNN haricindeki modeller gürültüsüz, 20 ve 10 dB de birbirine oldukça yakın performans sergilemişlerdir. Sonuç olarak önerilen Lanczos2-CNN modeli uyuşmayan koşulda da olduğu gibi 0 dB de yine en yüksek performansı gösterirken Lanczos3-CNN modeli ortalamada %99,05 ile oldukça iyi bir sınıflandırma performansı sergilemiştir.

Çizelge 3.8. Farklı yeniden boyutlandırma yöntemleri kullanılarak yapılan çoklu durum testleri (standart sapma ile birlikte).

Sistem	Gürültüsüz	20 dB	10 dB	0 dB	Ortalama
En Yakın Komşu-CNN	99,1(0,1)	99,1(0,1)	98,9(0,2)	97,2(0,9)	98,6(0,2)
Bilineer-CNN	98,9(0,2)	98,9(0,3)	98,8(0,2)	97,3(0,4)	98,5(0,3)
Bikübik-CNN	99,1(0,1)	99,1(0,3)	98,9 (0,2)	97,0(0,2)	98,5(0,2)
Lanczos2-CNN	99,1(0,1)	99,1(0,2)	98,9(0,3)	97,4(0,6)	98,6(0,3)
Lanczos3-CNN	99,1(0,2)	99,1(0,2)	99,0(0,2)	97,2(0,7)	98,6(0,3)

Tez çalışmasının bu bölümünde standart bir ses olayı sınıflandırma görevi üzerinde önerilen darbesel ses olayı tanıma yaklaşımı test edilmiştir. Önerilen yöntem ile uyuşmayan durum testlerinde yani sadece gürültüsüz ses verileri ile eğitilerek yapılan testlerde literatürdeki çalışmalara göre %63,4'lük bağıl hata azaltım oranına denk %4,5'lik bir performans artışı sağlayarak %97,4'lük bir sınıflandırma başarımı elde edilmiştir. Ayrıca eğitim setindeki verilerin bir kısmına 10 dB SNR oranında gürültü ekleyerek yapılan testlerde başarımları %98,63'lük bir değere ulaşmıştır. Ayrıca SIF'ler kullanılarak yapılmış olan daha önceki çalışmalarda rapor edilen gürültüsüz ortamdaki güncel tekniklere göre düşük sınıflandırma performansı problemi de tez çalışmasında önerilen yöntemde yaşanmamıştır ve gürültüsüz ortamdaki sınıflandırma başarımı güncel teknikleri yakalamıştır. Bununla birlikte tüm gürültü oranlarında bu veri seti ile yapılmış çalışmalara çok daha iyi bir sınıflandırma performansı elde edilmiştir.

Önerilen Lanczos2-CNN modeli ile diğer yöntemlere göre daha başarılı sonuçlar elde edilmesi nedeniyle silah sesine ait akustik olayların sınıflandırılması görevinde bu yöntemin kullanılması kararlaştırılmıştır.

BÖLÜM 4

SİLAH SESİNE AİT AKUSTİK BİLEŞENLERİN TESPİT EDİLMESİ

Silah sesine ait 2 temel bileşen olan ŞD ve NP son derece darbesel ses sinyalleridir. ŞD çok hızlı bir şekilde yükselen ve düşen kenarlara sahiptir. Mermi türüne göre değişmekle birlikte tipik olarak 200-300 µs kadar sürmektedir. Ayrıca ŞD'nin genlik değeri genellikle NP'nin üzerindedir. Literatürdeki birçok yaklaşımda mikrofonlar tarafından algılanan sinyalin genlik değerleri kontrol edilerek belirli bir eşik değerin üzerindeki sinyallerin silah sesine ait olup olmadıkları sinyalin süresi ve sıfır geçişli kodlama gibi teknikler kullanılarak değerlendirilmektedir. ŞD dalgasının kendine has bir şeklinin olması ve doğal olaylar tarafından taklit edilememesi gibi sebeplere bağlı literatürdeki çalışmalarda şok dalgalarının algılanmasında genellikle problem yaşanmamaktadır. Bununla birlikte ŞD'nin mikrofon dizileri tarafından algılanabilmesi için merminin mikrofona 50 m'nin altındaki bir mesafeden geçmesi gerekmektedir. Ayrıca tek bir sensör dizini ile yapılan atıcı konum tespit çalışmalarında sadece ŞD'nin algılanması durumunda atışın yönü ve konumu hesaplanamamaktadır. Tam tersi durumda sadece NP'nin algılanması halinde tek bir sensör dizini ile atıcı konumu hesaplanamasa da atışın yapıldığı doğrultu tespit edilebilmektedir. Ancak literatürdeki çalışmalar pek çok durumda NP'nin ŞD yankıları ile karıştırıldığını veya mesafenin artması ile birlikte SNR değerinin düşmesine bağlı olarak hiç tespit edilemediğini ortaya koymaktadır.

Atıcı konumunun tek sensör dizisi ile de başarılı bir şekilde tespit edilebilmesi için silaha ait akustik olayların yüksek gürültü oranları altında başarılı bir şekilde tespit edilmesi kritik öneme sahiptir. Ayrıca silah türünün tespit edilmesi ile Eşitlik 2.5'te ki mermiye ilişkin kalibre ve mermi boyu gibi parametrelerde belirlenmiş olacaktır. Bu durumda daha az sayıda bilinmeyen değişken bulunmaya çalışılacağı için konum tespiti doğruluğunu da olumlu etkileyecektir. Bu anlamda [92,93]'de yapılan çalışmalarda silaha ait ses kayıtları kullanılarak silah sınıflandırılması yapılmıştır.

Fakat bu çalışmalarda NP ve ŞD dalgası gibi akustik olayların tespit edilmesine ve sınıflandırılmasına odaklanılmamıştır. Buna karşılık [25]'de yapılan çalışmada NP, ŞD ve yankıların bulunduğu bir veri seti hazırlanmış ve öncelikli olarak ŞD'ler silah kalibresine göre 3 farklı kategoride sınıflandırılmıştır. Ayrıca NP sınıflandırılması yapılırken de 3 farklı kategori kullanılmıştır. Bunlardan iki tanesi silah türüne ilişki sınıflandırmadır. Bir diğer grup ise yankılardır. Önerilen yöntemde elde edilen ses pencereleri öncelikli olarak her bir pencerenin tepe gücüne göre değerlendirilmektedir. Böylelikle akustik olayın ŞD mi yoksa NP olabileceği değerlendirilmektedir. Sonrasında dalgacık özellikleri kullanılarak silah sesi sınıflandırılmaktadır. Toplamda 120 ŞD, 80 NP ve 140 yankıdan oluşan veri seti üzerinde sınıflandırma performansı değerlendirilmiştir. Bununla birlikte yapılan çalışmada yüksek gürültünün sınıflandırma performansını nasıl etkileyebileceği üzerine bir çalışma sunulmamıştır.

Tez çalışmasının bu bölümünde darbesel ses olayı tanımada kullanmış olduğumuz yaklaşım ile silah sesine ait akustik bileşenlerin sınıflandırılması yapılmaktadır. Sınıflandırma performansını kıyaslamak için 5 farklı silah türüne ait akustik bileşenler kullanılmaktadır. Darbesel ses olayı sınıflandırılmasında yapılan çalışmalara benzer şekilde sistemin yüksek gürültü altındaki performansı da değerlendirilmektedir.

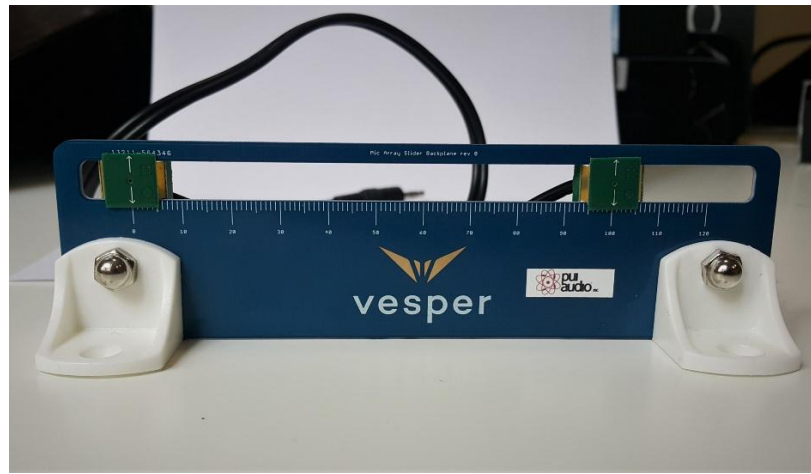
Silah sesine ait akustik olayların değerlendirilmesi aşamasında iki farklı tipte yaklaşım belirlenmiştir. Bunlardan birincisi mermi parametrelerinin ve silah türünün doğru bir şekilde sınıflandırılabilirliğini kontrol etmektir. Bu aşamada sistem performansı mermi ve silah türünü de dikkate alarak 10 farklı grup altında sınıflandırılmaktadır. İkinci bir yaklaşımda ise elde edilen sonuçlar silah ve mermi türünü dikkate almadan sadece ŞD, NP ve diğer olmak üzere üç farklı grup altında sınıflandırılmıştır. Böylelikle hem sistemin silah sesine ait akustik olayları diğer darbesel seslerden ve yankılardan ayırt edebilme kabiliyeti test edilmekte hem de her bir mermi ve silah türünü belirleyebilme kabiliyeti sınanmaktadır.

Böyle bir yaklaşım kullanılmasının temel sebebi özellikle sadece NP'nin algılandığı durumlarda sistemin silah türüne ilişkin bir parametreye ihtiyaç duymadan atışın yapıldığı doğrultuyu tespit edebilmesidir. Bu safhada ses olayının NP olduğunun tespit edilmesi yeterli bir veridir. Öte yandan ŞD'nin de tespit edilmesi durumunda Eşitlik

2.5'teki kalibre ve mermi boyu parametrelerin tespit edilebilmesi açısından mermi türünün sınıflandırılması önemli bir gerekliliktir. Ayrıca NP'lerin sınıflandırılması da özellikle birden fazla silah türüyle atış yapıldığı durumlarda ŞD'den sonra gelen doğru NP türünün belirlenmesi açısından son derece önemlidir. Böylelikle ŞD ve NP sınıflandırmalarının birbiri ile uyumu kontrol edilerek çapraz bir denetim yapılmasına olanak sağlamaktadır.

4.1. VERİ SETİ

Silah seslerinin elde edilmesi için Zoom H6 ses kayıt cihazı ve iki adet Vesper VM1000 mems mikrofon kullanılmıştır. Şekil 4.1'de silah atışlarının kayıt edildiği düzenek görülmektedir. Şekil 4.1'den de anlaşılacağı üzere iki mikrofon ölçekli bir kaydırıcı üzerine yerleştirilmiştir ve her iki mikrofon arasındaki mesafe 10 cm olarak ayarlanmıştır. Sonrasında sistem Şekil 4.3'de de görüldüğü gibi bir üç ayak sehpa üzerine monte edilerek atış kayıtları gerçekleştirilmiştir. Mikrofonlar atış doğrultusunun yan tarafına 1 ile 20 m arasında değişen mesafelerde yerleştirilerek kayıtlar yapılmıştır. Ayrıca mikrofonlar ile atışın yapıldığı konum arasındaki mesafe değerleri de 50-100 m arasında değiştirilerek yapılmıştır. Her silahla hem seri şekilde hem de tekli atışlar gerçekleştirilmiştir. Atışlar benzer iklim koşullarında farklı günlerde yapılarak kaydedilmiştir. Bunlara ilave olarak tüm atışlarda mikrofonlar yerden 1,5 m yükseğe yerleştirilmiştir ve atıcı tüm atışlarını ayakta gerçekleştirmiştir.



Şekil 4.1. Kaydırıcı üzerindeki iki adet Vesper VM1000 mems mikrofon.



Şekil 4.2. Zoom H6 ses kayıt cihazının ayarlanması.



(a)



(b)

Şekil 4.3. a) ve b) Kayıt ekipmanının farklı açılardan görünümü.

HK33, M4, AK47, G3 ve Sig Sauer 516 olmak üzere toplamda 2 farklı kalibrede 5 farklı silah ile atış gerçekleştirilmiştir. HK33, M4 ve Sig Sauer 516 ile yapılan atışlarda 5,56x45 mm standart NATO mühimmatı kullanılmıştır. G3 ile yapılan atışlarda 7,62x51 mm standart NATO mühimmatı kullanılırken AK47 ile yapılan atışlarda 7,62x39 mm mühimmat kullanılmıştır. Her bir silah ile 20 adet atış gerçekleştirilmiş ve atışlar 16 bit genişliğinde 96 KHz örnekleme hızında kayıt edilmiştir.

Atış konum tespit sistemlerinde mikrofonlardan elde edilen sinyallerin örnekleme hızı da oldukça önemli bir faktördür. 96 KHz'lik örnekleme hızında örnekleme periyodu 10,42 μ s'dir. Bu durum her bir akustik olayın tespiti için 0-10,42 μ s'lik düzenli bir hata olabileceği anlamına gelmektedir. [33]'de yapılan çalışmada örnekleme hızının konum tespiti üzerinde yapabileceği hata simüle edilmiş ve 51,2 KHz'lik bir örnekleme hızı kullanılarak elde edilen 19,61 μ s'lik bir örnekleme periyodu için tek bir algılayıcı dizinde yanca hata 13,76 m olarak tespit edilmiştir.

Veri seti toplamda:

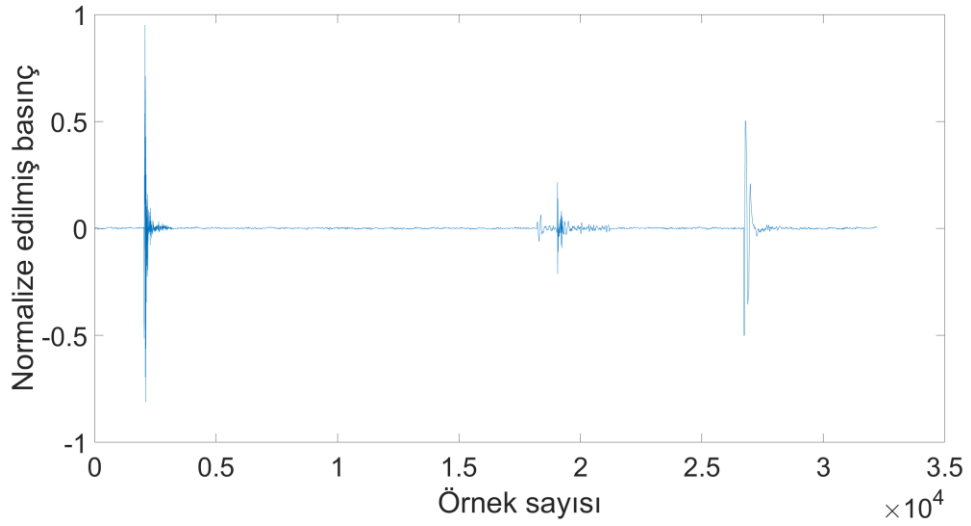
- 200 şok dalgası,
- 200 namlu patlaması,
- 260 yankıdan oluşmaktadır.

Ayrıca cam kırılması ve balon patlaması gibi 200 farklı sestem oluşan bir grupta veri setine dahil edilmiştir. 10 farklı grupta sınıflandırma işlemi yapılmaktadır ve gruplar şu şekildedir:

- Şok dalgası 7,62x51, 40 adet
- Şok dalgası 7,62x39, 40 adet
- Şok dalgası 5,56x45, 120 adet
- Namlu patlaması HK33, 40 adet
- Namlu patlaması Sig Sauer 516, 40 adet
- Namlu patlaması AK47, 40 adet
- Namlu patlaması G3, 40 adet
- Namlu patlaması M4, 40 adet
- Yankı, 260 adet

- Diğer darbesel ses, 200 adet

Şekil 4.4’de M4 silahına ait örnek bir atış kaydı görülmektedir. Soldan sağa doğru ilk sinyal mermiminin oluşturduğu ŞD’dir. ŞD’yi takip eden sinyal ise ŞD’ye ait yankıdır ve son olarak NP sinyali görülmektedir. Silah sesine ait her bir akustik olay başlangıç noktasından itibaren 6 ms’lik parça alınarak el ile manuel olarak etiketlenmektedir. Veri seti her biri 6 ms’lik toplam 860 kayıt içermektedir. Bunların 688 tanesi eğitim setinde ve kalan 172 tanesi de test setinde kullanılmıştır.



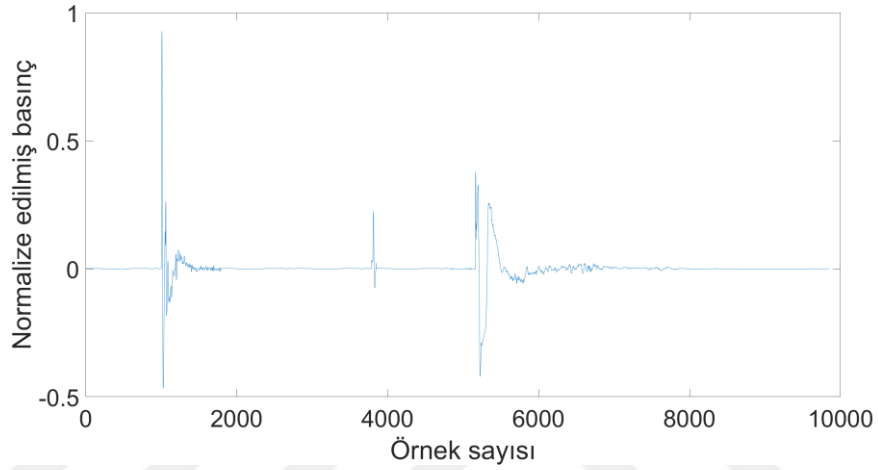
Şekil 4.4. M4’ e ait ses kaydının genlik-zaman grafiği.

4.2. DENEYSEL KURULUM

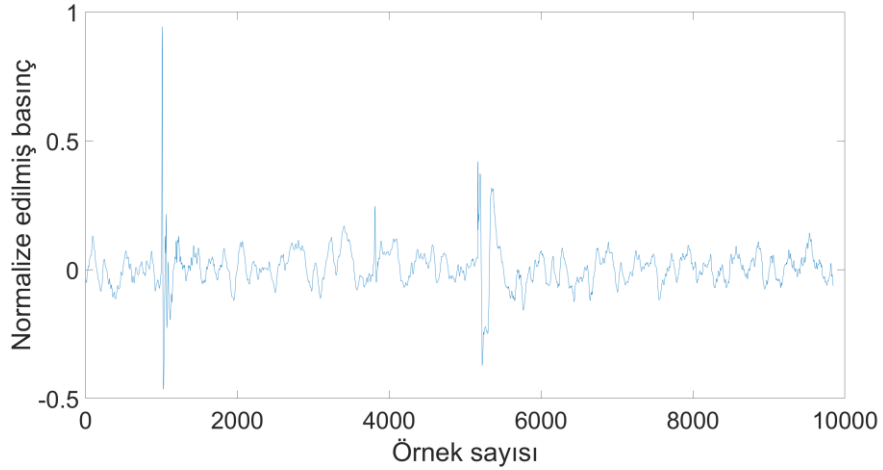
Darbesel ses tespitinde kullanılan RWCP veri seti farklı uzunluklarda ses kayıtları içermektedir ve bunun neticesi olarak elde edilen spektrogram görüntülerin boyutları birbirinden farklıdır. Bununla birlikte CNN’ler giriş katmanında sabit boyutlarda resimlere ihtiyaç duyarlar. Bu tez çalışmasında önerilen darbesel ses tespiti yaklaşımında spektrogram resimler yeniden boyutlandırma yöntemi ile sabit bir boyuta ayarlanarak sınıflandırma işlemi yapılmıştır.

Bunun aksine silah sesine ait akustik bileşenlerin tespit edilmesi aşamasında her biri 6 ms uzunluğunda olan sabit uzunluklu bir veri seti kullanılmaktadır. Bunun nedeni ŞD’ler genellikle 500 μ s aşmayan kısa sinyallerdir. Öte yandan NP’lerin süresi de

genellikle 5 ms'den daha kısadır ve bu süreden daha uzun olan kısımları genellikle atışın yapıldığı ortamdan kaynaklanan yansımalarıdır. Ayrıca algılayıcı dizin üzerindeki her bir mikrofon arasındaki ses sinyalinin alabileceği en yüksek değerin üzerindeki bir yayılma gecikmesi de dahil edilerek toplam kayıt uzunluğu belirlenmiştir. Bu doğrultuda her bir kaydın süresini 6 ms'lik sabit bir değere ayarlayarak spektrogram resimler elde edilmektedir.



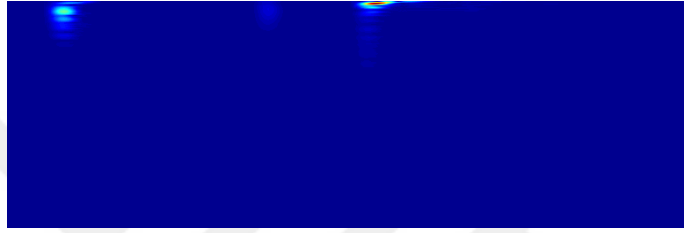
Şekil 4.5. AK47'ye ait ses kaydının genlik-zaman grafiği.



Şekil 4.6. AK47'ye ait ses kaydına 0 dB SNR değerinde insan konuşmalarına ait gürültü eklenmiş halinin genlik zaman grafiği.

Şekil 4.5'de bir AK47'ye ait ses kaydının genlik-zaman grafiği görülmektedir. Öte yandan Şekil 4.6'de aynı ses kaydına 0 dB SNR değerinde insan konuşmalarına ait gürültü eklenmiş hali görülmektedir. Şekil 4.6'den anlaşılacağı üzere gürültü genlik-

zaman grafiğinin tamamında bir bozulmaya neden olmaktadır. Bu nedenle genlik-zaman alanındaki ses özelliklerini kullanan yaklaşımlar yüksek gürültü altında genellikle düşük bir sınıflandırma performansı göstermektedirler. Şekil 4.7 ve 4.8’de ise aynı ses kaydının sırasıyla gürültüsüz koşullar altındaki spektrogram görüntüsü ve 0 dB SNR altındaki spektrom görüntüleri görülmektedir. Gürültü Şekil 4.7’de spektrogramın alt bandında bozulmaya neden olmuşken üst bantlarında herhangi bir bozulma olmamıştır. Bu durumda spektrogramın gürültü altındaki dayanımını büyük ölçüde artırmaktadır.



Şekil 4.7. AK47’ye ait silah sesinin gürültüsüz halinin spektrogram görüntüsü.



Şekil 4.8. AK47’ye ait ses kaydına 0 dB SNR değerinde insan konuşmalarına ait gürültü eklenmiş haline ait spektrogram görüntüsü.

Tez çalışmasının bu kısmında spektrogramlar 96 kHz örnekleme hızında 48 nokta hamming çerçevesi kullanılarak, 36 nokta ile yüksek derecede örtüşecek şekilde elde edilmiş ve yukarıdaki bölümde de belirtildiği gibi [59]’daki çalışmada tarif edilen şekilde lineer ölçekli spektrogram görüntülere dönüştürülmüştür. Her bir ses kaydından 25x45 boyutlarında spektrogram resimler elde edilmiştir. Spektrogram resimlerin elde edilmesinde Bölüm 3’deki çalışmaya benzer şekilde “summer” renk haritası kullanılmıştır. Ayrıca elde edilen her bir spektrogram resim CNN işlemlerinin hesaplama maliyetini azaltmak adına 25x45 boyutlarından 25x14’lük ebatlara yeniden boyutlandırılarak kullanılmıştır. Yeniden boyutlandırma işleminde Lanczos yeniden boyutlandırma metodu ($a = 2$ olmak üzere) kullanılmıştır.

4.3. CNN MİMARİSİ

Bu bölümde darbesel ses sınıflandırılmasında kullanılan mimariden farklı bir CNN mimarisi kullanılmıştır. Bunun başlıca sebepleri diğer mimariye göre daha az sayıda sınıflandırılma yapılacak grup bulunması ve eğitim için kullanılan veri miktarının az olmasıdır. Bir diğer önemli sebepte burada kullanılan mimarinin FPGA uyarlanacak olması nedeniyle diğer mimariye göre daha düşük miktarda donanım kaynağı kullanacak bir mimarinin tercih edilmek istenmesidir.

CNN'ler sağlamlıkları ve yüksek başarı oranları nedeniyle son yıllarda çok yaygın bir şekilde kullanılmaya başlanmıştır. Bu sağlamlık ve yüksek başarı oranları beraberinde yüksek hesaplama maliyetlerini de getirmiştir. Bazı CNN modellerinde tek bir çerçevenin sınıflandırma için bile 38 GOP işleme ihtiyaç duyulabilmektedir. Bu duruma bağlı olarak ihtiyaç duyulan donanım kaynağı da hatırı sayılır miktarda artmaktadır. FPGA kaynaklarını daha verimli kullanabilmek adına sınıflandırma performansından en az düzeyde kayıp verilecek şekilde daha sığ bir mimari kullanılmıştır.

CNN'ler başarılarının yanında çok yüksek hesaplama maliyeti getirirler. Bu hesaplama maliyetinin üstesinden gelebilmek adına genellikle yüksek hesaplama gücüne sahip GPU'lar kullanılmaktadır. Bununla birlikte FPGA'ların bağımsız donanım yapısını, CNN'lerin eşzamanlılık kabiliyetleri ile birleştirmek isteyen yaklaşımlar ortaya çıkmaktadır. Bunun yanı sıra daha düşük enerji sarfiyatının sağlanması kullanıcıların ilgisini çekmektedir. Bu doğrultuda CNN'lerin ileri yayılım aşamalarının FPGA üzerinde kullanılması işlemi literatürde yeni çalışılmaya başlanılan bir konudur.

İleri yayılım aşamasının kullanılması için farklı yaklaşımlar mevcuttur. Bu çalışmada doğrudan donanıma eşleme yaklaşımı benimsenmiştir.

Doğrudan donanıma eşleme yaklaşımı hız anlamında çok büyük bir fayda sağlarken, FPGA kaynaklarının sınırlı olması nedeniyle karmaşık mimarilerin uyarlanması için pek uygun bir yaklaşım değildir.

Tez çalışmasının bu kısmında silah sesine ait akustik olayların sınıflandırılması için 2'si konvolüsyon ve 2'si de havuzlama katmanı olmak üzere 4 katmandan oluşan bir CNN mimarisi kullanılmıştır. Herbir konvolüsyon katmanını bir havuzlama katmanı takip etmektedir. Havuzlama katmanında maksimum havuzlama yöntemi tercih edilmiştir. Bu katmanlar sırasıyla konvolüsyon, havuzlama, konvolüsyon ve havuzlama katmanı şeklindedir. Daha basit bir mimari kullanılarak FPGA üzerine uyarlanması aşamasında daha etkin bir donanım kaynağı kullanımının sağlanması amaçlanmaktadır.

4.4. SONUÇLAR VE TARTIŞMA

Bu bölümde öncelikli olarak hazırlanan ver seti darbesel seslerin tespit edilmesinde yapılan çalışmaya benzer şekilde orijinal kayıtların üzerine farklı SNR değerlerinde gürültüler eklenerek test edilmektedir. Böylelikle yüksek gürültü değerleri altındaki sınıflandırma performansı ve orijinal kayıtlara ait sınıflandırma performansı kıyaslanmaktadır. Sonrasında çoklu durum testleri ile gürültü eklenmiş ve orijinal ses kayıtları bir arada kullanılarak CNN eğitilmiş ve farklı gürültü koşulları altında testler gerçekleştirilmiştir. Tüm testlerde 5 parça çapraz doğrulama işlemi uygulanarak ortalama değerleri verilmiştir.

Bunlara ilave olarak önerilen yöntemin sonuçları 10 sınıf altında değerlendirilmesinin yanı sıra ŞD, NP ve diğer sesler olmak üzere de bir değerlendirme sunularak silaha ait akustik olayları diğer olaylardan ayırt etme performansı da kıyaslanmıştır.

Son aşamada her iki eğitim yönteminin nihai sonuçları değerlendirilerek atıcı konum tespit sisteminde kullanılacak yönteme karar verilmiştir.

Çizelge 4.1'de orijinal ses kayıtları ile, ilave gürültü eklenilmeden CNN ağının eğitilmesi ile elde sonuçlar görülmektedir. Gürültüsüz koşullarda yapılan testlerde ortalama %98,8'lik bir sınıflandırma performansı elde edilirken, test setine 20 dB gürültü eklenerek yapılan testlerde gürültüsüz koşullara aynı sınıflandırma başarımları elde edilmiştir. Bununla birlikte 10 dB ve 0 dB gürültü altında sınıflandırma performansı gürültüsüz koşullara göre sırasıyla %0,9 ve %11,1 olmak üzere daha

düşük bir başarıml göstermişlerdir. Tüm testlerin neticesinde ortalama %95,8'lik bir başarıml elde edilmiştir.

Bunlara ilave olarak G3 silahına ait NP ve 5,56x45 boyutlarındaki merminin oluşturduğu ŞD tüm koşullarda doğru bir biçimde sınıflandırılarak en yüksek başarımlı elde etmişlerdir. Bunun tam aksine HK33 ve M4'e ait NP seslerinde bir miktar daha düşük sınıflandırma performansı elde edilmiştir. Bu iki silah için sırasıyla ortalama %89,4 ve %91,3'lük bir sınıflandırma başarımlı elde edilmiştir.

Çizelge 4.1. Farklı SNR değerleri altında silah sesine ait akustik olayların ve diğer ses olaylarının sınıflandırma başarımlı.

	Gürültüsüz	20 dB	10 dB	0 dB	Ortalama
7,62x51	100,0	100,0	100,0	92,5	98,1
7,62x39	100,0	100,0	100,0	90,0	97,5
5,56x45	100,0	100,0	100,0	100,0	100,0
HK33	95,0	95,0	87,5	80,0	89,4
M4	95,0	95,0	90,0	85,0	91,3
SİG516	100,0	100,0	100,0	90,0	97,5
G3	100,0	100,0	100,0	100,0	100,0
AK47	100,0	100,0	100,0	97,5	99,4
Yankı	98,8	98,8	98,1	92,3	97,0
Diğer	98,5	98,5	98,0	70,0	91,3
Ortalama	98,8	98,8	97,9	87,7	95,8

Çizelge 4.2'de ise silah sesine ait akustik bileşenlerin önerilen yöntem ile tespit edilebilirliği kontrol edilmiştir. Çizelge 4.1 ile sonuçlar bir arada değerlendirildiğinde HK33 ve M4 silahlarına ait NP seslerinin gürültüsüz ve 20 dB gürültü altında yankı ve diğer sesler ile karıştırılmadığı görülmektedir. Bununla birlikte bu iki sesin tüm koşullarda birbiri ile karıştığı görülmektedir. Ayrıca önerilen yöntem ile ŞD tüm test koşullarında başarılı bir şekilde ayırt edilirken 0 dB gürültü altında 7,62x51 mm ve 7,62x39 mm mühimmatlarına ait ŞD'lerin birbiri ile karıştırılarak yanlış sınıflandırıldığı görülmektedir. 5,56 mm ile 7,62 mm çaplarındaki mühimmatların ise birbirinden başarılı bir şekilde ayırt edilebildikleri görülmektedir. Çizelge 4.2 NP'ler açısından değerlendirilecek olursa gürültüsüz ve 20 dB gürültü koşullarında NP tespitinde sorun yaşanmazken, 10 dB ve 0 dB gürültü altında NP'lerin zaman zaman

diğer sesler ile karıştırılmakla birlikte tüm koşullarda tatmin edici bir sınıflandırma başarımı elde edilmektedir.

Çizelge 4.2. Silah sesine ait akustik olayların tespit edilebilme başarımı.

	Gürültüsüz	20 dB	10 dB	0 dB	Ortalama
ŞD	100,0	100,0	100,0	100,0	100,0
NP	100,0	100,0	97,5	95,0	98,1
Diğer	100,0	100,0	99,3	98,9	99,6

Ayrıca uyuşmayan durum testlerinde gürültüsüz ve 20 dB gürültü koşullarında hiçbir hatalı alarm durumu yaşanmadı. Yani NP ve ŞD'lerin hiçbiri diğer grubunda algılanmadığı gibi diğer grubundaki hiçbir ses olayı da silah sesine ait olarak sınıflandırılmadı. Bununla birlikte 10 dB SNR değerleri altında 200 NP kaydından 5 tanesi algılanamadı ve hatalı olarak diğer grubunda değerlendirildi. 0 dB SNR değerinde ise 10 adet NP kaydı algılanamazken silah sesine ait akustik olay olmayan 5 ses kaydı NP olarak değerlendirildi. Öte yandan ŞD tüm durumlarda hatasız olarak tespit edilebildi.

Çizelge 4.3'te çoklu durum testlerine ilişkin sonuçlar görülmektedir. Çoklu durum testlerinde eğitim setindeki ses kayıtlarının %20'sine 0 dB SNR seviyesinde gürültü eklenerek, hem orijinal ses kayıtları ile hem de gürültü eklenmiş ses kayıtları ile ağırlı eğitimi gerçekleştirilmiştir. Çoklu durum testlerinde uyuşmayan durum testlerine göre ortalama sınıflandırma performansı %0,7'lik bir artış göstererek %96,5 oranında gerçekleşmiştir. Gürültüsüz koşul testlerinde uyuşmayan durum testleri ile aynı sınıflandırma başarımı elde edilirken diğer tüm koşullar için çoklu durum testlerinde daha iyi bir skor elde edilmiştir. 5,56x45 mm mühimmata ait ŞD'ler ve G3 silahına ait NP tüm durumlarda hatasız olarak tespit edilmiştir. Ayrıca sonuçlar silah sesine ait akustik olayların sınıflandırılması açısından değerlendirilecek olursa genel sınıflandırma performansının bir miktar daha üzerinde %97,7'lik bir sonuç elde edilmektedir. Bu durumda diğer grubundaki sesler ile yankıların sınıflandırılması esnasında birbiri ile karıştığı anlamına gelmektedir ki bu olay silah sesine ait akustik olayların tespit edilmesine engel olan bir durum değildir.

Çizelge 4.3. Çoklu durum testlerinde farklı SNR değerleri altında silah sesine ait akustik olayların ve diğer ses olaylarının sınıflandırma başarımı.

	Gürültüsüz	20 dB	10 dB	0 dB	Ortalama
7,62x51	100,0	100,0	100,0	92,5	98,125
7,62x39	100,0	100,0	100,0	90,0	98,75
5,56x45	100,0	100,0	100,0	100,0	100
HK33	95,0	95,0	90,0	85,0	91,3
M4	95,0	95,0	90,0	85,0	91,9
SİG516	100,0	100,0	100,0	90,0	97,5
G3	100,0	100,0	100,0	100,0	100
AK47	100,0	100,0	100,0	97,5	99,4
Yankı	98,1	98,8	98,5	95,4	97,9
Diğer	98,0	98,5	98,0	75,0	92,5
Ortalama	98,8	99,0	98,1	90,2	96,5

Son olarak uyuşmayan koşullardaki testlerde olduğu gibi elde edilen sonuçlar akustik olayların ayırt edilebilirliği açısından kontrol edilmiştir. Çizelge 4.4'te de görüldüğü gibi çoklu durum testlerinde de ŞD'ler her durumda başarılı bir şekilde ayırt edilmiştir. Bununla birlikte NP'lerin gürültüsüz ve 20 dB gürültü koşullarındaki ayırt edilebilirliklerinde herhangi bir problem yaşanmamıştır. 10 dB'deki tespit performansı uyuşmayan durum testleri ile aynı sonuç elde edilirken 0 dB gürültü altında sınıflandırma performansında %1,5'lik bir atış yaşanmıştır. Benzer şekilde NP'nin ortalama tespit performansı da %98,1'den %98,5'e yükselerek artış kaydetmiştir.

Çizelge 4.4. Çoklu durum testlerinde silah sesine ait akustik olayların tespit edilebilme başarımı.

	Gürültüsüz	20 dB	10 dB	0 dB	Ortalama
ŞD	100,0	100,0	100,0	100,0	100,0
NP	100	100,0	97,5	96,5	98,5
Diğer	100,0	100,0	99,3	99,3	99,6

Bunlara ilave olarak uyuşmayan durum testlerine benzer şekilde gürültüsüz ve 20 dB gürültü koşulları altında herhangi bir hatalı alarm durumu yaşanmamıştır. 10 dB gürültü altında ise 5 adet NP kaydı silah sesine ait akustik olay olarak tespit edilememiştir. Ayrıca yine aynı gürültü değerinde diğer grubunda bulunan 3 adet atış

kaydı hatalı bir biçimde NP olarak değerlendirilmiştir. 0 dB SNR değerinde ise 7 NP kaydı tespit edilemezken aynı şekilde 3 adet diğer grubundaki kayıt NP olarak algılanmıştır.

Literatürdeki birçok çalışmada [8,9,13,14,16,26,27] NP'lerin özellikle yüksek gürültü oranları altında tutarlı bir biçimde algılanamadığını ortaya koymaktadır. Bu durumda özellikle tek algılayıcı dizin kullanılarak yapılan çalışmaları son derece olumsuz bir şekilde etkilemektedir. Çizelge 5.1'de de belirtildiği gibi tek başına ŞD'nin algılanması durumunda atıcının konumu ve atışın yapıldığı doğrultu hesaplanamaz. Öte yanda NP'lerin tek başına tespit edilmesi durumunda atışın yapıldığı doğrultu hesaplanabilmektedir. ŞD ile birlikte tespit edilmesi durumunda ise tek algılayıcı dizin ile bile atışın yapıldığı konum hesaplanabilmektedir.

Bu tez çalışmasında elde edilen sonuçlar değerlendirildiğinde NP'lerin ortalama %98,5'lik bir tespit oranı ile son derece yüksek oranda algılanabildiğini ortaya koymaktadır. Bu durum özellikle yerel durum farkındalığının artırılmasını hedefleyen sistemler için kritik bir öneme sahiptir. Ayrıca çoklu durum testlerinde elde edilen sonuçlara göre 5 farklı silah farklı gürültü koşullarında yapılan testlerde %96'lık bir sınıflandırma başarımı ile tespit edilmiştir. Silahın türü her ne kadar NP'nin varış doğrultusunun hesaplanması açısından gerekli bir parametre olmasa da birden fazla silah türü ile yapılan atışlarda ŞD'den sonra doğru NP türü ile işlem yapılabilmesi açısından önemlidir. Yani ŞD sınıflandırılması ve NP arasında bir kontrol yaparak gelen sinyalin başka bir silahtan mı yoksa aynı silahtan mı geldiği tespit edilebilmektedir. Bu durumda atış konum tespit sistemleri açısından önemli bir kabiliyet kazanımıdır. Ayrıca NP'lerin yüksek oranda tespit edilebilirliklerinin de sağlanması tek algılayıcı dizinden oluşan maliyet etkin çözümlerin önünü açmaktadır.

Öte yandan bu tez çalışmasında elde edilen sonuçlar ŞD'ler açısından değerlendirilecek olursa tespit edilebilirlik bakımından tüm koşullarda hiçbir problem yaşanmadan başarılı bir şekilde algılanmıştır. ŞD'lerin sınıflandırılması tarafından bakıldığında ise tüm gürültü koşullarında kalibre başarılı bir şekilde tespit edilmiştir. Bu başarıım ŞD modeli üzerindeki eksik parametrelerin belirlenebilirliği açısından son

derece önemlidir. Mermi boyu ile ilgili olarak ise sadece 0 dB gürültü altında bir miktar hata yaşanmıştır.

Sonuç olarak elde edilen değerler atış konum tespit sistemleri açısından son derece ümit vericidir. Sınıflandırma yöntemi ile ŞD'si modeline ilişkin parametreler yüksek doğruluk oranlarında tespit edilebilmektedir. Ayrıca NP'lerinde yüksek gürültü altında başarılı bir şekilde tespit edilebildiği ortaya konmuştur. Ayrıca sınıflandırılma bakımından ŞD'lerin bir miktar altında olmakla birlikte oldukça iyi bir performans ortaya koymuştur.



BÖLÜM 5

ATICI KONUMUNUN BULUNMASI

5.1. GENELLEŞTİRİLMİŞ ÇAPRAZ KORELASYON

Atıcı konumunun hesaplanabilmesi için öncelikli olarak tespit edilen bir akustik olaya ilişkin mikrofonların algılama zamanları arasındaki farkın bulunması gerekmektedir. Bu tez çalışmasında iki mikrofon arasındaki zaman farkının bulunabilmesi için literatürde çok yaygın olarak kullanılan genelleştirilmiş çapraz korelasyon (GÇK) algoritması [94] kullanılmıştır.

k ve l mikrofonlarına ait sinyallerin sırasıyla $x_k(t)$, $x_l(t)$ şu şekilde olduğu varsayılmak üzere:

$$x_k = s(t) * h(d_s, t) + n_k(t) \quad (5.1a)$$

$$x_l(t) = s(t - \tau_{kl}) * h(d_s, t) + n_l(t) \quad (5.1b)$$

İki mikrofon arasındaki varış zaman farkına ihtiyaç duyulduğu için basitlik açısından Eşitlik 5.1’de ses kaynağı ile k mikrofonu arasındaki yayılma gecikmesi eşitliğe dahil edilmemiş ve sadece k ve l mikrofonları arasındaki gecikme τ_{kl} dikkate alınmıştır [95].

Çapraz korelasyon fonksiyonu, herhangi bir zaman kaymasına karşılık iki sinyal arasındaki benzerliğin bir ölçüsünü verir ve ideal olarak TDAO karşılık gelen sürede bir tepe değeri göstermesi beklenir [96]. $x_k(t)$ ve $x_l(t)$ ’nin çapraz korelasyonu şu şekilde tanımlanır:

$$C_{kl}(\tau) = \int_{-\infty}^{\infty} x_k(k)x_l(t + \tau)dt \quad (5.2)$$

Bir çapraz spektrumda çapraz korelasyonun Fourier dönüşümü alındığında:

$$C_{kl}(\omega) = \int_{-\infty}^{\infty} c_{kl}(t) e^{j\omega\tau} d\tau \quad (5.3)$$

$X_l(\omega)$, $x_l(t)$ sinyalinin Fourier dönüşümü, “*” kompleks eşleniği (conjugate) belirtmek üzere Fourier dönüşümünün konvolüsyon özelliği uygulandığında aşağıdaki ifade elde edilir:

$$C_{kl}(\omega) = X_k(\omega)X_l^*(\omega) \quad (5.4)$$

Eşitlik 5.4’ün ters Fourier dönüşümü, mikrofon sinyallerinin Fourier Dönüşümü açısından çapraz korelasyon fonksiyonunu verir:

$$C_{kl}(\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} X_k(\omega) X_l^*(\omega) e^{j\omega\tau} d\omega \quad (5.5)$$

$x_k(t)$ ve $x_l(t)$ sinyallerinin GÇK’sı bu sinyallerin filtre edilmiş hallerinin çapraz korelasyonudur. $W_k(\omega)$ ve $W_l(\omega)$ bu filtrelerin Fourier dönüşümlerini belirtmek üzere, GÇK $R_{kl}(\tau)$ ’yu şu şekilde ifade etmek mümkündür:

$$R_{kl}(\tau) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} (W_k(\omega)X_k(\omega)) (W_l(\omega)X_l(\omega))^* e^{j\omega\tau} d\omega \quad (5.6)$$

Bu durumda birleştirilmiş ağırlık fonksiyonu $\Psi_{kl}(\omega)$ şu şekilde ifade edilebilir:

$$\Psi_{kl}(\omega) \equiv W_k(\omega)W_l^*(\omega) \quad (5.7)$$

5.7 eşitliği 5.6 eşitliğinde yerine konulursa, GÇK ifadesi şu şekle gelir:

$$R_{kl}(\tau) \equiv \int_{-\infty}^{\infty} \Psi_{kl}(\omega)X_k(\omega)X_l^*(\omega)e^{j\omega\tau} d\omega \quad (5.8)$$

k ve l mikrofonları arasındaki TDOA GÇK $R_{kl}(\tau)$ ’yu maksimize eden zaman gecikmesi değeridir ve bu durumda şu şekilde ifade edilebilir:

$$\hat{\tau}_{k,l} = \operatorname{argmax} R_{k,l}(\tau) \quad (5.9)$$

Öte yandan literatürdeki çalışmalar gürültü altında faz dönüşümü ağırlıklandırma fonksiyonun (PHAT) daha iyi sonuçlar verdiğini ortaya koymaktadır. PHAT şu şekilde ifade edilebilir:

$$\Psi_{kl}(\omega) \equiv \frac{1}{|X_k(\omega)X_l^*(\omega)|} \quad (5.10)$$

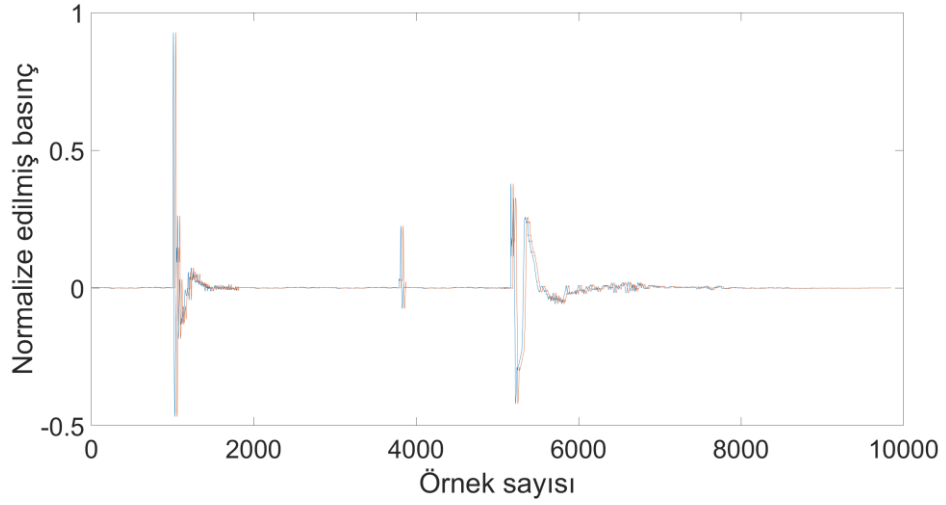
Eşitlik 5.10'daki ağırlıklandırma fonksiyonu Eşitlik 5.8'de yerine konulduğunda GCC-PHAT ifadesi şu şekilde elde edilir:

$$R_{k,l}(\tau) \equiv \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{|X(\omega)X(\omega)|} X(\omega)X^*(\omega) e^{j\omega\tau} d\omega \quad (5.11)$$

5.2. VARIŞ DOĞRULTUSU HESAPLANMASI

Atıcı konum tespit sistemlerinde doğru bir konum hesaplaması yapılabilmesi için öncelikli olarak silaha ait akustik olayların doğru bir biçimde ayırt edilmesi gereklidir. Yukarıdaki bölümlerde de belirtildiği gibi ŞD'nin kendine has bir şeklinin olması ve doğal fenomenler aracılığıyla taklit edilememesi nedeniyle literatürdeki çalışmalarda ŞD'nin tespit edilmesinde genellikle problem yaşanmamaktadır. Ayrıca [25]'deki çalışmada 50 m'nin üzerindeki mesafelerde ŞD'nin tepe gücünün NP'nin 13 dB kadar üzerinde olduğu gözlemlenmiştir. Öte yandan özellikle atıcı konumu ile sensör dizini arasındaki mesafenin arttığı durumlarda SNR değerine bağlı olarak NP çoğu zaman algılanamamaktadır. Ayrıca yeryüzü şekilleri ve engeller gibi faktörlere bağlı olarak pek çok durumda ŞD'ye ait yankılar NP olarak algılanmaktadır. Oysaki konum hesabı için akustik olaylar arasındaki zaman farkı kritik öneme sahip konulardan biridir. Buna ilave olarak tek bir sensör dizini ile atıcı konumunun hesaplanabilmesi için hep ŞD'nin hem de NP'nin tespit edilmesi gerekir. Aksi takdirde konum hesaplaması için birden fazla sensör dizisine ihtiyaç duyulmaktadır. Birden fazla sensör dizisi konum tespit doğruluğunu artırmakla birlikte özellikle arazide harekât halindeki birliklerin her zaman birden fazla sensör dizisi bulundurması operasyonel açıdan zor olabilir. Bu durumda yüksek gürültü altında akustik olayların doğru bir şekilde ayırt edilmesi önem kazanmaktadır. Çizelge 5.1'de de görüldüğü gibi tek bir sensör dizisi tarafından sadece ŞD'nin tespit edilmesi durumunda atıcının konumu ve yönü hesaplanamamaktadır.

Öte yandan tek bir sensör dizisinin sadece NP'yi algıladığı durumlarda atışın yapıldığı yön tespit edilebilirken mesafe bilgisi hesaplanamadığı için konum bilgisi elde edilememektedir. Diğer tüm durumlar için konum hesabı yapılabilmektedir. Bu tez çalışmasında yerel durum farkındalığının artırılmasına odaklanılmıştır ve bu doğrultuda mümkün mertebe tek bir sensör dizisi ile atıcı konumunun hesaplanması hedeflenmiştir. Bununla birlikte önerilen yöntem geniş alan savunması içinde uyarlanabilir şekilde tasarlanmıştır.



Şekil 5.1. İki farklı mikrofondan kaydedilen silah sesine ait akustik olaylar arasındaki zaman farkı.

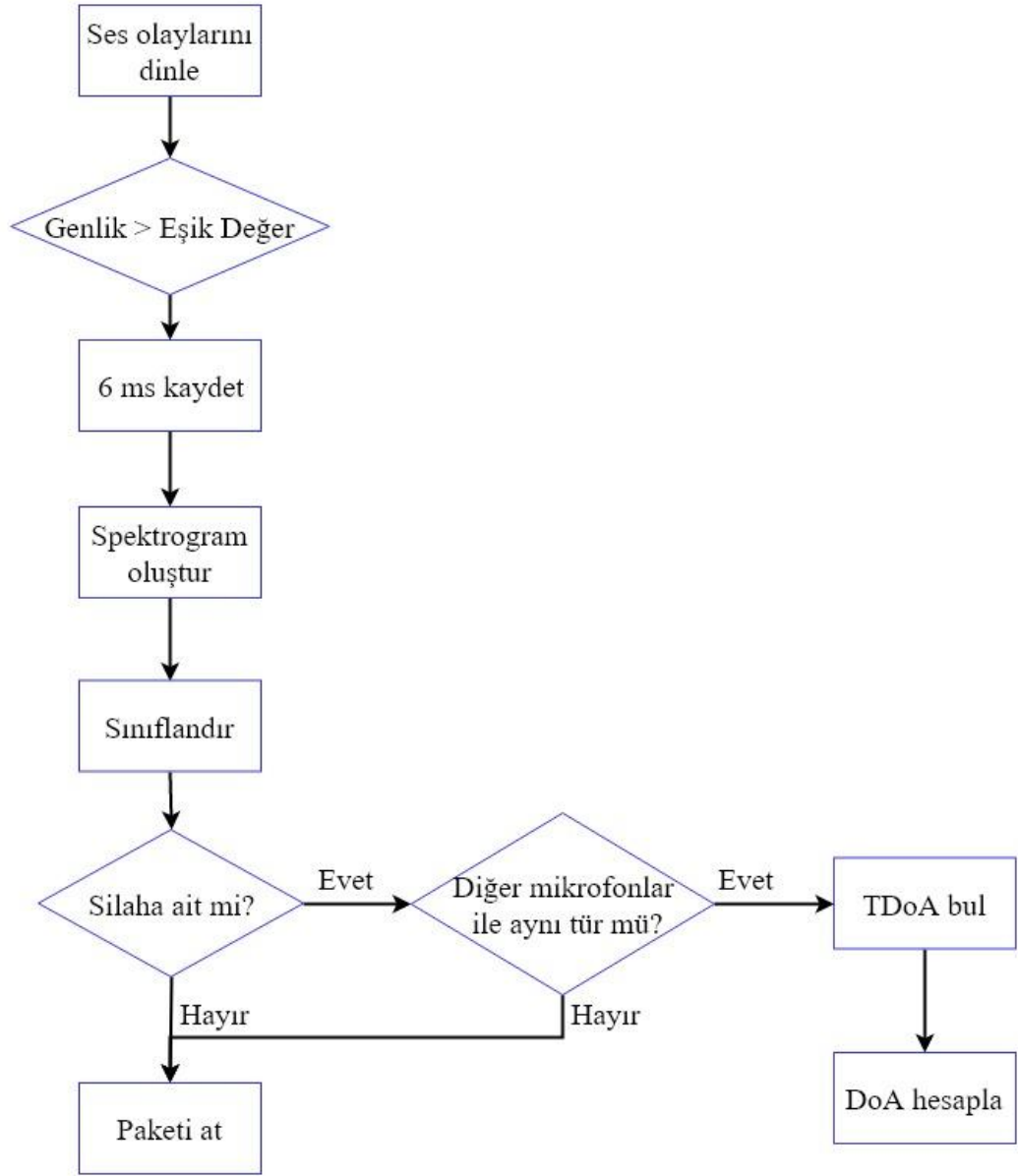
Silah sesine ait akustik olaylar bir dizi mikrofondan gözlemlenmektedir ve bilinen konumlardaki mikrofonların akustik olayı algılama zamanları arasındaki fark tespit edilerek akustik kaynağın varış yönü (DoA) belirlenmektedir. Şekil 5.1'de iki farklı mikrofondan algılanan akustik olaylar arasındaki zaman farkı görülmektedir. Bu gözlemin doğru bir biçimde yapılabilmesi adına en az üç mikrofondan izleme yapılmalıdır. Daha az sayıdaki mikrofondan yapılacak DoA hesaplamalarında hata payı oldukça yüksek olacaktır [12,97]. Literatürde DoA hesaplanması için kullanılan birçok yöntem bulunmaktadır. [98]'de yapılan çalışmada hüzme şekillendirmeye dayalı bir yaklaşım ve TDOA tabanlı iki farklı yaklaşım kullanılarak kıyaslanmıştır. Hüzme şekillendirmeye dayalı yaklaşım daha yüksek hesaplama maliyeti gerektirmekle birlikte DoA tespit hassasiyeti çok daha yüksektir. Bununla birlikte TDOA tabanlı yaklaşımlarda hesaplama maliyeti daha düşüktür.

Fakat DoA'yı hesaplamada ki başarısı da hüzme şekillendirmeye göre daha düşüktür. Bu nedenle hesaplama zamanının ön planda tutulduğu çalışmalarda TDOA yaklaşımı tercih edilirken yüksek konum hassasiyeti gerektiren uygulamalar içinde hüzme şekillendirme yaklaşımının daha uygun olabileceğini söylemek mümkündür.

Çizelge 5.1. Algılayıcı dizin sayısına bağlı olarak konum, atış doğrultusu ve mermi yörüngesinin hesaplanabilirliği ([11], [32]).

Algılanan Sinyal Tipi	Algılayıcı Dizin Sayısı (1)	Algılayıcı Dizin Sayısı (2)	Algılayıcı Dizin Sayısı (3)
Yalnızca şok dalgası	Konum ve yön hesaplanamam aktadır	Konum hesaplanabilir. Merminin izlediği yol da hesaplanabilir	Konum hesaplanabilir. Merminin izlediği yol da hesaplanabilir.
Yalnızca namlu sesi	Yön hesaplanabilir. Uzaklık bilgisi hesaplanamaz.	Algılayıcı füzyonuyla konum hesaplanabilir.	Algılayıcı füzyonuyla konum hesaplanabilir. (daha yüksek hassasiyette)
Hem şok dalgası hem de namlu sesi	Konum Hesaplanabilir	Algılayıcı füzyonuyla konum hesaplanabilir.	Algılayıcı füzyonuyla konum hesaplanabilir. (daha yüksek hassasiyette)

Bu tez çalışmasında önerilen yöntemde silah sesine ait akustik olaylar sürekli olarak mikrofon dizisi tarafından dinlenmektedir. Bu aşamada silah sesine ait her iki akustik olayın da darbesel ses olayları olmasından faydalanılarak önceden belirlenmiş eşik değerinde genliğe sahip bir darbesel ses olayının meydana gelip gelmediği takip edilir. Eğer darbesel bir ses olayı meydana gelmişse bunun ŞD veya NP olup olmadığını tespit etmek amacıyla SIF özellikleri ve CNN kullanılarak sınıflandırılır. ŞD veya NP tespit edilmemişse paket direk olarak atılır. Öte yandan tüm mikrofonlar silah sesine ait aynı akustik olayı algılamışsa TDOA kullanılarak sensör dizisinin bilinen geometrisi üzerinde DoA hesabı yapılır. Şekil 5.2'de önerilen modelin nasıl işlediği görülmektedir.



Şekil 5.2. Önerilen modelin akış diyagramı.

Herhangi bir akustik olay mikrofonların birbirlerinden farklı konumlarda bulunmaları nedeniyle aynı anda algılamayacaktır ve her bir mikrofon için ses hızına bağlı olarak bir yayılma gecikmesi meydana gelecektir. Bu noktada sinyalin her bir mikrofonu geliş arasındaki fark yukarıdaki bölümde belirtilen GCC-PHAT yöntemi kullanılarak tespit edilmektedir. TDOA değeri ölçülerek tespit edilebilirken DoA değerinin hesaplanması gerekir. Bu tez çalışmasında DoA hesabı için akustik kaynağın mikrofonlardan yeterince uzakta bulunduğu ve her iki akustik olayında düzlem dalgalar şeklinde yayıldığı varsayımı ile gerekli hesaplamalar yapılmıştır. DoA'nın

Eşitlik 5.12’de C_b değeri merminin tipine bağlı balistik sabiti ve η genellikle 0,5 kabul edilen üstel değerdir. Şekil 5.3 üzerinde mermi bir v hızıyla A noktasına kadar gelir bu noktadan sonra ŞD O noktasına kadar ses hızında algılayıcı dizine ulaşır. Öte yandan NP Z noktasından O noktasına kadar olan mesafeyi ses hızında alır.

Bu durumda t_0 Z noktasında bulunan silahın ateşlendiği an olmak üzere NP’nin algılayıcı dizine ulaştığı süre şu şekilde ifade edilebilir:

$$t_{MB} = t_{shot} + \frac{d_{Z,O}}{c} \quad (5.13)$$

Burada $d_{Z,O} = \|Z - O\|$ şeklinde ifade edilir ve Z noktası ile O noktası arasındaki Öklit uzaklığıdır. NP silahın patlaması sonucu meydana gelen bir ses iken ŞD havada ses üstü hızda hareket eden mermiden kaynaklanan sestir. Yani namludan kaynaklanmamaktadır. Bu durumda ŞD’nin algılayıcı dizine ulaşma süresi hesaplanmak istendiğinde v ve c olmak üzere iki farklı hız söz konusu olacaktır. Bu iki farklı hız dikkate alındığında t_{SW} şu şekilde ifade edilebilir:

$$t_{SW} = t_{shot} + \frac{d_{Z,A}}{v} + \frac{d_{A,O}}{c} \quad (5.14)$$

Farklı yayılma hızları nedeniyle bu iki sinyal sensör dizisine farklı zamanlarda ulaşacaktır. Bu durumda Eşitlik 5.13 ve Eşitlik 5.14 dikkate alınarak her iki sinyalin TDOA değeri şu şekilde gösterilebilir:

$$\Delta t = t_{MB} - t_{SW} = \frac{d_{Z,O}}{c} - \frac{d_{Z,A}}{v} - \frac{d_{A,O}}{c} \quad (5.15)$$

Şekil üzerinde B ve O noktaları arasındaki mesafe merminin sensör dizinine en yakın olarak geçtiği mesafedir ve kaçış mesafesi olarak adlandırılır. Mermi hızı ve kaçış mesafesi gibi faktörlere de bağlı olarak ŞD algılayıcı dizine genellikle NP’den önce ulaşmaktadır.

Ayrıca kaçış mesafesi 0-50 m arasında değişebilmektedir. Bununla birlikte zemin etkileşimi olmadan güvenli bir şekilde tespit edebilmek için ideal mesafe 20 m'dir [100].

Bu bilgiler doğrultusunda tek bir algılayıcı dizin kullanılarak atıcı konumu hesaplamak şu şekilde hesaplanabilir:

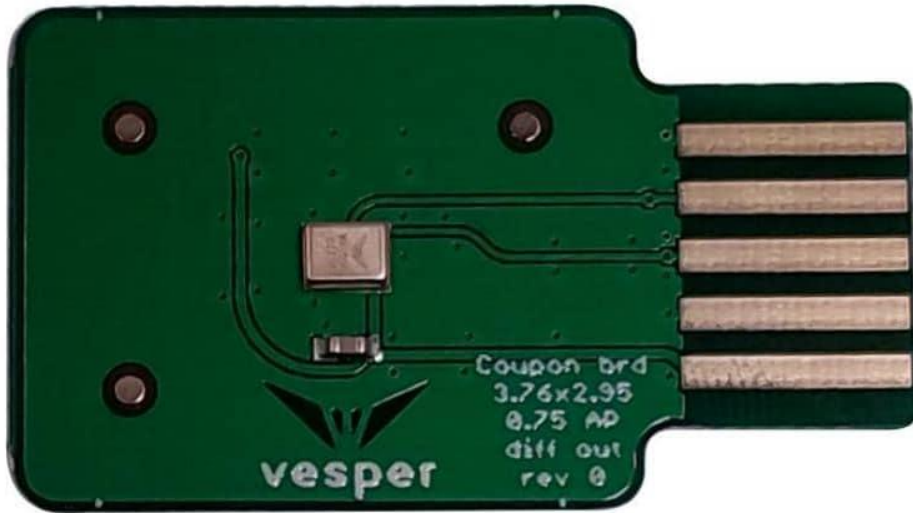
$$x_0 = s + c \left(\frac{t_{MB} - t_{SW}}{1 - (u_{MB} \cdot u_{SW})} \right) u_{MB} \quad (5.16)$$

$$t_0 = \frac{t_{SW} - t_{MB}(u_{MB} \cdot u_{SW})}{1 - (u_{MB} \cdot u_{SW})} \quad (5.17)$$

BÖLÜM 6

SİSTEM DİZAYNI

Bu tez çalışmasında hazırlanan atıcı konum tespit sisteminde donanım olarak MEMS mikrofonlar, analog dijital dönüştürücüler (ADC), GPS alıcısı, manyotometre, ivme ölçer, jiroskop, sıcaklık sensörü, basınç sensörü, mikroişlemci ve FPGA kullanılmıştır. Tüm ekipmanlarda hazır ticari ürünler tercih edilmiştir. Sistemi oluşturan temel bileşenler mikrofonlar, ADC'ler ve FPGA kartıdır. Diğer bileşenler cihazın kendi konumunu bulması, yükseltisini hesaplaması, yönelimini belirlemesi ve çevre sıcaklığına bağlı olarak ses hızının hesaplandığı yardımcı birimlerdir. Yardımcı birimlere ait tüm bileşenler ARM7 mikroişlemci üzerinde çalışmaktadır. Ses sinyallerinin toplanması, işlenmesi, sınıflandırılması ve atıcı konumunun hesaplanması için paralel çalışma kabiliyetlerinin yüksekliği nedeniyle FPGA tercih edilmiştir.

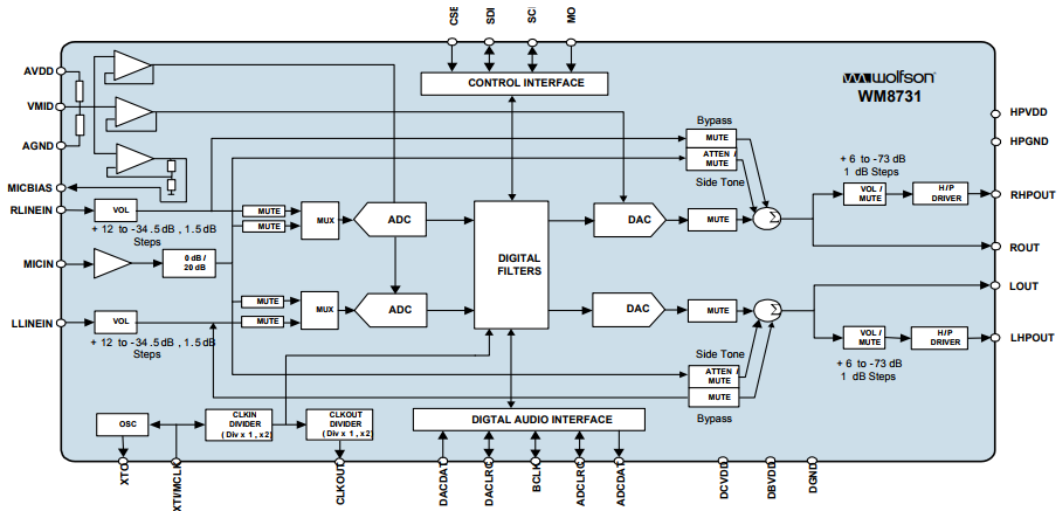


Şekil 6.1. Tez çalışmasında kullanılan Vesper VM1000 mikro elektro mekanik mikrofon.

Literatürdeki atıcı konum tespit sistemlerinde electret, condenser ve MEMS olmak üzere farklı türlerde mikrofonlar kullanılmıştır. Bu tez çalışmasında Vesper firmasına

ait VM1000 analog piezoelektrik MEMS mikrofonları kullanılmıştır. Piezoelektrik mikrofonlar basınç dayanımı yönünden iyi bir performans sergilemese de saha testlerinde toz ve yağmur gibi durumlara karşı iyi bir dayanım ortaya koymaktadır. Ayrıca MEMS mikrofonların boyutlarının diğer mikrofon türlerine kıyasla oldukça küçük olması ve buna ilave olarak düşük güç gereksinimleri nedeniyle saha uygulamaları için iyi bir ürün niteliğindedir. Şekil 6.1’de tez çalışmasında kullanılan mikrofon görülmektedir. Tez çalışmasında üretilen algılayıcı dizin altı adet mikrofondan oluşmaktadır.

Ayrıca bu tez çalışmasında analog mikrofon tercih edilmesi nedeniyle bu mikrofonlardan elde edilen sinyallerin sayısal sinyallere dönüştürmek amacıyla ADC’lere ihtiyaç duyulmaktadır. Bu doğrultuda “Wolfson WM8731 Audio Codec” tercih edilmiştir. Şekil 6.2’de ses kodeğine ilişkin blok diyagramı görülmektedir. Kullanılan ekipman 2 veya 3 kablo arayüz aracılığıyla kontrol edilebilmektedir. Bu tez çalışmasında ses kodeği I2C arayüzü kullanılarak konfigüre edilmiştir. Ekipman ADC’nin yan sıra yükselteç ve dijital filtre de barındırmaktadır. Bu tez çalışmasında her bir mikrofondan alınan analog sesler 16 bit uzunluğunda ve 96 KHz örnekleme hızında sayısal sinyallere dönüştürülerek kullanılmıştır. Ayrıca her bir mikrofon bir adet ses kodeğine bağlanmıştır.



Şekil 6.2. WM8731 ses kodeğine ait blok diyagramı.

Tasarlanan atıcı konum tespit sisteminde ses sinyallerinden spektrogram resim elde edilmesi, bu resimlerin sınıflandırılması, GCC-PHAT değerlerinin bulunması, DoA hesaplanması ve atıcı konumun belirlenmesi FPGA tabanlı bir sistem üzerinde gerçekleştirilmiştir. Diğer ekipmanlara benzer şekilde burada da hazır geliştirme kartı tercih edilmiştir. Geliştirme kartı olarak Digilent firmasına ait Genesys 2 kartı tercih edilmiştir. Kart üzerinde Xilinx firmasına ait her biri 6 girişli 4LUT ve 8 flip-flop içeren toplam 50 950 mantık ve 840 DSP biriminde oluşan XC7K325T-2FFG900C Kintex 7 FPGA bulundurmaktadır. Ayrıca geliştirme kartının üzerinde 8 adet anahtar ve led, oled ekran, kullanıcı giriş çıkışları için 400 pinli FMC HPC konnektör, VGA ve HDMI bağlantıları, ethernet bağlantısı, usb bağlantısı ve PMOD konnektörleri gibi kullanıcı giriş çıkışlarına olanak sağlayan donanım bulunmaktadır. Buna ilave olarak kart üzerinde birkaç tane osilatör ve kristal bulunmaktadır ve bunlardan 2 tanesi doğrudan FPGA'nın pinlerine bağlanmış durumdadır. Bu osilatörlerden biri 135 MHz hızında diğeri de 200 MHz hızındadır, ayrıca her ikisi de diferansiyel osilatördür. Bu osilatörlerden 135 MHz hızında olan gigabit alıcı vericiye ve görüntüleme portlarına bağlıdır. Öte yandan 200 MHz hızında olan ise FPGA için gerekli olan ana saat girişini sağlamaktadır. Farklı saat girişlerine ihtiyaç duyulması durumunda Xilinx firmasına ait saat sihirbazı kullanılarak farklı çalışma hızlarında saatler elde edilebilmektedir. Ayrıca Bölüm 7'de de belirtildiği FPGA'ların kullandıkları saat girişleri mikroşlemcilerle kıyasla çok daha düşük olmakla birlikte yüksek paralel çalışma kabiliyetleri bu açığı iyi bir şekilde kapatmaktadır.

Bu tez çalışmasında geliştirme kartının üzerinde bulunan 5 adet PMOD konnektör ses kodeklerinin bağlantısı için kullanılmıştır. Her bir konnektör toplam 12 adet pin bağlantısı içermektedir. Bunlardan 8 tanesi FPGA'ya bağlıdır dijital giriş çıkış işlemleri için kullanılabilir. Bununla birlikte Şekil 6.3'te de görüldüğü gibi geri kalan 4 pinin 2 tanesi toprak ve 2 tanesi de güç pinleridir. Her bir güç ve toprak pini 3,3 voltluk bir gerilim ve 1 amperin üzerinde bir akım değeri sağlayabilmektedir. Bu da birçok donanım için ilave güç kaynağına ihtiyaç duymadan doğrudan geliştirme kartı üzerinden beslenmesine imkan sağlamaktadır. Her bir ses kodeğinin kontrol arayüzü ve veri girişi bu pinler aracılığıyla yapılmaktadır. Ayrıca ses kodeğinin senkronizasyonu için FPGA tarafından üretilen bit saati bu pinler vasıtasıyla ses kodeğine iletilmektedir.



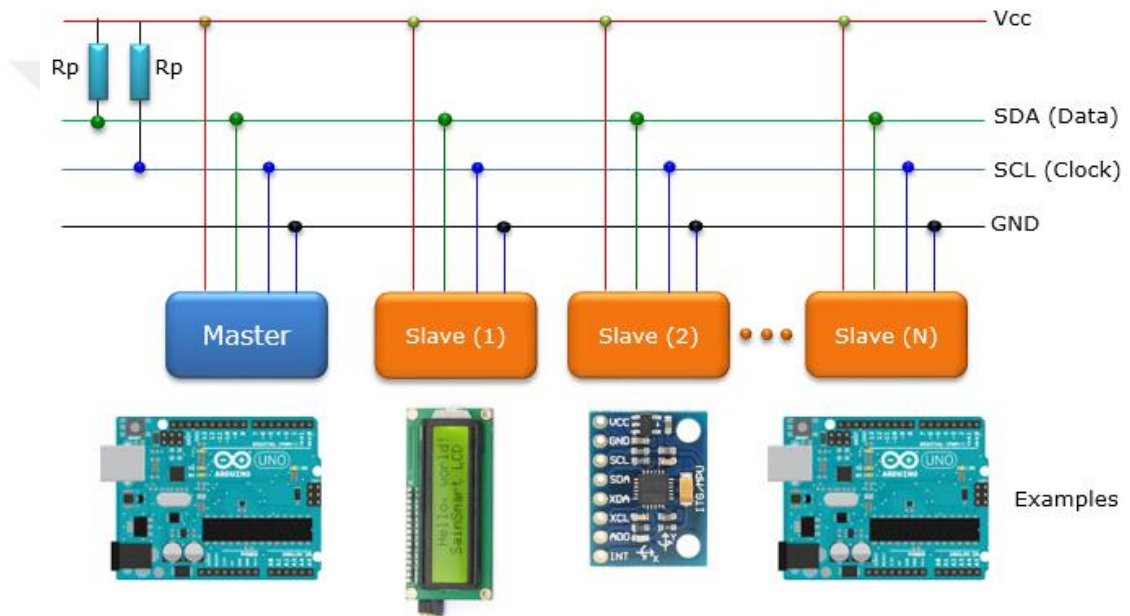
Şekil 6.3. Genesys2 geliştirme kartı üzerindeki PMOD konnektörlerinin pin bağlantısı.

Bu tez çalışmasında GPS alıcısı dışındaki tüm sensörlerin kontrolü I2C (Inter-Integrated Circuit) seri veri yolu ile sağlanmaktadır. GPS alıcının kontrolü için ise RS232 seri haberleşme protokolü kullanılmaktadır. I2C seri veri yolu senkron seri haberleşmenin bir örneğidir. 1982 yılında, orijinal 100 kHz I2C sistemi, çeşitli Philips çipleriyle kontrol elektroniği oluşturmak için basit bir dahili veri yolu sistemi olarak oluşturulmuştur ve en yaygın kullanılan hali 2014 yılında yayınlanmış version 6 halindedir. Düşük hıza sahip cihazların haberleşmesi için kullanılan toğrağa referanslı veri yoludur. Biri verinin taşındığı hat (Serial Line, SDA) ve bir diğeri de saat sinyali (Serial Clock, SCL) olmak üzere 2 tane çift yönlü sinyal hattı kullanılır. Her bir hat pull-up dirençleri vasıtası ile pozitif gerilime bağlanmış open-drain hattır. Hatların pull-up dirençleri ile besleme gerilimine bağlı olması nedeniyle hat üzerinde haberleşmenin olmadığı durumlarda hem SDA hem de SCL lojik 1 seviyesinde bulunmaktadır. Bu protokolde haberleşme için SDA, SCL yanı sıra bir adet referans toprak hattına ihtiyaç duyulmaktadır. Dolayısıyla haberleşme için 3 farklı hat gerekmektedir. Cihazların beslemelerinin aynı hattan gelmesine ihtiyaç yoktur fakat haberleşme için toprak hattının ortak olması gerekmektedir. İhtiyaç duyulan hat sayısının fazlalığına bağlı olarak genellikle kısa mesafede ve düşük hız gerektiren uygulamalarda tercih edilmektedir. Şekil 6.4'ten de anlaşıldığı gibi haberleşme için bir adet master cihaza ve en az bir adet slave cihaza ihtiyaç duyulmaktadır. Öte yandan slave cihaz sayısı ihtiyaç doğrultusunda birden fazla da olabilmektedir. Master cihaz hattı hangi donanımın kullanacağına karar verdiği için aynı hat üzerinden birden fazla cihaz kullanılabilir. Bu tez çalışmasında FPGA master cihaz olarak kullanılmaktadır.

Atıcı konum tespit sistemi her başlatıldığında ses kodekleri üzerindeki kaydedicilerin doğru konfigürasyona göre ayarlanması gerekmektedir. Bunun için ses kodeklerinin pmod konnektörlerine bağlı SDA ve SCL hatları vasıtası ile kodek üzerindeki 11 farklı

kaydedici önceden belirlenmiş konfigürasyona ayarlanmaktadır. Temelde yapılan işlem cihazın 16 bit genişliğinde ve 96 KHz hızda çalışmasını sağlayacak şekilde ayarlamaktır. Buna ilave olarak kodlar üzerindeki yükselteç seviyeleri de aynı değerde olacak şekilde konfigüre edilmektedir.

Bu tez çalışmasında kullanılan ses kodeği +3,3 voltluk bir besleme gerilimine ihtiyaç duymaktadır. Kodeklerin besleme gerilimi için harici güç kaynağı kullanılmamış doğrudan geliştirme kartı üzerindeki pmod konnektörlerinin güç pinleri kullanılmıştır.



Şekil 6.4. I2C seri veri yoluna ait örnek bir bağlantı.

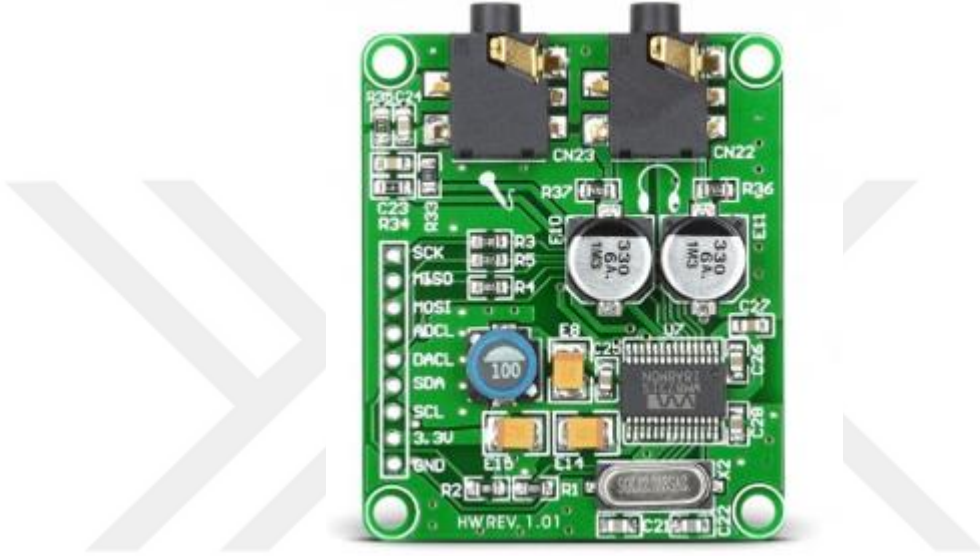
Atıcı konum tespit sistemlerinde son derece kritik olan konulardan bir tanesi de cihaza bağlı mikrofonlardan elde edilen sinyaller arasında zaman senkronizasyonun sağlanmasıdır. Zaman senkronizasyonunda yaşanacak hatalar DoA değerlerinin hesaplanmasında yanlışlıklara neden olacaktır. Bu durumda konum tespitindeki hata paylarını artıracaktır. Bu nedenle özellikle geniş alan savunması için kullanılan ağ merkezli atıcı konum tespit sistemlerinde senkronizasyon sağlanması oldukça zorlu bir konudur. Öte yandan bu tez çalışmasında tüm mikrofonların aynı FPGA tarafından tek bir saat sinyali ile kontrol edilmesi nedeniyle cihazlar arasında senkronizasyon problem yaşanmamaktadır.

Çizelge 6.1. Ses kodeğine ait bağlantıların listesi.

Sinyal İsmi	FPGA Pin	Pin Fonksiyonu
ADC_SDATA	AH19	Analog dijital dönüşüm sonrasında elde edilen seri sinyal
DAC_SDATA	AJ19	Dijital analog dönüşümü için FPGA'dan alınan seri sinyal
BCLK	AG18	Seri veri port clock
LRCLK	AJ18	Seri veri port çerçeve clk
MCLK	AK19	Master clk
SDA	AF18	I2C konfigürasyon arayüzü
SCL	AE19	I2C konfigürasyon arayüzü

Her bir ses kodeği ile FPGA arasında I2C arayüz bağlantılarının haricinde veri akışı ve senkronizasyon saatleri gibi bir takım başka bağlantılarda bulunmaktadır. Çizelge 6.1’de her bir kodeğin sahip olduğu bağlantıların bir listesi verilmiştir. SDA ve SCL yukarıda da belirtildiği gibi I2C seri veri yolunun kontrolü için kullanılan bağlantılardır. Bu bağlantılar aracılığıyla ses kodeğinin hangi modda çalışacağı belirlenir ve gerekli diğer ayarlamalar yapılır. MCLK ses kodeğinin çalışması için gerekli olan ana çalışma saatinin sağlandığı giriştir. Bu saat bilgisi doğrudan FPGA tarafından üretilip kodeğe gönderilebildiği gibi, harici bir osilatör aracılığıyla da üretilebilir. Bu tez çalışmasında 12,288 MHz çalışma hızında harici bir osilatör kullanılarak MCLK bilgisi sağlanmıştır. LRCLK elde edilen pencerenin sağ kanala mı yoksa sol kanala mı ait olduğunu gösteren bağlantıdır. Bu tez çalışmasında her bir kodek üzerinde sadece bir mikrofon bağlı olması nedeniyle bu bağlantıya ihtiyaç duyulmamaktadır. BCLK kodek ile FPGA arasında veri senkronizasyonunun sağlanması için kullanılan bağlantıdır. İki farklı konfigürasyonda kullanılabilir. Bunlardan birincisi BCLK bilgisinin kodek tarafından üretilerek bağlı olduğu cihaza gönderilmesidir. Bir diğeri de bu bilginin kodeğin bağlı olduğu cihaz tarafından üretilerek senkronun sağlanmasıdır. Bu doğrultuda bu çalışmada BCLK bilgisi FPGA tarafından üretilerek tüm kodeklere gönderilmektedir. Böylelikle mikrofonlar arasında TDOA ve DoA hesaplamaları için senkronizasyon problem yaşanmamaktadır. Bir diğeri bağlantı da ADC_SDATA bağlantısıdır. Yani kodek içerisindeki ADC’lerden edilen bilgilerin seri bir şekilde dışarı aktarıldığı bağlantılardır. Ses kodeğinin ADC_SDATA bağlantısından elde edilen bilgiler FPGA’nın pmod konnektörlerine giriş olarak uygulanıp sonrasında FPGA tarafından işlenmektedir. Ses kodeğine ilişkin son bağlantı DAC_SDATA bağlantısıdır. Ses kodeği bu girişinden aldığı sayısal

verileri işleyerek analog sinyallere dönüştürmektedir. Bu bağlantı tez çalışması esnasında mikrofonlardan doğru bir şekilde sinyal alınıp alınmadığını kontrol etmek için kullanılmıştır. Ses kodeğinin ADC_SDATA bağlantısından elde edilen bilgiler tekrardan DAC_SDATA üzerinden gönderilerek ses sinyallerinin kontrol edilmesi sağlanmıştır. Atıcı konum tespit sisteminin normal çalışması esnasında bu bağlantıya ihtiyaç duyulmamaktadır. Şekil 6.5'te ses kodeğinin bağlı olduğu kart görülmektedir.



Şekil 6.5. WM8731 ses kodeğinin monte edildiği ses kartı.

Bu tez çalışmasından tasarlanan sistem ile sabit bir noktaya kurularak atıcı konumunun belirlenmesi üzerine çalışmalar yapılmış olmakla birlikte gelecekteki çalışmalarda sistemin hareketli araçlar üzerinden de atıcı konumu tespiti yapabilmesi adına bir dizi ilave sensör ile donatılmıştır. Bu sensörlerin kontrolü bir ARM7 mikroişlemci tarafından kontrol edilmektedir. Burada bu sensörlerin doğrudan FPGA bağlanmamasının temel nedeni CNN'lerin, ses sınıflandırma için kullanılan diğer özelliklerin ve konum hesaplama için kullanılan yöntemlerin FPGA donanımının büyük kısmını kullanmasına bağlı olarak FPGA üzerinde kısıtlı bir alanın kalması ve ısınma problemlerinin oluşmasıdır. Bu nedenle bu sensörler bir mikroişlemci üzerinde çalışmakta ve FPGA ile mikroişlemcinin bağlantısı bir seri arayüz aracılığı ile sağlanmaktadır. FPGA tarafından bir istek paketi geldiğinde gerekli sensör verisi mikroişlemci tarafından FPGA'ya gönderilmektedir. Böylelikle FPGA üzerindeki iş yükü azaltılmaktadır.



(a)



(b)

Şekil 6.6. a) Harici anten ve b) GPS alıcısı.

Sabit bir noktaya kurulu olan atıcı konum tespit sistemlerinin kendi konum bilgilerini tespit etmesine genellikle ihtiyaç duyulmamaktadır. Bu bilgiler operatör tarafından cihazın kurulumu esnasında cihaza girilerek konum hesabı için gerekli bilgiler elde edilebilmektedir. Öte yandan mobil sistemler kendi konum bilgilerini de yüksek hassasiyetle bulunmasına ihtiyaç duyar. Bu tez çalışmasında GPS alıcısı olarak Antenova firmasının 22 farklı uyduyu takip kalitesine sahip olan M10478-A2-1 alıcısı kullanılmıştır. Ürün GPS alıcısının ve anten birimin aynı ünite içerisinde yer aldığı kompakt bir çözüm sunmakla birlikte bu çalışmada 28 dB kazanca sahip suya dayanımlı ilave bir harici GPS anteni de kullanılmıştır. Şekil 6.6'da GPS alıcısı ve harici anten görülmektedir.

Özellikle geniş alan savunmasında kullanılan sistemlerde önemli olan konulardan bir tanesi de cihazları birbirleriyle olan yönelimlerinin doğru bir şekilde ayarlanmasıdır. Ayrıca mobil atıcı konum tespit sistemlerinde cihazın ivmelenmesi ve hareket doğrultusu gibi bilgilere ihtiyaç duyulmaktadır. Bu kapsamda bu çalışmada içerisinde 3 eksen manyometre, 3 eksen jiroskop ve 3 eksen ivme ölçer barındıran LSM9DS0 sensörü kullanılmıştır. Bunlara ilave olarak sensörün bünyesinde bir adet gömülü sıcaklık sensörü bulunmaktadır. Eşitlik 2.3'te ses hızının ortam sıcaklığına bağlı olarak değişen bir parametre olduğu gösterilmiştir. Ortam sıcaklığının tespit edilmesi ses hızının belirlenmesi açısından önemlidir. Dolayısıyla konum tespitinde de hatalara neden olmaktadır. [98]'de yapılan çalışmanın bir bölümünde ses hızının tespit

edilmesinde yaşanan belirsizliklerinin atıcı konumunun belirlenmesinde meydana getirdiği hatalar ortaya konulmuştur. Bu nedenle bu tez çalışmasında sabit bir ses hızı kabulü ile işlem yapmak yerine ses hızını atmosferik koşullara göre en doğru şekilde hesaplamayı baz alan bir yaklaşım ortaya konulmuştur. Bununla birlikte atışın yapıldığı konum ile algılayıcı sensör dizini arasındaki mesafeye bağlı olarak farklı atmosferik koşullarda gözlemlenebilmektedir. Fakat bu etkileri tespit etmenin zorluğu nedeniyle bunlar göz ardı edilmiştir.

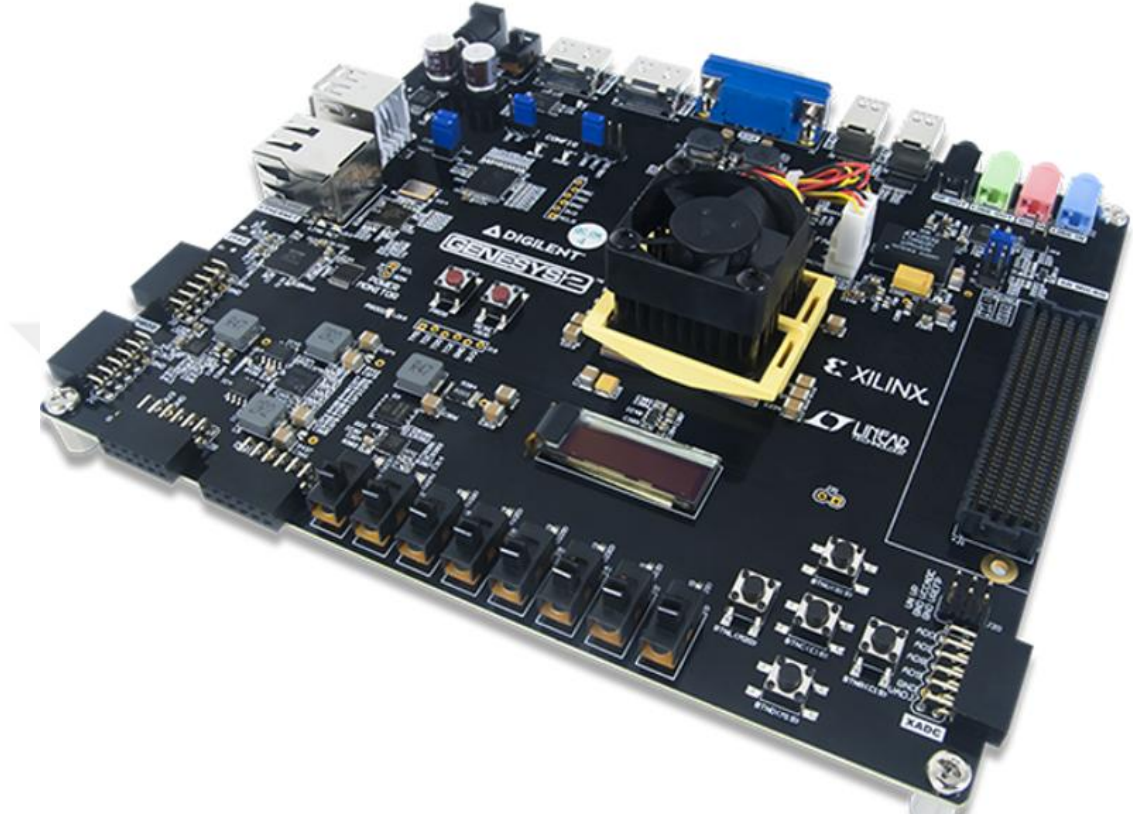
Bunlara ilave olarak Eşitlik 2.5'te de belirtildiği gibi ŞD modelinde ihtiyaç duyulan parametrelerden bir tanesi de atmosfer basıncıdır. ŞD dalgasının genlik değeri ve dalga boyu merminin kaçış mesafesi yani algılayıcı dizine en yakın olarak geçtiği mesafe hakkında bilgi sağlamaktadır. Fakat buradaki eşitliğin bir kısmı atmosfer basıncına bağlıdır. Bu doğrultuda Bosch firmasına ait BMP280 dijital basınç sensörü kullanılarak basınç ölçümleri yapılmıştır. Çalışmada kullanılan sensör Şekil 6.7'de görülmektedir.



Şekil 6.7. BMP280 dijital basınç sensörü.

Tüm sensörlerden elde edilen veriler silaha ait akustik bileşenlerin tespit edilmesi ve atıcı konumunun hesaplanması için nihai olarak FPGA'ya gönderilmektedir. FPGA üzerindeki devrelerin göreve özel bir şekilde tanımlanmış olması nedeniyle güç tüketimi yönünden geleneksel işlemciler ile kıyaslandığında daha düşük güç sarfiyatında buldukları görülmektedir [101,102]. Bu durum sahada yapılacak uygulamalar için önemli bir avantaj sağlamaktadır. Ayrıca bünyesinde birbirinden bağımsız çok sayıda donanımın bulunuyor olması görüntü ve sinyal işleme gibi yüksek

hesaplama maliyeti gerektiren uygulamalarda daha az çevrim ile sonucu elde etme kabiliyeti kazandırmaktadır [103]. Şekil 6.8’de FPGA’nın bağlı olduğu Genesys2 kartı görülmektedir.



Şekil 6.8. Genesys2 kartı.

Bu tez çalışmasında algılayıcı dizin olarak 6 mikrofondan oluşan bir sistem kullanılmıştır (Şekil 6.9 ve Şekil 6.10). Bir tane mikrofona daha yüksekte bulunurken diğer 5 mikrofona bir çember oluşturacak şekilde eşit aralıklarla yerleştirilmiştir. Tez çalışmasında geliştirilen sistemin görüntüsü Şekil 6.11’de görülmektedir.



Şekil 6.9. Algılayıcı dizin.



Şekil 6.10. Sistemde kullanılan tüm donanım ekipmanları.



(a)



(b)

Şekil 6.11. a) ve b) Tez çalışmasında geliştirilen sistemin görüntüsü.

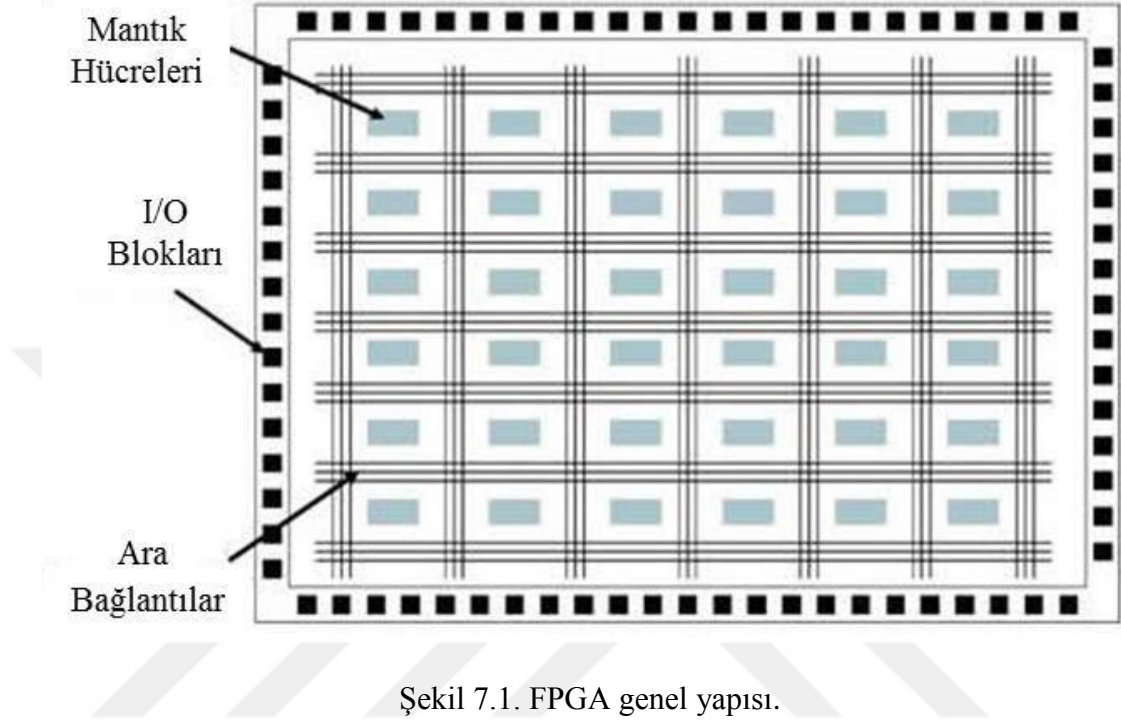
BÖLÜM 7

FPGA

“Field Programmable Gate Array” kelimelerinin kısaltılması ile elde edilen FPGA terimini Türkçeye “Alanda Programlanabilir Kapı Dizileri” olarak çevirmek mümkündür. FPGA çipleri kullanıcılara, üretim aşamasından sonra içerisinde barındırdığı ara bağlantıları ve mantık bloklarını programlayarak yeniden tanımlama imkânı sağlamaktadır [104]. Bu özellik kullanıcılara ihtiyaca göre mantık blokları arasındaki yolları yeniden tasarlayarak FPGA'nın kapasitesi doğrultusunda istenilen tümleşik devrenin gerçekleştirilmesini sağlar [105-106]. FPGA'lar birçok konuda mikroişlemcilerin açıklarını ve eksiklerini kapatırken [107], Uygulamaya Özgü Tümleşik Devre (ASIC) teknolojilerinin uzun süren üretim aşamaları ve maliyetleri gibi dezavantajlarını da tasarım kolaylıkları ve yeniden programlanabilirlik gibi avantajları sayesinde büyük ölçüde ortadan kaldırmaktadır [108-109].

ASIC'ler uygun kaynak kullanım oranları ve güçlü hesaplama kabiliyetlerine sahip olmakla birlikte göreve özgün bir şekilde tasarlanıp üretilmeleri sebebiyle sonrasında değiştirilemezler. Bu durumda kullanımları açısından önemli bir kısıtlama getirir. Öte yandan mikroişlemci üzerinde kullanılan yazılımlar ciddi anlamda esnekliğe sahiptirler, yeniden tasarlanıp geliştirilebilirler. Fakat ASIC'ler ile bir karşılaştırma yapılacak olursa güç gereksinimleri ve performans kriterleri açısından daha kötü bir performans ortaya koyarlar [110]. Donanım platformlarını hız açısından değerlendirirken göz ardı edilemeyecek iki temel unsur performans ve esneklik faktörleridir [111]. Bu anlamda ASIC'ler ve genel amaçlı işlemcilerin iki ayrı uç noktayı temsil ettiklerini söylemek mümkündür. ASIC'ler yüksek performansa sahip pahalı ve esnek olmayan ürünlerdir. Öte yandan genel amaçlı mikroişlemciler düşük fiyatlı, yeniden programlanabilir ve kolay temin edilebilir ürünlerdir. Bununla birlikte FPGA'lar ise düşük maliyetleri, yüksek performansları ve sağladıkları esneklik ile iki temel hesaplama çözümü sayılabilecek yazılım ve donanımın birleşimi

niteliğindedirler. Dolayısıyla ASIC ve genel amaçlı işlemciler arasında orta noktada yer alırlar. FPGA'lar kullanıcılarına istedikleri tümleşik devreyi yazılım ortamında tasarlayarak donanımsal olarak uygulamalarına olanak sağlarlar.



Şekil 7.1’de de görüldüğü gibi, giriş-çıkış blokları, konfigüre edilebilir lojik bloklar ve ara bağlantılar olmak üzere bir FPGA temelde üç ana birimden meydana gelmektedir [112]. Mantık hücreleri arasındaki bağlantılar ayarlanarak ihtiyaç duyulan tümleşik devrenin elde edilmesinin yanı sıra FPGA’nın yapısı gereği lojik blokların bir kısmı bir fonksiyonu icra ederken başka bir kısmı da bütünüyle farklı bir görevi yerine getirebilmektedir. Bu durumda FPGA’lara çok ciddi bir paralel çalışma kabiliyeti kazandırmaktadır. Genellikle kişisel bilgisayarlara kıyasla çok daha düşük çalışma frekanslarına sahip olmakla birlikte üstün paralel çalışma kabiliyetleri FPGA’ların birçok durumda kişisel bilgisayarlardan daha hızlı işlem yapmasına imkan tanımaktadır [113].

Konfigüre edilebilir lojik bloklar genellikle flip-floplar ve buna ilave olarak doğruluk tablolarından (look-up table) meydana gelirler. Öte yandan giriş-çıkış blokları FPGA ile dış ortam arasında bağlantı kurulmasına olanak sağlarlar. Yani FPGA içerisinde üretilen bir sinyalin dışarı taşınmasına veya dış ortamdan alınan bir sinyalin içeride

işlenmesine aracılık ederler. Ara bağlantılar ise uygun lojik blok konfigürasyonun oluşturulmasının yanı sıra bu bloklar ile giriş-çıkış blokları arasındaki bağlantının da yapılmasını sağlarlar.

Öte yandan Grafik İşlemci Birimi (GPU) gibi yüksek performanslı donanım araçlarının bir tasarımı hızlandırma yaklaşımı ile FPGA'nın hızlandırma yaklaşımı arasında mantıksal farklılıklar bulunmaktadır. GPU gibi sabit bir mimariye sahip olan donanımlarda bağımsız hesaplama birimleri üzerinde görevleri paralel olarak çalıştıracak bir yazılım icra etme modeli izlenir. Derin öğrenme açısından değerlendirecek olursak amaç hesaplamaların paralel olarak yapıldığı ve veri bağıllığının sağlandığı bir algoritma gerçekleştirmektir. Bundan farklı olarak FPGA mimarisi ise uygulama için uyarlanmıştır. Üst düzey yazılım dillerinde uygulanması oldukça zor ve karmaşık olan birçok düşük seviye donanım kontrol işlemine ihtiyaç duyan teknikler özellikle FPGA uygulamaları için caziptir [111].

Bu avantajlarının yanı sıra yüksek derleme zamanları ve yaygın olarak bir Donanım Tanımlama Diline (HDL) ihtiyaç duyulması FPGA'ların başlıca dezavantajlarındanır. C gibi geleneksel bir programlama dilinde donanımı anlamaya gerek kalmadan icra edilecek komutları sıralı bir şekilde yazmak yeterlidir. Öte yandan Verilog ve VHDL gibi dillerde kullanıcının dijital tasarım bilgisine sahip olmasına ihtiyaç vardır [111].

FPGA programlanması için yaygın olarak Verilog veya VHDL gibi HDL dilleri kullanılır. Bu tez çalışmasında donanım tanımlanması için VHDL dili tercih edilmiştir. VHDL dilinin geliştirilmesine ilk kez Amerika Birleşik Devletleri Savunma Bakanlığı tarafından başlanılmıştır ve ilk sürümü VHDL-87'dir. Sonrasında birtakım geliştirmeler yapılarak VHDL-93 sürümü yayınlanmıştır. Bu tez çalışmasında VHDL-2008 sürümü kullanılmıştır.

7.1. FPGA TABANLI YAZILIMLAR

Akustik tabanlı atıcı konum tespit sistemlerinin mikrofon dizileri aracılığıyla çevredeki ses olaylarını sürekli dinleyerek silah sesine ait akustik olayları hızlı ve en

dođru bir şekilde ayırt etmesi son derece önemlidir. Ayrıca bu tez çalışmasında önerilen silah sesine ait akustik olayların SIF ve CNN ile sınıflandırılması yaklaşımı çok yüksek miktarda hesaplama maliyeti gerektirmektedir. Bu amaç doğrultusunda kısa sürede yüksek hesaplama maliyetlerinin üstesinden gelebilmek adına yüksek paralel işlem yapabilme kabiliyetleri nedeniyle donanım olarak FPGA tercih edilmiştir. Donanımın yanı sıra FPGA üzerinde çalışacak gömülü yazılımın da kaynak kullanımını açısından verimli olması ve hızlı çalışması gerekmektedir. FPGA cihazlarının programlanmasında çeşitli yazılımlar kullanılmaktadır. Bu tez çalışmasında her ne kadar yazılım geliştirme süreçleri muadillerine göre daha uzun olsa da donanım seviyesine yakınlığı, cihazlar üzerinde etkin bir kontrol sağlaması ve hata ayıklama süreçlerinde sağladığı faydalar nedeniyle VHDL dili tercih edilmiştir. Mikrofon sinyallerinden spektrogram görüntü elde edilmesi, CNN tabanlı sınıflandırma yapılması ve GCC-phat tabanlı atıcı konum hesabı olmak üzere bu tez çalışması kapsamında VHDL ile hazırlanan yazılımı üç ana parçaya atında toplamak mümkündür ve Şekil 7.2’de tez çalışması kapsamında hazırlanan yazılımın blok diyagramı görülmektedir.

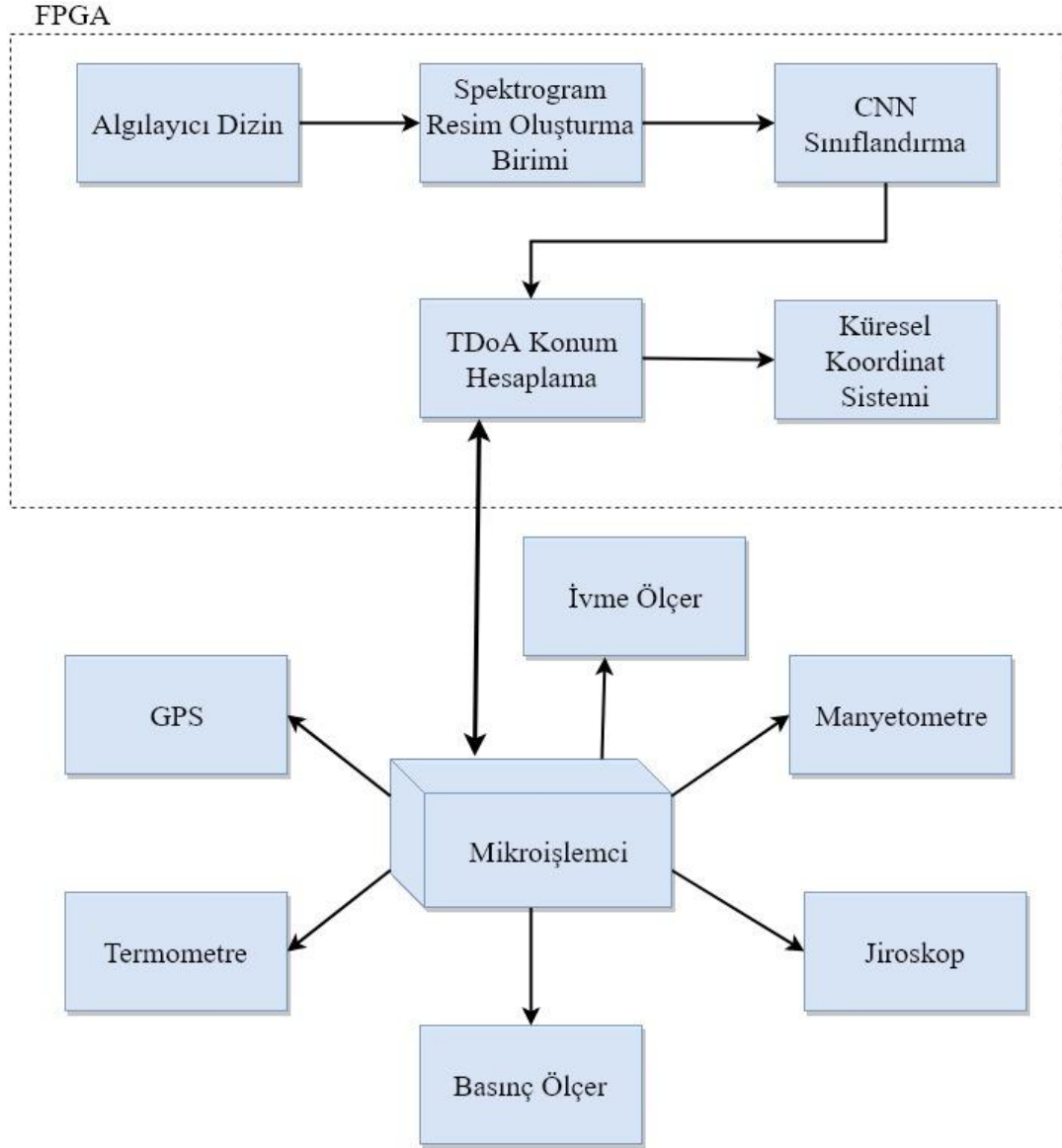
7.1.1. Spektrum Analizörü

Ses, sürekli değişken ve analog bir sinyaldir. Genellikle periyodik değildir ve bu nedenle dönem başına analiz edilemez. FPGA tarafından işlenebilmesi için dijital sinyallere dönüştürülmesi gerekir. Bu doğrultuda bir ADC vasıtasıyla dijital olarak örneklenip FPGA’nın girişine aktarılmalıdır.

7.1.1.1. Hamming Penceresi

Sinyal işlemede pencereleme işlemi kullanmanın temel sebebi spektral sızıntıları önlemektir. Sinyaller işlenmeden önce pencere adı verilen belirli uzunlukta parçalara ayrılmaktadır. Pencere fonksiyonları ise bu pencereler ile çarpım işlemi yapılan fonksiyonlardır. Amacı sinyalin orta kısmına belirginleştirirken kenar kısımlarını sönmölemektedir. Gerçek zamanlı olarak bir FFT uygulamak istendiđi için, sinyali parçalara bölünmelidir. Bu bölüm ne kadar büyük olursa, düşük frekanslar da o kadar iyi tespit edilebilir, ancak spektrumun çıkışında o kadar çok gecikme olur. FFT

algoritmasının doğası gereği bu parça periyodik olarak tekrarlanan bir parça olarak kabul edilecektir. Genlikteki bir değişimin birçok yüksek spektral bileşen içereceği bilinmektedir. Ayrıca örnekleme belirli bir zamanda yapılabileceğinden, ilk ve son örneğin eşit olmamasına neden olabilir.



Şekil 7.2. FPGA ortamında hazırlanan yazılımların genel yapısı.

Bu tez çalışmasında spektral sızıntıları önlemek için bir hamming penceresi kullanılmaktadır. Bu, uçların çok yavaş yavaş sifira getirilmesini sağlar, böylece ani çıkıntı elde etmeyi mümkün kılar. Hamming penceresinin kabaca bir kosinüs şekli vardır:

$$w(n) = \alpha - \beta \cos\left(\frac{2\pi n}{N-1}\right) \quad (7.1)$$

Numuneler, hamming penceresinin noktaları ile çarpım işlemine tabi tutularak, bitim noktalarında uçların neredeyse sıfır olması sağlanmaktadır. Hamming işlevinin değerleri 0'dan 255'e tamsayı aralığına eşlenip kullanılmaktadır. Bunun nedeni, FPGA üzerinde tamsayılarla işlem yapmanın kayan noktalı sayılarla işlem yapmaya göre daha kolay olmasıdır.

Her bir mikrofondan alınan 16 bit genişlikli sinyaller 8 bit genişliğindeki Hamming çerçevesi ile çarpılır ve sonuç olarak her bir kanal için 24 bit genişliğinde sinyaller elde edilir. Hamming çerçevesine ait veriler ROM bellekte tutulmaktadır.

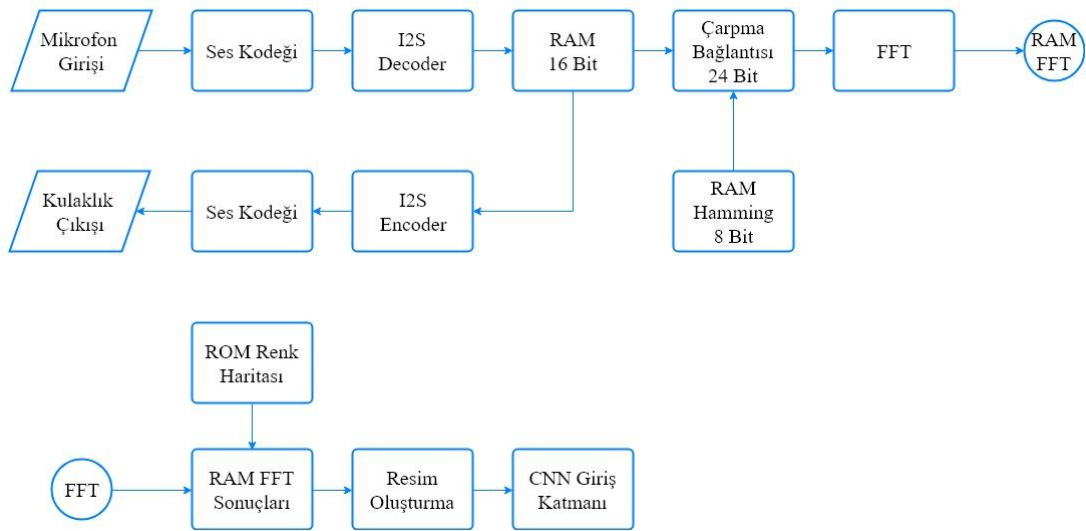
7.1.1.2. Spektrogram Resim Elde Edilmesi

Spektrogram resim elde edilmesi için Fourier dönüşümün yapılması gerekir Fourier dönüşümü, bir sinyalin farklı genlik, frekans ve fazlarda kosinüs ve sinüs temel bileşenlerinin toplamı olarak ifade edilebilir. Sesin doğası gereği burada bir DFT işleminin yapılması gerekir. DFT işlemi analiz için kısıtlı bir zamana ihtiyaç duyan bir dönüşümdür. Sabit bir süre boyunca örnekleme yapılarak, bu dönüşüm için girdi olarak kullanılabilen bir pencere oluşturulur. Devrenin dijital doğası ve işlemin gerçekleştirilmesi için gereken hız göz önüne alındığında, 2'nin üssü olan birkaç ardışık örnek seçmek daha iyi olacaktır. Sonuç, daha hızlı bir algoritma olan, FFT ile hesaplanabilir. FFT işlemi için Xilinx firmasının IP (Inteligent Property) çekirdeği kullanılmıştır.

Şekil 7.3'te spektrogram görüntü elde edilmesi testlerinde kullanılan sistemin blok diyagramı görülmektedir. Spektrogram görüntü elde basamakları şu şekildedir:

- WM8731'in mikrofon girişine bir sinyal uygulanır.
- Pmod konnektörleri kullanılarak WM8731 ses kodeğinden gelen sinyaller okunur.
- Ses sinyali bir I2S kod çözücü vasıtasıyla çözülür.

- Genlik deęerinin belirlenen eşik deęerinin üzerinde olup olmadığı kontrol edilir.
- Genlik eşik deęerinin üzerinde ise 6 ms boyunca belleęe kaydedilir. Her örnek kanal başına 16 bit genişliğinde sinyallerden oluşur.
- Ses örneęi hafızası tamamen doldurulduğunda, bir hazır bayraęı üretilir ve numuneler teker teker okunarak bir hamming penceresi deęeri ile çarpılır. Bu işlem her kanal için ayrı ayrı yapılır.
- Ses örnekleri daha sonra FFT işlemine tabi tutulur.
- Elde edilen sonuçlar belleęe kaydedilir.
- Son olarak Genesys2 geliştirme kartı üzerindeki anahtar konumuna göre renk haritası seçilerek görüntü elde edilir.



Şekil 7.3. Test aşamasında kullanılan spektrogram resim oluşturma yönteminin blok diyagramı.

7.1.1.3. Ses Arabirimi

Ses arabirimi bileşeni (audio_interface), WM8731'i I2C veri yolunu kullanarak yapılandırma imkanı sağlar. Buna ek olarak, I2S veri yolunu kullanarak ses örneklerini alıp kodeęe göndermek de mümkündür. Bunun için takip edilen durum makinesine genel bir bakış Ek B.2'de bulunabilir.

I2C veri yolu üzerinden ayarlanabilen konfigürasyon kayıtları sayesinde ses kodek bileşeninin konfigürasyonu gerçekleştirilir. Doğru kaydedici ayarlarını bulabilmek için kodeğe ait veri yaprağına bakmak gerekir. Ses sinyallerinin düzgün bir konfigürasyonda istenilen genişlikte ve örnekleme hızlarında alınabilmesi için doğru bir konfigürasyonun yapılması gerekir. Ayrıca kodeğin master modunda mı yoksa slave modunda mı çalışacağına karar verilir. Yukarıdaki bölümlerde de belirtildiği gibi TDOA belirlenmesinde problem yaşanmaması için FPGA tarafı master modunda çalışmaktadır. Çizelge 7.1’de ses arabiriminde kullanılan varlık tanımlamasına ait her bir port tanımlaması ve işlevleri belirtilmiştir. Ayrıca ilgili portun bir saat sinyali, giriş portu mu veya bir çıkış portu mu olduğu görülmektedir. 6 kanalın tamamında da aynı port bilgileri kullanılmaktadır. Ayrıca buradaki örnekte mikrofonlardan alınan ses sinyalleri WM8731’in DAC girişine tekrar gönderilmektedir. Bu işlem sadece sistem testleri safhasında kullanılmaktadır.

Çizelge 7.1. Ses arabiriminde kullanılan varlık tanımlamasına ait port bilgileri ve işlevleri.

Sinyal	Tip	Açıklama
Clk_100M in	Saat	100 Mhz saat girişi
Clk_2M288	Saat	12M288 saat girişi
m_clk	Saat	WM8731 için üretilen ana saat.
lr_clk	Saat	WM8731 için sağ / sol kanal seçimi
b_clk	Saat	WM8731 için üretilen bit saati.
sdata	Giriş	WM8731’den gelen seri veri girişi
audio_sample_in	Giriş	Ses kanalı için 16 bit genişlikte bir ses örneği
I2C_addr	Çıkış	Yapılandırma kayıtlarını ayarlamak için üretilen I2C adresinin çıktısı
sample_clk	Çıkış	Her bir örnek hazır olduğunda bir bayrak üretilir
sample_l	Çıkış	Ses kanalı için 16 bit genişlikte bir ses çıkışı
Serial_data_out	Çıkış	Seri verilerin WM8731’e gönderilmesi.
sda	Çıkış	I2C arabirimi için seri veri hattı.
scl	Çıkış	I2C arayüzü için seri saat hattı.

WM8731 - Arayüzü

WM8731 arabirimi, ses örnekleri gönderme ve alma olanağı işlemlerini yürütmektedir. Ses kodek bileşeninden gelen seri veri, 16 bit genişliğinde bir sinyal haline dönüştürülecektir. Buna ek olarak, test işlemleri esnasında DAC'ye gönderilecek 16 bit genişlikte verilerde bir seri sinyal haline dönüştürülmektedir.

Çizelge 7.2. WM8731 kullanılan varlık tanımlamasına ait port bilgileri ve işlevleri.

Sinyal	Tip	Açıklama
lr_clk	Saat	Sol/sağ saat girişi
b_clk	Saat	Bit saat girişi
sdata	Giriş	WM8731'den seri veri girişi
audio_sample_in	Giriş	Ses kanalı için 16 bit genişlikte bir ses örneği
sample_clk	Çıkış	Her bir örnek hazır olduğunda bir bayrak üretilir
sample_1	Çıkış	Ses kanalı için 16 bit genişlikte bir ses çıkışı
serial data out	Çıkış	Seri verilerin WM8731'e çıkışı.

7.1.1.4. Bellek Kontrolü

Bellek kontrolü bileşeni, adreslerin üretilmesinden sorumludur. Bu adresler bellek bölgelerine veri yazmak ve almak için kullanılır. Yazma adresleri, yeni bir örnek okunduğunda oluşturulur. Eşik değerinin üzerinde genliğe sahip bir sinyal alındığında bunların sınıflandırma işlemi için 6 ms süreyle kayıt edilmesi gerekir. Bellekte her bir kanal için 16 bit genişliğinde 576 örnek hazır olduğunda, bir bayrak üretilir bu adresler örnekleri okumak ve onları hamming penceresine ve ardından FFT'ye iletmek için kullanılır.

7.1.1.5. Renk Haritası

Bu bileşen, aslında her piksel için renkleri RGB formatında oluşturmaktan sorumludur. Bu renk değerleri bir ROM belleğinden gerçek zamanlı olarak alınır. Ayrıca, resmin geçerli x ve y konumunu da takip eder. Geçerli x pozisyonunu veri haritalama ROM'unun adresi olarak kullanır. Bu değer daha sonra geçerli y konumu

ile karşılaştırılır. Buna bağlı olarak, rengin ne olacağına karar vermektedir. Çizelge 7.3'te renk haritasına ilişkin varlık tanımlaması görünmektedir.

Kullanıcı isteği doğrultusunda “summer” veya “hot” renk haritalarından birini FPGA üzerindeki anahtarı kullanarak seçebilmektedir.

ROM summer renk haritası: “summer” renk haritası için renk bilgisini içeren ROM belleğidir.

ROM hot renk haritası: “hot” renk haritası için renk bilgisini içeren ROM belleğidir.

Çizelge 7.3. Renk haritası.

Sinyal	Tip	Açıklama
clk	Saat	75 MHz saat girişi.
switch_color	Giriş	Renk haritası seçimi
data_data	Giriş	FFT RAM'a veri yolu.
r	Çıkış	Kırmızı kanal sinyali
g	Çıkış	Yeşil kanal sinyali
b	Çıkış	Mavi kanal sinyali
data_address	Output	FFT RAM'ına adres veri yolu.

7.1.2. CNN FPGA

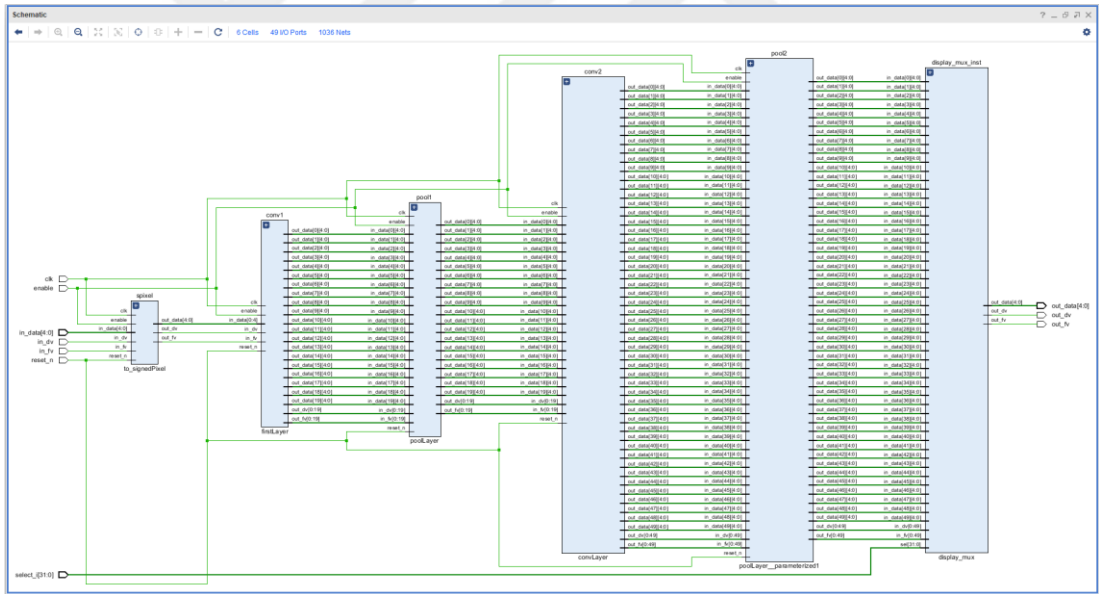
Yukarıdaki bölümlerde de belirtildiği gibi CNN'ler görüntü işleme uygulamalarında çok yüksek bir performans göstermektedirler. Fakat bu performans beraberinde çok yüksek bir hesaplama maliyeti getirmektedir. Literatürdeki pek çok yaklaşımda bu hesaplama maliyetinin üstesinden gelebilmek adına GPU tabanlı çalışmalar ortaya konmaktadır [114-116]. Bununla birlikte son yıllarda düşük enerji gereksinimleri, ve ızgara tabanlı bağımsız donanım yapılarının CNN tabanlı algoritmaların sergilediği kapsamlı eşzamanlılıktan faydalanmaktan için oldukça müsait olmaları nedeniyle CNN'lerin FPGA'ya uyarlanması yönünde çalışmalar yapılmaya başlanmıştır.

Birçok çalışmada CNN'lerin ileriye doğru yayılımını uygulamak için FPGA hesaplama gücünden yararlanılmaktadır. Bu çalışmalara ilişkin ayrıntılı bir değerlendirme [117]'de bulunmaktadır. Birçok yaklaşımda ara sonuçların bellekte tutulması temeline

dayanmaktadır [118]. Bu yaklaşımda bellekte okuma yazma işlemleri performans kaybına neden olmaktadır. Bu tez çalışmasında CNN'lerin ileri yayılım aşamasını FPGA üzerinde kullanabilmek için [119,120]'de önerilen veri akış modeli ve doğrudan donanıma eşleştirme yaklaşımı kullanılmıştır.

Doğrudan donanıma eşleştirme yaklaşımı karmaşık CNN mimarilerinin FPGA üzerinde uygulanmasında yüksek donanım maliyeti gerektirse de ara değerlerin bellekten okunmasına ve yazılmasına ihtiyaç duyulmaması nedeni ile gerçek zamanlı uygulamalardaki hız gereksinimleri açısından oldukça ideal bir yaklaşımdır. Doğrudan donanıma eşleştirme yaklaşımında tüm ağ bir donanım elemanı ile eşleştirilir ve ilave bellek kullanımı gerektirmeyen bir modeldir.

Şekil 7.4'de kullanılan CNN mimarisine ilişkin şematik görünüm bulunmaktadır.



Şekil 7.4. Silah sesine ait akustik olayların sınıflandırılmasında kullanılan CNN mimarisinin şematik görünümü.

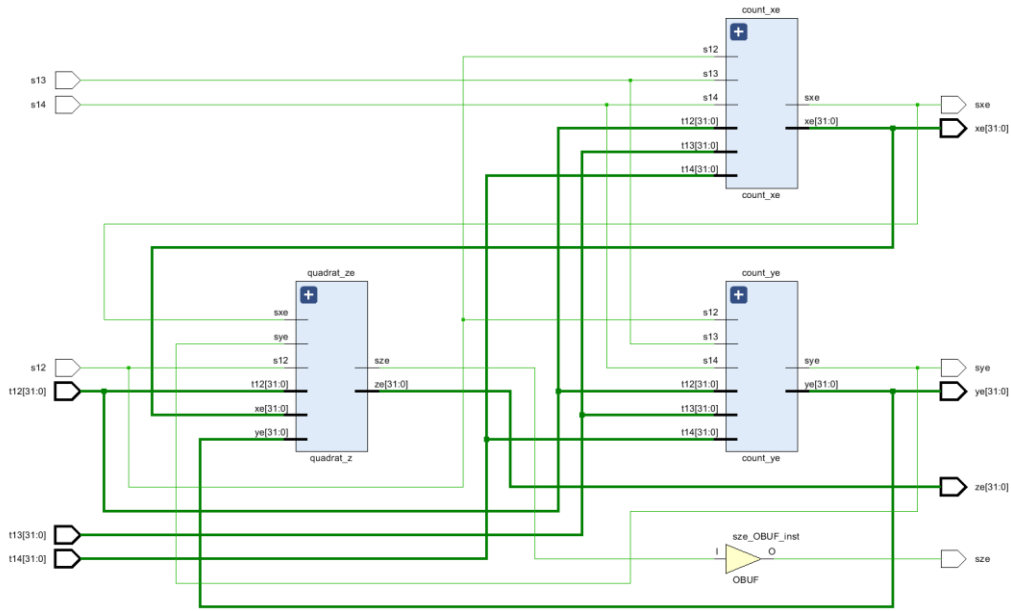
7.1.3. Konum Hesaplama

Sınıflandırma işlemi sonrasında aynı akustik olay tespit edilmişse konum belirleme adımına geçilmektedir. Bu aşamada öncelikli olarak her bir mikrofon çifti arasındaki

TDOA farkı bulunur. Bu işlem için GCC-Phat algoritması kullanılmaktadır. Sonrasında ŞD ve NP için DoA hesaplaması yapılır.

Konum hesaplama işlemi için ses hızı bilgisine ihtiyaç duyulmaktadır. Çevre koşullarını sürekli olarak değişmeyeceği varsayılarak bu işlem her 5 dakikada bir yapılmaktadır. Ses hızını hesaplayabilmek için FPGA her 5 dakikada bir RS232-USB (Universal Serial Bus) arayüzü üzerinden mikroişlemciye bir istek paketi gönderir. Mikroişlemci sensör verilerini belirli bir sıra halinde gönderir. FPGA aldığı her bilgi için bir doğrulama mesajı gönderir. Mikroişlemci doğrulama mesajını alamamış ise aynı paketi tekrar gönderir. Mikroişlemci ve FPGA arasında paket alışverişinde 5 denemeden fazla problem yaşanırsa ses hızını 340 m/s alarak konum hesaplamasını yapar.

Konum hesaplama işlemi sonucunda x-y-z uzay gösteriminde konum bilgisi elde edilir. Daha sonra konum bilgisi küresel koordinat sistemine çevrilir. Şekil 7.5'te küresel koordinat sistemine dönüşüm aşamasının şematik gösterimi görülmektedir.



Şekil 7.5. Kartezyen koordinat sistemindeki konum bilgisinin küresel koordinat sistemine dönüştüren sistemin şematik görünümü.

BÖLÜM 8

ATIŞ TESTLERİ

Atış testleri aşamasında G3 ve AK47 silahları kullanılarak her biri ile 20'şer atış olmak üzere toplam 40 atış gerçekleştirildi. Tüm atış testleri aynı gün içerisinde aynı atmosferik koşullar altında gerçekleştirildi. Tüm atışlarda atıcı konumu ile algılayıcı dizin arasındaki mesafe 100-150 m arasında değişecek şekilde atışlar gerçekleştirildi. Tek algılayıcı dizin kullanılması nedeniyle tüm atışlar mikrofonların ŞD'leri algılayacağı şekilde kaçış mesafesi 1-20 metre olacak şekilde gerçekleştirildi. Bunlara ilave olarak tüm atışlar tekli atış modunda gerçekleştirildi.

Yukarıdaki bölümlerde de belirtildiği gibi atış konum tespit sistemlerinin en önemli aşamalarından bir tanesi silah sesine ait akustik olayların başarılı bir şekilde tespit edilmesidir. Bununla birlikte literatürdeki birçok çalışma NP'lerin tutarlı bir şekilde tespit edilemediğini ortaya koymuşlardır [8,9,13,14,16,26,27]. Oysaki NP'nin algılanması tek algılayıcı dizin ile yapılan çalışmalar için olmazsa olmaz bir durumdur. Ayrıca sınıflandırma işlemi ile mermiye ilişkin parametrelerin tespit edilmesi hassas konum tespiti açısından son derece önemlidir. Her iki sinyalin başarılı bir şekilde sınıflandırılması ile birden fazla silah türü ile atış yapılan ortamlar için ciddi bir kabiliyet kazanımı sağlamaktadır. Aynı zamanda NP'nin yüksek gürültü altında da tespit edilebilirliğinin sağlanması ağ gereksinimlerinden bağımsız atıcı konum sistemlerinin kurulabilmesi açısından önemlidir. Bunlara ilave olarak tek algılayıcı dizinden oluşan sistemler mobilite açısından da askeri personele ciddi bir avantaj sağlamaktadır. Ayrıca geniş alan savunması için kullanılan ağ merkezli sistemlerde de farklı yönlerden aynı anda atış yapılması durumunda her bir algılayıcı dizinin aynı anda farklı silah türlerini algılayarak füzyon merkezine göndermesi muhtemeldir. Bu durumda silah sesine ait akustik olayların sınıflandırması ve veri füzyon birimine silah ve mermi türüne ilişkin bir parametre ile gönderilmesi konum tespit doğruluğunun

artırılması ve farklı silah türleriyle hatalı konum hesaplama işlemi yapılmaması açısından oldukça önemlidir.

Öte yandan literatürde kullanılan atıcı konum tespit sistemlerinde silah sesinin sınıflandırılması genellikle geniş bir algılayıcı dizin ağı kullanarak mermi hızına ilişkin parametrelerin elde edilmesi ile yapılmaktadır. Bu tez çalışmasında ise tek algılayıcı dizinden oluşan bir sistem sunulmaktadır. Tüm testler tek algılayıcı dizin ile gerçekleştirilmiş olmakla birlikte sistem geniş alan savunmasına da uyarlanabilecek durumdadır.

Tez çalışmasının bu kısmında Bölüm 3 ve Bölüm 4'te önerilen akustik olayların sınıflandırılması yaklaşımı gerçek atışlarla yapılan saha testleri ile kontrol edildi. Bölüm 4'te benzer şekilde gerçek atış kayıtlarından elde edilen sonuçlar sınıflandırma ve silah seslerinden elde edilen akustik olayların tespit edilebilirlikleri açısından değerlendirilmişti ve özellikle NP'lerin tespit edilebilirlikleri açısından literatüre kıyasla oldukça iyi sonuçlar elde edilmişti. Bu bölümde iki farklı silah türü ile yapılan 40 atışın her biri algılayıcı dizin üzerinde bulunan 6 mikrofonun tamamı tarafından başarılı bir şekilde tespit edildi ve ŞD ve NP'ler birbirinden başarılı bir şekilde ayırt edildi. Ayrıca sistem hiç bir yankıyı NP olarak algılamadı ve buna bağlı olarak hiç bir atışta yüksek hata oranı rapor edilmedi.

Bunlara ilave olarak tüm atış durumlarında ŞD'den mermi türü doğru bir biçimde sınıflandırıldı ve mermi boyu ve kalibre bilgisi tespit edilebildi. Benzer şekilde tüm atış koşullarında NP başarılı bir şekilde sınıflandırıldı ve silah türüne ilişkin bilgi elde edilebildi.

Bu tez çalışmasında önerilen akustik tabanlı atıcı konum tespit sisteminin performansı öncelikli olarak silah sesine ait akustik olayları tespit edilebilirliği ve sınıflandırma performansı açısından değerlendirildi. Tüm atışlarda hem ŞD'ler hem de NP'ler başarılı bir şekilde sınıflandırıldı. Her iki akustik olayında tek bir algılayıcı dizin ile tespit edilebilmesine bağlı olarak 40 atışın tamamından atıcı konumu hesaplanabildi.

Bu tez çalışmasında genel hatları itibariyle silah seslerine ait akustik olayların sınıflandırılması temeline dayanan bir yaklaşım önerilmiştir. Fakat konum hesaplama algoritmalarının geliştirilmesine odaklanılmamıştır. Önerilen yöntem silah sesine ait akustik olayların tespiti ve sınıflandırılması aşaması için literatüre kıyasla oldukça iyi sonuçlar vermiştir. Bu sonuçlar doğrultusunda konum tespit algoritmalarının geliştirilmesi faydalı olacaktır.

Yukarıda da belirtildiği gibi tüm atışlar en yüksek kaçış mesafesi 20 m olacak şekilde atışlar yapılmıştır. Kaçış mesafesinin hesaplanmasında ortalama mutlak hata 0,9 metrelik hata oranı ile literatüre benzer bir değer elde edilmiştir. Burada önerilen yaklaşım ile ŞD'nin tespit edilmesinde ve sınıflandırılmasında problem yaşanmazken. Önerilen yöntem ŞD'nin keskin kenarlarını tam olarak tespit ederek ŞD'ye ait dalga boyunun hesaplanması noktasında bir yenilik sunmadığı için sonuçlar literatür ile paralellik göstermektedir.

Bunun yanı sıra G3'e ait kaçış ortalama mutlak kaçış mesafelerinin AK47'ye oranla daha iyi bir değer sergilediği görülmüştür. G3 ile yapılan atışlarda ortalama 0,7 metrelik bir kaçış mesafesi hatası tespit edilmiştir. Bunun sonucu 7,62x51 mm'lik mühimmatın daha karakteristik bir sinyale sahip olması olarak değerlendirilmiştir.

Ayrıca konum tespit oranları açısından değerlendirilecek olursa sistem yanca ortalama 2.1°'lik bir hata oranı tespit edilmiştir. Ayrıca ortalama mesafe hatası %7,9 olarak gerçekleşmiştir.

Konum hesaplama testlerine ilave olarak G3 ve AK47 silahları ile sadece NP'nin tespit edilmesi durumunda varış doğrultusunun başarılı bir şekilde tespit edebildiğini kontrol etmek için algılayıcı dizin sadece NP'yi algılayabilecek şekilde konumlandırılarak her bir silah ile 5'er atış gerçekleştirildi. Bu testlerde aynı zamanda NP'nin uzak mesafelerden tespit edilebilirliğini de kontrol edebilmek adına atışlar 200 m mesafeden gerçekleştirildi. Atış sahasının kısıtlamaları nedeniyle daha uzun mesafeden atış gerçekleştirilemedi.

200 m mesafeden yapılan atışlarda da algılayıcı dizin üzerindeki her bir mikrofon tarafından NP'ler başarılı bir şekilde tespit edilebildi ve her bir mikrofon tarafından silah türü de başarılı bir şekilde sınıflandırıldı. Ayrıca varış doğrultusu hesabında yanca hata mesafe artışına ve ŞD'nin algılanmamasına bağlı olarak bir miktar artış kaydederek 3.5^o derece olarak gerçekleşti.

Sonuç olarak önerilen bu tez çalışmasında önerilen ses olayı tanıma yaklaşımı silah sesine ait akustik olayların sınıflandırılması açısından oldukça başarılı bir performans sergilemiştir. Ses kayıtları üzerinden yapılan tespit ve algılama başarısı saha testleri ile doğrulanmıştır.

Önerilen yöntem ile ŞD'ye ilişkin parametreler başarılı bir şekilde belirlenebilmektedir. Aynı zamanda NP'lerin orta ve uzun mesafeden yapılan atışlarda tek bir algılayıcı dizin ile algılanmasının sağlanması yerel durum farkındalığının artırılması açısından çok ciddi bir fayda sağlamaktadır.

BÖLÜM 9

SONUÇ VE ÖNERİLER

Bu tez çalışmasında silah sesine ait akustik bileşenleri SIF ve CNN'leri bir arada kullanarak sınıflandıran bir yaklaşım ortaya konulmuştur. Bu sayede özellikle SNR değerlerinin düşmesine bağlı olarak NP'lerin algılanamaması veya yankılar ile NP'lerin karıştırılmasından kaynaklanan yüksek konum belirleme hatalarının önüne geçilmesi veya hiç konum hesaplayamama durumunun ortadan kaldırılması hedeflenmiştir.

Bu doğrultuda 5 farklı silah ile 20'şer atış gerçekleştirilmiş ve her bir atış 2 farklı mikrofon ile kayıt edilerek, toplamda 200 ŞD, 200 NP ve 260 adet yankı içeren bir veri seti hazırlanmıştır. Ayrıca önerilen yöntemin performansı öncelikli olarak standart bir kıyaslama görevi kullanılarak literatürdeki diğer çalışmalar ile kıyaslanmıştır. Sonuç olarak önerilen yöntemin gürültü altındaki performansı literatürdeki diğer çalışmalardan çok daha iyi bir performans sergilemiştir. Farklı gürültü koşulları altında standart bir kıyaslama görevi üzerinde yapılan testlerde sınıflandırma başarımında ortalama %63,4'lük bağıl hata azaltma oranına eşdeğer %4,5'lik artış sağlamıştır. Öte yandan SIF özelliklerinin gürültü altındaki performansı güncel diğer tekniklere kıyasla çok daha iyi iken literatürde yapılan diğer çalışmalarda gürültüsüz ortamdaki sınıflandırma performansının literatürdeki güncel tekniklerde %7-8'e kadar daha düşük olabildiği rapor edilmiştir. Bu tez çalışmasında önerilen darbesel ses olayı tanıma yaklaşımı ile sınıflandırma performansı güncel teknikleri yakalamıştır. İlave olarak tüm gürültü koşullarındaki performansı ise aynı kıyaslama görevini kullanan çalışmaların üzerindedir. Hazırlanan veri seti ile yapılan silah sesine ait akustik olayların tespit edilmesi ve sınıflandırılması çalışmasında ŞD'ler her durumda başarılı bir şekilde tespit edilebilmiştir. ŞD ile yapılan sınıflandırma sonuçlarında ise %99,4'lük bir başarıım elde edilmiştir. Bu sonuçlar hem atıcı konum hesabının yapılabilmesi açısından hem de balistik ŞD modeli üzerinde mermi parametrelerinin

belirlenmesi açısından oldukça iyidir. Literatürdeki birçok çalışmada ŞD'ler başarılı bir şekilde tespit edilebilmektedir. Bununla birlikte sınıflandırma yaklaşımı ile mermi parametrelerinin de ön kabullere bağlı kalmadan tespit edilebilmesi sağlanmıştır. ŞD aksine NP'ler birçok çalışmada tutarlı bir biçimde tespit edilememiştir. Bu tez çalışmasında önerilen yaklaşım ile ise farklı gürültü koşulları altında ortalama %98,5'lik bir tespit edilebilme başarımı elde edilmiştir. Ayrıca sınıflandırma başarımında ise %96'luk bir performans sergilemiştir. Elde edilen sonuçlar NP'lerinde orta ve uzun mesafeli atışlarda başarılı bir şekilde tespit edilebileceğini ortaya koymaktadır. Ayrıca sonuçlar tek bir algılayıcı dizin ile konum tespiti ve varış doğrultusu hesaplanabilmesi açısından literatüre önemli bir katkı sağlamaktadır. Buna ilave olarak her iki akustik olayın başarılı şekilde sınıflandırılması ile birden fazla silah türü ile atış yapılan ortamlar için ciddi bir kabiliyet kazanımı sağlamaktadır. Ayrıca ağ gereksinimlerinden bağımsız tek algılayıcı dizinden oluşan atıcı konum sistemlerinin kurulabilmesi açısından önemlidir. Ağ merkezli geniş alan savunma sistemleri açısından da ilave doğrulama parametreleri sağlanması nedeniyle ciddi bir kabiliyet artışı sağlanmaktadır.

Son aşamada önerilen yaklaşım ile FPGA tabanlı bir donanım hazırlanmış ve sistem performansı tek sensor dizisi kullanılarak G3 ve AK47 silahları ile 20'şer atış yapmak suretiyle toplamda 40 atış ile test edilmiştir. Farklı mesafelerden yapılan tüm atışlar algılayıcı dizin üzerindeki tüm mikrofonlar ile başarılı bir şekilde tespit edilmiş ve doğru bir şekilde sınıflandırılmıştır.

Ayrıca 200 m mesafeden ŞD algılanamayacak şekilde G3 ve AK47 ile 5'er atış gerçekleştirilmiş ve sistemin uzun mesafeli atışlardaki NP tespit edebilme kabiliyeti gözlemlenmiştir. Her bir mikrofon NP'leri başarılı bir şekilde tespit etmiş ve doğru bir şekilde sınıflandırmıştır.

Bu tez çalışmasında literatürde özellikle NP'lerin tespit edilmesinde yaşanan problemler nedeniyle silah sesine ait akustik olayların tespit edilene odaklanılmıştır. Yeni bir konum tespit algoritması üzerinde durulmamıştır. Sınıflandırma yaklaşımı ile silah ve mermi türüne ilişkin parametreler tespit edilebilmesinden dolayı mermi

yavaşlama hızı gibi parametreleri de dikkate alan yeni bir konum hesaplama yaklaşımı geliştirilebilir.

Ayrıca sistem performansının daha fazla atış testiyle doğrulanması, yeni silah ve mermi türlerinin ilave edilmesi, çok daha uzun mesafelerden atışlarla çalışmanın değerlendirilmesi faydalı olacaktır. Buna ilave olarak gelecek çalışmalarda çoklu sensor dizileriyle veri füzyonuna odaklanması konum hesaplama doğruluğu açısından avantaj sağlayacaktır.

TDOA değerlerinin bulunması aşamasında da GCC-Phat algoritması yerine gürültüye karşı daha dirençli bir metot tercih edilebilir.

KAYNAKLAR

1. Reid, W. P., "Microphone array location from sounding made by a passing projectile", *Naval Surface Weapons Center, Silver Spring, MD*, (1975).
2. Fischer, J., Loges, W. and Meuser, W., "Method for determining artillery position", *Translation from German Patent no. DE-3239403-A1*, (1984).
3. Lahr, J. C. and Fischer, F. G., "Location of acoustic sources using seismological techniques and software", *US Geological Survey*, (1993).
4. Aguilar, J. R., "Gunshot location systems the transfer of the sniper detection technology from military to civilian applications", *In Security Technology (ICCST), 2013 47th International Carnahan Conference on IEEE*, 1-6, (2013).
5. Hillis, D. B., "Infrared sniper detection enhancement", *Google Patents*, U.S. Patent No. 5 686 889, (1997).
6. Moroz, S. A., Pierson, R. B., Ertem, M. C. and Burchick Sr, D. A. "Airborne deployment of and recent improvements to the viper counter sniper system", *Naval Research Lab Washington Dc*, (1999).
7. Agurok, I., Falicoff, W., Alvarez, R. and Shatford, W., "Passive electro-optical projectiles tracker", *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense XI*, Vol. 8359. International Society for Optics and Photonics, (2012).
8. Volgyesi, P., Balogh, G., Nadas, A., Nash, C. B. and Ledeczki, A., "Shooter localization and weapon classification with soldier-wearable networked sensors", *In Proceedings of the 5th international conference on Mobile systems, applications and services* (pp. 113-126), ACM, (2007)
9. Sallai, J., Hedgecock, W., Volgyesi, P., Nadas, A., Balogh, G. and Ledeczki, A., "Weapon classification and shooter localization using distributed multichannel acoustic sensors", *Journal of Systems Architecture*, 57.10: 869-885, (2011).
10. Ozugur, O., Sonmez, T., Basli, A. B. and Leblebicioğlu, M. K., "Atış tespit sistemleri için duyarlılık analizi", *In Signal Processing and Communications Applications Conference (SIU), 2015 23th*, 1709-1716, IEEE, (2015).
11. Özüğür, Ö. and Leblebicioğlu, M. K., "Akustik Algılayıcı Ağının Optimizasyonu ile Ateşli Silah Konumunun Tespit Edilmesi", *Journal of Defense Sciences/Savunma Bilimleri Dergisi*, 15.2, (2016).

12. Mäkinen, T., Pertila, P., "Shooter localization and bullet trajectory, caliber, and speed estimation based on detected firing sounds", *Applied acoustics*, 71.10: 902-913, (2010).
13. Lédeczi, Á., Völgyesi, P., Maróti, M., Simon, G., Balogh, G., Nádas, A. and Pap, G., "Multiple simultaneous acoustic source localization in urban terrain", *In Proceedings of the 4th international symposium on Information processing in sensor networks (p. 69)*, IEEE Press, (2005).
14. Valenzise, G., Gerosa, L., Tagliasacchi, M., Antonacci, F. and Sarti, A., "Scream and gunshot detection and localization for audio-surveillance systems", *Advanced Video and Signal Based Surveillance, AVSS 2007, IEEE Conference on. IEEE*, (2007).
15. Osborne, R. W., Bar-Shalom, Y., George, J. and Kaplan, L., "Data fusion from multiple passive sensors for multiple shooter localization via assignment", *Information Fusion (FUSION), 2014 17th International Conference on. IEEE*, (2014).
16. Sallai, J., Völgyesi, P., Lédeczi, Á., Pence, K., Bapty, T., Neema, S. and Davis, J. R., "Acoustic shockwave-based bearing estimation", *Proceedings of the 12th international conference on Information processing in sensor networks, ACM*, (2013).
17. Danicki, E., "The shock wave-based acoustic sniper localization", *Nonlinear analysis: theory, methods & applications*, 65(5), 956-962, (2006).
18. Agirman, H., Karakutuk, S. and Gonendik, E. "MIKES gunshot detection system prototype", *Signal Processing and Communications Applications Conference (SIU), 2014 22nd. IEEE*, (2014).
19. Page, E., Sharkey, B., "SECURES: system for reporting gunshots in urban environments", *Proc. SPIE*, 2497, 160-172, (1995).
20. Aguilar, J. R., Salinas, R. A. and Abidi, M. A., "Acoustical model of small calibre ballistic shock waves in air for automatic sniper localization applications", *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense VI., Vol. 6538, International Society for Optics and Photonics*, (2007).
21. Sinha, P., Kumar, S. and Hegde, R. M., "Sniper localization using passive acoustic measurements over an ad-hoc sensor array", *In Communications (NCC), 2015 Twenty First National Conference on (pp. 1-6)*, IEEE, (2015).
22. Volgyesi, P., Balogh, G., Nadas, A., Nash, C. B. and Ledeczi, A., "Shooter localization and weapon classification with soldier-wearable networked sensors", *In Proceedings of the 5th international conference on Mobile systems, applications and services (pp. 113-126)*, ACM, (2007).

23. Mazurek, J. A., Barger, J. E., Brinn, M., Mullen, R. J., Price, D., Ritter, S. E. and Schmitt, D., "Boomerang mobile counter shooter detection system", Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IV., Vol. 5778, *International Society for Optics and Photonics*, (2005).
24. Mäkinen, T., Pertilä, P. and Auranen, P., "Supersonic bullet state estimation using particle filtering", *In Signal and Image Processing Applications (ICSIPA), 2009 IEEE International Conference on (pp. 150-155)*, IEEE, (2009).
25. Libal, U. and Spyra, K., "Wavelet based shock wave and muzzle blast classification for different supersonic projectiles", *Expert Systems with Applications*, 41.11: 5097-5104, (2014).
26. Astapov, S., Berdnikova, J., Ehala, J., Kaugerand, J. and Preden, J. S., "Gunshot acoustic event identification and shooter localization in a WSN of asynchronous multichannel acoustic ground sensors", *Multidimensional Systems and Signal Processing*, 29.2: 563-595, (2018).
27. Ash, J. N., Whipps, G. T. and Kozick, R. J., "Performance of shockwave-based shooter localization under model misspecification", *Acoustics Speech and Signal Processing (ICASSP)*, 2010 IEEE International Conference on. IEEE, (2010).
28. Freire, I. L. and Apolinário, J. A., "GCC-based DoA estimation of overlapping muzzleblast and shockwave components of gunshot signals", *Circuits and Systems (LASCAS)*, 2011 IEEE Second Latin American Symposium on. IEEE, (2011).
29. George, J. and Kaplan, L. M., "Shooter localization using soldier-worn gunfire detection systems", Information Fusion (FUSION), *2011 Proceedings of the 14th International Conference on. IEEE*, (2011).
30. George, J., Kaplan, L. M., Deligeorges, S. and Cakiades, G., "Multi-shooter localization using finite point process", Information Fusion (FUSION), *2014 17th International Conference on. IEEE*, (2014).
31. Damarla, T., Kaplan, L. M. and Whipps, G. T., "Sniper localization using acoustic asynchronous sensors", *IEEE Sensors Journal*, 10.9: 1469-1478, (2010).
32. Millet, J. and Baligand, B., "Latest achievements in gunfire detection systems", France, (2006).
33. Akman, Ç., Sönmez, T., Özüğür, Ö., Başlı, A. B. and Leblebicioğlu, M. K., "Sensor fusion, sensitivity analysis and calibration in shooter localization systems", *Sensors and Actuators A: Physical*, 271: 66-75, (2018).
34. Maher, Robert C., "Acoustical characterization of gunshots.", *Signal Processing Applications for Public Security and Forensics 2007 (SAFE'07)*, IEEE Workshop on. IEEE, (2007).

35. Khalid, M. A., Babar, M. I. K., Zafar, M. H. and Zuhairi, M. F., "Gunshot detection and localization using sensor networks", *Smart Instrumentation, Measurement and Applications (ICSIMA)*, 2013 IEEE International Conference on. IEEE, (2013).
36. Bédard, J., "Performance metrics for acoustic small arms localization systems", *Defence Research And Development Canada Valcartier (QUEBEC)*, (2006).
37. Brustad, B. M. and Freytag, J. C., "A survey of audio forensic gunshot investigations", *Audio Engineering Society Conference: 26th International Conference: Audio Forensics in the Digital Age*, Audio Engineering Society, (2005).
38. Weissler, P. G. and Kobal, M. T., "Noise of police firearms", *The Journal Of The Acoustical Society Of America*, 56 (5): 1515–1522, (1974).
39. Whitham, G. B., "The flow pattern of a supersonic projectile", *Communications On Pure And Applied Mathematics*, 5 (3): 301–348, (1952).
40. Sadler, B. M., Pham, T. and Sadler, L. C., "Optimal and wavelet-based shock wave detection and estimation", *The Journal Of The Acoustical Society Of America*, 104 (2): 955–963, (1998).
41. Duckworth, G. L., Gilbert, D. C. and Barger, J. E., "Acoustic counter-sniper system", *Command, control, communications, and intelligence systems for law enforcement.*, Vol. 2938, International Society for Optics and Photonics, (1997).
42. Bandi, A. K., Rizkalla, M. and Salama, P., "A novel approach for the detection of gunshot events using sound source localization techniques", *Circuits and Systems (MWSCAS), 2012 IEEE 55th International Midwest Symposium on*. IEEE, (2012).
43. Lindgren, D., Wilsson, O., Gustafsson, F. and Habberstad, H., "Shooter localization in wireless microphone networks.", *Eurasip journal on advances in signal processing*, (2010).
44. Yuldashev, P., Ollivier, S., Averiyarov, M., Sapozhnikov, O., Khokhlova, V. and Blanc-Benon, P., "Nonlinear propagation of spark-generated N -waves in air: Modeling and measurements using acoustical and optical methods", *The Journal Of The Acoustical Society Of America*, 128 (6): 3321–3333, (2010).
45. Maher, R. C., "Modeling and signal processing of acoustic gunshot recordings." *Digital Signal Processing Workshop, 12th-Signal Processing Education Workshop, 4th. IEEE*, (2006).
46. Wold, E., Blum, T., Keislar, D. and Wheaton, J., "Content-based classification, search, and retrieval of audio", *IEEE multimedia*, 3(3), 27-36, (1996).

47. Weninger, F. and Schuller, B., "Audio recognition in the wild: Static and dynamic classification on a real-world database of animal vocalizations", *Acoustics, Speech And Signal Processing (ICASSP), 2011 IEEE international conference on. IEEE*, (2011).
48. Ghiurcau, M. V., Rusu, C., Bilcu, R. C. and Astola, J., "Audio based solutions for detecting intruders in wild areas", *Signal Processing*, 92(3), 829-840, (2012).
49. Rabaoui, A., Davy, M., Rossignol, S. and Ellouze, N., "Using one-class SVMs and wavelets for audio surveillance", *IEEE Transactions on information forensics and security*, 3(4), 763-775, (2008).
50. Chu, S., Narayanan, S. and Kuo, C. C. J., "Environmental sound recognition with time–frequency audio features", *IEEE Transactions on Audio, Speech, and Language Processing*, 17(6), 1142-1158, (2009).
51. Cowling, M. and Sitte, R., "Comparison of techniques for environmental sound recognition", *Pattern recognition letters*, 24(15), 2895-2907, (2003).
52. Walters, T., "Auditory-based processing of communication sounds", PhD Thesis, *University of Cambridge*, (2011).
53. Dennis, J., Tran, H. D. and Li, H., "Spectrogram image feature for sound event classification in mismatched conditions", *IEEE signal processing letters*, 18(2), 130-133, (2011).
54. Guo, G. and Li, S. Z., "Content-based audio classification and retrieval by support vector machines", *IEEE transactions on Neural Networks*, 14(1), 209-215, (2003).
55. Portelo, J., Bugalho, M., Trancoso, I., Neto, J., Abad, A. and Serralheiro, A., "Non-speech audio event detection", *In Acoustics, Speech and Signal Processing, 2009. ICASSP 2009. IEEE International Conference on* (pp. 1973-1976), IEEE, (2009).
56. Atrey, P. K., Maddage, N. C. and Kankanhalli, M. S., "Audio based event detection for multimedia surveillance", *In Acoustics, Speech and Signal Processing, 2006. ICASSP 2006 Proceedings. 2006 IEEE International Conference on* (Vol. 5, pp. V-V), IEEE, (2006).
57. Sidiropoulos, P., Mezaris, V., Kompatsiaris, I., Meinedo, H., Bugalho, M. and Trancoso, I., "On the use of audio events for improving video scene segmentation.", *In Image Analysis for Multimedia Interactive Services (WIAMIS), 2010 11th International Workshop on* (pp. 1-4), IEEE, (2010).
58. McLoughlin, I., Zhang, H., Xie, Z., Song, Y. and Xiao, W., "Robust sound event classification using deep neural networks", *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, 23(3), 540-552, (2015).

59. Dennis, J. W., "Sound event recognition in unstructured environments using spectrogram image processing", *Nanyang Technological University*, Singapore, (2014).
60. Sharan, R. V. and Moir, T. J., "Noise robust audio surveillance using reduced spectrogram image feature and one-against-all SVM", *Neurocomputing*, 158, 90-99, (2015).
61. Xie, Z., McLoughlin, I., Zhang, H., Song, Y. and Xiao, W., "A new variance-based approach for discriminative feature extraction in machine hearing classification using spectrogram features", *Digital Signal Processing*, 54, 119-128, (2016).
62. Ferrari, A., Lombardi, S. and Signoroni, A., "Bacterial colony counting with convolutional neural networks in digital microbiology imaging", *Pattern Recognition*, 61, 629-640, (2017).
63. Segalin, C., Cheng, D. S. and Cristani, M., "Social profiling through image understanding: Personality inference using convolutional neural networks", *Computer Vision and Image Understanding*, 156, 34-50, (2017).
64. Li, H., Li, Y. and Porikli, F., "Convolutional neural net bagging for online visual tracking", *Computer Vision and Image Understanding*, 153, 120-129, (2016).
65. Sun, J., Xiao, Z. and Xie, Y., "Automatic multi-fault recognition in TFDS based on convolutional neural network", *Neurocomputing*, 222, 127-136, (2017).
66. Krizhevsky, A., Sutskever, I. and Hinton, G. E., "Imagenet classification with deep convolutional neural networks", *In Advances in neural information processing systems (pp. 1097-1105)*, (2012).
67. Hinton, G. E. and Sejnowski, T. J., "Learning and relearning in Boltzmann machines", *Parallel distributed processing: Explorations in the microstructure of cognition*, 1, 282-317, (1986).
68. Ozer, I., Ozer, Z. and Findik, O., "Lanczos kernel based spectrogram image features for sound classification", *Procedia computer science*, 111, 137-144, (2017).
69. LeCun, Y., Bottou, L., Bengio, Y. and Haffner, P., "Gradient-based learning applied to document recognition", *Proceedings of the IEEE*, 86(11), 2278-2324, (1998).
70. Fukushima, K., Miyake, S. and Ito, T., "Neocognitron: A neural network model for a mechanism of visual pattern recognition", *IEEE transactions on systems, man, and cybernetics*, (5), 826-834, (1983).
71. Cireşan, D., Meier, U. and Schmidhuber, J., "Multi-column deep neural networks for image classification", *arXiv preprint arXiv:1202.2745*, (2012).

72. Ma, J., Wu, F., Zhu, J., Xu, D. and Kong, D., "A pre-trained convolutional neural network based method for thyroid nodule diagnosis", *Ultrasonics*, 73, 221-230, (2017).
73. Nair, V. and Hinton, G. E., "Rectified linear units improve restricted boltzmann machines", *In Proceedings of the 27th international conference on machine learning (ICML-10)*, 807-814, (2010).
74. Ciresan, D. C., Meier, U., Masci, J., Maria Gambardella, L. and Schmidhuber, J., "Flexible, high performance convolutional neural networks for image classification", *In IJCAI Proceedings-International Joint Conference on Artificial Intelligence*, Vol. 22, No. 1, p. 1237, (2011).
75. Yu, J., Tao, D. and Wang, M., "Adaptive hypergraph learning and its application in image classification", *IEEE Trans. Image Processing*, 21(7), 3262-3272, (2012).
76. Nian, F., Li, T., Wang, Y., Xu, M. and Wu, J., "Pornographic image detection utilizing deep convolutional neural networks", *Neurocomputing*, 210, 283-293, (2016).
77. Yu, S., Jia, S. and Xu, C., "Convolutional neural networks for hyperspectral image classification", *Neurocomputing*, 219, 88-98, (2017).
78. Lopes, A. T., de Aguiar, E., De Souza, A. F. and Oliveira-Santos, T., "Facial expression recognition with convolutional neural networks: coping with few data and the training sample order", *Pattern Recognition*, 61, 610-628, (2017).
79. Pinheiro, P. H. and Collobert, R., "Recurrent convolutional neural networks for scene labeling", *In 31st International Conference on Machine Learning (ICML)* (No. EPFL-CONF-199822), (2014).
80. Abdel-Hamid, O., Mohamed, A. R., Jiang, H. and Penn, G., "Applying convolutional neural networks concepts to hybrid NN-HMM model for speech recognition", *In Acoustics, Speech and Signal Processing (ICASSP), 2012 IEEE International Conference on* (pp. 4277-4280). IEEE, (2012).
81. Sainath, T. N., Mohamed, A. R., Kingsbury, B. and Ramabhadran, B., "Deep convolutional neural networks for LVCSR", *In Acoustics, speech and signal processing (ICASSP), 2013 IEEE international conference on* (pp. 8614-8618). IEEE, (2013).
82. Zhang, H., McLoughlin, I. and Song, Y., "Robust sound event recognition using convolutional neural networks", *In Acoustics, Speech and Signal Processing (ICASSP), 2015 IEEE International Conference on* (pp. 559-563). IEEE, (2015).

83. Nakamura, S., Hiyane, K., Asano, F., Nishiura, T. and Yamada, T., "Acoustical sound database in real environments for sound scene understanding and hands-free speech recognition", (2000).
84. Varga, A. and Steeneken, H. J., "Assessment for automatic speech recognition: II. NOISEX-92: A database and an experiment to study the effect of additive noise on speech recognition systems", *Speech communication*, 12(3), 247-251, (1993).
85. Szegedy, C., Liu, W., Jia, Y., Sermanet, P., Reed, S., Anguelov, D. and Rabinovich, A., "Going deeper with convolutions", *In Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1-9), (2015).
86. Russakovsky, O., Deng, J., Su, H., Krause, J., Satheesh, S., Ma, S. and Berg, A. C., "Imagenet large scale visual recognition challenge", *International Journal of Computer Vision*, 115(3), 211-252, (2015).
87. Erhan, D., Manzagol, P. A., Bengio, Y., Bengio, S. and Vincent, P., "The difficulty of training deep architectures and the effect of unsupervised pre-training", *In Artificial Intelligence and Statistics* (pp. 153-160), (2009).
88. Ozer, I., Ozer, Z. and Findik, O., "Noise robust sound event classification with convolutional neural network", *Neurocomputing*, 272, 505-512, (2018).
89. Madhukar, B. N. and Narendra, R., "Lanczos resampling for the digital processing of remotely sensed images", *In Proceedings of International Conference on VLSI, Communication, Advanced Devices, Signals & Systems and Networking (VCASAN-2013)* (pp. 403-411). Springer, India, (2013).
90. Turkowski, K., "Filters for common resampling tasks", *In Graphics gems* (pp. 147-165). Academic Press Professional, Inc., (1990, August).
91. Glassner, A. S., Ed., "Graphics gems", *Elsevier*, (2013).
92. Kiktova, E., Lojka, M., Pleva, M., Juhar, J. and Cizmar, A., "Gun type recognition from gunshot audio recordings", *In Biometrics and Forensics (IWBF), 2015 International Workshop on* (pp. 1-6), IEEE, (2015).
93. Özseven, T. and Düğenci, M., "Silah Sesleri Kullanılarak Ateşli Silahların Sınıflandırılmasında Akustik Parametrelerin Etkisi", *Gazi Üniversitesi Fen Bilimleri Dergisi Part C: Tasarım ve Teknoloji*, 6(2), 462-469.
94. John, S. M., "Acoustic speech source localization in a constrained search space", (2010).
95. Scola, C. F. and Ortega, M. D. B., "Direction of arrival estimation: A two microphones approach", (2010).

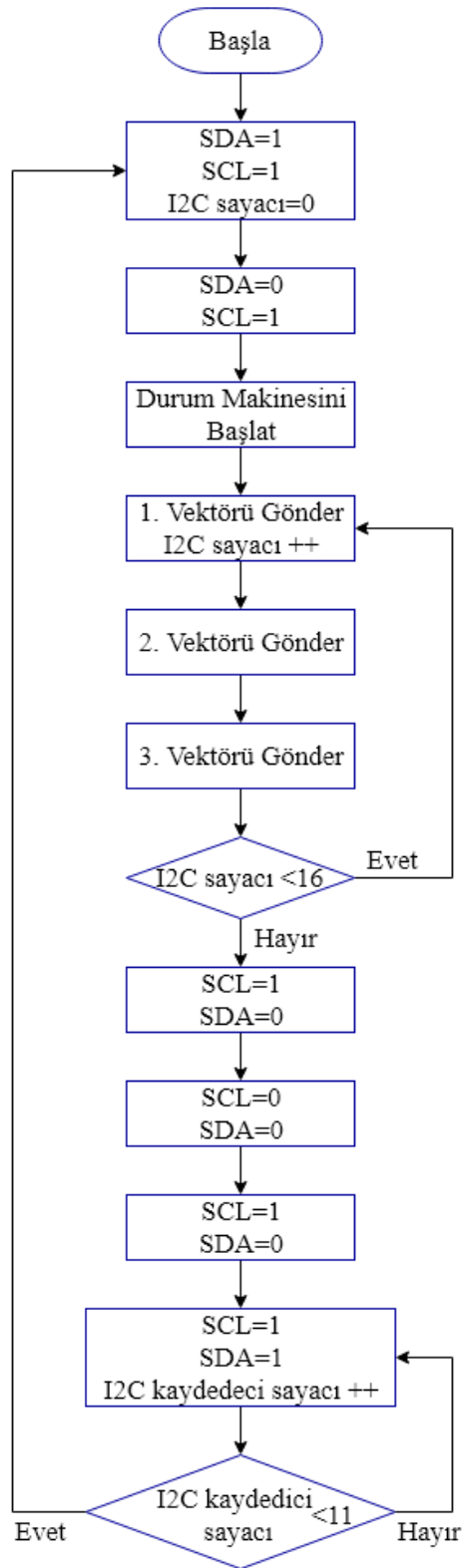
96. Svaizer, P., Matassoni, M. and Omologo, M., "Acoustic source location in a three-dimensional space using crosspower spectrum phase", *In Acoustics, Speech, and Signal Processing, 1997. ICASSP-97., 1997 IEEE International Conference on* (Vol. 1, pp. 231-234), IEEE, (1997).
97. Alameda-Pineda, X., Horaud, R. and Mourrain, B., "The geometry of sound-source localization using non-coplanar microphone arrays", *In Applications of Signal Processing to Audio and Acoustics (WASPAA), 2013 IEEE Workshop on* (pp. 1-4). IEEE, (2013).
98. Akman, Ç., "Multi Shooter Localization With Acoustic Sensors", PhD Thesis, *The Graduate School Of Natural and Applied Sciences of Middle East Technical University*, (2017).
99. Carlucci, D. E. and Jacobson, S. S., "Ballistics: theory and design of guns and ammunition", *CRC Press*, (2018).
100. Stoughton, R., "Measurements of small-caliber ballistic shock waves in air", *The Journal of the Acoustical Society of America*, 102(2), 781-787, (1997).
101. Todman, T. J., Constantinides, G. A., Wilton, S. J., Mencer, O., Luk, W. and Cheung, P. Y., "Reconfigurable computing: architectures and design methods", *IEE Proceedings-Computers and Digital Techniques*, 152(2), 193-207, (2005).
102. Erkol, H. O., "Robot Kinematik Denklemlerinin FPGA ile Çözülmesi ve çok Eklemlili Bir Robota Uygulanması", Doktora Tezi, *Karabük Üniversitesi Fen Bilimleri Enstitüsü*, Karabük, (2015).
103. Tessier, R. and Burleson, W., "Reconfigurable computing for digital signal processing: A survey", *Journal of VLSI signal processing systems for signal, image and video technology*, 28(1-2), 7-27, (2001).
104. Chu, P. P., "FPGA prototyping by Verilog examples: Xilinx Spartan-3 version", *John Wiley & Sons*, (2011).
105. Smith, G., "FPGAs 101: Everything you need to know to get started", *Newnes*, (2010).
106. Kuon, I., Tessier, R. and Rose, J., "FPGA architecture: Survey and challenges", *Foundations and Trends® in Electronic Design Automation*, 2(2), 135-253, (2008).
107. Wolf, W. H., "FPGA-Based System Design", *Prentice Hall PTR*, New Jersey, (2004).

108. Yıldırım, Ö., Erişti, B., Erişti, H. ve Demir, Y., "Gerçek zamanlı akım ve gerilim sinyallerinin izlenmesi için FPGA tabanlı bir sistem tasarımı", *ELOCO, 2012 Elektrik-Elektronik ve Bilgisayar Mühendisliği Sempozyumu*, ss. 654-658, Bursa, (2012).
109. Carter, W. S., "The future of programmable logic and its impact on digital system design", *In Proceedings 1994 IEEE International Conference on Computer Design: VLSI in Computers and Processors*, 10-16, IEEE, (1994).
110. Hauck, S. and DeHon, A., "Reconfigurable computing: the theory and practice of FPGA-based computation (Vol. 1)", *Elsevier*, (2010).
111. Lacey, G., Taylor, G. W. and Areibi, S., "Deep learning on fpgas: Past, present, and future", *arXiv preprint arXiv:1602.04283*, (2016).
112. Güneren, H., "FPGA Üzerinde Kayan Nokta Sayı Formatı Kullanılarak Yapay Sinir Ağı Tabanlı Sınıflandırma İşlemi", Bitirme Projesi, *Yıldız Teknik Üniversitesi*, 44 s, İstanbul, (2010).
113. Şahin, İ., Koyuncu, İ., "A new module design for 3D graphic transformations using generated floating-point core units.", *I. Rev. On Modelling and Sim.*, 4(2):691-698, (2011).
114. Canziani, A., Paszke, A. and Culurciello, E., "An analysis of deep neural network models for practical applications", *arXiv preprint arXiv:1605.07678*, (2016).
115. Zhang, C., Wu, D., Sun, J., Sun, G., Luo, G. and Cong, J., "Energy-efficient CNN implementation on a deeply pipelined FPGA cluster", *In Proceedings of the 2016 International Symposium on Low Power Electronics and Design*, 326-331, ACM, (2016).
116. Chetlur, S., Woolley, C., Vandermersch, P., Cohen, J., Tran, J., Catanzaro, B. and Shelhamer, E., "cudnn: Efficient primitives for deep learning", *arXiv preprint arXiv:1410.0759*, (2014).
117. Coates, A., Huval, B., Wang, T., Wu, D., Catanzaro, B. and Andrew, N., "Deep learning with COTS HPC systems", *In International Conference on Machine Learning*, 1337-1345, (2013).
118. Abdelouahab, K., Pelcat, M., Sérot, J., Bourrasset, C. and Berry, F., "Tactics to directly map CNN graphs on embedded FPGAs", *IEEE Embedded Systems Letters*, 9(4), 113-116, (2017).
119. Abdelouahab, K., Pelcat, M., Sérot, J., Bourrasset, C., Quinton, J. C. and Berry, F., "Hardware Automated Dataflow Deployment of CNNs", *arXiv preprint arXiv:1705.04543*, (2017).



EK AÇIKLAMALAR A.

AKIŞ DİYAGRAMI



Şekil Ek A.1. Ses ara yüzüne ait akış diyagramı.

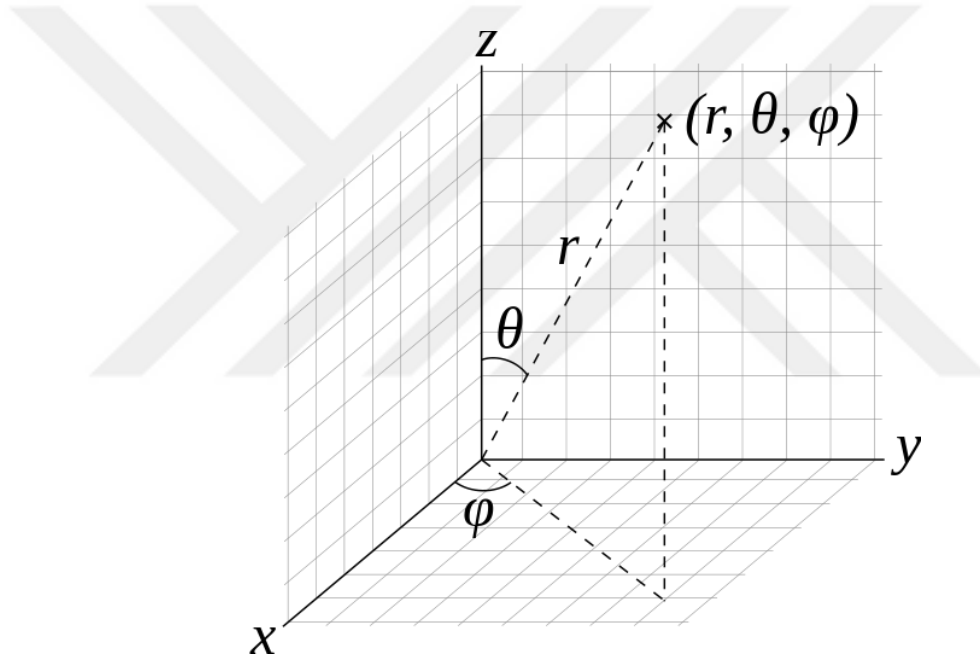


EK AÇIKLAMALAR B.

KÜRESEL KOORDINAT SİSTEMİ

Küresel Koordinat Sistemi

Tez çalışmasının bu kısmında çalışma boyunca kullanılan küresel koordinat sistemi tanıtılmaktadır. Şekil Ek 2.1’de de görüleceği üzere küresel koordinat sistemi üç boyutlu koordinat sisteminde her bir noktanın üç farklı değer ile ifade edildiği bir koordinat sistemidir. “r” terimi merkez noktaya olan radyal mesafeyi ifade ederken θ terimi sabit bir zenit doğrultusundan ölçülen kutupsal açıyı ifade eder. Son olarak φ terimi ise azimut açısını ifade etmektedir. φ açısının değeri $[0 \ 2\pi]$ aralığında değişmektedir. Öte yandan θ açısının değeri ise $[-\pi/2 \ \pi/2]$ aralığında değişmektedir.



Şekil Ek B.1. Koordinat sistemi.

Tez çalışmasında elde edilen konum bilgileri küresel koordinat sistemi ile gösterilirken DoA birim vektörleri ise Kartezyen koordinat sistemi ile temsil edilmektedirler. Bu nedenden dolayı her iki koordinat sistemi arasında dönüşüm yapılması gerekliliği doğmaktadır. Eşitlik 1 ve Eşitlik 2 sırasıyla Kartezyen koordinat sisteminden küresel koordinat sistemine ve küresel koordinat sisteminden Kartezyen koordinat sistemine olan dönüşümleri göstermektedir.

$$\vec{P}(x_p, y_p, z_p) \rightarrow \vec{P}(\varphi_p, \theta_p, R_p)$$

$$\begin{aligned}
 \varphi_p &= \tan^{-1}\left(\frac{y_p}{x_p}\right), \quad \varphi_p \in [0, 2\pi] \\
 \theta_p &= \sin^{-1}\left(\frac{z_p}{\sqrt{x_p^2 + y_p^2 + z_p^2}}\right), \quad \theta_p \in [-\pi/2, \pi/2] \\
 R &= \sqrt{x_p^2 + y_p^2 + z_p^2}
 \end{aligned}
 \tag{1}$$

$$\begin{aligned}
 \vec{P}(\varphi_p, \theta_p, R_p) &\rightarrow \vec{P}(x_p, y_p, z_p) \\
 x_p &= R \cos(\theta_p) \cos(\varphi_p) \\
 y_p &= R \cos(\theta_p) \sin(\varphi_p) \\
 z_p &= R \sin(\theta_p)
 \end{aligned}
 \tag{2}$$

ÖZGEÇMİŞ

İlyas ÖZER 1984 yılında Yozgat'ta doğdu; ilk ve orta öğrenimini Ankara'da tamamladı. Çağrıbey Anadolu Lisesi'nden mezun oldu. 2002 yılında Dumlupınar Üniversitesi Mühendislik-Mimarlık Fakültesi Elektrik-Elektronik Mühendisliği Bölümü'nde öğrenime başlayıp 2006 yılında mezun oldu. 2007-2008 yılları arasında askerlik hizmetini tamamlayıp. 2008 yılında Ülker Grubu şirketlerinden Della Gıda bünyesinde üretim mühendisi olarak göreve başladı. 2010 yılında Ülker Grubu'ndaki görevinden ayrılarak aynı yıl Türktelekom A.Ş. Ankara Merkez I Bölge Müdürlüğü'nde santral ve NGN mühendisi olarak göreve başladı. 2011 yılında Sakarya Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar ve Bilişim Mühendisliği Anabilim Dalı'nda yüksek lisans eğitimini tamamladı. 2013 yılından itibaren Türktelekom A.Ş. bünyesindeki görevini Karabük Telekom Müdürlüğü bünyesinde sırasıyla santral ve NGN mühendisi ve erişim ekip lideri olarak 2018 yılı Ağustos ayına kadar sürdürdü.

ADRES BİLGİLERİ

Adres : Kartaltepe Mah. Cevizkent Toki Konutları
C4-1 Blok, 14. Kat, No:62
Merkez / KARABÜK
Tel : (507) 751 7576
E-posta : ilyasozer@outlook.com