T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

TEK MİKROFON DİZİSİ İLE YAPAY ZEKA DESTEKLİ SİLAH SESİ KONUMLAMA YÖNTEMİ

Kazım ZENGİN

DOKTORA TEZİ
Mekatronik Mühendisliği Anabilim Dalı
Mekatronik Mühendisliği Programı

Danışman
Prof. Dr. Aydın YEŞİLDİREK

Ekim, 2024

T.C.
YILDIZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

TEK MİKROFON DİZİSİ İLE YAPAY ZEKA DESTEKLİ SİLAH SESİ KONUMLAMA YÖNTEMİ

Kazım ZENGİN tarafından hazırlanan tez çalışması 24.10.2024 tarihinde aşağıdaki jüri tarafından Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü Mekatronik Mühendisliği Anabilim Dalı Mekatronik Mühendisliği Programı **DOKTORA TEZİ** olarak kabul edilmiştir.

Prof. Dr. Aydın YEŞİLDİREK
Yıldız Teknik Üniversitesi
Danışman

Jüri Üyeleri

Prof. Dr. Aydın YEŞİLDİREK, Danışman
Yıldız Teknik Üniversitesi _____

Prof. Dr. Hüseyin Metin ERTUNÇ, Üye
Kocaeli Üniversitesi _____

Prof. Dr. Cenk ULU, Üye
Yıldız Teknik Üniversitesi _____

Prof. Dr. Duygun EROL BARKANA, Üye
Yeditepe Üniversitesi _____

Prof. Dr. Ahmet KOYUN, Üye
Yıldız Teknik Üniversitesi _____

Danışmanım Prof. Dr. Aydın YEŞİLDİREK sorumluluğunda tarafımca hazırlanan Tek Mikrofon Dizisi İle Yapay Zeka Destekli Silah Sesi Konumlama Yöntemi başlıklı çalışmada veri toplama ve veri kullanımında gerekli yasal izinleri aldığımı, diğer kaynaklardan aldığım bilgileri ana metin ve referanslarda eksiksiz gösterdiğimi, araştırma verilerine ve sonuçlarına ilişkin çarpıtma ve/veya sahtecilik yapmadığımı, çalışmam süresince bilimsel araştırma ve etik ilkelerine uygun davrandığımı beyan ederim. Beyanımın aksinin ispatı halinde her türlü yasal sonucu kabul ederim.

Kazım ZENGİN

İmza

Aileme
ve
Eşime

TEŞEKKÜR

Ders dönemimden, tezimin tamamlanmasına kadar geçen sıkıntılı süreçlerimde benimle hoşgörü ve sabırla ilgilenenen her açıdan desteğini esirgemeyen beni daha başarılı bir akademisyen olmam yönünde eğiten, tezim süresince bilgi ve tecrübelerinden sürekli faydalandığım, değerli danışman hocam Prof. Dr. Aydın YEŞİLDİREK'e,

Lisans eğitimimden bugüne kadar bilgi ve desteklerini esirgemeyen, tezimin gelişimine katkılar sunan değerli hocalarım Prof. Dr. Hasan OCAK'a, Prof. Dr. H. Metin ERTUNÇ'a,

Desteğini esirgemeyen, tezimin gelişimine katkılar sunan değerli hocalarım Prof. Dr. Duygun EROL BARKANA'ya, Prof. Dr. Cenk ULU'ya, Prof. Dr. Ahmet KOYUN'a,

Doktora sürecimde idari desteğini esirgemeyen bölüm başkanım Doç. Dr. Hayrettin TOYLAN'a,

Veri toplama sürecimde bana destek olan Çağdaş Detay GEZGİN'e ve kıymetli dostum Aykut HAVSUT'a,

Beni sevgi ve saygı içinde büyüten maddi ve manevi olarak her zaman yanımda olan değerli, annem Hülya ZENGİN, babam İsmail ZENGİN, ve ablam Şule Azime YENİÇERİ'ye

Tez sürecimde desteğini ve sabrını hiçbir zaman esirgemeyen değerli eşim Arzu ZENGİN'e Sonsuz teşekkürü bir borç bilirim.

Kazım ZENGİN

İÇİNDEKİLER

SİMGE LİSTESİ viii

KISALTMA LİSTESİ x

ŞEKİL LİSTESİ xii

TABLO LİSTESİ xv

ÖZET xvii

ABSTRACT xix

1 GİRİŞ 1

1.1 Literatür Özeti . . . 1

1.1.1 Akustik Konumlamada ve Silah Patlaması Tespit Sistemlerinde Yapay Zeka . . . 7

1.2 Tezin Amacı . . . 9

1.3 Hipotez . . . 10

1.4 Tezin İçeriği . . . 11

2 SİLAH SESİ KONUMLAMA YÖNTEMLERİ 12

2.1 Ses Fiziği . . . 12

2.1.1 Sesin Şiddeti ve Algılanması . . . 13

2.2 Geliş Açısı Tespit Yöntemleri . . . 14

2.2.1 Direkt Yön Hesaplama Yöntemleri . . . 15

2.2.2 Uzak Alan Yaklaşımı ve TDoA Hesaplama . . . 18

2.3 Konum Belirleme Yöntemleri . . . 22

2.3.1 Üçgenleme Yöntemi . . . 22

2.3.2 Multilaterasyon Yöntemi . . . 24

2.3.3 Balistik Akustik Model . . . 26

3 SES TANIMLAMADA KULLANILAN ÖZELLİKLER 30

3.1 Zamansal Özellikler . . . 30

3.1.1 Genlik Zarfı . . . 30

3.1.2 RMS Enerjisi . . . 31
3.1.3 Sıfır Geçiş Oranı . . . 32
3.2 Spektral Özellikler . . . 32
3.2.1 Band Enerji Oranı . . . 34
3.2.2 Spektral Merkez . . . 34
3.2.3 Spektral Entropi . . . 35
3.2.4 Spektral Yayılım-Band Genişliği . . . 36
3.2.5 Spektral Skewness . . . 36
3.2.6 Sprektral Kurtosis . . . 37
3.3 Zaman-Frekans Özellikleri . . . 38
3.3.1 Duyusal Özellikler . . . 39
3.3.2 Mel-Ölçeği . . . 39
3.3.3 Mel Spektrogram . . . 40
3.3.4 Mel Frekans ve Gamaton Kepstal Katsayıları . . . 41
3.3.5 3.3.5 Bark Ölçeği ve Bark Spektrogram . . . 42
3.3.6 ERB Ölçeği ve Spektrumu . . . 43
3.3.7 Delta ve Delta Delta Özelliği . . . 43
3.4 En Küçük Fazlalık En Büyük Alaka (MRMR) Özellik Seçim Yöntemi 44

4 DERİN ÖĞRENME YÖNTEMLERİ 45
4.1 Çok Katmanlı Algılayıcı . . . 46
4.2 Evrişimsel Sinir Ağları . . . 47

5 MULTİLATERASYON YÖNTEMİNDE DUYARLILIK ANALİZİ VE KAPALI FORM DoA FORMÜLLERİ 49
5.1 Eş Merkezli ve Dikey Yerleşimli Mikrofon Çiftleri için Kapalı Geliş Yönü Tespit Çözümü . . . 53
5.2 Önerilen Yöntemin Diğer Yöntemlere Göre Başarısı . . . 56

6 DENEYSEL ÇALIŞMALAR 62
6.1 Deney Sahasının Oluşturulması ve Kullanılan Donanımlar . . . 62
6.2 Deneysel Verilerle DoA Elde Etme . . . 63
6.3 Deneysel Veriler İle Multilaterasyon Çözümü Sonuçları . . . 66
6.4 Derin Öğrenme Verilerinin Oluşturulması . . . 67
6.5 Çok Katmanlı Algılayıcı İle Mesafe Tahmini . . . 69
6.6 Konvolisyonel Ağlar İle Mesafe Tahmini . . . 71
6.6.1 Özellik Seçiminin Tahminleme Doğruluğuna Etkisi . . . 73
6.7 Makine Öğrenmesi Yöntemlerinin Başarısı . . . 76

7 SONUÇ 78

KAYNAKÇA **81**

TEZDEN ÜRETİLMİŞ YAYINLAR **91**

SİMGE LİSTESİ

Pr_{atm}	Atmosfer Basıncı
θ	Azimuth Açısı
Pr	Basınç
φ	Elevasyon Açısı
f	Frekans
τ	Geliş Zamanı Farkı
P_i	i. Mikrofon Konumu
Pr_0	İnsan Kulağının Duyabileceği En Düşük Basınç Değeri
P_k	Kaynak Konumu
P	Konum
M	Mach Sayısı
v_m	Mermi Hızı
ϕ_b	Mermi Kalibresi
l_b	Mermi Uzunluğu
d_m	Mikrofonlar Arası Mesafe
x_n	n. Mikrofondan Alınan Ses Sinyali
W	Nöronlar Arası Ağırlık
Pr_s	Patlamanın Maksimum Basınç Değeri
T_0	Pozitif Faz Gecikme Süresi
v_s	Sesin Havadaki Yayıyım Hızı
T	Sıcaklık
v_0	Sıfır Derece Sıcaklıkta Sesin Havadaki Yayılım Hızı
α	Şok Dalgası Tepe Açısı

α_r	Öğrenme Oranı
N	Örnek Sayısı
F_s	Örnekleme Frekansı
t_s	Örnekleme Zamanı
b	Üstel Zayıflama Oranı
t	Zaman

KISALTMA LİSTESİ

ANN	Yapay Sinir Ağı (Artificial Neural Network)
AOA	Geliş Açısı
CNN	Konvolisyonel Sinir Ağı
DOA	Geliş Açısı Doğrultusu
ESPRIT	Dönme Değişmezliği Teknikleri ile Sinyal Parametreleri Tahmini
FFT	Hızlı Fourier Dönüşümü
GCC	Genelleştirilmiş Çapraz Korelasyon
GMM	Gauss Karışım Modelleri
GNNS	Küresel Konum Belirleme Sistemi
GPR	Gauss Süreci Tahminlemesi
GPS	Küresel Konumlama Sistemi
GPU	Grafik İşleme Birimi
HMM	Saklı Markov Model
ILD	Kulaklararası Seviye Farkı
IPD	Kulaklararası Faz Farkı
ITD	Kulaklararası Zaman Farkı
LSE	En Küçük Karesel Hata
k-NN	K En Yakın Komşu
MAE	Ortalama Mutlak Hata
MFCC	Mel Frekansı Keptral Katsayıları
ML	Maksimum Benzerlik
MLP	Çok Katmanlı Algılayıcı
MSE	Ortalama Karesel Hata (Mean Square Error)

MUSIC	Çoklu İşaret Sınıflama
PDA	Kişisel Veri Yardımcısı (Personal Data Assistant)
PHAT	Faz Dönüşümü
RBFNN	Radyal Taban Fonksiyonlu Sinir Ağı
RMS	Ortalama Karekök
RMSE	Ortalama Kareköksel Hata
RNN	Tekrarlayan Sinir Ağı
SCOT	Yumuşatılmış Tutarlılık Dönüşümü
SGD	Olasılıksal Değişim Düşümü
SGO	Sıfır Geçiş Oranı
SNR	Sinyal Gürültü Oranı
SPL	Ses Basınçı Seviyesi
SRP	Yönlendirilmiş Yanıt Gücü (Streed Response Power)
STFT	Kısa Süreli Fourier Dönüşümü
SVM	Destek Vektör Makinası
TDOA	Geliş Süreleri Farkı
TOA	Geliş Süresi

ŞEKİL LİSTESİ

Şekil 1.1 Dünya savaşında kullanılan Sound Ranging sistemi 1
Şekil 1.2 Sound Ranging kullanılarak elde edilen çizim 2
Şekil 1.3 (a) Namlu patlaması, (b) şok dalgası 3
Şekil 2.1 Mikrofonların algılama yönleri 14
Şekil 2.2 Mikrofonlar ile geliş açısı tespiti ve mikrofon dizisi yapısı 19
Şekil 2.3 (a) Düzlemde kullanılan mikrofon dizisi, (b) Uzayda kullanılan mikrofon dizisi 20
Şekil 2.4 Mikrofonlar arası TDoA simülasyonu 23
Şekil 2.5 Üçgenleme yöntemi ile konumlama 23
Şekil 2.6 AK-47 silahından 1600 m mesafeden yapılan atışın ses grafiği . . 26
Şekil 2.7 Ses üstü merminin şok ve namlu uçu patlaması dalgaları ve yayılım geometrisi 27
Şekil 3.1 Faklı sesler için genlik zarfı grafikleri 31
Şekil 3.2 Farklı sesler için RMS enerji grafikleri 31
Şekil 3.3 Farklı sesler için SGO grafikleri 32
Şekil 3.4 Kısa zamanlı Fourier dönüşümü 33
Şekil 3.5 Farklı sesler için band enerji oranı grafikleri 34
Şekil 3.6 Farklı sesler için spektral merkez grafikleri 35
Şekil 3.7 Spektral entropi grafiği 36
Şekil 3.8 Farklı sesler için speaktral yayılım grafikleri 36
Şekil 3.9 Farklı sesler için spektral skewness grafikleri 37
Şekil 3.10 Farklı sesler için speaktral kurtosis grafikleri 37
Şekil 3.11 Namlu patlama sesinin sprektrogram resmi 38
Şekil 3.12 Mel ölçeği-frekans grafiği 39
Şekil 3.13 0-7000 Hz arası oluşturulmuş 18 adet Mel filtreleri 40
Şekil 3.14 Namlu patlaması sesinin Mel sprektrogram resmi 41
Şekil 3.15 MFCC oluşturma aşamaları 41
Şekil 3.16 14 Katsayılı MFCC grafiği 42
Şekil 4.1 Tek algılayıcılı ağ modeli 45
Şekil 4.2 Çok katmanlı algılayıcı ağ modeli 46
Şekil 4.3 İlk defa önerilen evrişimsel ağ modeli 48
Şekil 5.1 Simulasyon çalışmalarında kullanılan TDoA süresi modeli 49

Şekil 5.2 Benzetim çalışmalarında kullanılan mikrofon dizisi yapısı 50
Şekil 5.3 Benzetim çalışmasında kullanılan mikrofon dizisi konumu ve patlama noktaları dağılımı 50
Şekil 5.4 Analitik hesaplanan TDoA süresiyle yapılan simulasyon ile elde edilen konumlama ve hata grafiği 51
Şekil 5.5 1 MHz F_s ile hesaplanan TDoA süresiyle yapılan simulasyon ile konumlama ve hata grafiği 51
Şekil 5.6 Mikrofon arası mesafe artışına göre konumlama tahmini değişimi 53
Şekil 5.7 Ortogonal yerleşimli sensör çiflerinden oluşan mikrofon dizileri . 56
Şekil 5.8 d_k=100 m,F_s=44.1 kHz durumunda mikrofon çiftleri arasındaki mesafenin etkisi (a) d_m=0.1 m, (b) d_m=0.5 m, (c) d_m=1 m, (d) d_m=1.5 m 57
Şekil 5.9 d_m=0.5 m,F_s=44.1 kHz durumunda kaynak uzaklığının etkisi-(a) d_k=10 m, (b) d_k=500 m 57
Şekil 5.10 $d_m \in [0.1, 1]$ m, $d_k \in [5, 100]$ m F_s=44.1 kHz durumunda 0-90 derece arası ortalama hatanın değişimi 58
Şekil 5.11 d_m=0.5 m, d_k=100 m durumunda örnekleme frekansının etkisi-(a) F_s=22.05 kHz, (b) F_s= 96 kHz 58
Şekil 5.12 d_m=0.5 m,d_k=100 m F_s=44.1 kHz durumunda 0-90 derece Arası tahmin hataları-(a) hatası eklemesiz, (b) hata eklemeli 59
Şekil 5.13 Eşit olmayan Mikrofonlar arası Mesafenin Etkisi (a) Hata Eklemesiz, (b) Hata Eklemeli 60
Şekil 6.1 Önerilen hibrit model 62
Şekil 6.2 Deney sahası ve kullanılan mikrofon dizisinin yerleşimi 63
Şekil 6.3 (a) Mikrofon dizisi açı ayarlama düzeneyi, (b) patlama yüksekliğinin ayarlanması 63
Şekil 6.4 Deney sahasından ölçülen azimuth açısının yayılım grafiği (a) kapalı form, (b) uzak alan 65
Şekil 6.5 Deney sahasından ölçülen elevasyon açısının yayılım grafiği (a) kapalı form, (b) uzak alan 66
Şekil 6.6 Deneysel TDoA değerleri ile elde edilen multilaterasyon yöntemi sonuçları 67
Şekil 6.7 Sıfır derece atışında +Z yönündeki mikrofondan kayıt edilen kırpılmış farklı mesafe sesleri 68
Şekil 6.8 Farklı mesafelerdeki seslerin oluşturduğu Mel sprektrogram resimleri 69
Şekil 6.9 Ses kayıtlarından veri seti dönüşümü lle MLP ağının giriş vektörü 69
Şekil 6.10 Geliştirilen MLP mimarisi 69
Şekil 6.11 Seçilen ağın tahminleme grafikleri (a) eğitim, (b) test 70

Şekil 6.12 Ses kayıtlarından veri seti dönüşümü ile CNN ağının giriş matrisi 71
Şekil 6.13 Matlab programında oluşturulan ve eğitilen CNN ağı 71
Şekil 6.14 Seçilen ağın tahminleme grafikleri (a) eğitim, (b) test 72
Şekil 6.15 Oluşturulan konvolüsyonel ağın %RMSE tahmin hatasının mesafelere göre değişimi 73
Şekil 6.16 MRMR algoritması ile elde edilen en etkili 16 özelllik 74
Şekil 6.17 Özellik seçimi ile eğitilen CNN ağı 75
Şekil 6.18 Özellik seçimi ile eğitilen CNN ağının tahminleri (a)- eğitim, (b)-test 75
Şekil 6.19 Özellik seçimi ile eğitilen CNN ağının mesafelere göre ortalama tahminleme hataları 76
Şekil 6.20 En başarılı makine öğrenmesi yöntemi tahmin grafikleri 76
Şekil 6.21 En başarılı makine öğrenmesi yönteminin tahmin başarısı 77

TABLO LİSTESİ

Tablo 1.1 BNN Sistem tarafından geliştirilen sistem performansı 3
Tablo 1.2 Ölçülenbilen ve tahminlenen büyüklükler 4
Tablo 1.3 Algılanan dalga türleri ve tahminlenen konum değişkenleri . . . 4
Tablo 1.4 Dalga çeşiti ve mikrofon dizisi sayısına göre tahmin edilen konum değişkenleri 5
Tablo 1.5 Ticari sistemler 6
Tablo 1.6 Balistik akustik model parametreleri hataları 7
Tablo 2.1 Sesler ve gürültü değerleri 13
Tablo 2.2 Ağırlık fonksiyonları ve etkileri 22
Tablo 5.1 Farklı mikrofon dizisi yarıçapları ve örnekleme frekansları için elde edilen % RMS hata oranları 52
Tablo 5.2 Normal dağılımlı hata eklenmiş gecikme süreleriyle elde edilen %RMS hata oranları 52
Tablo 5.3 d_m=0.5 m, d_k=100 m F_s=44.1 kHz 0-90 derece arası tahmin hatalarının ortalamala değerleri 59
Tablo 5.4 Farklı mikrofon dizisi geometrilerinin 0-90 derece arası tahmin hatalarının ortalamala değerleri 60
Tablo 6.1 Deneysel ölçümlerde elde edilen TDoA değerleri [τ Örnek Adeti, θ Derece] 64
Tablo 6.2 Azimuth açısı için Şekil 6.4-a'da gösterilen her mikrofon açısında elde edilen hata değerleri 65
Tablo 6.3 Azimuth açısı için Şekil 6.4-b'de gösterilen her mikrofon açısında elde edilen hata değerleri 65
Tablo 6.4 Elevasyon açısı için Şekil 6.5-a'da gösterilen her mikrofon açısında elde edilen hata değerleri 66
Tablo 6.5 Elevasyon açısı için Şekil 6.5-b'de gösterilen her mikrofon açısında elde edilen hata değerleri 66
Tablo 6.6 Şekil 6.6'de gösterilen konumlama verileri ve hata yüzdeleri . . . 67
Tablo 6.7 Farklı nöron sayılarına göre elde edilen ortalama mesafe hatası . 70
Tablo 6.8 En başarılı sonuçları elde ettiğimiz ağın mesafelere göre hata değerleri 71
Tablo 6.9 Ağ ve eğitim parametreleri 72

Tablo 6.10 Oluşturduğumuz konvolüsyonel ağın mesafelere göre hata değerleri 73
Tablo 6.11 Seçilen özellikler ile eğitilen CNN ağının mesafelere göre hata değerleri 74
Tablo 6.12 Seçilen özellikler ile eğitilen makine öğrenmesi yönteminin sonuçları 77
Tablo 7.1 Mesafe tahmini % başarı oranları 79

ÖZET

Tek Mikrofon Dizisi İle Yapay Zeka Destekli Silah Sesi Konumlama Yöntemi

Kazım ZENGİN

Mekatronik Mühendisliği Anabilim Dalı
Doktora Tezi

Danışman: Prof. Dr. Aydın YEŞİLDİREK

Silah konumlama sistemleri sivil ve askeri alanlarda sıklıkla kullanılmakta ve önemi giderek artmaktadır. Akustik tabanlı silah sesi konumlama sistemleri mikrofon dizileri vasıtasıyla algılanan patlama seslerini kullanarak patlamanın yönü, uzaklığı ya da kartezyen koordinatları gibi sonuçlar üretmektedir. Mevcut yöntemler patlama sırasında oluşan namlu patlaması ve şok dalgası seslerinin yayılımlarını modelleyerek sonuçlar üretmektedir. İnsanlar kulakları ile yön tespiti yapmasının yanında öğrenilmiş seslerin mesafelerini de tecrübeye dayalı belirleyebilmektedir. Bu gerçeklikten yola çıkarak akustik tabanlı derin öğrenme destekli hibrit bir silah konumlama yöntemi geliştirilmiştir. Ayrıca eşmerkezli ve dik yerleşimli mikrofon çiftlerinden oluşan sensor dizileri için kapalı-form geliş açısı hesaplama formülleri türetilmiştir. Tez, simulasyon çalışmaları ile geliş açısı formüllerinin türetilmesini, yapay zeka için kullanılacak gerçek seslerden elde edilen veri setinin oluşturulmasını ve geliştirilen yapay sinir ağları, özellik seçimi ve makine öğrenmesi ile elde edilen sonuçları kapsar. Elde edilen sonuçlar mevcut konumlama yöntemlerine göre 100 m'den daha uzak mesafelerdeki patlama sesleri için mevcut yöntemlerden çok daha başarılı mesafe tahminin yapılabildiğini mesafe arttıkça başarı oranının da arttığını göstermiştir. Mevcut literature göre; eşmerkezli ve dik yerleşimli sensor dizileri için yön tahminini algoritmik çıkartıma itiyaç duymadan yapabilen kapalı form formüllerin türetilmesi, sadece namlu patlaması sesi kullanarak tek mikrofon dizisi mesafe tahminin yapılabilmesi, ses yayılım modellerinden farklı olarak derin öğrenme destekli konumlama yapılması, mevcut sistemlerde en az iki mikrofon dizisi kullanılarak elde edilen

konumlama başarısından daha iyi sonuçların tek mikrofon dizisi ile elde edilmesi ve çalışmalarda kullanılmak üzere veri seti oluşturulması bu çalışmanın literature katkılarını oluşturmaktadır.

Anahtar Kelimeler: Silah Sesi Konumlama, Kapalı-Form Geliş Yönü, Hibrit Konumlama Modeli, Derin Öğrenme.

ABSTRACT

A Single Microphone Array-Based AI-Assisted Gunshot Sound Localization Method

Kazım ZENGİN

Department of Mechatronics Engineering
Doctor of Philosophy Thesis

Supervisor: Prof. Dr. Aydın YEŞİLDİREK

Weapon location systems are widely used in civilian and military applications and are becoming increasingly important. Acoustic-based gunfire localization systems use the sound of explosions detected by microphone arrays to produce results such as the direction, distance or Cartesian coordinates of the explosion. Existing methods generate results by modeling the propagation of muzzle blast and shockwave sounds during the explosion. Humans can determine the distances of learned sounds based on experience as well as direction determination with their ears. Based on this reality, an acoustic-based deep learning supported hybrid weapon positioning method has been developed. In addition, closed-form angle of incidence calculation formulas are derived for sensor arrays consisting of concentric and orthogonal microphone pairs. The thesis covers the derivation of the angle of incidence formulas through simulation studies, the creation of a dataset obtained from real sounds to be used for artificial intelligence, and the results obtained with the developed neural networks, feature selection and machine learning. The obtained results show that the distance estimation is much more successful than the existing positioning methods for explosion sounds at distances greater than 100 m, and the success rate increases as the distance increases. According to the liteatüre, the contributions of this study include; derivation of closed form formulas for direction estimation for concentric and orthogonal sensor arrays without the need for algorithmic inference, distance estimation with single microphone array using only muzzle blast sound, deep learning supported positioning unlike sound propagation models, achieving better positioning results with a single microphone

array than is achieved with existing systems by using at least two microphone arrays, and creating a data set to be used in studies.

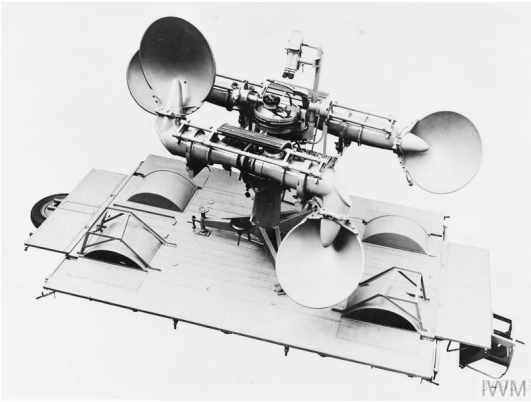
Keywords: Gunfire Localization, Closed-Form Direction of Arrival, Hybrid Localization Model, Deep Learning.

1
GİRİŞ

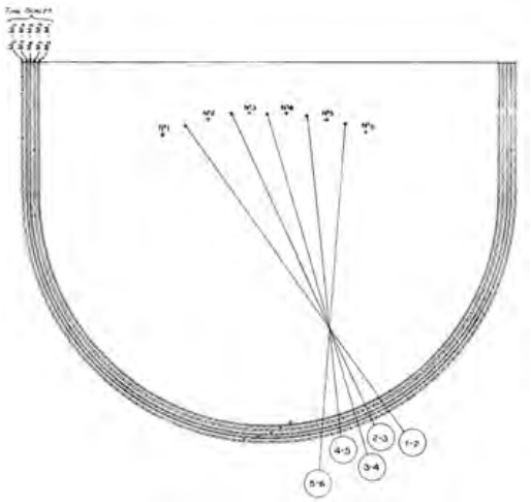
1.1 Literatür Özeti

Ses lokalizasyonu ilk olarak askeri alanda kullanılmak üzere ortaya çıkmış ve daha sonra ve görüntüleme sistemleri günümüzde başta askeri alanda, şehir içi güvenlik sistemleri, konferans sistemleri, mobil robot sistemleri gibi alanlarda sıklıkla kullanılmaktadır. Özellikle askeri alanlarda açılan ateş yönünün tespiti, coğrafi bilginin ya da konuçlanma bilgisinin olmadığı alanlarda veya gözetlemenin zor olduğu bölgelerde oluşabilecek saldırıların konumlandırılması açısından giderek artan bir öneme sahiptir. Şekil 1.1'de görülen farklı konumlara yerleştirilmiş mikrofonlara sesin gelme süresini kullanarak konum bulma sistemini ilk olarak 1914 yılında Charles Nordmann tasarlamış ve birinci dünya savaşında kullanılmıştır [1]. "Sound Ranging" adı verilen sistem ikili mikrofon çiftleri arasındaki zaman gecikmesine göre çizilen hiperbollerin kesişimine göre Şekil 1.2'de gösterildiği gibi pozisyon bulmaktaydı ve top bataryalarının yerlerini tespit etmede kullanılmıştır [2, 3] .

1990'lı yılların başlarından sonra ise gelişen teknoloji ile birlikte silah patlama sesi ve sniper konum tespiti çalışmaları başlamıştır. Bu konumlandırma sistemlerinin ilk uygulamalarından birisi Secures şehir içi güvenlik sistemidir [4]. Bu sistem



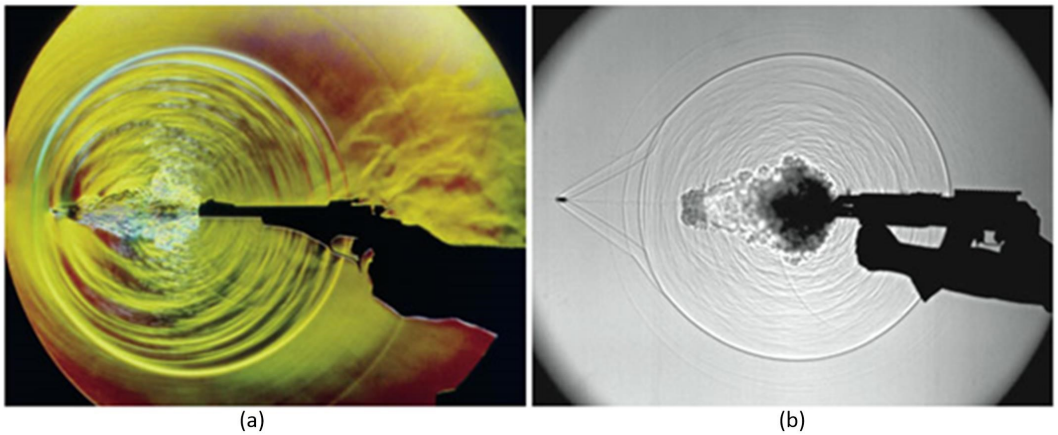
Şekil 1.1 Dünya savaşında kullanılan Sound Ranging sistemi



Şekil 1.2 Sound Ranging kullanılarak elde edilen çizim

binaların konulan akustik sensörler ile silah sesini diğer seslerden ayırarak her hangi alarm durumunda polis merkezine haber vermektedir. Benzer şekilde Shotspotter sistemi de geliştirilmiştir. 1995'den beri kullanılan ve geliştirilen bu sistemler 25 m doğrulukta konumlama yapabilmektedir [5, 6]. Keskin nişancı konumlama sistemleri de 90'lı yılların başlarında çalışılmaya başlanmıştır. Bu alanda ilk çalışmalar Amerika, Kanada, Fransa gibi ülkelerin savunma bakanlıklarının sponsorluklarında ilgili ülkelerin savunma sanayi teknolojileri üreten şirketleri tarafından yapılmıştır. Silahlar yapısına ve kullanılan mühimmata göre farklı patlama sesleri üretirler. Bunlardan ilki barut patlaması sonucu oluşan basınçlı gazların namlu ağzından dışarı çıkarken oluşturdukları namlu patlaması sesidir ayrıca silahın kurma yayı çekiç mekanizması sesi gibi mekanik seslerini de içermektedir ve tüm silahlar bu sesi üretirler [7]. Susturucu takılmadıkça baskılanması zor bir sestir. Diğeri ise ses üstü hızlarda namludan çıkan merminin ses duvarını delmesi ile oluşturduğu şok dalgası sesidir. Bu ses mermi hızı ses hızının altına düşene kadar oluşmaktadır. Şekil 1.3-a'da Sclieren tekniği ile elde edilmiş namlu patlaması ses dalgası yayılımı, Şekil 1.3-b'de ise gölge tekniği ile elde edilmiş şok dalgası sesi yayılımı görülmektedir [8].

DARPA'nın 1995 yılında keskin nişancı tespit sistemlerine sponsor oldu ve 1997 biri Israil'li, beşi ABD'li olmak üzere 6 şirketin geliştirdiği sistemlerin testlerini yayınladığı çalışma ve ürün geliştiren şirketlerin çalışma raporları ilk keskin nişancı konumlama üzerine yayınlanmış çalışmalardandır [9–12]. Tek ya da çoklu mikrofon dizileri ile namlu patlaması ve şok dalgası sesleri kullanılarak atıcının kullandığı silahın kalibresi, atış noktasının azimuth ve elevasyon açıları ile mesafesinin belirlenmesi ve bu verilere göre atıcı konumunun iki boyutlu saha haritasında gösterilmesi hedeflenmiştir. Katılan altı firmadan BNN firmasının



Şekil 1.3 (a) Namlu patlaması, (b) şok dalgası

prototipi doğruluk, boyut, maliyet gibi başarı kriterlerine göre ilk sırada yer almıştır [9]. BNN Systems and Technologies firması aynı yıl yayımladıkları çalışma raporunda parametrik mermi yayılım modelini kullanarak mesafe tahmini yaptıklarını belirtmiştir. İki adet dört yüzlü sensor düğümü yerden 1.5 m yüksekliğe yerleştirilerek ve aralarındaki mesafe 40 m olacak şekilde ayarlanarak yapılan saha testlerinde 3 farklı kalibredeki (22, 35, 50) silah ile 167 toplam atış sonucu geliştirdikleri sistemin başarı oranları Tablo 1.1'de verilmiştir [10].

Tablo 1.1 BNN Sistem tarafından geliştirilen sistem performansı

Parametre	Başarı Koşulu	Başarı Oranı
Mesafe Tahmini Hatası	<%1 <%5 <%10	%28 %60 %70
Azimuth Hatası	<1° <5° <10°	%72 %93 %96
Elevasyon Hatası	<1° <5°	%38 %91
Kalibre Doğruluğu	-	%90

Sistem geliştirilerek üretilen Boomerang ve Bullet Ears isimli ticari sistemler Afganistan ve Irak harekatların sıklıkla kullanılmıştır [13, 14]. Firma askeri kompozit başlık üzerine üzerine 12 küçük mikrofon, gps, yönelim sensörü, kablosuz haberleşme modülü ekleyerek asker üzerine giyilebilir konumlama sistemi de geliştirmiştir. Farklı askerler üzerinde bulunan giyilebilir sistemler merkez ünitede birleştirilerek çoklu sistem olarak da kullanıldığında sabit sisteme yakın performans

sergilemiştir [15]. Aynı testlerde ikinci sırada olan SAIC firması da aynı yıl kendi çalışmalarını raporlamışlardır [11]. Aralarında 1 m açıklık bulunan toplam 16 kondanser mikrofondan oluşan iki adet volimetrik sensor dizisi kullanılmıştır. Bu sensor dizileri ile ölçülebilecek büyüklükler ve yapılabilecek kestirimler Tablo 1.2'de verilmiştir.

Tablo 1.2 Ölçülenbilen ve tahminlenen büyüklükler

Gözlemler	Türetilen Büyüklükler
Şok dalgası DoA Şok dalgası geliş zamanı Namlu patlaması DoA Namlu patlaması geliş zamanı Şok dalgası periyodu Şok dalgası genliği	Atıcı azimuth açısı Atıcı elevasyon açısı Mermi yörüngesi Mermi kalibresi Namlu çıkış hızı

İki ayrı sensor dizilerinden algılanan dalga şekilleri ve türetilen büyüklükler ise Tablo 1.3'de verilmiştir.

Tablo 1.3 Algılanan dalga türleri ve tahminlenen konum değişkenleri

Algılanan Dalga Şekli	Türetilen Sonuçlar
Namlu patlaması 1	DoA1
Namlu patlaması 1 ve 2	DoA1, DoA2
Şok 1 ve 2	Mermi yörüngesi ve kalibresi
Şok 1, namlu patlaması 1 ve şok 2	DoA, mesafe, kalibre, namlu çıkış hızı, yörünge

SenTech firmasıda iki adet dört mikrofonlu dörtyüzlü (tetrahedral) sensör dizisi kullanmıştır ve atıcı mesafesinin, balistik dalga ile namlu patlaması arasındaki süre, mermi hızı ve mikrofonları ıskalama uzaklığının fonksiyonu olduğunu belirtmiştir [12]. Defence R&D Canada ve MacDonald Dettwiler & Associates arasında yedi yıl süren ortak çalışma sonucu geliştirilen ve 2003 yılında raporlanan Ferret silah konumlama sistemi diğer yöntemlerden farklı olarak şok dalgasının basıncını algılayan sensörler de içermektedir. Bu sayede ıskalama mesafesi ve mermi kalibresi hakkında ek bilgi elde edilmiştir. Raporda tek mikrofon dizisi ile mesafe tahmini için formül önermesi yapılmıştır [16]. Bu sistem atmosferik türbülansın olmadığı iyi yayılımlı bir atmosferde sensor dizisine 200 m'ye kadar uzaklıktan geçen merminin şok dalgasını yakalayabilmektedir. Bu sistem araç üzerinede takılabilir tek mikrofon dizisinden oluşmaktadır ve doğruluk değerleri oldukça yüksektir. Sistem deniz aracının hareketi sırasında oluşabilecek 15 farklı akustik gürültü kaynağına göre ve 5 farklı elektronik gürültü kaynağına göre de test

edilmiştir ve elektronik gürültü kaynaklarından etkilenmediği görülmüştür [17]. Çalışma ekibinden Jacques Bedard 2006 yılında atıcı tespit sistemleri için 9 adet performans metriğini belirlemiştir [14]. Fransa savunma bakanlığı kurulusu DGA desteği ile Metravib defence firması tarafından da PILAR sistemi geliştirilmiş Bosna ve Kosova'daki çatışmalarda başarılı sonuçlar vermiştir [18] . Keskin nişancı konumlama sistemlerinin kullanılan mikrofon dizisine göre tespit edilen dalga türüne göre elde edebileceği sonuçlar Tablo 1.4'te sunulmuştur [19]. Burada atıcı yönü azimuth ve elevasyon açılarını, atıcı konumu ise atıcı yönü ile birlikte mesafeyi de kapsamaktadır.

Tablo 1.4 Dalga çeşiti ve mikrofon dizisi sayısına göre tahmin edilen konum değişkenleri

Algılanan Dalga Türü	Olası Durumlar	Sistemin Ürettiği Sonuçlar	
		Tek Dizi	Çift Dizi
Şok	Susturucu ile atış Çok uzak mesafeden atış Ses sönümlemesi yüksek ortam (Karlı vb.)	iki olası azimuth açısı. (Elevasyon bilgisi yok)	(iki dizi arasından geçerse)
Namlu Patlaması	Ses altı mühimmat Tespit bölgesine yapılmayan atış	Atıcı yönü	Atıcı konumu (üçgenleme ile)
Şok+Namlu Patlaması	Susturucusuz yapılan ses üstü mühimmat atışları	Atıcı konumu. (Sensör dizisini 30m'den yakın geçerse)	Atıcı konumu (Fizyon ile)

Daha sonra yapılan sistem geliştirmeleri ile PILAR sistemine kablosuz bağlantı eklenerek çoklu mikrofon dizileri ile konumlama başarısı daha da arttırıldı. Ayrıca sisteme entegre edilen pan-tilt kulesine gündüz kamerası, kızıl ötesi kamera ya da lazer ölçüm donanımları eklenebilir hale getirildi [20]. Akustik sistemlerin yanında WeaponWatch, Viper gibi Infared sensörler kullanan sistemler de geliştirilmiştir [21, 22]. Bu sistemler yöneltildikleri bölgeden oluşan silah patlama ışıklarını algılayarak derecenin beşte biri hassasiyette yön tayini yapabilmektedir. Sadece IR kamera ile geliştirilen ilk sistemlerde güneş yansımaları gibi sebeplerle yanlış alarm oranı oldukça fazla olmaktaydı. Bu sebeple daha sonra akustik sistemler de eklenerek ilk oluşan tetiklenmeden sonra patlama sesi algılanması eklenerek patlama sesinin de algılandığı durumda gerçek alarma dönüştürülmesi sağlanarak yanlış alarm oranları azaltılmıştır [22]. Akustik sistem entegresi ile algılanan ışık ve tespit edilen ses arasındaki süreden mesafe ölçümü de kazandırılmıştır. Boston üniversitesinde geliştirilen REDOWL sisteminde ise mikrofon ve 300x kamera, IR ve termal kamera ve lazer mesafe ölçer gibi sistemler bir arada kullanılıştır [23]. QinetiQ şirketi ise askerlerin omuzlarına takılabilen 427gr toplam ağırlığa sahip, sesli olarak açı ve mesafe bilgisi verebilen SWARTs sistemini geliştirmiştir [24]. Askeri kasklara yerleştirilen mikrofonlar kullanan sistemler de sıkca çalışılmıştır [25, 26]. Ticari olarak sunulan bazı sistemler Tablo 1.5'de gösterilmiştir.

Daniki (2005) şok dalgasının daha belirgin ve gizlenemez bir ses dalgası

olduğunu belirterek sadece şok dalgasına dayalı bir konumlama yöntemi önermiştir. Tüm yönlü mikrofonların kullanıldığı sistemlere yönlü mikrofonlar de ilave ederek %1 yön hatasına kadar olan ölçümler için başarılı sonuçlar veren kendini düzeltebilme yeteneği eklemiş ve şok dalgasının geliş açının tahminleme doğruluğunu arttırmıştır. Altı tüm yönlü, üç yönlü mikforon yerleşimi ile yaptığı simulasyon çalışmasında ise %1-2 hatanın altında açı tespiti yapıldığında elde ettiği doğrusal olmayan denklemlerin iteratif çözümü ile açısal hatayı azaltıldığını göstermiştir [27, 28].

Tablo 1.5 Ticari sistemler

Üretici Firma	Sistem	Uygulama Sahası
DARPA - Raytheon BBN Technologies	Boomerang	Askeri
Metravib Defence	PILAR	Askeri
QinetiQ	SWATS	Askeri
Boston Üniversitesi	REDOWL	Askeri
Radiance Technologies	WeoponWatch	Askeri
Aselsan	Knats	Askeri
SoundThinking	Shotspotter	Şehir Güvenliği
Alliant Techsystems Inc.	SECURES	Şehir Güvenliği

Makinen ve Pertila (2010) şok dalgası yayılım modeline ve sayısal bir yaklaşım modeli uygulayarak mermi yörüngesi, kalibresi ve hızı için istatistiksel tabanlı yeni bir yöntem önermiştir. Burada GCC-Phat ve MSE olarak iki adet kıyaslama parametresi kullanmıştır. Bu parametrelere göre elde ettiği uzamsal olabilirlik fonksiyonlarını SA (simulated annealing) algoritması ile çözdürmüştür. Önerdiği yöntem ile tahmin edilen parametreye bağlı olarak konumlama ve mermi durumu hakkında %2 ile %10 arasında değişen hata oranları elde etmiştir [29]. Damarla ve ark. (2010) dağınık yerleştirilmiş tek mikrofonlardan ya da PDA cihazlarından oluşan asenkron bir ağ ile sadece şok dalgası ve namlu patlaması sesi arasındaki TDoA sürelerini kullanarak yeni bir konumlama yöntemi önermişlerdir. Bu yöntem mevcut yöntemlerden farklı olarak senkron örnekleme zamanına ihtiyaç duymamaktadır [30]. Cheinet ve Broglin (2015) atmosfer durumunun ve çevre koşullarının ses lokalizasyonuna etkisini incelemişlerdir. Açık hava koşullarında açı tespitinin birkaç derece hassasiyette yapılabildiğini fakat olumsuz hava koşullarında ve orman kenarlarında şok dalgası yankısı sebebiyle ölçüm hatalarının arttığını ayrıca atıcı konumu tespitinde sapma ve saçılma oluştuğunu gözlemlemişlerdir [31]. Calderon ve Apolinario (2015) yaptıkları çalışmada akustik algılama sistemine, tapografik verilerden oluşan digital harita bilgisini entegre ederek yeni bir konumlama sistemi önermişlerdir. Hakim bir noktaya konulan 7 mikrofonlu sensor dizisi ile namlu patlaması sesini algılayarak azimuth ve elevasyon açılarını belirledikten sonra mikrofon dizisinden geliş yönüne doğru çizgi çekerek

konumlama yapmışlardır. DoA tahminini geliştirmek için kullandıkları mikrofon düzeneğine göre TDoA süreleri toplamının sıfır olması gerekliliği ile hata düzeltimi önermişlerdir [32]. Akman ve ark. (2018) yaptıkları çalışmada mermi hızındaki, ses hızındaki, sensör dizisinde bulunan mikrofonlar arası mesafelerdeki ve sesin geliş süresindeki (TOA) belirsizlikler olmak üzere 4 farklı parametrenin mesafe tahminine etkisini incelemişlerdir.Bu hataların kümilatif olarak tek mikrofon dizisi kullanılarsa 640 metre uzaklıktaki atıcı konumu için 76.1 metre hata oluşturduğu belirtilmiştir. Birinci mikrofona 223 metre uzaklıkta ikinci bir mikrofon dizisi eklendiği de kümilatif hatanın 45.72 metre olduğu belirtilmiştir. Sonuçlar Tablo 1.6'da sunulmuştur. Aynı çalışmada 50 metre ve 120 metre mesafeden G3 ve AK47 silahları kullanılarak yapılan atışlarda, aralarında 20 metre bulunan iki sensör dizisi kullanılarak %10 ile %30 arasında değişen mesafe hataları elde edilmiştir [33].

Tablo 1.6 Balistik akustik model parametreleri hataları

Sistem Parametreleri	Etkisi Ölçülen Hata Miktarı	1 Mik. Dizisi	2 Mik. Dizisi
Mermi Hızı	%10	12.34 m	18.35 m
Ses Hızı	%10	31.82 m	19.34 m
Mikrofon Konumu	1cm	28.05 m	12.13 m
Mikrofon Konumu	2cm	51.23 m	23.93 m
Geliş Zamanı	Bir Örnek (@51.2kHz)	13.76 m	6.05 m
Toplam Hata	Tüm Hatalar	76.10 m	45.92 m

Abiri ve Parsayan (2020) yalnızca şok dalgası kullanarak düşük SNR'lı ortamlar için yeni konumlama bir sistem önermişlerdir. Önerdikleri sistemde araça monte edilebilen 3 mikrofonlu diziden iki adet kullanmışlardır. Şok dalgasından elde edilen mermi yolu ve diğer özelliklere, mermi oranı dahil edilerek konumlama hatasının iyileştiği belirtilmiştir ve deneysel sonuçlarda 400 m mesafedeki atıcı konumu için 30 m konum hatası elde edilmiştir [34].

1.1.1 Akustik Konumlamada ve Silah Patlaması Tespit Sistemlerinde Yapay Zeka

Son yıllarda yapay zeka yöntemleri kullanılarak yapılan kapalı ortam ses konumlama çalışmaları oldukça fazladır [35–38]. 2011 ile 2016 yılları arasında MLP yoğun şekilde kullanılırken, 2017'den sonra CNN tabanlı sistemler daha yoğun şekilde kullanılmaktadır [35]. Zhang ve ark. (2018) yaptıkları iç mekan konumlama çalışmasında CNN ağı kullanılarak 9 eşit bölgeye ayrılan bir odada 10000 farklı konum için üretilen sentetik ses verilerinin spektrogramları kullanılarak farklı gürültü oranları ile yapılan benzetim çalışmasında bölgesel konum sınıflama başarısının %99 seviyelerine kadar çıktığı gösterilmiştir [37]. Vera-Diaz ve ark. (2018) iç mekan konuşmacı konumla üzerine yaptıkları çalışmada

yarı sentetik verilerle eğitildikten sonra gerçek verilerle ince ayarı yapılan CNN ağı ile yapılan üçboyutlu konumlamanın SRP-PHAT yönteminden daha başarılı olduğu ortaya konmuştur [38]. Silah konumlama sistemleri üzerine yapılan yapay zeka uygulamaları ise ağırlıklı olarak farklı ses sınıfları içerisinden silah sesi sınıflama, farklı silahlara ait sesler içerisinden silah tipi sınıflama ve mermi kalibresi sınıflama gibi alanlarda yoğunlaşmıştır. 2015'li yıllarak kadar ses sınıflamada makine öğrenmesi yöntemleri sıklıkla kullanılırken 2012'den sonra konuşma tanımlama ve müzik sınıflama çalışmalarında konvolüsyonel ağların gelişmesiyle son on yılda derin öğrenme modelleri üzerinde çalışmalar artmıştır. Valenzise ve ark. (2007) yaptıkları çalışmada halka açık mekanlar gibi gürültülü ortamlarda çığlık ve silah seslerinin gürültüden ayırabilecek iki parelel GMM sınıflandırıcısını farklı ses özellikleri ile eğiterek %93 oranında başarılı sınıflama sağlamışlardır. Modern ve klasik toplam 47 ayrı özellik içinden en doğru sonucu verecek özellikler kümesini elde etme yöntemini açıklamışlardır. Alarm sesleri algıladıklarında ise multilaterasyon yöntemi ile konum tahmini yaparak kamera yönlendiren bir sistem uygulamışlardır [39]. Ayrıca Bayesian network [40], SVM [41, 42], HMM [43–45], ANN [46], RBFNN [42] gibi yapay zeka yöntemleride kullanılmıştır. Piczak (2015) yaptığı çalışmada 2 konvolüsyon katmanı 2 havuzlama katmanı ve iki tam bağlantılı katman kullanrak ESC10-ESC50 ve Urbansound8K very setlerini kullarak içerisinde silah patlama seslerinin de olduğu çevresel sesleri mel-spectrogram görüntülerini elde ederek sınıflandırmıştır [47]. Colangelo ve ark. yüksek gürültüde bile cam kırılması, silah sesi ve çıklık sesini tespit edebilen iki RNN'den oluşan hiyerarşik bir model önermişlerdir [48]. İnik (2023) CNN hiper parametlerelerini parçaçık sürü optimisasyonu (PSO) algoritması ile optimize ettiği çalışmasında aynı veri setleri kullanarak günümüze kadar geliştirilen bir çok derin öğrenme modelini incelmiştir ve hiper parametreleri optimize edilen ağı ile %98.5 doğruluğa ulaştığını raporlamıştır [49]. Silahlar tabanca (el silahı), tüfek (av) ya da tüfek (piyade) olarak üç ana sınıfa ayrılmaktadır. İlk silah sınıflama çalışmaları silahın hangi tipte olduğunu belirlemek üzere yapılan çalışmalardır. Khan ve ark. (2009) yaptıkları çalışmada tabanca ve tüfekten oluşan 10 silaha ait 100 patlama sesini tüfek ya da tabanca olarak %90 doğrulukta iki sınıfa ayırmıştır [50]. Daha sonra veri setleri geliştirilerek aynı sınıftaki farklı model silahlara ait tahminler yapılmaya başlanmıştır [50–54]. Son yıllarda oluşturulan veri setleri ile silah sesinin hangi silaha ait olduğunu belirlemeye yönelik çalışmalar artmaktadır. Tuncer ve ark. (2021) 8 ayrı silaha ait oluşturdukları 851 ses dosyası ile k-NN yöntemi ile sınıflama yapmışlar ve % 94.1 doğruluk elde etimşlerdir [55]. Qureshi ve ark. (2022) yaptıkları çalışmada 8 ayrı sınıfa ait 722 silah sesinden oluşan özel bir veri setine rüzgar etkisi ve azimuth ve elevasyon açıları için zaman gecikmeleri ekleyerek elde ettikleri sentetik data setini kullanarak silah sınıflayabilen ve silaha

atışının azimuth ve elevasyon açılarını tahmin edebilen bir derin öğrenme modeli önermiştir [56]. Raponi ve ark. (2022) youtube videolarından elde ederek oluşturdukları 3655 farklı ses dosyası içeren veri setlerini kullarak eğittikleri 4 konvolüsyon katmanından oluşan CNN ağları ile kategori, silah modeli, kalibresi %90'lara varan doğrulukla sınıflayabilmişlerdir. Veri setleri 30 farklı tabanca, 11 av tüfeği ve 18 askeri silah ile 3 kategori, 59 farklı silah sınıfı ve 7 farklı mermi kalibre sınıfından oluşmaktadır [57]. Park ve ark. (2022) PUBG oyunundan elde ettikleri 37 farklı silaha ait 2195 adet sesten BGG veri seti oluşturarak yaptıkları çalışmada CNN-Transformers ağ modeli ile %93.60 test doğruluğuna ulaşmıştır. Ayrıca diğer çalışmalardan faklı olarak ikinci bir veri setini 3 farklı silah ile 0 m, 50 m, 100 m, 200 m, 400 m ve 600 m'deki seslerinden oluşturmuş, mesafeleri farklı derin öğrenme yöntemleri ile sınıflandırmış ve CNN-Transform ağı ile 0.915 doğruluk değeri ve 0.9156 F1 değeri elde edilmiştir. Bu çalışma sentetik veri seti kullanarak mesafe sınıflaması yapan ilk çalışmadır [58]. Kartezyen koordinatlarda konum tespiti çalışmalarında spektrogram temelli ses özelliklerinin yaygın kullanıldığı görülmektedir [35]. Silah sesi konumlama üzerine yapılan çalışmalar ise genel olarak geliş yönü tahmini üzerine yoğunlaşmıştır [56, 59]. Kartezyen koordinatlara göre konumlama çalışmalarında ise çalışma ekipleri kendi veri setlerini oluşturmaktadır [60].

1.2 Tezin Amacı

Silah sesi tespit sistemlerine olan ihtiyaç gerek şehir içi güvenlik sistemleri, gerek askeri alanda gerekse de yasadışı avlanma tespitinde giderek artmaktadır. Mevcut silah sesi konumlama sistemleri üç farklı matematiksel modele dayanmaktadır. Şehir içi güvenlik sistemlerinde genel olarak birçok noktaya yerleştirilmiş algılayıcılar kullanılmakta ve bu algılayıcılar arasında metrelerce mesafe bulunmaktadır. Benzer şekilde özel bölge korumalarında da aralarında metrelerce uzaklık bulunan belirli noktalara yerleştirilmiş mikrofon dizileri kullanılmaktadır. Bu sistemlerde üçgenleme ya da multilaterasyon yöntemleri başarılı sonuçlar üretmektedir. Fakat bu sistemlerde bütün mikrofonların senkron olarak kayıt edilmeli bu sebeple tek bir kontrol noktasında verilerin toparlanarak karar verilmesi gerekmektedir. Bu sistemler hareketli askeri birlikler için uygun değildir. Akustik balistik model ise tek mikrofon dizisi ile konumlama yapabilmektedir fakat bu sistemin de birçok duyarlılık paremetresi vardır. Ayrıca tek mikrofon dizisi ile elde edilen mesafe tahmin hataları kısa mesafeler için %10, uzak mesafeler için ise %30'a kadar yükselmektedir. Sistemin mesafe tahmini yapabilmesi için aynı mermiye ait hem şok dalgası sesinin hemde namlu patlaması sesinin algılanması sistemin ana gereksinimidir. Sistem sadece namlu patlaması sesini

ya da sadece şok dalgası sesini algılayarak konumlama yapamamaktadır. Sesin azimuth açısını ölçmek için düzlem oluşturacak şekilde yerleştirilmiş en az üç mikrofona gerek duyulmaktadır. Elevasyon açısı da ölçülmek isteniyorsa üç mikrofonun oluşturduğu düzlemsel yerleşimin dışında uzay oluşturacak şekilde bir 4. mikrofona ihtiyaç duyulmaktadır. Mevcut yöntemlerde geliş açısı tespitinde uzak alan yaklaşımı sıklıkla kullanılmakta ve algoritmik kontroller sonucunda doğru açı değerleri elde edilmektedir. Mikrofon sayılarının arttırılması konumun daha doğru tespit edilmesinde yardımcı olmaktadır. Günümüzde kullanılan sistemlerde açı tahmin hataları tek mikrofon dizisi ile 3 derecenin altına inmişdir. Silah sesi konumlama sistemlerinde konumlama başarısının arttırılması için derin öğrenme destekli sistemler konusunda literatürde eksiklik vardır. Bu eksikliklerin ana sebeplerinden birisi, bu çalışmalar için gerekli veri setlerin mevcut olmayışıdır. Farklı mesafelerde ve açılarda oluşturulacak veri seti üzerinde derin öğrenme çalışmaları yapılarak derin öğrenmenin bu alanda katıları incelenmelidir. Bu tez de derin öğrenme ile silah sesi lokalizasyonunda tek mikrofon dizisi ile mevcut yöntemlerin elde ettiği doğrulukların iyileştirilmesi amaçlanmıştır.

1.3 Hipotez

Günümüzde modellemesi zor ya da karmaşık olan akustik problemler, yeterli veri seti oluşturulduktan sonra fiziksel modele derin öğrenme desteği eklenerek geliştirilen hidrit modeller ile çözülebilmektedir [36]. Kulaklarımızın bize sağladığı uzamsal işitmenin iki ana görevi ses kaynaklarının yerini belirlemek ve ses kaynaklarının yerlerini ayırmaktır [61]. İnsanların sesi algılarken kullandığı iki binaural özellik sesin kulaklararası zaman farkı (ITD) ya da kulaklararası faz farkı (IPD) ve kulaklararası seviye farkı (ILD)'dir [62]. ITD iki kulak arasında sesin TDoA farkını ifade etmektedir. ILD ise ses şiddetinin kulaklarda oluşturduğu algılama yoğunluğu farkıdır. Azimuth açısı algılanırken 1-2kHz'den düşük frekanstaki sesler için ITD özelliği kullanırken 1-2kHz üzerindeki sesler için ise ILD özelliği ile algılama yapılmaktadır [63]. İnsan için en zor algılan büyüklük sesin mesafesidir ve bunun için genellikle monaural özelliklik olan ulaşan ses şiddetini kullanır [64, 65]. Bu özelliği değerlendirebilmesi için ise duyulan ses şiddetine ait önceden bilgi sahibi olması gereklidir [66]. İlgili tez çalışması; "İnsan konumlama yeteneğini paralelinde, patlama sesi geliş açılarını mikrofonlar arası mesafe ile algılayan, patlama mesafesini ise iyi eğitilmiş derin öğrenme yöntemleri ile tahmin ederek konumlama yapabilen mevcut matematiksel yöntemlerin dışında hibrit bir yöntem gerçekleştirilebilirse, mevcut sistemlerden daha küçük boyutlu taşınabilir silah atışı konumlama sistemi geliştirilebilir" hipotezinden doğmuştur. Bu amaca yönelik eş merkezli ve dik yerleşimli 6 mikrofondan oluşan ve istenilen

açı kadar döndürülebilen bir mikrofon dizisi üretilmiş, 50m'den 500 m'ye kadar 50 m artımlarla ses toplamaya imkan sağlayan deney sahası oluşturulmuştur. Kuru sıkı tabancadan elde edilen patlama seslerinden veri seti oluşturularak MLP ve CNN ağları eğitilmiştir. Mesafe tahmininde etkili olan özellikler belirlenmiş ve etkileri ortaya konuştur. Makine öğrenmesi yöntemlerinin sonuçları incelenmiştir. Ayrıca kullanılan mikrofon dizisine özel geliş açısı formülü elde edilmiştir. Deney sonuçları değerlendirilmiştir.

1.4 Tezin İçeriği

Tezin ikinci bölümünde mevcut akustik konumlama modelleri, üçüncü bölümde yapay zeka uygulamalarında sıklıkla kullanılan akustik özellik çıkartma metodları ve MRMR özellik seçim algoritması açıklanmıştır. Dördüncü bölümde literatürde rastlamadığımız multilaterasyon yöntemi duyarlılık analizi yapılmış ve bu analizler sonucunda elde ettiğimiz kapalı-form DoA açı formüllerinin çıkarımı anlatılmıştır. Beşinci bölümde kullanılan derin öğrenme yöntemleri açıklanmıştır. Altıncı bölümde veri seti için oluşturulan deney sahası ve veri seti oluşturma adımları, önerilen konumlama yöntemi, geliştirilen derin öğrenme ağları açıklanmış, derin öğrenme ve makine öğrenmesi ile elde edilen değerler raporlanmıştır. Yedinci bölümde elde edilen sonuçlar ve öneriler sunulmuştur.

2
SİLAH SESİ KONUMLAMA YÖNTEMLERİ

2.1 Ses Fiziği

Sesin hızı yayıldığı ortamın özelliğine göre değişmektedir. Sesin yayılım hızı ile hava sıcaklığı arasında denklem (2.1) ile verilen bağlantı vardır.

$$v_s = v_0\sqrt{1+\frac{T}{273}} \approx (331+0.6T) \tag{2.1}$$

Burada c_0 sesin T= 0° C hava sıcaklığındaki hızıdır ve 331 m/s'dir. Bu formüle göre 20° C'de sesin yayılım hızı 343 m/s olarak elde edilmektedir. Akustik enerji havada her yönde yayılım gösterse de en yüksek ses basıncı merminin ilerleme yönünde oluşmaktadır. Bunun nedeninin silahı kullanan kişinin vücudunun ve silahın yapısal özelliklerinin etkisi olabileceği düşünülmektedir [67]. Silah sesi merminin ilerleme doğrultusuna göre oluşan azimuth açısına bağlı olarak yayılım göstermektedir. Ses basıncının azimuth açısına bağlı değişimi için denklem (2.2) önerilmiştir [68].

$$\mathrm{Pr}(\theta) = \mathrm{Pr}_{0^\circ}\left[1-\frac{Pr_{0^\circ}-Pr_{180^\circ}}{\mathrm{Pr}_{0^\circ}}\sin\left(\frac{\theta}{2}\right)\right] \tag{2.2}$$

Burada Pr_{0^o} ileri yöndeki maksimum basınç, Pr_{180^o} ise geri yöndeki maksimum basınç değeridir. Sesin basıncı zaman göre eksponansiyel olarak azalma göstermektedir ve namlu ucu patlama sesi çevresel gürültülere karşı giderek hassaslaşmaktadır. Bu azalma oranı kaynağın akış karakteristiğine göre değişse de, patlama gibi sesler için genel olarak denklem (2.3)'te verilen Friedlander denklemi model olarak kullanılabilir [67].

$$\mathrm{Pr}(t) = \mathrm{Pr}_{atm} + \mathrm{Pr}_s\left(1-\frac{t}{T_0}\right)e^{-bt/T_0} \tag{2.3}$$

Burada Pr_{atm} atmosfer basıncı, Pr_s patlamanın maksimum basınç değeri, T_0

pozitif faz gecikme süresi ve b ise üstel zayıflama oranıdır.

2.1.1 Sesin Şiddeti ve Algılanması

Mikrofonlar ses dalgalarını algılayıp elektriksel sinyale çeviren sensörlerdir. Sesin basınç seviyesine göre mikrofonların ürettiği gerilim değişmektedir. Sesin basınç seviyesi SPL olarak birimlendirilmektedir. Bu ölçümlerde 1 kHz'lik işaret referans alınmaktadır. Pascal değeri olarak bu basınç seviyesi çok geniş bir aralığa sahip olduğu için dB SPL birimi kullanılmaktadır ve denklem (2.4) ile hesaplanır.

$$dbSPL = 20 \log \left(\frac{P}{Pr_0} \right) \tag{2.4}$$

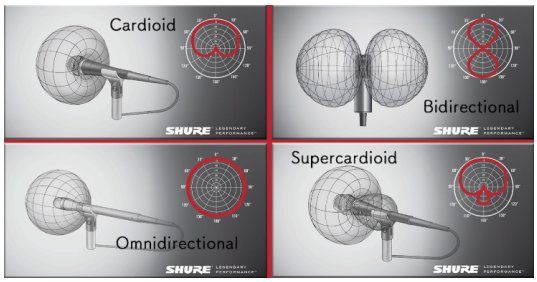
Burada Pr_0 değeri havada yayılan ses için, insan kulağının duyabildiği en küçük basınç değeri olan 20µPa, suda yayılan ses için ise 1µPa olarak alınmaktadır. Bir tüfekten çıkan patlama sesi yakınında 200Pa basınç oluşturmaktadır bu yayılım havada gerçekleştiği için 140 dB SPL re 20µPa olarak ifade edilmektedir. Tablo 2.1'de bazı seslerin gürültü değerleri verilmiştir.

Tablo 2.1 Sesler ve gürültü değerleri

Ölçüm Kaynağı	Basınç Değeri	Gürültü Değeri
Kulağın Duyma Eşiği	20µPa	0 dB
Araba İçi Gürültüsü	20 mPa	60 dB
Mağaza Gürültüsü	0.2 Pa	80 dB
Tır Gürültüsü	1 Pa	94 dB
Darbeli Matkap Gürültüsü	2 Pa	100 dB
Uçak Kalkış Gürültüsü	20 Pa	120 dB
Kulağın Ağrı Eşiği	200 Pa	140 dB

Ses basıncını elektriksel sinyale çevirmede farklı yapıda mikrofonlar üretilmiştir. Yaygın olarak karbon mikrofonlar, kondanser mikrofonlar, manyetik mikrofonlar ve piezoelektrik mikrofonlar kullanılmaktadır. Karbon mikrofonlar düşük kalitede yapılacak algılamalarda kullanılır. Piezoelektrik mikrofonlar düşük hassasiyete sahiptir ve yüksek ses basıncının ölçüleceği uygulamalarda öne çıkmaktadır. Manyetik mikrofonlar orta kalitedeki mikrofonlardır. Kondanser mikrofonlar ise kapasitif mikrofonlardır ve yüksek hassasiyette algılama yapabilmektedir. Uygulama alanlarına göre 0.125" ,0.25" 0.5" ve 1" gibi özel kapsül boyutlarında üretilirler.

Seslerin istenilen şekilde ölçülebilmesi için doğru mikrofon seçimlerinin yapılması



Şekil 2.1 Mikrofonların algılama yönleri

önem kazanmaktadır. Mikrofonların algılama alanları, dinamik tepkileri, frekans tepkileri, polarizasyon tipi, çalışma sıcaklığı parametreleri uygun seçimi yapmakta yardımcı olmaktadır.

Algılama alanı Şekil 2.1'de görüldüğü gibi mikrofonun gelen sesin yönüne gösterdiği hassasiyeti ifade etmektedir [8]. Yapılan çalışmalarda heryönlü (omni-directional) ve Supercadioid (shutgun) mikrofonlar ses yönü tespit sistemlerinde kullanılmıştır. Mikrofon tepkilerinin gelen sesin frekansına göre farklılık göstermemesi istenir. Bu sebeple frekans yanıtı grafikleri oldukça önem kazanmaktadır.

2.2 Geliş Açısı Tespit Yöntemleri

Ses kaynağından gelen sesin bir alıcıya ulaşma açısının tespit edilmesi ses konumlama sistemlerinde en sık kullanılan metodlardan birisidir. Ses konumlama yöntemleri direkt ve endirekt yöntemler olarak ikiye ayrılmaktadır. Direkt yöntemlerde tek adımda ses kaynağı pozisyonu tahmin edilirken, endirekt yöntemlerde ise mikrofonlar arası TDoA süreleri hesaplandıktan sonra bilinen mikrofon dizisi geometrisine göre konumlama yapılmaktadır [69]. Direkt yöntemler ise spektral tabanlı ve hüzmeleyici(beamformer) tabanlı olarak iki ana sınıfa ayrılmıştır. Mikrofondan alınan işaretin filtrelenip ağırlıklandırıldıktan sonra toplanmasıyla direkt olarak konumlama yapan çalışmalar ilk sınıfı oluşturmaktadır. Bir hüzmeleyicinin yönlendirilmiş yanıt gücünün (Steered response power of a beamformer) maksimizasyonu bu sınıfın temelini oluşturur ve SRP-PHAT [70] bu sınıfta en çok kullanılan yöntemdir [71, 72]. İkincisi ise sinyal korelasyon matrisine dayalı, yüksek çözünürlüklü sprektral tahmine uygulamalarını kapsar. Bu sınıfın en çok sık kullanılanları MUSIC [73], ESPRIT [74] yöntemleridir. Endirekt yöntemler ise mikrofonlar arasındaki Geliş Süreleri Farkı (TDoA) [75, 76] bilgilerinden konum hesaplaması yapan yöntemleri kapsar [77]. Son yıllarda ise bu yöntemlerin dışında derin öğrenme tabanlı yöntemler de geliştirilmektedir [78–80]. Direkt konumlama yöntemleri oluşan sesin konumu bilinen aday çoklu ses kaynaklarından

hangisine ait olduğunun belirlendiği uygulamalarda sıklıkla kullanılmaktadır ve bu yöntemlerin hesap yükü oldukça fazladır. Konumu bilinmeyen ve mikrofon dizisi referans noktasına uzaklığı, mikrofon dizisindeki mikrofonlar arası mesafeye göre oldukça fazla olan bir ses kaynağının geliş açısı tespiti uzak alan yaklaşımı ile hesaplanmaktadır. Uzak alan yaklaşımında sesin paralel doğrular şeklinde yayılarak mikrofonlara ulaştığı varsayılmaktadır [69]. Konumu bilinmeyen ses kaynaklarının geliş yönü tespitinde sıklıkla kullanılmaktadır. Özel olarak Mach sayısı 1'den büyük olan mühimmatlar atabilen silahların konum tespiti için kullanılan modelde geliş açısı ve şok dalgası ile namlu patlaması arasındaki süre farkı kullanılarak mesafe tahmini, dolayısı ile konumlama yapılabilmektedir [16]. Ses kaynağının mikrofon dizisi referans noktasına uzaklığı ile mikrofonlar arası uzaklık birbirine yakın değerler ise TDoA değerleri kullanılarak ses kaynağı için üçgenleme yöntemi ya da multilaterasyon yöntemi kullanılarak konumlama yapılabilmektedir [81–84]. Uzak alan yaklaşımı yönteminde mikrofonlar arası TDoA süreleri kullanılmaktadır. TDoA değerlerinin elde etmek için literatürde en sık kullanılan metot fazdönüşüm (PHAT) ağırlık fonksiyonunun kullanıldığı Genelleştirilmiş Çapraz Korelasyon (GCC) yöntemidir [85]. Yapılan çalışmalarda farklı mikrofon geometrileri kullanılsa da doğrusal dağılımlı, eş dağılımlı dairesel, L şekilli ve küresel mikrofon dizilerine daha sık rastlanmaktadır. Bu mikrofon dizilerinden aynı düzlemde 4'ün katı miktarda eş dağılımlı mikrofona sahip düzlemsel [71, 86], küresel [32, 87] ya da uzaysal(spatial) [86] ve L şekilli [88, 89] mikrofon dizileri eş merkezli ve dik yerleşimli mikrofon çiftlerine sahiptir.

2.2.1 Direkt Yön Hesaplama Yöntemleri

2.2.1.1 SRP-PHAT Yöntemi

Hüzmeleme yöntemine dayanan bir geliş yönü tespit algoritmasıdır. Hüzmeleme çıkış gücünü maksimum yapan yön geliş yönü olarak gelirlenmektedir. Mikrofon dizisinde yer alan n'inci mikrofondan alınan ses denklem (2.5) olarak yazılabilir.

$$x_n(t) = a_s(t) * h_n\left(\theta_s, t\right) + v_n(t) \tag{2.5}$$

Burada $a_s(t)$ ses kaynağının oluşturduğu işaret, θ_s kaynak geliş açısı, $v_n(t)$ sensörde oluşan beyaz gürültü, $h_n(\theta_s, t)$ ise kaynak pozisyonundan n'inci sensöre darbe yanıtıdır. Filtrele ve topla hüzmeleme yapısındaki yönlendirilmiş tepki gücü (Steered Response Power) denklem (2.6) olarak yazılabilir.

$$\hat{\theta}_s = \underset{\theta}{\operatorname{argmax}} P(\theta) \tag{2.6}$$

Burada $P(\theta)$ denklem (2.7) ile elde edilmektedir.

$$P(\theta) = \int_{-\infty}^{+\infty} \left| \sum_{n=0}^{N-1} G_n(\omega) X_n(\omega) e^{-j\omega\tau_n(\theta)} \right|^2 d\omega \tag{2.7}$$

Burada N sensör dizisinde bulunan mikrofon sayısı, $G_n(\omega)$ n'inci sensör kanalındaki filtrenin frekans cevabı, $\tau_n(\omega)$ sesin oluştuğu θ noktasından n. mikrofona ulaşana kadar geçen süre, $X_n(\omega)$ ise $x(t)$ işaretinin Fourier dönüşümüdür. Bu formulun hesap yükü oldukca fazla olduğu için GCC fonksiyonu olarak denklem (2.8) yazılarak hesaplanabilmektedir.

$$P(\theta) = 2\pi \sum_{n=0}^{N-1} \sum_{n=0}^{M-1} R_{nm}\left(\Delta\tau_{nm}(\theta)\right) \tag{2.8}$$

Burada $\Delta\tau_{nm} = \tau_n(\theta) - \tau_m(\theta)$ ve R_{nm} ise n'ici mikrofon ile m'inci mikrofon arasındaki GCC değeridir. GCC değerini hesaplamak için PHAT ağırlık fonksiyonu kullanılmaktadır [90].

2.2.1.2 MUSIC Yöntemi

MUSIC yöntemi, 1979 yılında Schmidt ve arkadaşları tarafından önerilen, çoklu sinyal kaynaklarının konumlarını tespit etmek için kullanılan bir spektral analiz yöntemidir. Bu, özellikle radyo frekansı alanında, radar sistemlerinde ve kablosuz iletişim sistemlerinde yaygın olarak kullanılır. İkinci dereceden istatistiksel ifadelere dayalı bir uzaysal spektrum tahminleyicisidir. Kullanılan sensör dizisi ile çoklu ses kaynaklarının sayısı ve seslerin geliş yönü belirlenebilmektedir. M tane sensörden oluşan diziye D farklı kaynaktan gelen işaret ve görültüler denklem (2.9) gibi modellenebilmektedir.

$$\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_M \end{bmatrix} = \begin{bmatrix} a_1(\theta_1) & a_1(\theta_2) & \cdots & a_1(\theta_D) \\ a_2(\theta_1) & a_2(\theta_2) & \cdots & a_2(\theta_D) \\ \vdots & \vdots & \cdots & \vdots \\ a_M(\theta_1) & a_M(\theta_2) & \cdots & a_M(\theta_D) \end{bmatrix} \begin{bmatrix} s_1 \\ s_2 \\ \vdots \\ s_D \end{bmatrix} + \begin{bmatrix} n_1 \\ n_2 \\ \vdots \\ n_M \end{bmatrix} \tag{2.9}$$

Burada x_i i'inci sensörden alınan işaret olmak üzere, $a_i(\theta_j)$; j'inci kaynağın i'inci sensör pozisyonuna göre oluşan DoA açısına karşılık gelen yönelmesi, s_j;

j'inci kaynağın genlik fonksiyonu ve n_i; i'inci sensöre gelen gürültü şeklinde modellenebilir. Daha kısa yazarsak;

$$x(t) = \sum_{i=1}^{D} A\left(\theta_i\right) s_i(t) + n(t) \tag{2.10}$$

Sensör dizisinden elde edilen toplam işaret matrisi kullanılarak, kovaryans matrisi denklem (2.11) olarak oluşturulur.

$$R_{xx} = E[x(t)x(t+\tau)]^* = AR_{ss}A^* + \sigma_n^2 I \tag{2.11}$$

Burada R_{ss}, DxD boyutlu kaynak korelasyon matrisi, σ_n^2 I ise MxM boyutlu gürültü korelasyon matrisidir. Elde edilen kovaryans matrisine özdeğer ayrışımı yapılarak özdeğerleri ve özvektörleri elde edilir. Elde edilen özdeğer vektöründe sıfırdan farklı her değer o kaynağa ait sinyal olduğunu ifade etmektedir. İşaret elde edilen kaynağın geliş açısı ise özvektör matrisinden kestirilerek elde edilir. Oluşan özvektörler $\boldsymbol{A}$ matrisindeki her sütun vektörüne dik olmalıdır. Buna göre $\boldsymbol{A}$ matrisindeki her yönelim vektörü elde edilerek DoA kestirimi denklem (2.12) ile yapılmaktadır.

$$P_{MU}(\theta) = \frac{1}{a^*(\theta)E_n E_n^* a(\theta)} \tag{2.12}$$

Burada E_n, Mx(M-D) boyutlu özvektörler matrisidir [91, 92].

2.2.1.3 ESPRİT Yöntemi

Roy ve arkadaşları tarafından 1989 yılında önerilen bu yöntem de MUSIC algoritması gibi alt uzay tabanlı bir yöntemdir fakat DoA açısı tahmini için MUSIC algoritmasına göre hesaplama ve depolama gereksinimlerini azaltmaktadır. Bu yöntemde sensör dizisi en az iki alt diziye ayrılır. MUSIC algoritmasındakine benzer olarak $x(t)$ işaretini iki alt sensör dizisinden elde edildiği şeklinde yazarsak denklem (2.13);

$$x(t) = \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} A \\ A\varphi \end{bmatrix} s(t) + n(t) \tag{2.13}$$

Burada $\phi = \text{diag}\left[e^{j\frac{2\pi}{\lambda}d\sin\theta_1}, e^{j\frac{2\pi}{\lambda}d\sin\theta_2}, \ldots, e^{j\frac{2\pi}{\lambda}d\sin\theta_D}\right]$ şeklinde DxD boyutlu bir matristir. Sensör dizisinin korelasyon matrisi denklem (2.14) elde edilir.

$$R_{xx} = \begin{bmatrix} A \\ A\Phi \end{bmatrix} R_{ss} \begin{bmatrix} A^H \\ {}^H A^H \end{bmatrix} + \sigma^2 I \tag{2.14}$$

Burada $A = [a(\theta_1), a(\theta_2), \ldots, a(\theta_M)]$ şekilden vektörlerden oluşan MxD'lik matristir. R_{xx} matrisi ise $(\lambda_1 \geq \lambda_2 \geq \cdots \lambda_D \geq \lambda_{D+1} = \cdots = \sigma^2)$ şeklinde gerçek özdeğerlere sahip Hermityan matristir. Burada en yüksek N adet öz değer uzaydaki N adet kaynaktan üretilen sesi ifade eder. E_s alt uzay matrisini E_1 ve E_2 olacak şekilde iki alt uzaya ayırırsak ve doğrusal dönüşüm uygularsak denklem (2.15);

$$E_s = \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \begin{bmatrix} AT \\ A\Phi T \end{bmatrix} \tag{2.15}$$

elde edilir. Burada $T = R_{11}^{-1} R_{21}$ ve $R_1 1 = 1/N X_1 X_1^H$ ve $R_2 1 = 1/N X_2 X_1^H$ olacak şekilde iki alt sensör uzayı arasıdaki kovaryans matrisleridir. E_2'yi E_1 cinsinden denklem (2.16) olarak yazabiliriz.

$$E_2 = E_1 T^{-1} \Phi T = E_1 \Psi \tag{2.16}$$

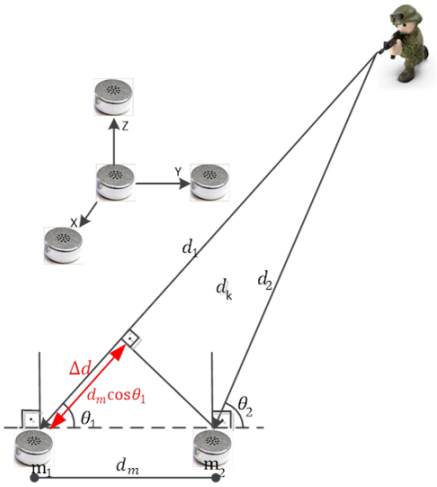
Burada $\Psi = T^{-1}\Phi T$'ye eşit DxD'lik bir martistir. Φ ve Ψ'nin öz değerleri aynıdır ve $\lambda_i = e_i^{jkdsin\theta_i}$'dir. Buna göre geliş açıları denklem (2.17) ile bulunur [93].

$$\theta_i = \sin^{-1}\left(\frac{\arg(\lambda_i)}{kd}\right) \tag{2.17}$$

2.2.2 Uzak Alan Yaklaşımı ve TDoA Hesaplama

Aralarında d_m kadar mesafe ile yerleştirilmiş mikrofonlardan oluşan bir sensör dizisiyle d_k kadar uzaklıkta oluşan patlama sesinin algılanması modeli Şekil 2.2'de gösterilmiştir.

Burada $d_k \gg d_m$ olduğu durumda ses dalgasının doğrusal bir ışın demeti gibi geleceği düşünülür. Buna göre mikrofonlar arasındaki ses gecikmesini oluşturan mesafe ve geliş açısı denklem (2.18) ve (2.19) ile hesaplanır.



Şekil 2.2 Mikrofonlar ile geliş açısı tespiti ve mikrofon dizisi yapısı

$$\Delta d = v_s * \tau_{12} \tag{2.18}$$

$$\theta = \text{acos}\left(\frac{v_s * \tau_{12}}{d_m}\right) \tag{2.19}$$

İki mikrofonla geliş açısı hesaplandığında mikrofonlar arası doğrultuyla simetrik iki açı değeri elde edilir. Bunun önüne geçmek için Şekil 2.3-a'daki gibi iki mikrofon doğrultusuna dik üçüncü bir mikrofon yerleştirilir. Diğer iki mikrofonla elde edilecek sonuç da iki açı değeri üretmektedir. Fakat iki çözümün kesişimi doğru geliş açısını vermektedir. Yükseliş açısı da hesaplanmak istenirse, Şekil 2.3-b'deki gibi üç mikrofonun oluşturduğu düzleme dik dördüncü bir mikrofon yerleştirilerek hem azimuth hemde elevasyon açıları tespit edilir, fakat mesafe tahmini yapılamaz [94].

Yapılan bir çalışmada Şekil 2.2'de gösterilen mikrofonların orta noktası orjin kabul edilerek, geliş açısının orijine göre hesaplanmış ve Şekil 2.2'de elde edilen dik üçgenin yarı ölçeklisi kullanılarak modellenmiştir [95]. Bu modele göre önerilen geliş açısı formülü denklem (2.20)'dır.

$$\theta = 90 - \text{acos}\left(\sqrt{1 - \left(\frac{v_s * \tau_{12}}{d_m}\right)^2}\right) \tag{2.20}$$



Şekil 2.3 (a) Düzlemde kullanılan mikrofon dizisi, (b) Uzayda kullanılan mikrofon dizisi

2.2.2.1 Çapraz Korelasyon

Mikrofonlar arasında sesin gelme süreleri arasındaki farkı algılamak için farklı yöntemler kullanılmaktadır. Bunlardan en sık kullanılanı çarpraz korelasyon yönteminden türetilen GCC yöntemidir. İki mikrofondan alınan seslerden birinci mikrofondan alınan ses $x_1(n)$ denklem (2.21) ve ikinci mikrofondan alınan ses $x_2(n)$ denklem (2.22) olarak gösterilmektedir.

$$x_1(n) = s_1\left(n - \tau_1\right) + n_1(n) \tag{2.21}$$

$$x_2(n) = s_1\left(n - \tau_2\right) + n_2(n) \tag{2.22}$$

(2.21) ve (2.22) denklemlerinde $s(n)$ kaynaktan çıkan ses sinyalini $n_1(n)$ ve $n_2(n)$ ise ses sinyaline karışan beyaz Gauss gürültüsünü ifade etmektedir. Sesin mikrofonlara ulaşma süreleri ise sırasıyla τ_1 ve τ_2 olarak gösterilmektedir. Buna göre mikrofonlara ses sinyalinin geliş zaman farkı denklem (2.23) ile gösterilir.

$$\tau_{12} = \tau_1 - \tau_2 \tag{2.23}$$

$x_1(n)$ ve $x_2(n)$ sinyalleri arasındaki korelasyon fonksiyonu denklem (2.24)'de olduğu gibi tanımlanır.

$$R_{12}(\tau) = E\left[x_1(n)x_2(n-\tau)\right] \tag{2.24}$$

(2.24) denkleminde (2.21) denklemini yerine yazarsak (2.25) denklemini elde ederiz.

$$\begin{aligned} R_{12}(\tau) = E\left[s\left(n-\tau_1\right)s\left(n-\tau_1-\tau\right)\right] + E\left[\left(s\left(n-\tau_1\right)n_2(n-\tau)\right)\right] \\ +E\left[s\left(n-\tau_2-\tau\right)n_1(n)\right] + E\left[n_1(n)n_2(n-\tau)\right] \end{aligned} \tag{2.25}$$

(2.25) denkleminde $s(n)$, $n_1(n)$ ve $n_2(n)$ sinyalleri birbiri ile ilişkisiz olduğu için (2.26) denklemi aşağıdaki gibi elde edilir.

$$R_{12}(\tau) = E\left[s\left(n-\tau_1\right)s\left(n-\tau_1-\tau\right)\right] = R_s\left(\tau-\left(\tau_1-\tau_2\right)\right) \tag{2.26}$$

(2.26) denklemini $R_s(0)$ olduğunda maksimum değeri almaktadır ve bu τ değeri bizim mikrofonlar arasındaki geliş zaman farkı değerini üretmektedir. Çapraz korelasyon yöntemi hızlı çözüm sunmaktadır, fakat gürültüye ve yankılanmalara karşı oldukça duyarlıdır. Denklem çözümlerinde de ifade edildiği üzere gürültülerin ilişkisiz olduğu düşünülmektedir, fakat gerçek uygulamalarda bu varsayım bazen doğru olmamaktadır. Bu sebeple genelleştirilmiş çapraz korelasyon yöntemi geliştirilmiştir.

2.2.2.2 Genelleştirilmiş Çapraz Korelasyon

Bu yöntem C.H.Knapp ve G.C. Carter tarafından oluşturuşmuş bir maksimum olabilirlik kestirimcisidir [85]. Çapraz korelasyon fonksiyonu, $G_{12}(\omega)$ çapraz güç sprektrum fonksiyonunun ters Fourier dönüşümü denklem (2.27) ile de elde edilebilmektedir.

$$R_{12}(\tau) = \int_0^{\pi} G_{12}(\omega)e^{j\omega\tau}d\omega \tag{2.27}$$

Çapraz güç spektrum fonksiyonu ise frekans bölgesinde iki sinyal arasındaki benzerliği tanımlamaktadır ve $x_1(n)$ işaretinin Fourier dönüşümü $X_1(\omega)$, $x_2(n)$ işaretinin Fourier dönüşümünün karmaşık eşleniği $X_2^*(\omega)$, örnek sayısı N olmak üzere denklem (2.28) ile ifade edilmektedir.

$$G_{12}(\omega) = \frac{X_1(\omega)X_2^*(\omega)}{N^2} \tag{2.28}$$

Tablo 2.2 Ağırlık fonksiyonları ve etkileri

Ağırlık	Ağırlık Fonksiyonu	Karakteristiği
Çapraz Korelasyon	1	Dış ortam gürültülerine ve yankılara karşı hassastır
ROTH	$\frac{1}{G_{11}(\omega)}$	Etkili bir şekilde büyük gürültü bandını bastıran Wiener filtresine eşdeğerdir.
SCOT	$\frac{1}{\sqrt{G_{11}(\omega)G_{22}(\omega)}}$	SCOT iki kanalın etkisini de ele alır. Tepe değerini genişletir ve istenmeyen gürültüleri bastırır.
PHAT	$\frac{1}{\lvert G_{12}(\omega)\rvert}$	Düşük sinyal enerjisinde hata artar. Sabit eklenerek bu sorun azaltılabilir.
Wiener	$\gamma_{12}^2(\omega) = (\frac{G_{12}(\omega)}{\sqrt{G_{11}(\omega)G_{22}(\omega)}})^2$	İki sinyalin coherens fonksiyonu ile çarpılır.
ML	$\frac{\gamma_{12}^2(\omega)}{\lvert 1-\gamma_{12}^2(\omega)\rvert G_{12}(\omega)\rvert}$	Zaman gecikmesinin doğruluğunu arttırmaktadır.

Genelleştirilmiş çapraz ilişki fonksiyonunda ise frekans bölgesinde çapraz güç spektrumu, $\Psi_{12}(\omega)$ ağırlık fonksiyonu ile çarpılarak sinyaller üzerinde beyazlatma işlemi yapılmaktadır. Özellikle SNR değeri yüksek olan sinyallerde gürültü etkisi azaltılmış olmaktadır. Bu çarpım sonucu elde edilen fonksiyonun ters Fourier dönüşümü alınarak (2.29) genelleştirilmiş çapraz korelasyon denklemi elde edilmektedir.

$$R_{12}^g(\tau) = \int_0^\pi \Psi_{12}(\omega)G_{12}(\omega)e^{j\omega\tau}d\omega \tag{2.29}$$

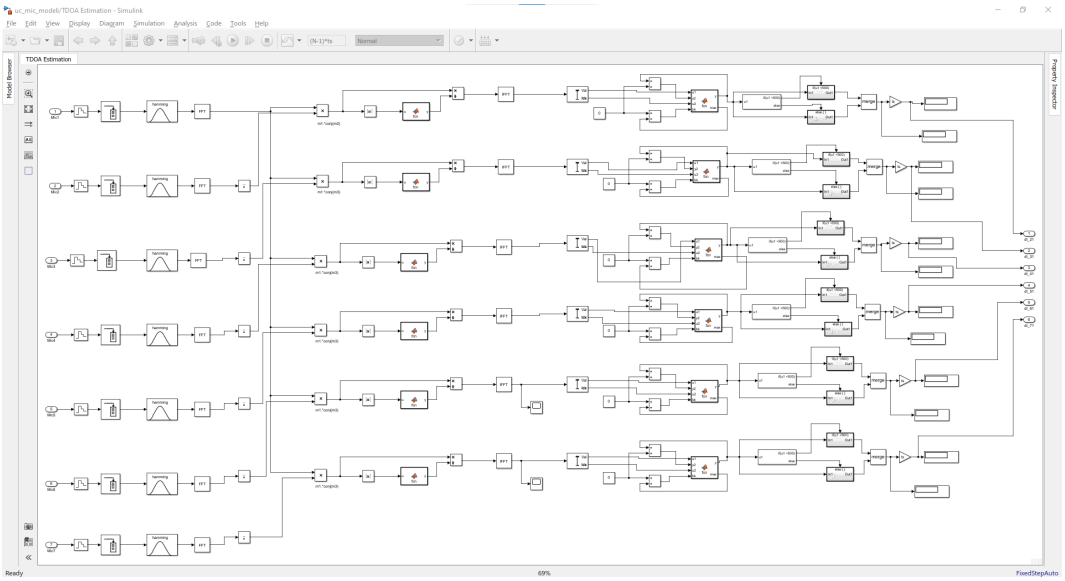
Genelleştirilmiş çapraz ilişki fonksiyonunda kullanılan bazı ağırlık fonksiyonları ve onların etkileri Tablo 2.2'de verilmiştir [96].

Şekil 2.4'te görüldüğü gibi Matlab/Simulink programında hazırlanan simülasyona 7 ayrı mikrofondan senkron olarak kayıt edilmiş sesler giriş olarak verildiğinde, genelleştirilmiş çapraz korelasyon yöntemi ve PHAT filtresi kullanılarak aralarındaki gecikme süreleri doğru olarak hesaplanmıştır.

2.3 Konum Belirleme Yöntemleri

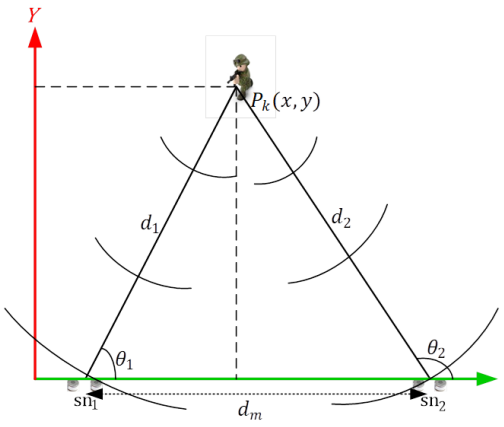
2.3.1 Üçgenleme Yöntemi

Bir nesnenin en az üç referans alıcıya uzaklığını belirleyerek o nesnenin konumunu belirlemeye yarayan yaygın kullanıma sahip bir yöntemdir. Wi-Fi erişim noktalarının belirleme, radar istasyonları ile uçak konumu belirleme ve GPS sinyalleri ile konum belirleme uygulamaları bu yöntemin kullanıldığı en önemli uygulama alanlarıdır. Akustik alanda ise referans mikrofonlara sesin gelme süreleri



Şekil 2.4 Mikrofonlar arası TDoA simülasyonu

farkı kullanılarak bir açı bilgisi elde edilir. Aralarında belirli mesafe bulunan en az iki sensör dizisine ses sinyalinin geliş açısı hesaplanarak üçgenleme yöntemi ile konum hesaplanır [97, 98]. Şekil 2.5'de konum belirlemek için kullanılan geometrik yapı görülmektedir. Buna göre konum hesaplama formulleri denklem (2.30)-(2.33)'de verilmiştir.



Şekil 2.5 Üçgenleme yöntemi ile konumlama

$$\theta_i = \tan^{-1}\left(\frac{P_{ky} - P_{sn_1 y}}{P_{kx} - P_{sn_1 x}}\right) \tag{2.30}$$

$$d_m = \frac{P_{ky}}{\tan\theta_1} + \frac{P_{ky}}{\tan\theta_2} \tag{2.31}$$

$$P_{ky} = \frac{d_m}{\cot\theta_1 + \cot\theta_2} \tag{2.32}$$

$$P_{kx} = \frac{P_{ky}}{\tan\theta_1} \tag{2.33}$$

Burada d_1 ve d_2 sırasıyla ses kaynağının birinci ve ikinci mikrofon dizi ile uzaklığını, P_{sn_i} i'inci sensör dizisinin konumunu, P_k ses kaynağının konumunu, d_m ise mikrofonlar arası uzaklığı ifade etmektedir. Bu yöntemde konumun doğru belirlenebilmesi için ses kaynağının mikrofonlara uzaklığı, mikrofon dizileri arası mesafeye yakın büyüklükte olmalıdır.

2.3.2 Multilaterasyon Yöntemi

İki boyutlu kartezyen düzlemde konumlama yapmak için en az üç farklı konuma yerleştirilmiş mikrofonlardan oluşan sensör dizisine ihtiyaç vardır. Şekil 2.3-a'da gösterilen 3 mikrofonlu sensör dizisi için mikrofonların düzlemsel konumunu $p_i = [x_i, y_i]^T$ olarak ve ses kaynağının konumunu ise $p_k = [x_k, y_k]^T$ olarak gösterelim.

Mikrofonlara gelen ses sinyallerinin ulaşma zamanlarını τ_i ile ve iki mikrofon arasında ses sinyali geliş zaman farkını $\tau_{ij} = \tau_i - \tau_j$ olarak gösterelim. Bu durumda sesin havadaki yayılım hızına bağlı olarak ses kaynağı ile mikrofonlar arasındaki mesafe

$$d_i = \tau_i \times v_{\text{ses}} = \|\boldsymbol{p_k} - \boldsymbol{p}_i\| \tag{2.34}$$

v_{ses} sesin nominal hızıdır. Buna göre i. mikrofon ile kaynak arasındaki mesafe

$$d_i = \sqrt{(x_k - y_i)^2 + (y_k - y_i)^2} \tag{2.35}$$

Birinci mikrofonu referans alarak iki mikrofon arasındaki mesafenin, ses hızına bağlı fark denklemi ile geometrik fark denklemi eşitlenerek

$$\tau_{j1} \times v_{\text{ses}} = \left\|\boldsymbol{p_k} - \boldsymbol{p}_j\right\| - \|\boldsymbol{p_k} - \boldsymbol{p}_1\| \tag{2.36}$$

Denklem (2.36) de iki tarafın karesi alınarak

$$\tau_{j1}{}^2 \times v_{\text{ses}}{}^2 + \|\boldsymbol{p_k} - \boldsymbol{p}_1\|^2 + 2 \times v_{\text{ses}} \times \tau_{j1} \times \|\boldsymbol{p_k} - \boldsymbol{p}_1\| = \left\|\boldsymbol{p_k} - \boldsymbol{p}_j\right\|^2 \quad j = 2, 3, \ldots, N \tag{2.37}$$

Denklem (2.37)'de oluşan karesel ifadelerin açılımıyla

$$\tau_{j1}{}^2 \times v_{ses}{}^2 + \boldsymbol{p}_k^T \boldsymbol{p}_k - 2\boldsymbol{p}_1^T \boldsymbol{p}_k + \boldsymbol{p}_1^T \boldsymbol{p}_1 + 2 \times v_{\text{ses}} \times \tau_{j1} \times \|\boldsymbol{p}_k - \boldsymbol{p}_1\| = \boldsymbol{p}_k^T p_k - 2p_j^T p_k + p_j^T p_j \tag{2.38}$$

Denklem (2.38)'de gerekli sadeleşmeler yapılıp, bilinmeyenli çarpımlar denklemin sol tarafına bilinenler sağ tarafına alınıp düzenlenerek

$$2\boldsymbol{p}_j^T \boldsymbol{p_k} - 2\boldsymbol{p}_1^T \boldsymbol{p_k} + 2 \times v_{\text{ses}} \times \tau_{j1} \times \|\boldsymbol{p_k} - \boldsymbol{p}_1\| = \boldsymbol{p}_j^T \boldsymbol{p}_j - \boldsymbol{p}_1^T \boldsymbol{p}_1 - \tau_{j1}^2 \times v_{\text{ses}}{}^2 \tag{2.39}$$

Denklem (2.39)'da $d_1 = \|\boldsymbol{p_k} - \boldsymbol{p}_1\|$ eşitliği yerine yazılarak tüm mikrofonlar için doğrusal denklem takımını

$$2\boldsymbol{p}_j^T \boldsymbol{p_k} - 2\boldsymbol{p}_1^T \boldsymbol{p}_k + 2 \times v_{\text{ses}} \times \tau_{j1} \times d_1 = \boldsymbol{p}_j^T \boldsymbol{p}_j - \boldsymbol{p}_1^T \boldsymbol{p}_1 - \tau_{j1}{}^2 \times v_{\text{ses}}{}^2, j = 2, 3, \ldots, N \tag{2.40}$$

Denklem (2.40), referans kabul edilen birinci mikrofon ile j. mikrofonun arasındaki bağlantıyı tanımlamaktadır. N adet mikrofon için aynı referans mikrofonu kullanarak $N-1$ adet denklem yazılır. 5 adet mikrofondan oluşan bir sensör dizisi için düzenlenerek

$$\begin{bmatrix} 2\left(\boldsymbol{p_2} - \boldsymbol{p}_1\right)^T & 2v_s\tau_{21} \\ 2\left(\boldsymbol{p}_3 - \boldsymbol{p}_1\right)^T & 2v_s\tau_{31} \\ 2\left(\boldsymbol{p_4} - \boldsymbol{p}_1\right)^T & 2v_s\tau_{41} \\ 2\left(\boldsymbol{p}_5 - \boldsymbol{p}_1\right)^T & 2v_s\tau_{51} \end{bmatrix} \times \begin{bmatrix} \boldsymbol{p_k} \\ d_1 \end{bmatrix} = \begin{bmatrix} \boldsymbol{p}_2^T \boldsymbol{p_2} - \boldsymbol{p}_1^T \boldsymbol{p_1} - v_s^2\tau_{21}^2 \\ \boldsymbol{p}_3^T \boldsymbol{p}_3 - \boldsymbol{p}_1^T \boldsymbol{p_1} - v_s^2\tau_{31}^2 \\ \boldsymbol{p}_4^T \boldsymbol{p_4} - \boldsymbol{p}_1^T \boldsymbol{p_1} - v_s^2\tau_{41}^2 \\ \boldsymbol{p}_5^T \boldsymbol{p_5} - \boldsymbol{p}_1^T \boldsymbol{p_1} - v_s^2\tau_{51}^2 \end{bmatrix} \tag{2.41}$$

Genel bir çözüm için (2.41) matris denklemi şeklinde yazılarak

$$\boldsymbol{Ax} = \boldsymbol{b} \tag{2.42}$$

burada $\boldsymbol{A} \in R^{(N-1)\times 4}, \boldsymbol{b} \in R^{(N-1)}, \boldsymbol{x} \in R^4$. $\boldsymbol{A}$ ve $\boldsymbol{b}$ matrislerinin $(i-1)$. satırı $i = 2, 3, \ldots, N$ için

$$\boldsymbol{a_{i-1}} = [\boldsymbol{2(p_i - p_1)^T 2v_s\tau_{i1}}], \qquad b_{i-1} = [p_i^T p_i - p_1^T p_1 - v_s^2 \tau_{i1}^2]$$

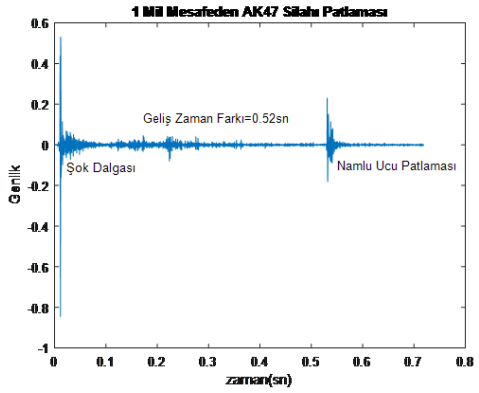
yeterli sayıda mikrofon ölçümü alınması halinde ($\boldsymbol{A^T A}$ tam rank bir matris olması durumunda) LSE metodu kullanılarak bilinmeyen x çözümü [99]

$$\begin{bmatrix} \boldsymbol{p}_k \\ d_1 \end{bmatrix} = \left(\boldsymbol{A}^T \boldsymbol{A}\right)^{-1} \boldsymbol{A}^T \boldsymbol{b} \tag{2.43}$$

(2.43) denklemi incelendiğinde **A** matrisi ve **b** vektörünün TDoA değerlerine bağlı olduğu görülmektedir. Bu sebeple yöntemin doğru sonuçlara yakın değerler üretmesi mikrofonlar arası mesafe ve örnekleme frekansı parametrelerine bağlıdır.

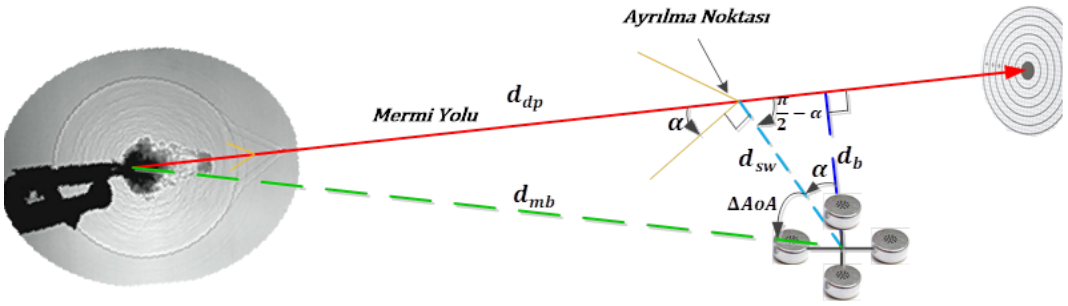
2.3.3 Balistik Akustik Model

Literatürde kullanılan bir diğer yöntem ise ses üstü hızlarda ilerleyen mühimmatları kullanan silahların konumunu belirlemektedir. Ses üstü hızda ilerleyen mühimmat şok dalgası oluşturmaktadır. Şekil 2.6'da görüldüğü gibi şok dalgasının sesi ile namlu ucu patlaması sesi arasındaki geliş süresi farkı ses kaynağının uzaklığı ile alakalı önemli bilgiler taşımaktadır. Bu yöntemle silah sesinin geliş yönü ve geliş mesafesi hesaplanabilmektedir.



Şekil 2.6 AK-47 silahından 1600 m mesafeden yapılan atışın ses grafiği

Şekil 2.7'den görüldüğü gibi silah patlamasını iki boyutta incelersek, namlu ucu patlaması dairesel olarak yayılırken şok dalgası ise tepe açışı α olan koni şeklinde



Şekil 2.7 Ses üstü merminin şok ve namlu uçu patlaması dalgaları ve yayılım geometrisi

yayılmaktadır. Mühimmatın kendi hızında d_{dp} mesafesi kadar ilerlediği noktada oluşan şok dalgası ses hızıyla yayılarak d_{sw} mesafesini aldıktan sonra mikrofona ulaşmaktadır. Namlu ucu patlaması ise ses hızında d_{mb} mesafesini aldıktan sonra mikrofona ulaşmaktadır. Buna göre şok dalgasının ve namlu ucu patlaması seslerinin mikrofona ulaşma süreleri sırasıyla denklem (2.44) ve denklem (2.45) ile elde edilmektedir. Burada v_m mühimmat hızıdır.

$$t_{sw} = \frac{d_{dp}}{v_m} + \frac{d_{sw}}{v_s} \tag{2.44}$$

$$t_{mb} = \frac{d_{mb}}{v_s} \tag{2.45}$$

Denklem (2.44) bulunan d_{sw} mesafesi aynı zamanda merminin mikrofonu ıskalama mesafesi olan d_b uzunluğunun $cos(\alpha)$'ya bölümüne eşittir. Buna göre denklem (2.44)'i düzenlersek denklem (2.46) elde edilir.

$$t_{sw} = \frac{d_{dp}}{v_m} + \frac{d_b}{v_s \cos(\alpha)} \tag{2.46}$$

Burada α açısı mühimmatın Mach hızıyla ters orantılıdır ve Mach hızı denklem (2.47) ile α açısı ise denklem (2.48) ile hesaplanır.

$$M = \frac{v_m}{v_s} \tag{2.47}$$

$$\alpha = \arcsin\left(\frac{1}{M}\right) \tag{2.48}$$

Şekil 2.7'deki geometrik yapı ve (2.44)-(2.48) denklemleri kullanılarak d atış mesafesi denklem (2.49) ile elde edilmektedir.

$$d = d_{mb} = \frac{v_s \left(t_{mb} - t_{sw}\right)}{1 - \cos(\triangle AoA)} \tag{2.49}$$

Burada t_{mb} namlu patlaması gelme süresi, t_{sw} şok dalgası gelme süresi, ΔAoA namlu patlaması geliş açısı AoA_{mb} ile şok dalgası geliş açısı AoA_{sw} arasındaki fark ve d silah patlama uzaklığıdır. AoA_{mb} ve AoA_{sw} değerleri uzak alan yaklaşımı ya da kapalı-form çözüm yöntemleri ile elde edilmektedir.

2.3.3.1 Ses Üstü Mühimmat İçin Kalibre, Namlu Çıkış Hızı ve Mermi Yolu Tahmini

Ses üstü mühimmatların algılanan şok dalgasının oluşturduğu sinyalin zaman grafiği mermiye ait kalibre, namlu çıkış hızı ve mermi yolu özelliklerinin tahmin edilmesini sağlamaktadır. Whitman'ın oluşturduğu matematiksel modele göre oluşan maksimum basınç genliği P_{max} ve p_i noktasındaki mikrofondan algılanan sinyalin uzunluğu Δt_{sw} olmak üzere aşağıdaki denklemler ile hesaplanmaktadır, burada birimler sırasıyla paskal ve saniyedir [100]. Bu model çalışmalarda sıklıkla kullanılmaktadır [25, 29, 31, 98, 101].

$$P_{\max} = \frac{0.53 Pr_{atm} \left(M^2 - 1\right)^{\frac{1}{8}} \phi_b}{d_b^{3/4} l_b^{1/4}} \tag{2.50}$$

$$\Delta t_{sw} = \frac{1.82 M d_b^{1/4} \phi_b}{v_s \left(M^2 - 1\right)^{3/8} l_b^{1/4}} \tag{2.51}$$

Burada Pr_{atm} ortam atmosfer basıncı, ϕ_b metre cinsinden mermi çapı, l_b merminin metre cinsinden uzunluğudur. Mikrofonlar ile ölçülen yukarıdaki değerleri kullanarak mermiye ait ıskalama mesafesi ve kalibresi ölçülmeye çalışılır, fakat Mach hızı da bilinmemektedir. Denklem (2.50) ve (2.51) oranlanırsa mermi boyutundan bağımsız olan denklem (2.52) elde edilir.

$$\frac{\Delta t_{sw}}{P_{\max}} = \frac{1.82 M d_b}{0.53 v_{\mathrm{s}} \, Pr_{atm} \sqrt{\left(M^2 - 1\right)}} \tag{2.52}$$

Mach hızının bilinmesi durumunda d_b uzaklığının elde edilmesi ile mermi yörüngesi tahmin edilebilmektedir. Bulunan d_b değeri denklem (2.50) ve (2.51)

kullanılarak mermi boyutu elde edilebilir. Genel olarak mermi üretiminde uzunluk ile çap arasında denklem (2.53) ile verilen oran kullanılmaktadır [16].

$$\frac{l_b}{\phi_b} \approx 4.35 \tag{2.53}$$

$$\phi_b = 0.45 \times 4.35^{\frac{1}{3}} \left(\frac{v_s \Delta t_{sw}}{M} \right)^{\frac{4}{3}} \frac{\sqrt{(M^2-1)}}{d_b^{1/3}} \tag{2.54}$$

$$d_b = \Delta t_{sw} \frac{v_s^2}{v_m} \frac{Pr_{atm}\sqrt{(M^2-1)}}{3.44 P_{\max}} \tag{2.55}$$

Buna göre patlama mesafesi d_{mb} için kapalı form çözümü denklem (2.56)-(2.58) olarak önerilmiştir [101].

$$d_{mb} = \frac{1}{2\left(v_{\mathrm{s}}^4 - v_m^4\right)}(K - 2\sqrt{M}) \tag{2.56}$$

$$K = -2v_m^3 b\sqrt{v_m^2 + v_s^2} - 2(\Delta\tau)v_s^3 v_m^2 + 2v_{\mathrm{s}}^2 b v_m \sqrt{v_m^2 + v_{\mathrm{s}}^2} - 2(\Delta\tau)v_{\mathrm{s}} v_m^4 \tag{2.57}$$

$$\begin{aligned} M = &-2v_{\mathrm{s}}^4 v_m^4 b^2 + 2(\Delta\tau)^2 v_s^6 v_m^4 + (\Delta\tau)^2 v_{\mathrm{s}}^4 v_m^6 - 2v_{\mathrm{s}}^7 b v_m \sqrt{v_m^2 + v_{\mathrm{s}}^2} \\ &+ v_{\mathrm{s}}^8 (\Delta\tau)^2 v_m^2 + 2v_{\mathrm{s}}^8 b^2 + 2v_m^5 b\sqrt{v_m^2 + v_{ses}^2}(\Delta\tau)v_{\mathrm{s}}^3 \end{aligned} \tag{2.58}$$

3
SES TANIMLAMADA KULLANILAN ÖZELLİKLER

Ses sinyali, üretildiği kaynak hakkında frekans, şiddet, tını gibi önemli bilgiler taşımaktadır. Bu bilgilere göre ses kaynağının ne olabileceği, ne kadar uzakta olabileceği gibi bilgileri sağlamaktadır. İnsan kulağının sesi algılayabilmesi için ses sinyalinin minimum 10 ms sürmesi ve 10 dB üzerinde bir gürültüye sahip olması gerekmektedir. Literatürde sesin tanımlanması ve sınıflanması için birçok yöntem kullanılmaktadır. Bunlar temelde zaman boyutundaki özellikler, frekans boyutundaki özellikler ve zaman-frekans boyutundaki özellikler olarak üçe ayrılmaktadır.

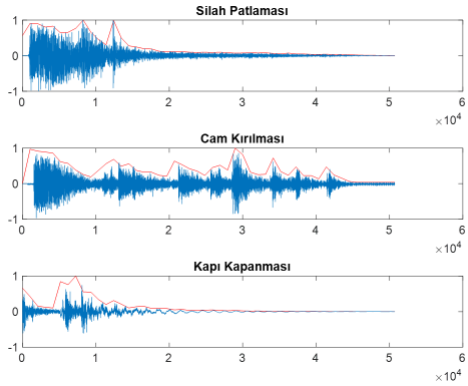
3.1 Zamansal Özellikler

Sesin sinyalinin örneklenmiş hali üzerinden herhangi dönüşüme uğratmadan elde edilen özelliklerdir. Bunlardan sıklıkla kullanılanları Genlik Zarfı, RMS enerjisi ve Sıfır Geçiş Oranıdır.

3.1.1 Genlik Zarfı

Sesin gürültüsel büyüklüğünün şeklini veren bir özelliktir. Ses sinyali belirli örnek sayısında küçük pencerelere ayrılır ve her bir penceredeki maksimum genlik değeri alınır, denklem (3.1) ile elde edilmektedir. Pencerede yer alan aykırı değerlere karşı duyarlıdır, pencere içinde sadece bir örnek değeri büyük olsa diğer değerler sıfır olsa değer varmış gibi sonuç üretir. Sesin başlama anında sinyal genliğinde yükselmeler oluştuğu için daha çok sesleri başlama noktalarını belirlemede kullanılır. Şekil 3.1'de farklı sesler için normalize edilmiş genlik zarfı grafikleri görülmektedir.

$$GZ_t = \max_{n.t \leq n \leq (p+1)N-1} x(n) \tag{3.1}$$

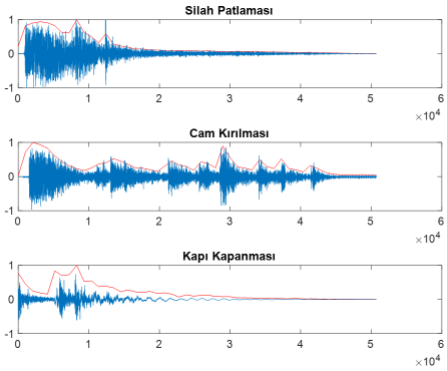


Şekil 3.1 Faklı sesler için genlik zarfı grafikleri

3.1.2 RMS Enerjisi

Genlik zarfına benzer şekilde RMS Enerji de sesin gürültüsel özelliği hakkında bilgi vermektedir. Denklem (3.2)'de görüldüğü gibi her bir çerçevedeki örneklerin enerjilerinin ortalamasıyla elde edilen bir sonuç olduğu için aykırı değerlere karşı daha duyarsızdır. SNR'nin yüksek olduğu durumlarda duyulabilir sesi sessizlikten ayırmak için kullanılmaktadır. Konuşma algılamada ise sesler arası ayrım için sıklıkla kullanılmaktadır [102]. Şekil 3.2'de farklı sesler için normalize edilmiş RMS grafikleri görülmektedir.

$$RMS_p = \sqrt{\frac{1}{N} \sum_{n=p,N}^{(p+1)N-1} x(n)^2} \tag{3.2}$$



Şekil 3.2 Farklı sesler için RMS enerji grafikleri

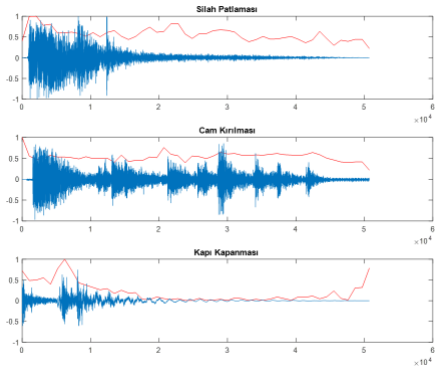
Şekil 3.1 ile Şekil 3.2 benzer görünse de RMS enerji grafiğinde seçilen veri penceresindeki en büyük değerin değil, o pencerede bulunan verilerin ortalamasının alındığı görülmektedir.

3.1.3 Sıfır Geçiş Oranı

Sıklıkla kullanılan ses özelliklerinden birisidir. Ses işaretinin ne sıklıkla işaret değiştirdiğinin dolayısıyla frekansın da bir ölçüsüdür. Denklem (3.3) ile elde edilir. Sıfır geçiş grafiklerinin düzenlilik, periyodiklik, kararlılık ve genlik aralığı gibi özellikleri incelenerek sesin sınıfı hakkında bilgi alınabilecektir. Ayrıca konuşma algılamada sessiz bölgeler daha yüksek SGO'ya sahipken konuşma olan bölgeler daha düşük SGO'ya sahiptir, bu sebeple bölge ayırmada kullanılmaktadır [102]. Şekil 3.3'de farklı sesler için normalize edilmiş sıfır geçiş oranı grafikleri görülmektedir.

$$SGO_p = \frac{1}{2} \sum_{n=p,N}^{(p+1)N-1} |\operatorname{sgn}(x(n)) - \operatorname{sgn}(x(n+1))| \tag{3.3}$$

$$\operatorname{sgn}(x(n)) = \begin{cases} 1, x(n) \geq 0 \\ -1, x(n) \leq 0 \end{cases} \tag{3.4}$$

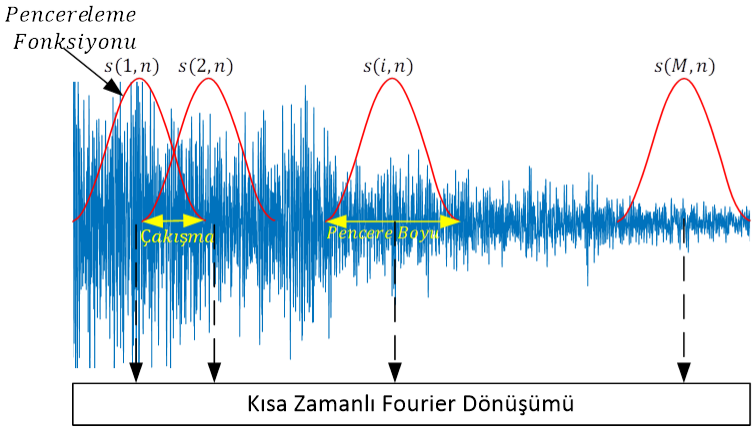


Şekil 3.3 Farklı sesler için SGO grafikleri

3.2 Spektral Özellikler

Zaman boyutundaki işaretler ayrık Fourier dönüşümü alınarak frekans boyutuna aktarılırlar. Bu dönüşümün hesap yükü işaretin uzunluğu ile üstel olarak artmaktadır. Ayrıca gerçek zamanlı sistemlerde sinyal sürekli olarak geldiği için sınırlı bir boyuta sahip değildir. Bu sebeple uygulamada işaretler 2^n örnek sayısındaki çerçevelere bölünerek bu kısa aralıklarla dönüşüme tabi tutulmaktadır. Kısa süreli Fourier dönüşümü de denilen bu işlem Şekil 3.4'deki gibi gösterilebilir.

Frekans boyutuna geçişlerde işaretin sınır bölgelerinde süreksizlik oluştuğu için bu bölgelerden kaynaklı frekans boyutunda istenmeyen spektral sızıntılar



Şekil 3.4 Kısa zamanlı Fourier dönüşümü

oluşmaktadır. Bunların önüne geçmek için Hamming penceresi gibi pencerelerle işaretin sınır bölgelerindeki değerleri bastırılır. Sınır bölgelerdeki işaretlerin bastırılmasının olumsuz etkisinin önüne geçmek içinse işaret çerçevelenirken iki çerçeve arası atlama boyutu çerçeve boyunun $\%25 - \%50 - \%75$'i gibi değerlerde seçilir. Ayrık Fourier dönüşümü bize sadece frekans değişkeniyle çıkış üretirken, kısa süreli fourier dönüşümü ise bize çerçeve numarası ve frekans olarak iki değişkenli sonuç üretir. Ayrık Fourier dönüşümü denklem (3.5), kısa süreli Fourier dönüşümü denklem (3.6) ile elde edilmektedir.

$$X(k) = \sum_{n=0}^{N_w-1} x(n)e^{-j\frac{2\pi nk}{N_w}} \tag{3.5}$$

$$S(p,k) = \sum_{n=0}^{N_W-1} x(n+pH)w(n)e^{-j\frac{2\pi nk}{N_W}} \tag{3.6}$$

Denklemlerde n örnek numarasını, H atlama boyutunu (hop size), N_w çerçevedeki örnek sayısını, p çerçeve numarasını, k ise frekans bölüntü numarasını (Bin number) ifade etmektedir. Toplam çerçeve sayısı denklem (3.7) ile toplam frekans bölümü sayısı da denklem (3.8) ile elde edilmektedir.

$$p_{\max} = \frac{N_w}{2} + 1 \tag{3.7}$$

$$k_{max} = \frac{toplam_ornek - N_w}{H} + 1 \tag{3.8}$$

Ayrık Fourier dönüşümü bize sadece frekans boyutunda sonuç üretirken kısa süreli

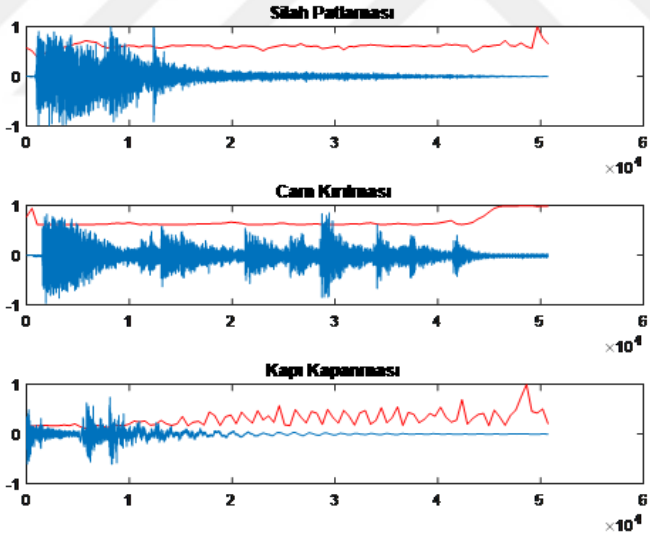
Fourier dönüşümü ise bize hem zaman hem frekans boyutunda sonuç üretmektedir. Burada pencere boyutunu uzun seçmek frekans çözünürlüğünü arttırken zaman çözünürlüğünü azaltmaktadır.

3.2.1 Band Enerji Oranı

Belirlenen bir frekans değerinden alçak olan frekanslardaki enerjinin, o frekanstan yüksek olan frekansların enerjilerinin oranıdır. Düşük frekanslarının baskınlığının seviyesini göstermektedir. Denklem (3.9) ile elde edilmektedir.

$$BEO_p = \frac{\sum_{i=0}^{F_b-1} m_p(i)^2}{\sum_{i=F_b}^{N} m_p(i)^2} \tag{3.9}$$

Denklemde p çerçeve numarası, N toplam frekans noktası sayısı, F_b belirlenen ayırma frekansı, i frekans noktasının numarası, $m_p(i)$ ise p numaralı çerçevenin frekans sprektrumunun i'nci frekans noktasındaki genlik değeridir. Şekil 3.5'de farklı seslere ait 8800 Hz frekansı sınır kabul edilerek elde edilmiş band enerji oranı grafikleri görülmektedir.



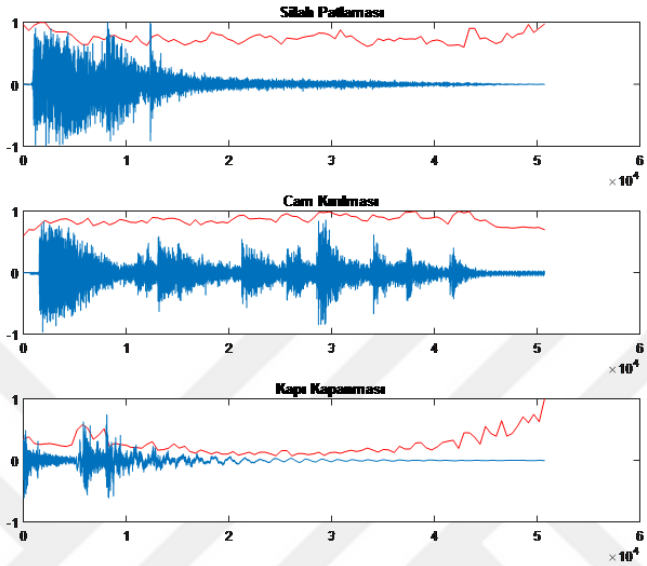
Şekil 3.5 Farklı sesler için band enerji oranı grafikleri

3.2.2 Spektral Merkez

İşaretin frekanslarının ağırlıklı ortalaması hesaplanarak elde edilmektedir ve genlik spektrumunun ağırlık merkezini vermektedir. İşaretteki baskın frekansların hangi bantta yoğunlaştığının göstergesidir. Denklem (3.10) ile elde edilmektedir.

$$SM_p = \boldsymbol{\mu}_p = \frac{\sum_{i=1}^{N} m_p(i) \cdot f_i}{\sum_{i=1}^{N} m_p(i)} \tag{3.10}$$

Denklemde f_i i numaralı frekans noktasının merkez frekansıdır [103]. Şekil 3.6'da farklı sesler için normalize edilmiş spektral merkez grafikleri görülmektedir.



Şekil 3.6 Farklı sesler için spektral merkez grafikleri

3.2.3 Spektral Entropi

Bu yöntemdeki temel hedef spektrumdaki tepe noktaları ve onların konumunu elde etmektir. Spektral dağılımdaki belli frekans bölgelerindeki enerji yoğunluklarının oluşturduğu sivrilikleri yakalamaktadır. Başka bir ifade ile de ses zarfındaki tepeleri yakalamaktadır. Denklem (3.11) ile frekans spektrumu, normalize edilmiş olasılık kütle fonksiyonuna dönüştürülür [103].

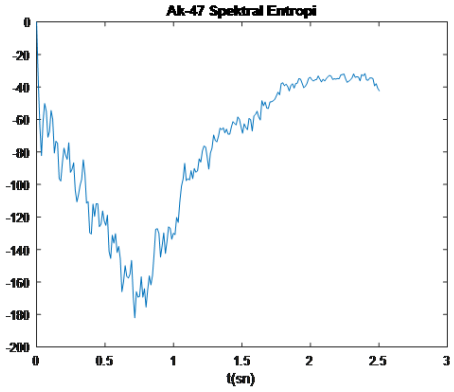
$$w_p(i) = \frac{m_p(i)}{\sum_{i=1}^{N} m_p(i)} \tag{3.11}$$

Burada $w_p(i)$ olasılık kütle fonksiyonudur. Bu olasılık kütle fonksiyonu denklem (3.12) ile spektral entropiye dönüştürülür.

$$H_p = \sum_{i=1}^{N} w_p(i) \log_2 w_p(i) \tag{3.12}$$

Burada H_p, p numaralı çerçevenin sprektal entropisidir. Şekil 3.7'de silah sesi

patlamasına ait spektral entropi grafiği görülmektedir.



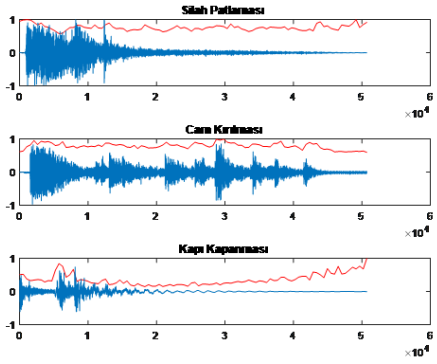
Şekil 3.7 Spektral entropi grafiği

3.2.4 Spektral Yayılım-Band Genişliği

İşaretin frekanslarının spektral merkez değerine göre nasıl yayıldığının bir ölçüsünü vermektedir, bu sebeple spektral yayılım olarak da isimlendirilmektedir. Sesin tınısıyla ilgi bilgi vermektedir. Spektral merkezden olan uzaklıkların ağırlıklı ortalamasıdır ve denklem (3.13) ile elde edilir [103].

$$BG_p = \sigma_p = \sqrt{\sum_{i=1}^{N} \left[(f_i - \mu_p)^2 \, w_p(i) \right]} \tag{3.13}$$

Burada μ_p p numaralı çerçevenin spektral merkezidir. Şekil 3.8'de farklı sesler için normalize edilmiş spektral yayılım grafikleri görülmektedir.



Şekil 3.8 Farklı sesler için speaktral yayılım grafikleri

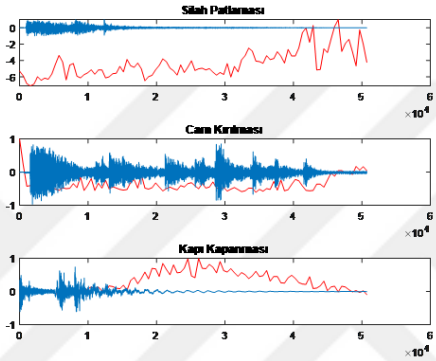
3.2.5 Spektral Skewness

Spektral genliğin ortalama değer etrafındaki yayılımının simetrikliğinin bir ölçüsüdür. Eğer simetrik dağılım varsa sıfır, dağılımda sağa doğru uzanan bir

kuyruk oluşuyorsa artı değerli, sola doğru bir kuyruk oluşuyorsa eksi değerli sonuç üretmektedir. Sinyaldeki sessiz bölgelerde bu özellik sıfır olacakken temel frekans değerine yakın yüksek enerjili bileşenlerin olduğu sesli bölgelerde ise yüksek olacaktır [103, 104]. Denklem (3.14) ile elde edilir.

$$SS_p = \frac{\sum_{i=1}^{N} \left[(f_i - \mu_p)^3 \, w_p(i) \right]}{\sigma_p^3} \tag{3.14}$$

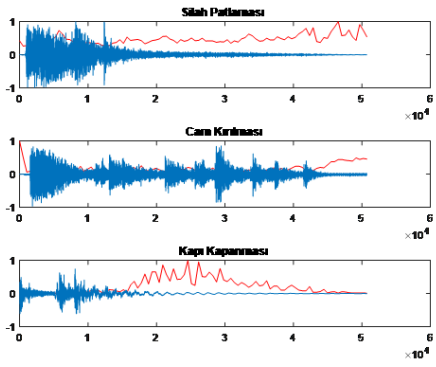
Burada σ_p p numaralı çerçevenin spektral yayılımıdır. Şekil 3.9'da farklı sesler için normalize edilmiş spektral skewness grafikleri görülmektedir.



Şekil 3.9 Farklı sesler için spektral skewness grafikleri

3.2.6 Sprektral Kurtosis

Spektral genlik dağılımının Gauss dağılımı şeklinde olup olmadığının bir ölçüsüdür ve normal dağılımlı olması durumunda 3 sonucunu verir. Eğer dağılım normal dağılıma göre daha düz ise 3'ten küçük daha dik ise 3'ten büyük değer üretir [103, 104]. Denklem (3.15) ile elde edilir. Şekil 3.10'da farklı sesler için normalize edilmiş spektral skewness grafikleri görülmektedir.



Şekil 3.10 Farklı sesler için speaktral kurtosis grafikleri

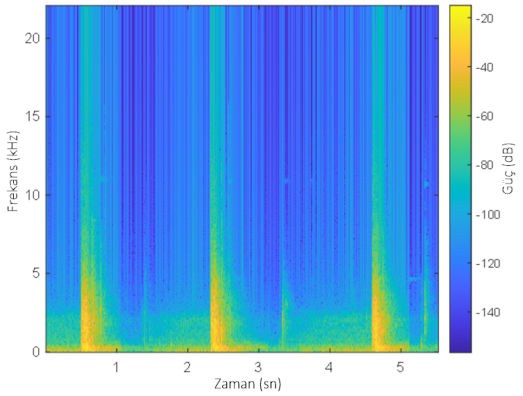
$$SK_p = \frac{\sum_{i=1}^{N} \left[(f_i - \mu_p)^4 \, w_p(i) \right]}{\sigma_p^4} \tag{3.15}$$

3.3 Zaman-Frekans Özellikleri

Derin öğrenme yöntemlerinde en sık kullanılan ses özellikleri zaman-frekans özellikleridir. Zaman özellikleri vektör boyutunda sonuçlar üretirken zaman-frekans özellikleri iki boyutlu matris olarak sonuçlar üretirler. Oluşan bu matrise sesin resmi de denilmektedir. Özellikle evrişimsel sinir ağları gibi derin öğrenme yapılarında kullanılmaya çok uygun özelliklerdir. Bunlarda en temel olanı spektrogram resmidir ve denklem (3.6) ile elde edilmektedir. Pencereleme her ses paketinin belirlenen noktalı FFT işlemleri yan yana konularak sesin resim matrisi oluşturulur. X ekseni genel olarak zaman eksenini oluştururken Y ekseni ise frekans eksenini oluşturmaktadır. Denklem (3.7)'den göründüğü gibi FFT işleminde kullanılan nokta sayısı frekans ekseni boyutunu ve çözünürlüğünü belirlemektedir. Denklem (3.8)'den göründüğü gibi pencere boyutu ve pencerenin kaç örnek ilerleme ile kaydırıldığı ise zaman ekseni çözünürlüğünü belirlemektedir. Zaman ekseni boyutu ilgilenilen sesin süresi, pencere boyutu ve pencere kaydırma örnek sayısına göre değişmektedir. Örnek sayısı N_s olan bir ses sinyali için zaman ekseni boyutu;

$$N_t = \frac{N_s - 1 - N_o}{N_w - N_o} \tag{3.16}$$

Burada N_o iki pencere arasıda çakışan örnek adeti, N_w penceredeki örnek sayısıdır. Şekil 3.11'de üç adet namlu patlaması sesinin spektrogram resmi görülmektedir.



Şekil 3.11 Namlu patlama sesinin sprektrogram resmi

3.3.1 Duyusal Özellikler

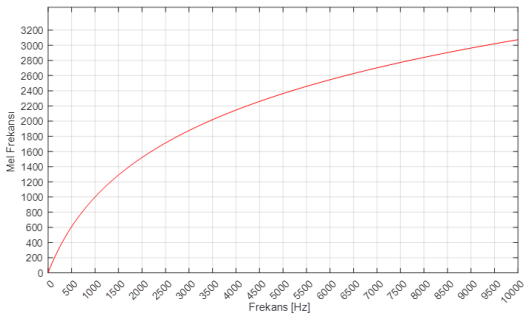
Araştırmalar insan kulağının frekansları lineer olarak algılayamadığını göstermektedir. Kulağın düşük frekanslarda ayırt edilebilme yeteneği doğrusallığa yakınlık göstermekteyken, yüksek frekanslara doğru ilerlendikçe ayırt edilebilme yeteneği doğrusallığını kaybetmektedir. Aynı frekans farkına sahip olsalar da insanın 500 Hz ile 1000 Hz frekansındaki iki sesin farkını algılaması, 12 kHz ile 12.5 kHz frekansındaki iki sesin farkını algılamasından daha kolaydır. Bu sebeple Mel [105], Bark ve ERB ölçekleri geliştirilmiştir.

3.3.2 Mel-Ölçeği

Mel ölçeği eşit mesafeler insan kulağında eşit algısal farklılık yaratan frekans aralığını üretmektedir. Normal frekanstan Mel frekansına dönüşüm denklem (3.17) ile yapılırken Mel frekansından normal frekansa dönüşüm denklem (3.18) ile elde yapılmaktadır. Koenig 0-10 kHz frekans aralığını, 1000 Hz altındaki frekanslar için doğrusal, üstündeki frekanslar için logaritmik algılama gölgesi olarak ikiye ayırmıştır. Fakat bu ayrımın doğruluğu konusunda net kanıtlar yoktur [106]. 1000 Hz frekans, 1000 Mel ölçeğine eşit olmaktadır. Şekil 3.12'de Mel Ölçeği grafiği görülmektedir.

$$f_m = 2595 \log\left(1 + \frac{f}{700}\right) \tag{3.17}$$

$$f = 700\left(10^{\frac{f_m}{2305}} - 1\right) \tag{3.18}$$



Şekil 3.12 Mel ölçeği-frekans grafiği

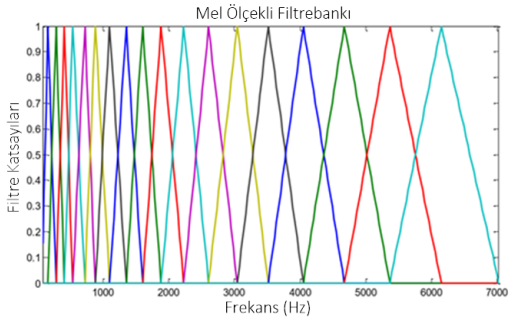
3.3.2.1 Mel Filtre Bankları

Özellikle konuşma tanıma sistemlerinde sıklıkla kullanılan özellik çıkartma yöntemlerinden birisidir. Kullanılması gereken filtre sayısı problemin türüne göre

değişse de oluşturulması benzer şekilde yapılmaktadır. Ses işaretinde incelenmek istenen en küçük ve en büyük frekans değerleri belirlendikten sonra bu sınır değerlerin Mel ölçeği değeri hesaplanarak logaritmik bölgeye geçilir. Logaritmik bölge kullanılmak istenen filtre adedince eşit aralıklara bölünür. Belirlenen her bir Mel frekansı noktası denklem (3.18) yardımıyla tekrar frekans bölgesine taşınır ve denklem (3.19)'deki f_c değerleri bulunur. Daha sonra her bir Mel frekans noktasında kendisine komşu bir önceki ve bir sonraki f_c frekanslarına doğrular çizilerek filtre bantları oluşturulur. Şekil 3.13'de 0 Hz ile 7 kHz arası 18 filtreden oluşan Mel filtre bantları görülmektedir. Mel filtre bankası denklem (3.19) ile elde edilmektedir.

$$H(k,m) = \begin{cases} 0 & f(k) < f_c(m-1) \\ \frac{f(k)-f_c(m-1)}{f_c(m)-f_c(m-1)} & f_c(m-1) \leq f(k) < f_c(m) \\ \frac{f(k)-f_c(m+1)}{f_c(m)-f_c(m+1)} & f_c(m) \leq f(k) < f_c(m+1) \\ 0 & f(k) \geq f_c(m+1) \end{cases} \tag{3.19}$$

Denklemde m Mel filtre numarasını, f_c Mel filtresi merkez frekansını ifade etmektedir. Sonuçta $k \times m$'lik bir H, Mel filtre bankası matrisi oluşmaktadır.



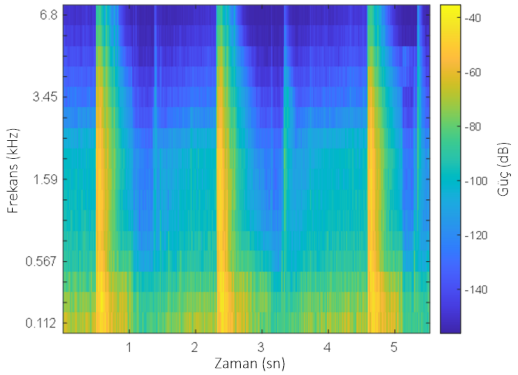
Şekil 3.13 0-7000 Hz arası oluşturulmuş 18 adet Mel filtreleri

3.3.3 Mel Spektrogram

Spektrogram resminin Mel filtre bankları ile çarpılması ile elde edilmektedir. Mel Spektrogram'ın frekans ekseni bank filtre adedi kadar çözünürlüğe sahiptir. Zaman boyutu spektrogram resmi ile aynıdır. Spektrograma göre daha düşük boyutlu ve frekans eksenini daraltan bir özelllik olmasına ragmen duyusal bir özelliktir. Yapılan bir çalışmada Mel sprektrogramın düşük veri setlerinde spektrogramdan daha belirgin sonuçlar ürettiği belirtilmiştir. Daha sonra yapılan başka bir çalışmada ise veri seti boyutu artsada iki özelliğinde yakın sonuçlar ürettiğini göstermiştir. Düşük boyutu sebebi ile eğitim ve tahmin işlemlerinin hesap yükünü azaltacağı

için işlem sürelerini hızlandırmaktadır. Denklem (3.20) ile hesaplanmaktadır. Şekil 3.14'de üç adet namlu patlaması sesinin Mel spektrogram resmi görülmektedir.

$$M_{\text{spe}}(p,m) = S(p,k) \times H(k,m) \tag{3.20}$$



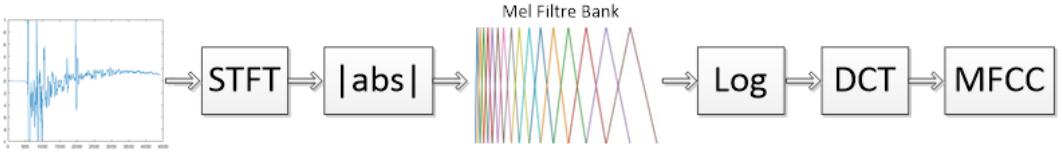
Şekil 3.14 Namlu patlaması sesinin Mel sprektrogram resmi

3.3.4 Mel Frekans ve Gamaton Kepstal Katsayıları

Mel Frekansı Keptral Katsayıları ses tanımlama en çok kullanılan özelliklerden birisidir. Şekil 3.15'de elde edilme adımları gösterildiği gibi ses sinyali çerçevelere ayrılıp hep bir çerçevenin kısa süreli Fourier dönüşümü alındıktan sonra, genlik değerleri Mel filtreleriyle çarpılarak her elde edilen toplamın logaritması alınmaktadır ve denklem (3.21) ile hesaplanmaktadır. Daha sonra ayrık kosinüs dönüşümü yapılarak gerçel değerlerden oluşan MFCC katsayıları elde edilmektedir. MFCC katsayıları denklem (3.22) ile hesaplanmaktadır.

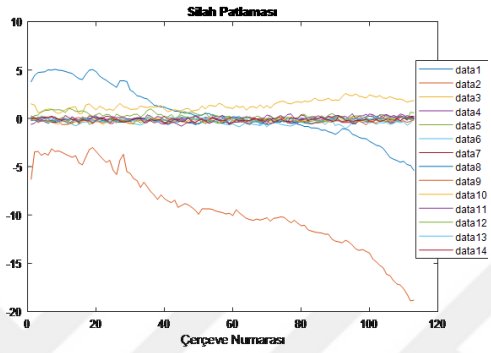
$$S_m = \ln\left(|S_p| \cdot H\right) \tag{3.21}$$

$$\text{MFCC}(i) = \sum_{m=1}^{M} S_m(m) \cdot \cos\left(i\frac{\pi}{M}\left(m - \frac{1}{2}\right)\right) \tag{3.22}$$



Şekil 3.15 MFCC oluşturma aşamaları

MFCC'de log basamağı konuşma sırasında ses yolunun yapmış olduğu filtrelemeyi modellemek için kullanılmıştır. Yapılan bir çalışmada ortam sesleri için MFCC katsayılarının elde edilmesinde log adımının kaldırılmasıyla elde edilen sonucun uyarı seslerini tanımlamada daha başarılı olduğunu saptanmıştır [107]. Şekil 3.16'da silah sesi için 14 katsayılı MFCC grafiği görülmektedir. MFCC'ye benzer şekilde elde edilen bir diğer katsayı ise gamaton kepstral katsayılarıdır(GTCC). Bu yöntemde gamaton filtre bankları kullanılmaktadır. Gamatone filtre banklarının kulaktaki basilar mebran frekans tepkisinin benzetimini yaptığı varsayılmaktadır.



Şekil 3.16 14 Katsayılı MFCC grafiği

3.3.5 3.3.5 Bark Ölçeği ve Bark Spektrogram

Bark Spektrogram, insan işitme sistemine daha uygun bir frekans çözünürlüğü sunan bir diğer analiz yöntemidir. Mel ölçeğine benzer şekilde insan kulağının frekansları logaritmik bir şekilde algılamasına temeline dayanmaktadır ve denklem (3.23) ile verilen formül ile ölçeklenir. 500 Hz altında daha doğrusal eğime sahipken, sonrasında logaritmik grafiğe yaklaşır. Bark spektrumu, bu işitme modelini esas alarak geliştirilmiştir ve frekansları 24 Bark bandına ayırır.

$$f_{\text{sark}} = \frac{26.81 \times f}{1960 + f} - 0.53 \tag{3.23}$$

$$f_{\text{Bark}} = \begin{cases} f_{\text{Bark}} + 0.15 \times (2 - f_{\text{Bark}}), & f_{\text{Bark}} < 2 \\ f_{\text{Bark}} + 0.22 \times (f_{\text{Bark}} - 20.1), & f_{\text{Bark}} > 20.1 \end{cases} \tag{3.24}$$

Bark spektrum kenar frekansları 0 Hz, 100 Hz, 200 Hz, 300 Hz, 400 Hz, 510 Hz, 630 Hz, 770 Hz, 920 Hz, 1080 Hz, 1270 Hz, 1480 Hz, 1720 Hz, 2000 Hz, 2320 Hz, 2700 Hz, 3150 Hz, 3700 Hz, 4400 Hz, 5300 Hz, 6400 Hz, 7700 Hz, 9500 Hz, 12000 Hz, 15500 Hz iken bu merkez frekansları ise 50 Hz, 150 Hz, 250 Hz, 350 Hz, 450 Hz, 570 Hz, 700 Hz, 840 Hz, 1000 Hz, 1170 Hz, 1370 Hz,

1600 Hz, 1850 Hz, 2150 Hz, 2500 Hz, 2900 Hz, 3400 Hz, 4000 Hz, 4800 Hz, 5800 Hz, 7000 Hz, 8500 Hz, 10500 Hz, 13500 Hz'dir. Bu bantlar, insan işitme sisteminin daha duyarlı olduğu bölgelerde daha yoğun, daha az duyarlı olduğu yerlerde daha geniştir. Bark spektrumu 15.5 kHz'e kadar ölçeklenmiştir. Bark ölçeğinden frekansa geçiş denklem (3.25) ile elde edilmektedir.

$$f_{\text{Bark}} = \begin{cases} \frac{f_{\text{Bark}} - 0.3}{0.85}, & f_{\text{Bark}} < 2 \\ \frac{f_{\text{Bark}} + 4.422}{1.22}, & f_{\text{Bark}} > 20.1 \end{cases} \tag{3.25}$$

$$f = 1960 \frac{f_{\text{Bark}} + 0.53}{26.28 - f_{\text{Bark}}} \tag{3.26}$$

3.3.6 ERB Ölçeği ve Spektrumu

Eşdeğer dikdörtgen bant genişliği spektrumu da insan kulağı frekans çözünürlüğünü modellemek için kullanılan yöntemlerden birisidir. Mel ve Bark spektrumlara benzer şekilde alt frekanslarda filtreler daha dar üst frekanslarda ise daha geniştir. Erb frekansı denklem (3.27) ile elde edilmektedir.

$$f_{ERB} = A \log_{10}(1 + f \times 0.00437) \tag{3.27}$$

$$A = \frac{1000 \log_e 10}{24.7 \times 4.37} \tag{3.28}$$

ERB frekansından, frekansa geçiş denklem (3.29) ile elde edilmektedir.

$$f = \frac{10^{\frac{f_{ERB}}{A}} - 1}{0.00437} \tag{3.29}$$

3.3.7 Delta ve Delta Delta Özelliği

Bir ses sinyalinin spektorgramları elde edilirken yapılan pencereleme sonucunda elde edilen değerler bir birinden bağımsız sonuçlar üretmektedir. Pencereler arasındaki spektrogram değişimi sese ait dinamik bilgiler taşımaktadır. Bu değişimi incelemek için bir pencerenin MFCC değerleri kendinden bir önceki pencerenin MFCC değerinden çıkartılarak Delta özelliği elde edilir. Delta özelliği kendinden bir önceki delta özelliğinden çıkartılarak ise deltaların değişimini gösteren delta delta özelliği elde edilir.

3.4 En Küçük Fazlalık En Büyük Alaka (MRMR) Özellik Seçim Yöntemi

Makine öğrenimi algoritmalarının daha hızlı ve doğru çözümler üretmesi için özellik seçim algoritmaları kullanılmaktadır. Bu algoritmalar alakasız özelliklerin kullanımının önüne geçerek giriş verilerinin boyutlarının azaltılmasını sağlamaktadır. En küçük fazlalık en büyük alaka (Minimum Redundancy Maximum Relevance-MRMR) yöntemi sınıflandırma için etkili özellik kümelerini çok hızlı şekilde üreten başarılı bir yöntemdir. Bu yöntem özellik seçerken gereksiz ve tekrarlayıcı (redündant) özellikleri elerken, sınıfladırma ya da tahminleme ile ilişkili (relevant) özellikleri tutmaktadır. Her bir özelliğin hedef ile olan ilişkisi karşılıklı bilgiye (mutual information) dayandırılarak ölçülmektedir. Yüksek karşılıklı bilgi değeri o özelliğin hedef ile güçlü ilişkide olduğunu göstermektedir. İki özellik arasındaki düşük karşılıklı bilgi ise bu özelliklerin birbirinden bağımsız olduğunu ve bu sebeple gereksiz tekrar oluşturmayacağını belirtmektedir. Bunu hesaplamak için iki özellik birer ranstlantı değişkeni (X, Y) olarak değerlendirilir ve karşılıklı bilgi değeri denklem (3.30) ile hesaplanır [108].

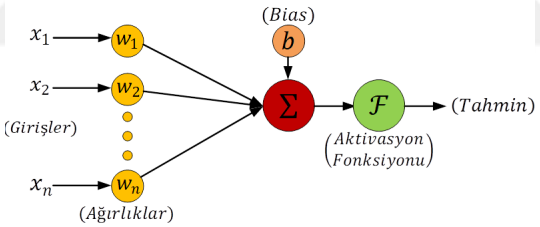
$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log \left(\frac{p(x,y)}{p_1(x) p_2(y)} \right) \tag{3.30}$$

Burada X ve Y rastlantı değişkenlerinin marjinal olasılık dağılımları sırasıyla $p_1(x)$ ve $p_2(y)$ olarak, aralarındaki birleşik olasılık dağılım fonksiyonu ise $p(x,y)$ olarak gösterilmiştir. Algoritma adımları aşağıda özetmenmiştir.

- Tüm özelliklerin hedef ile olan karşılıklı bilgisi elde edilir ve en yüksek değere sahip olan secilir.
- Seçilen özellik ile sonraki özelliğin birbirleri ile olan karşılıklı ilişkisine bakılır. Düşük ilişkiye sahip ise o özellik de secilir.
- Her özellik için ilişki ve tekrarlayıcılık değerleri elde edilip, bu değerler birbirinden çıkartılarak sıralama değeri elde edilir.
- Sıralama değeri en yüksek olan özellik bir sonraki seçim için kullanılır.

4
DERİN ÖĞRENME YÖNTEMLERİ

Makine öğrenmesi yöntemleri daha düşük boyutlarda veri setleri ile yapılan seçilen öz niteliklere göre eğitimi daha basit ve hızlı gerçekleşen yöntemlerdir. Genel olarak belirli görevleri yerine getirebilen derin öğrenme yöntemlerine göre daha basit algoritmalardan oluşmaktadır. Özellikle resim tanımlama gibi karmaşık uygulamalarda yetersiz kalmaktadır. Artan veri setleri, gelişen işlemci teknolojileri sonuçunda derin öğrenme modelleri giderek geliştirilmiştir. Derin öğrenme yöntemleri yapay sinir ağlarını kullanan yapılardır. Algılayıcı kavramı insan sinir hücresinin modellenmesi ile geliştirilmiştir ve öğrenebilen en küçük yapıdır. Şekil 4.1'de N adet girişe sahip tek algılayıcılı model görülmektedir.

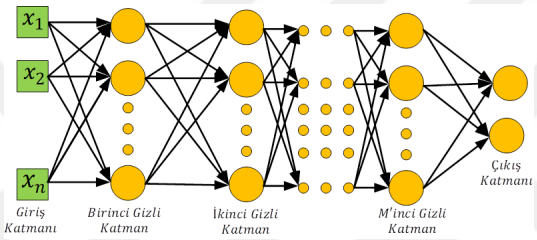


Şekil 4.1 Tek algılayıcılı ağ modeli

Bu yapıların bir araya getirilmesi ise ile ilk olarak tek katmanlı daha sonra ise çok katmanlı algılayıcı (multi layer perceptron) olarak adlandırılan derin öğrenme modelleri geliştirilmiştir. Daha sonra MLP modellerinin önüne evrişim katmanları eklenerek özellikle resim tanımlama çığır açan evrişimsel sinir ağları geliştirilmiştir. Derin öğrenme yöntemlerinin en önemli özelliklerinden birisi eğitildikleri verilerden anlamlı özellik çıkartımını gerçekleştirebilmeleridir. Bu özellik çıkarımının başarılı sonuçlar üretebilmesi için çok geniş veri setleri ile eğitilmeleri gerekmektedir. Derin öğrenme yöntemleri, makine öğrenmesi yöntemlerinden farklı olarak ağı ve eğitimi tanımlayan hiperparametrelere sahiptir. Aşağıda ses lokalizasyonu çalışmalarında en sık kullanılan derin öğrenme yöntemleri açıklanmıştır.

4.1 Çok Katmanlı Algılayıcı

MLP ağı Şekil 4.2'de görüldüğü gibi farklı sayılardaki nöronlardan oluşan katmanların ardışık olarak sıralanması ile elde edilen ağ yapısıdır. Ağın yapısı aynı zamanda hiperparametrelerini oluşturmaktadır. Öğrenme nöronlar arasındaki ağırlıkların ve nöron eşiklerinin (bias) değerlerinin doğru şekilde belirlenmesi ile sağlanır. Bunun için geri yayılım algoritması kullanılmaktadır. Ağ ileri yönlü beslendikten sonra oluşan sonuç ile olması gereken sonuç arasındaki hata belirlenir. Oluşan hata ile geriden başlanarak tüm ağırlık ve biasların türevleri alınarak hata üzerine etkileri belirlenir. Daha sonra seçilen optimizasyon algoritmasına göre her eğitim adımında (epoch), öğrenim oranına (learning rate) göre ağırlık ve eşik değerleri güncellenir değiştirilmesi sağlanır. Bu işlemler belirlenen eğitim adımına yada belirlenen bir doğruluk oranına ulaşıncaya kadar devam ettirilir.



Şekil 4.2 Çok katmanlı algılayıcı ağ modeli

En sık kullanılan optimizayon algoritması SGD (Stochastic Gradient Descent) algoritmasıdır. Denklem (4.1) ile

$$W_{ij}^{l}(n+1) = W_{ij}^{l}(n) - \alpha_r \frac{\partial E}{\partial W_{ij}^{l}(n)} \tag{4.1}$$

Burada $W_{ij}^{l}, (l-1)$. katmandaki i. nöron ile l. katmandaki j. nöron arasındaki güncellenecek ağırlık değeri, α_r öğreme oranı, E ise performans fonksiyonuna göre hesaplanan hata değeridir. Katman sayısı, her katmandaki nöron sayısı, her katmanda kullanılan nöronların aktivasyon fonksiyonları ağ yapısını oluşturan hiperparametreleridir. Öğrenme oranı, performance fonksiyonu, optimizasyon algoritması, eğitim adımı, kabul edilir hata değeri ise eğitimi belirleyen hiperparametrelerdir. Özellikle resim gibi iki boyutlu ve çok kanallı verilerin MLP gibi doğrusal dizilime sahip nöronlara giriş olarak verilmesi için resim yapısının doğrusal forma getirilmesi gerekmektedir. Resimlerde farklı sütunlarda yer alan pikseller bir birine desen oluşturma noktasında bağlıdır, fakat düzleştirildiklerinde bu desensel bağ ortadan kalkmaktadır. Ayrıca büyük boyutlu resimlerde her piksele bir nöron düşeceği için çok nöronlu giriş katmanına sebep olmaktadır. Tam bağlantılı nöron yapılarından dolayı da eğitim zorlaşmakta ve

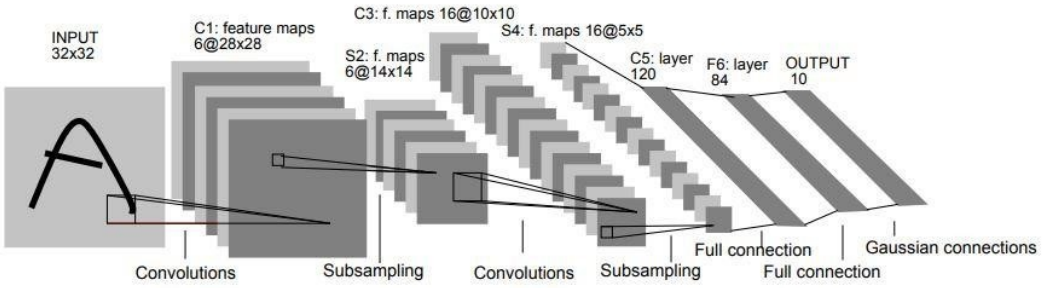
aşırı uyumun oluşmasına meydan vererek genelleşme yeteneğinin kaybolmasına neden olmaktadır. İlaveten resimlerde bir nesnenin yeri standart olmadığı için eğitilmiş bir MLP'ye resimde farklı yere konumlanmış objeyi tanımlayamayacaktır. Vektör haline getirilen bir resimde uzamsal bilgiler kayıp olmaktadır. Fakat ses sinyallerinden elde edilen spektrogram görüntüleri her zaman frekans ekseninde aynı uzamsal bilgiye sahip olduğu için spektrogram tabanlı uygulamalarda doğrusallaştırılan resmin etkisi azalmakta ve ses uygulamasında kullanılabilirliği devam etmektedir.

4.2 Evrişimsel Sinir Ağları

Evrişimsel (konvolüsyonel) ağlar özellikle 2 boyutlu matris yapısında giriş verileri olan durumlarda kullanılmaktadır. Bu giriş verilerinin en bilinenleri resimler ve spektrogramlardır. Evrişimsel ağlar belirlenen boyutlarda ve adette filtreleri oluşturarak veri üzerinde anlamlı tespitler yapan özellikler çıkartmaktadır. Konvolüsyon işlemi, işaret işlemede filtreleme sonrasında oluşacak filtrelenmiş sinyalin matematik modellenmesinde, ya da sayısal kontrolde birim dürtü yanıtı bilinen doğrusal zamanla değişmeyen (LTI) bir sisteme uygulanacak herhangi bir girişin oluşturacağı çıkış değerinin elde edilmesinde kullanılmaktadır. Matematiksel olarak ise bir fonksiyonun başka bir fonksiyon üzerinde yaptığı değişimi ifade eden integral alma işlemidir. Buna benzer olarak konvolüsyonel katmanlardaki filtrelerde bu veri üzerinde gezerek karşılıklı gelen pikselleri filtre ile nokta çarpım yapılarak konvolüsyon işlemi uygulanır. Eğitim süresi boyunca en anlamlı sonuçlar oluşturan filtre yapıları oluşturulmaktadır. Birden fazla konvolüsyonel katman kullanıldığında havuzlama işlemi yapılarak birbiri ile ilişkili desenlerin aynı noktada toplanması sağlanmaktadır. Bu şekilde bir özelliği tespit eden filtre özellik nerede olursa olsun onu yakalayıp konumlandırabilmektedir. Ayrıca boyut küçültülerek sonraki işlem hızları arttırılmaktadır. Şekil 4.3'da 1998 yılında sunulan LeNet-5 konvolüsyonel ağ yapısı görülmektedir [109]. Ağ 3 adet konvolüsyon katmanına, 2 adet havuzlama katmanına ve 2 adet tam bağlantılı katmana sahiptir.

Konvolüsyonel ağda kullanılan diğer katmanlardan bazıları ve görevleri;

- **Tam Bağlantılı Katman,** MLP'ye benzer olarak ağırlık ve eşik değerlerine sahip nöronlardan oluşur ve konvolüsyon katmanı ile çıkışı birbirine bağlar. Sınıflama işlemi bu katmanda yapılmaktadır. Konvolüsyon katmanından elde edilen özellik çıkartımları düzleştirilerek bu katmana girdi olarak verilir. Bu katmanda elde edilen sonuç çıkış katmanına taşınır.



Şekil 4.3 İlk defa önerilen evrişimsel ağ modeli

- **Aktivasyon Fonksiyonu,** Aktivasyon fonksiyonu bir katman değildir ve üzerinde herhangi bir ağırlık ya da öğrenme verisi barındırmaz. Genel olarak konvolüsyon katmanlarından sonra getirilir. Doğrusal olmayan fonksiyonlardan oluşur ve ağa doğrusal olmayan bir özellik kazandırır. Konvolüsyon işlemi sonrası elde edilen değerler aktivasyon fonksiyonundan geçirildikten sonra havuzlama işlemine tabi tutulurlar. Aktivasyon fonksiyonu özellik çıkartımı sırasında her katmandan elde edilen bilginin bir sonraki katmana ulaşıp ulaşmamasına karar verir. Neticede yapay sinir ağında öğrenilmesi gereken bir bilgi olup olmadığına karar veren bir fonksiyondur.

- **Bırakma (Dropout) Katmanı,** Bu katman tam bağlantılı katmanın önünde yer alır. Eğitim sırasında aşırı öğrenmenin önüne geçmek ve ağın genelleme yeteneğinin arttırılmasını sağlamak için kullanılmaktadır. Ağ eğitimi sırasında tam bağlantılı katmandaki nöronlardan bazılarını eğitimden çıkartır ve bu şekilde ağın ezberleme yapmasının önüne geçmek hedeflenir.

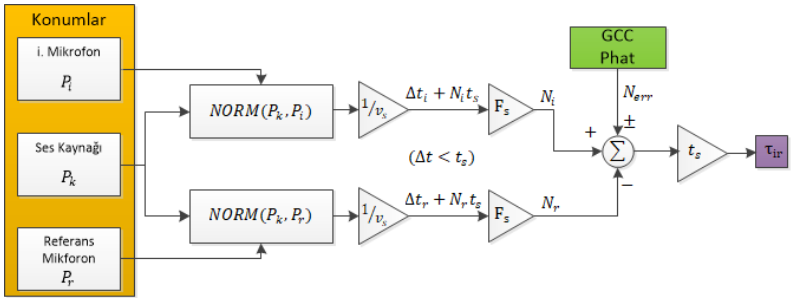
- **Toplu Normalleştirme,** Her bir katmana giren veriyi toplu olarak aynı ortalama ve standart sapma değerine göre normalleştiren bir katmandır [110]. Eğitimin daha kararlı olmasına yardımcı olur, öğrenme oranlarının belirlemesini kolaylaştırır ve özellikle eğitim adımının azalmasını sağlar. Öğrenmenin kararlı olması, kayıp eğrisinin daha kararlı olmasını sağlar. Üzerinde eğitilebilir bir veri bulundurmaz sadece matematiksel işlemleri yapan bir katmandır.

5
MULTİLATERASYON YÖNTEMİNDE DUYARLILIK ANALİZİ VE KAPALI FORM DoA FORMÜLLERİ

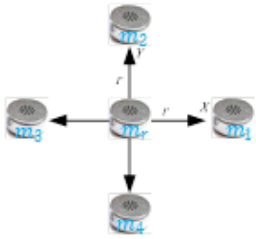
GCC-Phat algoritması kullanılarak referans mikrofona göre elde edilen TDoA değerleri Şekil 5.1'de gösterildiği gibi modellenerek benzetim çalışmalarında kullanılmıştır. Modelden de görüldüğü gibi TDoA değerleri mikrofonlar arası mesafeye ve örnekleme frekansına bağlıdır. Her örnek için kayıt edilen her zaman damgası (time stamp) örnekleme zamanı t_s'in tam katı olmaktadır. $\Delta t < t_s$ zaman aralıklarında zaman kuantalama hatası oluşmaktadır. Çevresel ve donanımsal gürültüler mikrofonlara gelen seslerde bozulmalar yaratarak GCC-Phat algoritmasında hatalı sonuçlar çıkmasına sebep olabilmektedir. Bu hatalar modelde N_{err} olarak modellenmiştir.

Benzetim çalışmalarında Şekil 5.2'de görüldüğü gibi bir adet merkezde ve dört adet belirli yarıçapta eşit açılarla dağılmış toplam 5 mikrofondan oluşan sensör dizisi kullanıldığı varsayılmıştır ve referans mikrofon merkezdeki mikrofon seçilmiştir. P_k konumundan gelecek bir patlama için elde edilecek çözüm denklemi (2.41) denklemine benzer oluşacaktır.

Atıcı konumun Şekil 5.3'de görüldüğü gibi değiştiği varsayılarak benzetim çalışmaları yapılmıştır. Atıcının orjine uzaklığı 50'şer metre aralıklarla arttırılarak

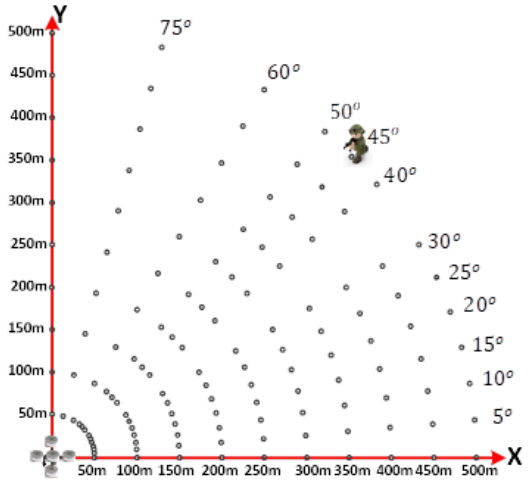


Şekil 5.1 Simulasyon çalışmalarında kullanılan TDoA süresi modeli



Şekil 5.2 Benzetim çalışmalarında kullanılan mikrofon dizisi yapısı

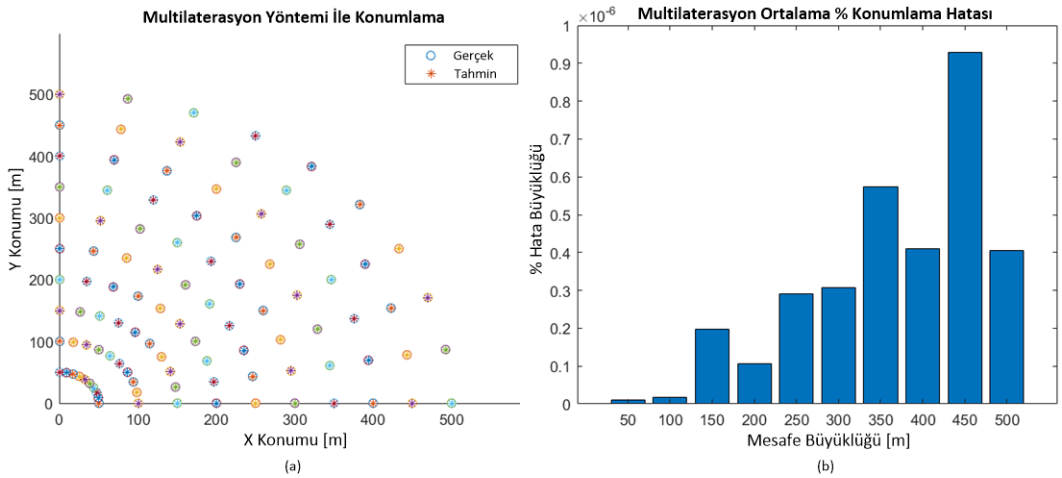
500 m mesafeye kadar ve geliş açılarının 0° , 5° , 10° , 15° , 20° , 25° , 30° , 40° , 45° , 50° , 60° , 75° , 90° derecelerde olduğu varsayılarak aynı düzlemde 130 farklı nokta için inceleme yapılmıştır.



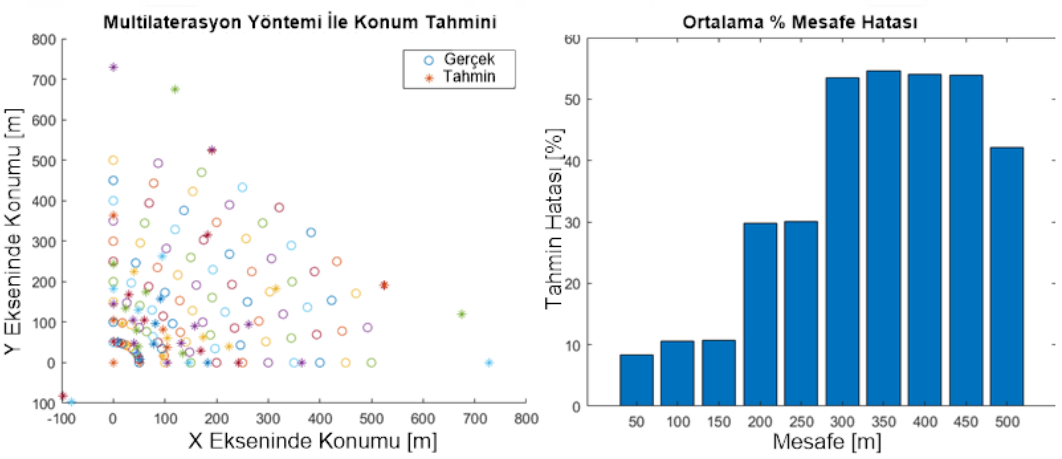
Şekil 5.3 Benzetim çalışmasında kullanılan mikrofon dizisi konumu ve patlama noktaları dağılımı

Ses hızının 343 m/s olduğu kabul edilmiştir. İlk incelemede örnekleme frekansından dolayı kayıp olan Δt süresinin ve N_{err} hatalarının olmadığı varsayılarak elde edildiğinde oluşan konumlama tahmini grafiği ile her mesafedeki ortalama hata grafikleri Şekil 5.4 olarak elde edilmiştir.

Analitik olarak hesaplanan TDoA değerlerine göre elde edilen çözümde hataların 10^{-6} değerlerinde olduğu ve çok başarılı tahminleme sonuçları elde edildiği görülmektedir. Matlab programında bir değişken virgülden sonra 16 basamak hassasiyete sahip olduğu için TDoA değerlerindeki bu analitik hassasiyetin örnekleme frekansıyla sağlanabilmesi için gerekli olan t_s süresi 1 fs (femtosaniye)'dir ve bu 1000 THz'lik bir örnekleme frekansına eşittir. Mevcut hızlı osiloskopların bile 1 GHz frekansında örneklediği düşünüldüğünde bu değerin bile bir milyon katında bir örnekleme değeridir. Örnekleme frekansını 1 MHz kabul ederek benzetim çalışması yaptığımızda t_s zamanı 1 μs değeri olmaktadır ve elde edilen konumlama tahmini grafiği ve her mesafedeki ortalama hata



Şekil 5.4 Analitik hesaplanan TDoA süresiyle yapılan simulasyon ile elde edilen konumlama ve hata grafiği



Şekil 5.5 1 MHz F_s ile hesaplanan TDoA süresiyle yapılan simulasyon ile konumlama ve hata grafiği

grafiği Şekil 5.5'deki gibi olmaktadır. Sonuçlar incelendiğinde gerçek zamanlı uygulamalarda mutlaka oluşacak N_{err} değeri olmadan bile hata değerlerinin %10 ile %55 arasında değiştiği görülmektedir. Bu bize 0.5 m mikrofon açıklığında bu yöntemin kabul edilebilir sonuçlar üretmesi için örnekleme frekansının 5 MHz üzerinde olması gerektiğini göstermektedir. Bu örnekleme frekansı gerçek ses kartları için ulaşılabilir değerler değildir. Düşük örnekleme frekansı ve yüksek örnekleme frekansı ses kartları için farklı yarıçap değerleri için elde edilen ortalama hata değerleri Tablo 5.1'de verilmiştir. Simulasyonlara Nyquist krtiterlerini sağlayacak minimum örnekleme frekansı 44.1 kHz ve hızlı ses kartlarından gerçekleştirilebilecek örnekleme frekansı olan 500 kHz ile devam ettiğimizde TDoA ölçüm hatasının olmadığı durum için mikrofonlar arası mesafe değiştirilerek her durum için denklem (2.41) elde edilip denklem (2.43) ile çözülerek mesafe hatasının %10'un altına inildiği yarıçap belirlenmiştir. Denklem (5.1)'de kullanılan

yüzde ortalama mutlak hata (MAE) fonksiyonu verilmiştir. Burada $\widehat{p}_k$ tahmin edilen konum vektördür. N_p ise bir konum için tekrarlanan patlama simülasyonu sayısıdır.

$$\%MAE = \left(\frac{1}{N_p} \sum \frac{\|p_k - \widehat{p}_k\|}{\|p_k\|} \right) \times 100 \tag{5.1}$$

Tablo 5.1 Farklı mikrofon dizisi yarıçapları ve örnekleme frekansları için elde edilen % RMS hata oranları

Fs [kHz]	r [m]	50m	100m	150m	200m	250m	300m	350m	400m	450m	500m	Mean
44.1	0.5	70.95	105.4	94.65	104	96.42	98.76	103.6	99.67	102.2	98.31	97.39
44.1	1	16.29	90.14	53.49	105.6	75.7	94.73	113	93.37	108.8	91.73	84.29
44.1	2	3.88	22.79	12.66	36	37.67	33.6	86.23	29.09	73.88	42.09	37.87
44.1	4.5	0.83	4.08	5.35	6.30	11.23	5.28	19.51	6.38	16.49	12.57	8.80
500	0.5	11.3	13.86	34.32	43.66	33.06	35.70	109.74	51.57	41.17	63.43	43.77
500	1	3.48	2.97	8.11	13.42	9.52	16.44	34.13	17.36	15.67	28.32	14.94
500	1.25	1.21	3.06	6.95	9.86	5.85	7.13	18.32	6.43	7.44	22.23	8.85

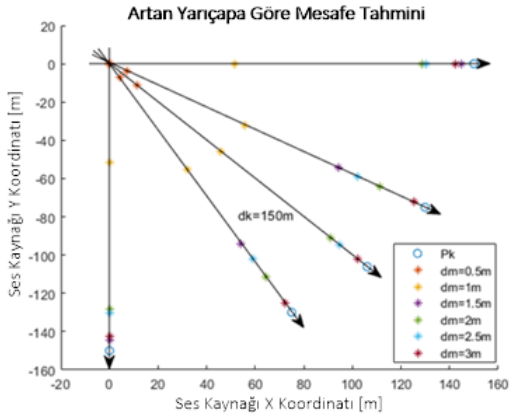
Ortalama hatanın %10'un altına düştüğü koşullarda her TDoA süresine Şekil 5.1'de görüldüğü gibi sıfır ortalamalı düzgün dağılımlı olarak en fazla 5 örnek olacak şekilde N_{err} ölçüm hataları eklenmiştir. Her konum için 1000 tekrarlı Monte Carlo simulasyonu yapılarak hatanın %10'un altına indiği yarıçap değerleri iki örnekleme frekansı için belirlenmiş ve sonuçları Tablo 5.2'de gösterilmiştir.

Tablo 5.2 Normal dağılımlı hata eklenmiş gecikme süreleriyle elde edilen %RMS hata oranları

Fs [kHz]	r [m]	50m	100m	150m	200m	250m	300m	350m	400m	450m	500m	Mean
44.1	4.5	8.26	18.33	23.26	42.35	40.42	56.78	76.12	71.86	82.23	78.58	49.82
44.1	10	1.63	3.47	4.54	7.05	7.78	9.69	12.89	13.37	15.85	16.28	9.26
500	1.25	9.56	18.59	26.57	50.92	54.29	59.33	82.17	80.97	78.74	77.61	53.87
500	3	1.61	2.98	4.33	6.85	8.16	9.04	12.89	13.75	13.49	15.34	8.85

Mikrofonlar arasındaki d_m uzaklığı 0.5 m'den 3 m'ye kadar 0.5 m adımlarla arttırılmıştır. F_s örnekleme frekansı 44.1 kHz olarak belirlenmiştir. Şekil 5.6'da her d_m uzaklığına göre denklem (15)'in çözümü ile elde edilen LSE çözümleri '*',gerçek kaynak pozisyonu ise 'o' ile gösterilmiştir. Mikrofonlar arası mesafe arttırıldıkça konum tahminlerinin gerçek konum değerine yaklaştığı gözlemlenmektedir.

Şekil 5.6 incelendiğinde, mikrofonlar arası mesafe doğru konum tahmini için yetersiz ve yüksek tahmin hataları olmasına rağmen konum tahminlerinin aynı doğru üzerinde değiştiği görülmüştür. Bu gözlem bize multilaterasyon yöntemine dayalı geliş açısı tespitini mikrofonlar arası mesafeden bağımsız yüksek doğrulukta yapılabileceği fikrini vermiştir.



Şekil 5.6 Mikrofon arası mesafe artışına göre konumlama tahmini değişimi

5.1 Eş Merkezli ve Dikey Yerleşimli Mikrofon Çiftleri için Kapalı Geliş Yönü Tespit Çözümü

Üç mikrofondan oluşan bir mikrofon dizi için birinci mikrofon referans alınarak multilaterasyon denklemleri denklem (2.41)'de gösterilmişti. Ses kaynağının birinci mikrofona olan uzaklığı d_1 mesafesinin bildiği varsayılıp denklem (2.41) yeniden düzenlenerek

$$\begin{bmatrix} 2\left(\boldsymbol{p_2}-\boldsymbol{p}_1\right)^T \\ 2\left(\boldsymbol{p}_3-\boldsymbol{p}_1\right)^T \end{bmatrix} \times [\boldsymbol{p_k}] = \begin{bmatrix} \boldsymbol{p}_2^T\boldsymbol{p_2}-\boldsymbol{p}_1^T\boldsymbol{p_1}-v_s\tau_{21}\left(v_s\tau_{21}+2d_1\right) \\ \boldsymbol{p}_3^T\boldsymbol{p}_3-\boldsymbol{p}_1^T\boldsymbol{p}_1-v_s\tau_{31}\left(v_s\tau_{31}+2d_1\right) \end{bmatrix} \tag{5.2}$$

(5.2) denkleminin basit gösterimi

$$\boldsymbol{A_1x_1}=\boldsymbol{b_1} \tag{5.3}$$

Düzlem üzerinde konumlama yapmak için Şekil 6-a'da gösterilen sensör dizisinde dik eksenlerde yer alan mikrofonların orjine göre mesafeleri d_{mx} ve d_{my} olarak ele alınmıştır. Birinci mikrofon orijinde yer aldığı için d_1=d_k olur. Bunlara göre A_1 matrisi

$$A_1 = 2 * \begin{bmatrix} d_{mx} & 0 \\ 0 & d_{my} \end{bmatrix} \tag{5.4}$$

b_1 vektörü

$$\boldsymbol{b_1} = \begin{bmatrix} d_{mx}^2 - v_s^2 * \tau_{21}^2 - 2v_s\tau_{21}d_k \\ d_{my}^2 - v_s^2 * \tau_{31}^2 - 2v_s\tau_{31}d_k \end{bmatrix} \tag{5.5}$$

İkinci mikrofonun x ekseninde, üçüncü mikrofonun ise y ekseninde olduğunu düşünürsek $\tau_{21} = \tau_x$ benzer şekilde $\tau_{31} = \tau_y$ yazabiliriz. Sesin geliş yönüne bağlı olarak $-d_{mx} \leq v_s\tau_x \leq d_{mx}$ ve $-d_{my} \leq v_s\tau_y \leq d_{my}$ olacağı açıktır. Buna göre b_1 vektörü iki bileşene ayrılıp düzenlenerek

$$\boldsymbol{b}_1 = \underbrace{\begin{bmatrix} d_{mx}^2 - v_s^2 * \tau_x^2 \\ d_{my}^2 - v_s^2 * \tau_y^2 \end{bmatrix}}_{\in \binom{[0, d_{mx}^2]}{[0, d_{my}^2]}} + \begin{bmatrix} -2v_s\tau_x d_k \\ -2v_s\tau_y d_k \end{bmatrix} \tag{5.6}$$

Denklem (5.6)'de birinci bileşen, ikinci bileşene göre çok küçük bir değere sahip olduğu için ihmal edilmiştir. Buna göre denklem (5.2) yeniden yazılıp ve düzenlenerek

$$\begin{bmatrix} d_{mx} & 0 \\ 0 & d_{my} \end{bmatrix} \times \begin{bmatrix} x_k \\ y_k \end{bmatrix} \approx \begin{bmatrix} -v_s\tau_x d_k \\ -v_s\tau_y d_k \end{bmatrix} \tag{5.7}$$

Denklem (5.7)'den p_k pozisyon vektörü yalnız bırakılarak

$$\begin{bmatrix} x_k \\ y_k \end{bmatrix} \approx v_s d_k \begin{bmatrix} -\tau_x/d_{mx} \\ -\tau_y/d_{my} \end{bmatrix} \tag{5.8}$$

Denklem (5.8)'de görüldüğü gibi ses kaynağının konum vektörü, mikrofonlar arası TDoA değerlerine göre değişmektedir. (1) denkleminde belirtilen x_k ve y_k konumları arasındaki oran (5.8) denklemine göre hesaplandığında TDoA süreleri arasındaki orana eşit olur ve θ azimuth açısı

$$\theta \approx \operatorname{atan} 2 \left(-\frac{\tau_y}{d_{my}}, -\frac{\tau_x}{d_{mx}} \right) \tag{5.9}$$

Üç boyutlu kartezyen uzayda Şekil 2.3-b'deki 4 mikrofonlu sensör dizisi kullanılarak hem yanca (azimuth), hem de yükseliş (elevasyon) açısı elde edilebilmektedir. Şekil 2.3-b'de gösterilen sensör dizisi için mikrofonların konumunu $p_i = [x_i, y_i, z_i]^T$ olarak ve ses kaynağının konumunu ise $p_k =$

$[x_k, y_k, z_k]^T$ olarak gösterelim. Buna göre (5.2) ve (5.3) denklemleri yeniden düzenlenerek

$$\begin{bmatrix} 2\left(\boldsymbol{p_2}-\boldsymbol{p}_1\right)^T \\ 2\left(\boldsymbol{p_3}-\boldsymbol{p}_1\right)^T \\ 2\left(\boldsymbol{p_4}-\boldsymbol{p}_1\right)^T \end{bmatrix} \times [\boldsymbol{p_k}] = \begin{bmatrix} \boldsymbol{p}_2^T\boldsymbol{p_2}-\boldsymbol{p}_1^T\boldsymbol{p_1}-v_s\tau_{21}\left(v_s\tau_{21}+2d_1\right) \\ \boldsymbol{p}_3^T\boldsymbol{p_3}-\boldsymbol{p}_1^T\boldsymbol{p_1}-v_s\tau_{31}\left(v_s\tau_{31}+2d_1\right) \\ \boldsymbol{p}_4^T\boldsymbol{p_4}-\boldsymbol{p}_1^T\boldsymbol{p_1}-v_s\tau_{41}\left(v_s\tau_{41}+2d_1\right) \end{bmatrix} \tag{5.10}$$

$$\boldsymbol{A_2x_2} = \boldsymbol{b_2} \tag{5.11}$$

Yeni durumda dördüncü mikrofon z yönünde kabul edilerek $\boldsymbol{A_2}$ ve $\boldsymbol{b_2}$

$$\boldsymbol{A}_2 = 2 * \begin{bmatrix} d_{mx} & 0 & 0 \\ 0 & d_{my} & 0 \\ 0 & 0 & d_{mz} \end{bmatrix} \tag{5.12}$$

$$\boldsymbol{b}_2 = \begin{bmatrix} d_{mx}^2 - v_s^2 * \tau_x^2 - 2v_s\tau_x d_k \\ d_{my}^2 - v_s^2 * \tau_y^2 - 2v_s\tau_y d_k \\ d_{mz}^2 - v_s^2 * \tau_z^2 - 2v_s\tau_z d_k \end{bmatrix} \tag{5.13}$$

(5.7) denkleminde yaptığımız benzer yaklaşımlarla (5.13) denklemini indirgedikten sonra p_k konum vektörü

$$\begin{bmatrix} x_k \\ y_k \\ z_k \end{bmatrix} \approx v_s d_k \begin{bmatrix} -\tau_x/d_{mx} \\ -\tau_y/d_{my} \\ -\tau_z/d_{mz} \end{bmatrix} \tag{5.14}$$

(5.14) denkleminde buluğumuz sonuçları (2) denkleminde yerine yazarsak φ elevasyon açısı

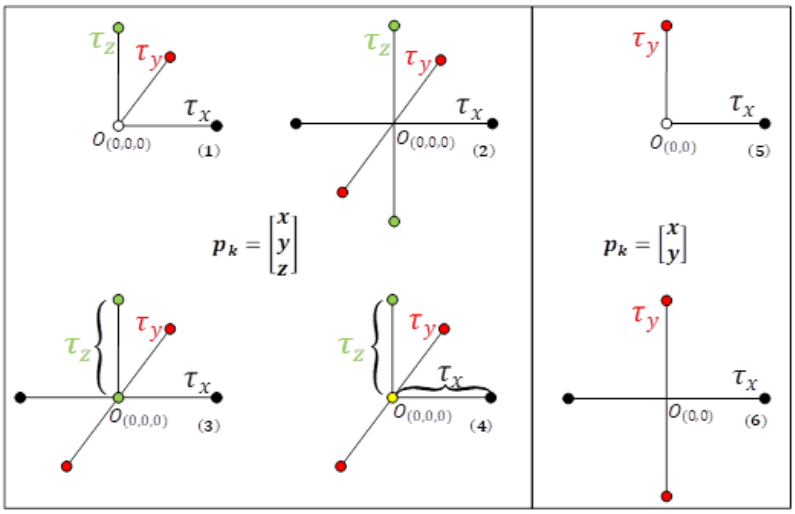
$$\varphi \approx \operatorname{atan} 2\left(-\left(\tau_z/d_{mz}\right), \sqrt{\left(-\tau_x/d_{mx}\right)^2 + \left(^{-\tau_y}/d_{my}\right)^2}\right) \tag{5.15}$$

Mikrofonlar arasındaki d_m mesafelerinin eşit olduğu durumda denklem (5.9) ve (5.15)'ün özel hali

$$\theta \approx \operatorname{atan2}\left(-\tau_y, -\tau_x\right) \tag{5.16}$$

$$\varphi = \operatorname{atan2}\left(-\tau_z, \sqrt{\tau_x^2 + \tau_y^2}\right) \tag{5.17}$$

Denklem (5.9) ve (5.15)'yi sağlayan sensör dizisi yerleşimleri Şekil 5.7'de verilmiştir [111].



Şekil 5.7 Ortogonal yerleşimli sensör çiflerinden oluşan mikrofon dizileri

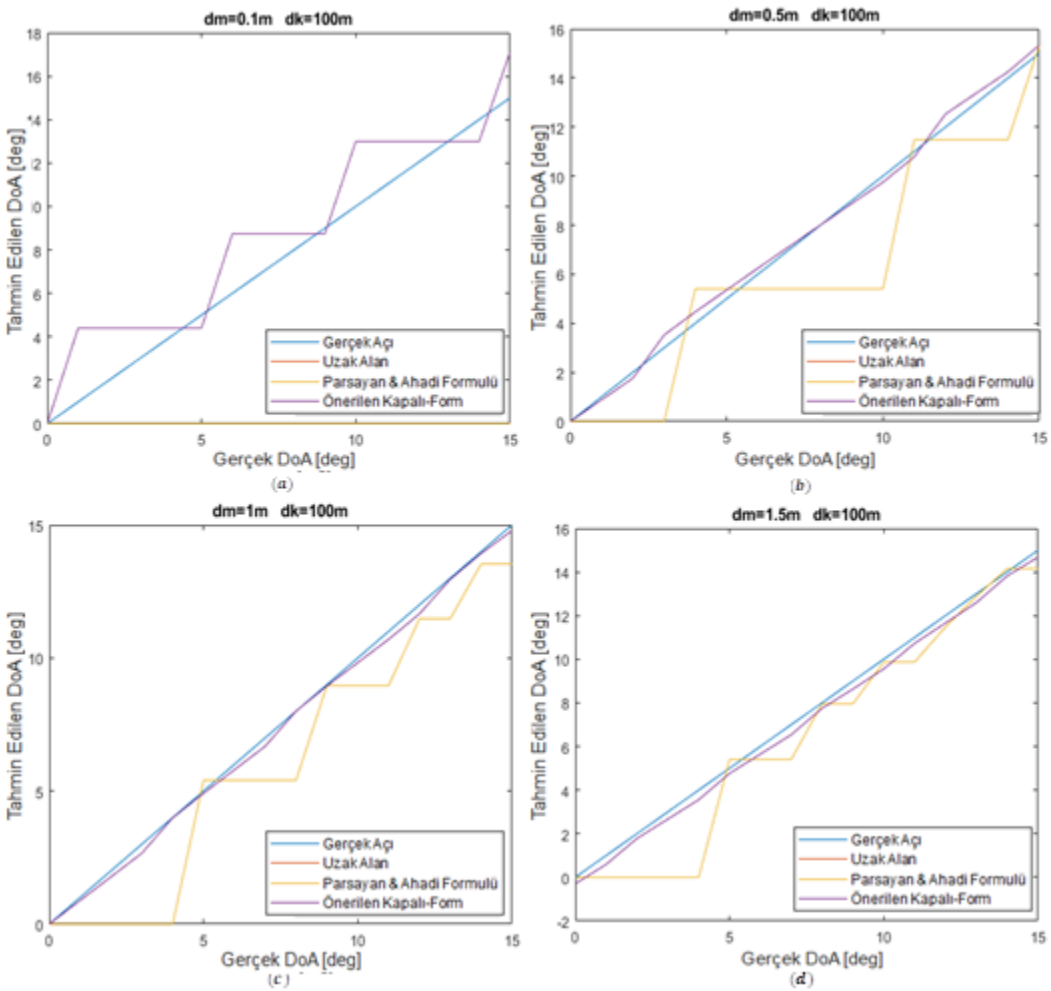
5.2 Önerilen Yöntemin Diğer Yöntemlere Göre Başarısı

XY düzleminde yapılan benzetim çalışmalarında ses kaynağının d_k uzaklığı 100 m olarak alınmıştır. Geliş yönü tespitinde en büyük ölçüm hataları 10° altındaki açılarda olmaktadır [99]. Şekil 5.8'de d_k=100 m ve F_s=44.1 kHz durumunda farklı d_m mesafeleri kullanılarak elde edilen sonuçlar görülmektedir.

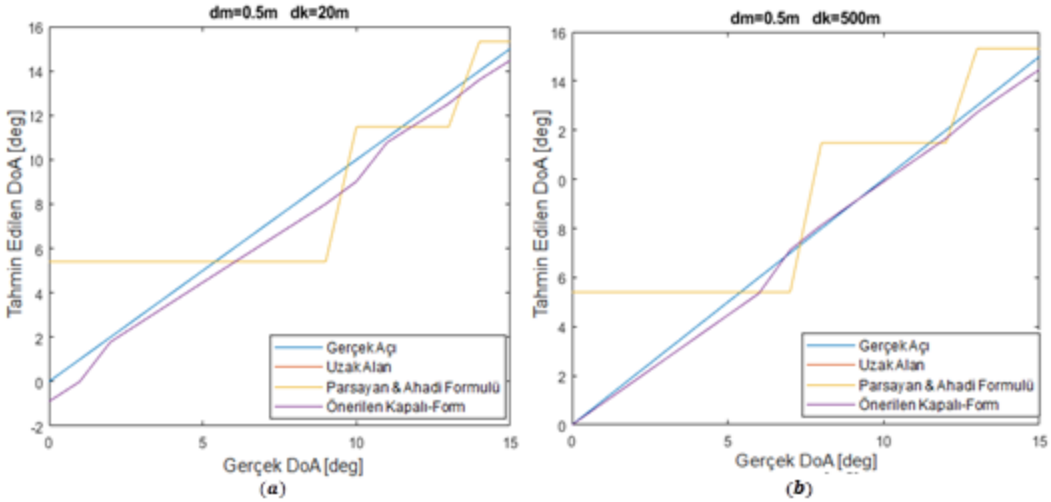
Şekil 5.9'da ise kaynak pozisyonunun yakın ve uzak mesafelere göre etkisi incelenmiştir. Örnekleme frekansı ve mikrofonlar arası açıklık sabit iken 10 m ve 500 m mesafedeki kaynak için sonuçlar görülmektedir.

d_k mesafesi 5 m'den 100 m'ye 1 m artımlarla, d_m mesafesi 0.1 m'den 1 m'ye 0.1 m artımlarla artarken 0-90 derece arasında 1° artımlarla yapılan tahminlerin hata değişimleri Şekil 5.10'da görülmektedir. Hata değerleri ortalama mutlak hata fonksiyonu

$$MAE = \frac{1}{N}\sum_{0}^{N} |\theta_r - \theta| \tag{5.18}$$

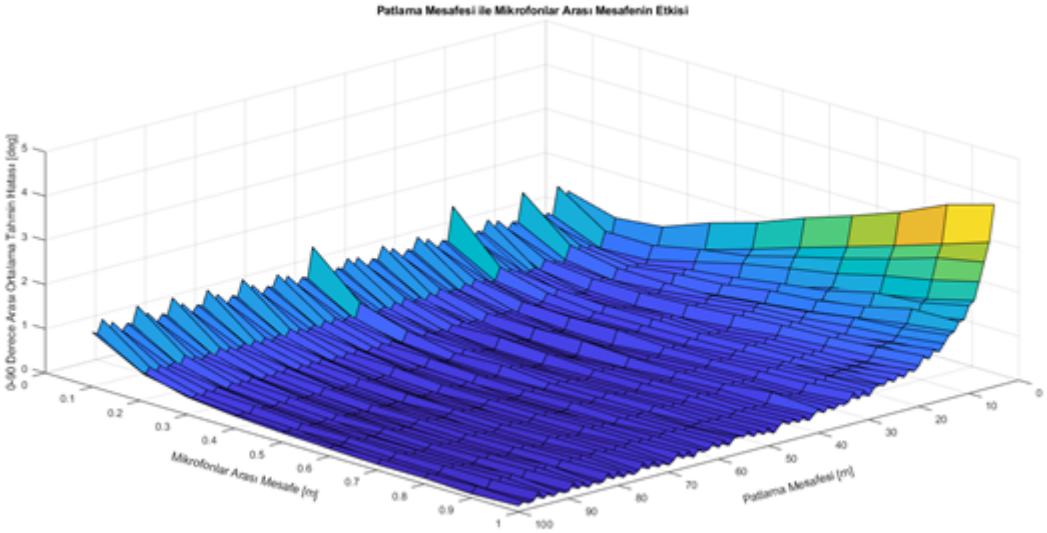


Şekil 5.8 d_k=100 m,F_s=44.1 kHz durumunda mikrofon çiftleri arasındaki mesafenin etkisi (a) d_m=0.1 m, (b) d_m=0.5 m, (c) d_m=1 m, (d) d_m=1.5 m



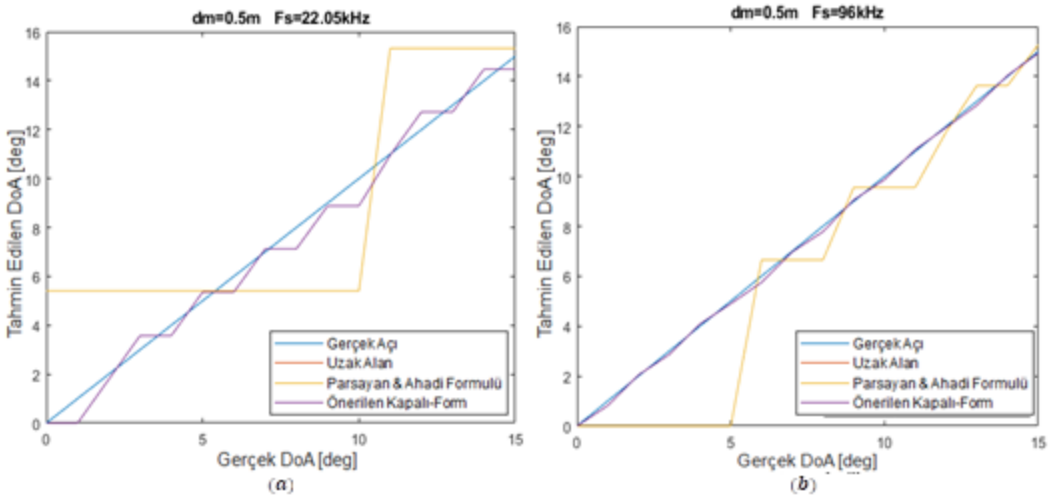
Şekil 5.9 d_m=0.5 m,F_s=44.1 kHz durumunda kaynak uzaklığının etkisi-(a) d_k=10 m, (b) d_k=500 m

ile hesaplanmıştır ve θ_r gerçek açı değeridir. d_m>0.3 m'den ve d_k>35 m'den sonra ortalama hatanın 0.5^o'nin altına inmektedir. Ayrıca d_m/d_k oranı bire yaklaştıkça en büyük hata değerlerine ulaşılmaktadır.



Şekil 5.10 $d_m \in [0.1, 1]$ m, $d_k \in [5, 100]$ m F_s=44.1 kHz durumunda 0-90 derece arası ortalama hatanın değişimi

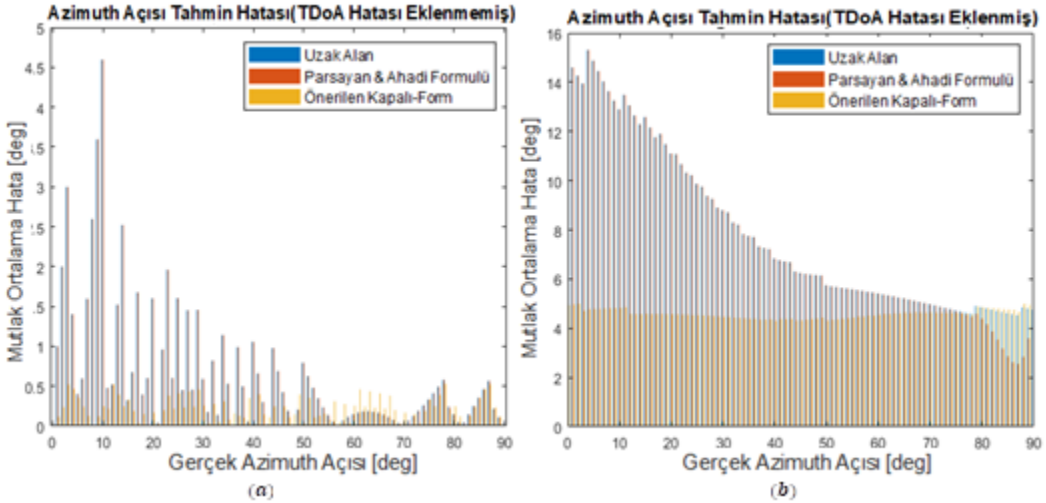
Mikrofonlar arası mesafe ve kaynak pozisyonu sabitken farklı örnekleme frekansları için elde edilen sonuçlar Şekil 5.11'de görülmektedir.



Şekil 5.11 d_m=0.5 m, d_k=100 m durumunda örnekleme frekansının etkisi-(a) F_s=22.05 kHz, (b) F_s= 96 kHz

Şekil 5.8-5.11'de görülen sonuçların elde edildiği benzetim çalışmaları Şekil 5.1'de gösterilen modelde GCC-Phat hatasının sıfır olduğu varsayımıyla incelenmiştir. Gerçek zamanlı uygulamalarda çevresel gürültüler, reverberasyon, mikrofon algılama paternlerinin farklılıkları gibi sebeplerle TDoA değerlerinin ölçülmesinde hatalar oluşmaktadır. Hata eklemesiz yapılan benzetim sonuçları Şekil 5.12-(a)'da,

her mikrofon çifti arasındaki TDoA değerlerine N(0,3) olacak şekilde gürültü eklenerek 1000 tekrarlı olarak yapılan benzetim sonuçları ise Şekil 5.12-(b)'de görülmektedir.



Şekil 5.12 d_m=0.5 m,d_k=100 m F_s=44.1 kHz durumunda 0-90 derece Arası tahmin hataları-(a) hatası eklemesiz, (b) hata eklemeli

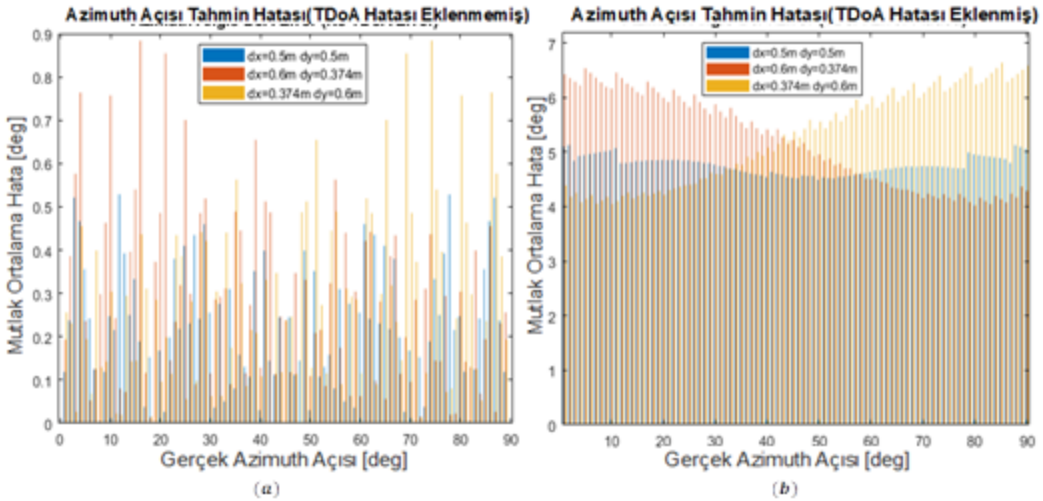
Şekil 5.12-(a) incelendiğinde önerdiğimiz yöntem 55-70 derece aralığı haricinde diğer yöntemlerden daha iyi sonuçlar üretmektedir. Şekil 5.12-(b)'ye göre ise önerdiğimiz yöntem 80 dereceye kadar olan açılarda diğer yöntemlere göre kararlı sonuçlar üretmektedir. Parsayan&Ahadi yöntemin 79-89 derece aralığında daha başarılı sonuçlar vermektedir. Tablo 5.3'de yöntemlerin 0-90 derece aralığında toplam ortalama hata değerleri görülmektedir.

Tablo 5.3 d_m=0.5 m, d_k=100 m F_s=44.1 kHz 0-90 derece arası tahmin hatalarının ortalamala değerleri

Hata	Uzak Alan	[Parsayan & Ahadi]	Önerilen
Yok	0.65	0.65	0.22
N(0,3)	8.01	7.85	4.72

Dik yerleştirilmiş mikrofon çiftlerinin aralarındaki mesafenin eşit olmadığında tahminleme hatasının değişimini incelemek için mikrofonların oluşturduğu üçgenlerin hipotenüslerinin eşit uzunlukta olması sağlanarak d_{mx}=0.5-d_{my}=0.5 , d_{mx}=0.6-d_{my}=0.374 , d_{mx}=0.374-d_{my}=0.6 olacak şekilde 3 farklı mikrofon dizisi geometrisi oluşturulmuştur. Hata eklemesi yapılmadan elde edilen benzetim sonuçları Şekil 5.13-(a)'da, N(0,3) olacak şekilde gürültü eklenerek 1000 tekrarlı olarak yapılan benzetim sonuçları Şekil 5.13-(b)'de görülmektedir.

Şekil 5.13-(a) incelendiğinde belirgin bir fark görülemese de Şekil 5.13-(b)'de



Şekil 5.13 Eşit olmayan Mikrofonlar arası Mesafenin Etkisi (a) Hata Eklemesiz, (b) Hata Eklemeli

değişen 3. durumun 0-30 derece arasında daha başarılı, 2. durumun ise 60-90 derece arasında daha başarılı sonuçlar verdiği görülmektedir. Tablo 5.4'de 3 farklı durumun 0-90 derece aralığında toplam ortalama hata değerleri görülmektedir.

Tablo 5.4 Farklı mikrofon dizisi geometrilerinin 0-90 derece arası tahmin hatalarının ortalamala değerleri

Hata	d_x**=0.5** d_y**=0.5**	d_x**=0.6** d_y**=0.374**	d_x**=0.374** d_y**=0.6**
Yok	0.2197	0.2776	0.2776
N(0,3)	4.7423	5.2387	5.2351

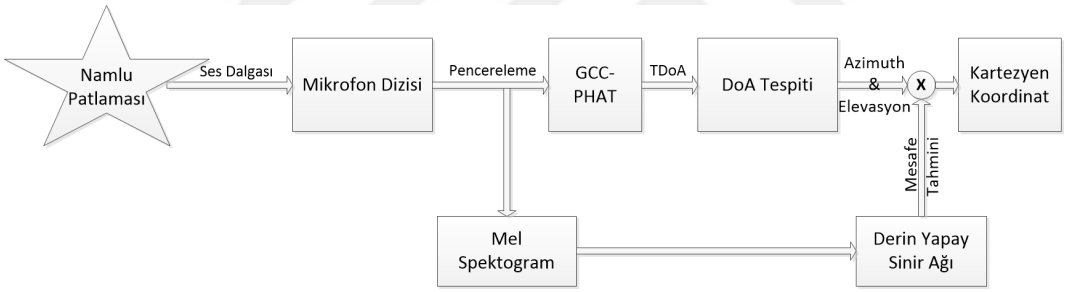
Uzak alan yaklaşımında özellikle dik açı değerlerinde TDoA değerindeki ölçüm hatalarından kaynaklı olarak tersi alınacak ifade $-1 \leq \cos^{-1}\left(\frac{v_s \cdot \tau}{d_m}\right) \leq 1$ aralığı dışına çıktığı durumda çözüm üretememektedir. Bu dik açı değerlerinde ölçüm duyarlılığını önemli ölçüde olumsuz etkilemektedir. Uzak alan yaklaşımında çözüm ses hızına bağlıdır. Ses hızı hava koşullarına göre değişmektedir ve ölçüm sonuçlarını etkilemektedir. Ayrıca uzak alan yaklaşımında GCC-Phat algoritması ile elde edilen örnek adeti tekrar t_s süresiyle çarpıldıktan sonra formüle eklenmektedir. Matematiksel ve algoritmik işlem sayısını arttırmaktadır. Önerdiğimiz geliş açısı tespit yöntemi, kaynak pozisyon denkleminden elde edildiği için Atan2 fonksiyonu ile geliş açısının hangi kuadrantta olduğu kesin olarak belirlenmektedir. Azimuth ve Elevasyon açısının tanjant fonksiyonu ile belirlendiği çalışmalar yapılmıştır, fakat kuadrant bilgisi yine algoritma ile elde edilmektedir [86, 88, 89]. Önerdiğimiz yöntem diğer yöntemler gibi koşula bağlı fonksiyonlara göre değil tek bir fonksiyonun sonucu olarak elde edilmektedir. Ayrıca tanjant fonksiyonuyla elde edilen çözümlerde kosinüs açısının sıfıra yakın

olacağı bölgelerde TDoA değerlerindeki ölçüm hatalarından kaynaklı tanımsızlık oluşmaktadır.

Önerdiğimiz yöntem direkt GCC-Phat algoritmasından elde edilen örnek farklarını kullanmaktadır. Bu sebeple v_s bağımsız sonuç üretmektedir. Atan2 fonksiyonu kosinüs değerinin sıfır olduğu durumda tanımsızlık oluşturmamaktadır. Önerdiğimiz yöntem daha hızlı ve özellikle ölçüm hatalarına karşı daha kararlı sonuçlar üretmektedir.

6
DENEYSEL ÇALIŞMALAR

Literatürde silah sesi konumlamada kullanılan yöntemler bölüm ikide anlatıldığı üzere geometrik yaklaşımlar ile elde edilmektedir. Mevcut yöntemlerin kısıtları oldukça fazla ve uzak mesafelerde hata oranları yüksektir. Mevcut yöntemlere alternatif olabilecek daha küçük boyutlarda mikrofon dizileri kullanarak konumlama yapabilen hibrit yöntem yöntem önerimiz Şekil 6.1'de sunulmuştur. Yöntemin başarısının ana unsurları TDoA değerlerinin doğru tahmini ve derin öğrenme ağlarının büyük veri setleri ile eğitilmesidir.



Şekil 6.1 Önerilen hibrit model

6.1 Deney Sahasının Oluşturulması ve Kullanılan Donanımlar

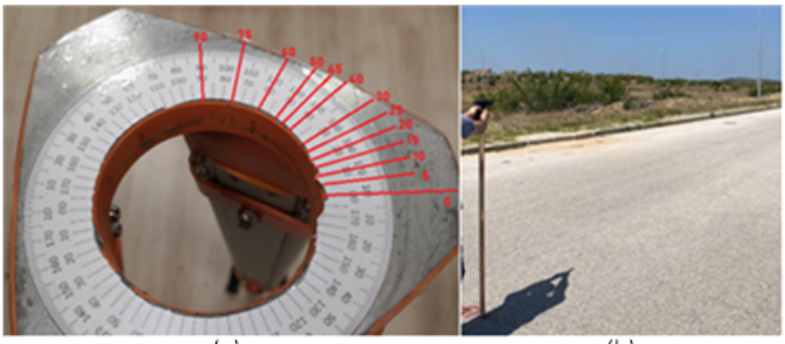
Patlama sesi geliş yönü tespiti dış ortamda en çok kullanılan uygulamalardan bir tanesidir. Bunu gerçekleştirmek için Şekil 6.2'deki gibi bir deney sahası oluşturulmuştur. Deney sahasındaki her nokta Topcon GR-5 GNSS cihazı ile ±2 cm hassasiyet ile işaretlenmiştir. Deney sahamızın 0 noktasına Şekil 5.7-(2)'de gösterilen geometride tasarlanan ve Şekil 6.2'de gösterilen mikron düzeneği orjini yerden 190 cm yukarıda olacak şekilde yerleştirilmiştir. Uzak mesafelerdeki patlama seslerini yakalayabilmek için düşük dip gürültüye sahip çok hassas kondanser elektret tümyönlü (omnidirectional) mikrofonlar kullanılmıştır. 6 mikrofonu senkron kayıt edebilmek için ZOOM H8 ses kaydedici kullanılmıştır. 50 m'den 500 m mesafeye kadar 50 m artımlarla 10 farklı uzaklıkta ve mikrofon dizisi Şekil 6.3-a'da gösterilen açı cetveli ile 0°, 5°, 10°, 15°, 20°, 25°, 30°, 40°,

45°, 50°, 60°, 75° ve 90° açılarında döndürülerek her durumdaki patlamalar kayıt edilmiştir. 13 farklı açıda ve 10 farklı mesafeden olacak şekilde 130 farklı noktadan iki tekrarlı toplam 260 adet atış yapılmıştır.



Şekil 6.2 Deney sahası ve kullanılan mikrofon dizisinin yerleşimi

Patlamalar Şekil 6.3-(b)'de görüldüğü gibi mastar yardımıyla yerden 1.5 m yükseklikten mikrofon dizisi yönüne doğru tek bir ses tabancası ile gerçekleştirilmiştir. Oluşturulan patlama sesleri 6 mikrofondan eş zamanlı olarak 44.1 kHz örnekleme frekansında Zoom H8 cihazı ile kayıt edilmiştir. Kayıtların yapıldığı gün hava sıcaklığı 28 derece olarak ölçülmüştür. Rüzgar hızı ise 2-6 km/h arasında değişmektedir.



Şekil 6.3 (a) Mikrofon dizisi açı ayarlama düzeneyi, (b) patlama yüksekliğinin ayarlanması

6.2 Deneysel Verilerle DoA Elde Etme

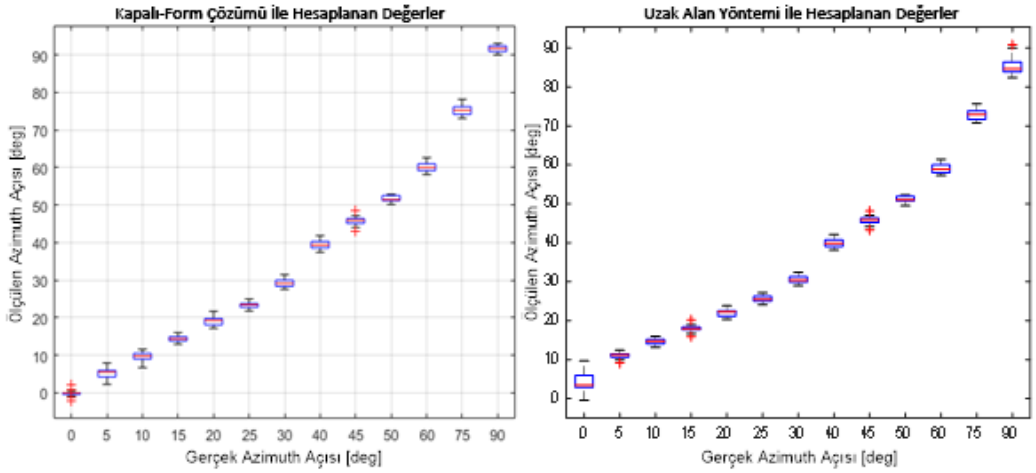
Her patlama için mikrofonlar arası TDoA süreleri GCC-Phat algoritmasıyla elde edilen değerleri ve bu değerlere göre hesaplanan açı değerleri Tablo 6.1'de

Tablo 6.1 Deneysel ölçümlerde elde edilen TDoA değerleri [τ Örnek Adeti, θ Derece]

θ	No	TDoA	50 m	100 m	150 m	200 m	250 m	300 m	350 m	400 m	450 m	500 m
0	1	τ_x	-128	-129	-134	-134	-135	-134	-133	-135	-134	-131
		τ_y	2	5	-1	-1	0	-5	0	-1	0	0
		$\hat{\theta}$	-0.89	-2.22	0.43	0.43	0	2.14	0	0.42	0	0
	2	τ_x	128	133	128	134	135	135	134	134	134	133
		τ_y	-1	2	3	1	2	0	0	-1	0	0
		$\hat{\theta}$	-0.45	0.86	1.34	0.43	0.85	0	0	-0.43	0	0
5	1	τ_x	127	130	129	129	129	130	129	128	129	129
		τ_y	-7	-12	-13	-18	-13	-14	-13	-13	-12	-9
		$\hat{\theta}$	-3.16	-5.27	-5.76	-7.94	-5.76	-6.15	-5.76	-5.80	-5.32	-4.00
	2	τ_x	127	129	129	129	128	129	130	129	129	130
		τ_y	-5	-13	-12	-14	-15	-15	-10	-12	-9	-6
		$\hat{\theta}$	-2.26	-5.76	-5.32	-6.19	-6.68	-6.63	-4.40	-5.32	-3.99	-2.64
10	1	τ_x	126	125	126	128	128	127	127	127	126	126
		τ_y	-25	-23	-23	-18	-22	-26	-20	-20	-24	-20
		$\hat{\theta}$	-11.22	-10.43	-10.35	-8.01	-9.75	-11.57	-8.95	-8.95	-10.78	-9.02
	2	τ_x	126	127	127	127	128	128	127	127	128	126
		τ_y	-21	-22	-23	-18	-24	-24	-23	-15	-21	-21
		$\hat{\theta}$	-9.46	-9.83	-10.27	-8.07	-10.62	-10.62	-10.27	-6.74	-9.32	-9.46
15	1	τ_x	128	125	125	125	125	125	125	123	122	126
		τ_y	-31	-32	-34	-32	-33	-31	-33	-30	-35	-29
		$\hat{\theta}$	-13.61	-14.36	-15.22	-14.36	-14.79	-13.93	-14.79	-13.71	-16.01	-12.96
	2	τ_x	126	124	124	126	125	125	126	124	126	124
		τ_y	-31	-31	-34	-33	-32	-33	-31	-34	-34	-29
		$\hat{\theta}$	-13.82	-14.04	-15.33	-14.68	-14.36	-14.79	-13.82	-15.33	-15.10	-13.16
30	1	τ_x	111	113	115	112	113	111	115	113	115	114
		τ_y	-62	-65	-60	-65	-61	-68	-62	-62	-64	-61
		$\hat{\theta}$	-29.19	-29.91	-27.55	-30.13	-28.36	-31.49	-28.33	-28.75	-29.10	-28.15
	2	τ_x	113	111	115	113	113	112	112	113	113	111
		τ_y	-61	-66	-60	-63	-63	-64	-65	-62	-60	-66
		$\hat{\theta}$	-28.36	-30.74	-27.55	-29.14	-29.14	-29.75	-30.13	-28.75	-27.97	-30.74
45	1	τ_x	91	89	90	91	90	90	94	89	90	89
		τ_y	-94	-96	-93	-91	-93	-93	-88	-93	-92	-94
		$\hat{\theta}$	-45.93	-47.17	-45.94	-45.00	-45.94	-45.94	-43.11	-46.26	-45.63	-46.57
	2	τ_x	92	92	91	93	91	89	92	91	89	85
		τ_y	-95	-92	-92	-90	-92	-94	-89	-92	-94	-96
		$\hat{\theta}$	-45.92	-45.00	-45.31	-44.06	-45.31	-46.57	-44.05	-45.31	-46.57	-48.48
60	1	τ_x	65	64	67	60	68	61	65	64	60	66
		τ_y	-112	-114	-112	-116	-111	-115	-113	-113	-116	-112
		$\hat{\theta}$	-59.87	-60.69	-59.11	-62.65	-58.51	-62.06	-60.09	-60.47	-62.65	-59.49
	2	τ_x	65	65	66	66	68	61	68	64	63	69
		τ_y	-109	-113	-113	-113	-112	-114	-112	-114	-115	-111
		$\hat{\theta}$	-59.19	-60.09	-59.71	-59.71	-58.74	-61.85	-58.74	-60.69	-61.29	-58.13
75	1	τ_x	34	33	32	27	30	38	37	32	36	33
		τ_y	-128	-127	-127	-129	-128	-125	-125	-127	-125	-126
		$\hat{\theta}$	-75.12	-75.43	-75.86	-78.18	-76.81	-73.09	-73.51	-75.86	-73.93	-75.32
	2	τ_x	32	34	29	30	31	34	37	34	35	36
		τ_y	-129	-126	-128	-128	-127	-126	-126	-127	-124	-126
		$\hat{\theta}$	-76.07	-74.90	-77.23	-76.81	-76.28	-74.90	-73.64	-75.01	-74.24	-74.06
90	1	τ_x	-2	-4	-7	-4	0.01	-6	-5	-4	-2	-7
		τ_y	-131	-132	-132	-134	-135	-131	-133	-131	-131	-131
		$\hat{\theta}$	-90.88	-91.74	-93.04	-91.71	-90.00	-92.62	-92.15	-91.75	-90.88	-93.06
	2	τ_x	-2	-2	-4	-1	0.01	-6	-6	-4	-4	-1
		τ_y	-130	-132	-132	-134	-135	-131	-132	-135	-131	-130
		$\hat{\theta}$	-90.88	-90.87	-91.74	-90.43	-90.00	-92.62	-92.60	-91.70	-91.75	-90.44

verilmiştir. Tablo 6.1'deki değerlere göre denklem (5.16) ve (5.17) ile geliş açıları hesaplanmıştır. Deney sahasında eğim çok az olduğu için elevasyon açıları farkı 1°'nin altındadır.

Tek noktaya konumlandırdığımız mikrofon dizisinde X, Y ve Z eksenlerinde yer alan dik yerleşimli mikrofon çiftleri arasındaki τ_x, τ_y ve τ_z değerleri GCC-Phat yöntemiyle elde edilmiştir. Elde edilen TDoA değerlerine göre azimuth açısı denklem (5.16), elevasyon açısı ise denklem (5.17) ile hesaplanmıştır. Buna göre elde ettiğimiz ortalama azimuth ve elevasyon açısı grafikleri Şekil 6.4 ve Şekil 6.5'de görülmektedir.



Şekil 6.4 Deney sahasından ölçülen azimuth açısının yayılım grafiği (a) kapalı form, (b) uzak alan

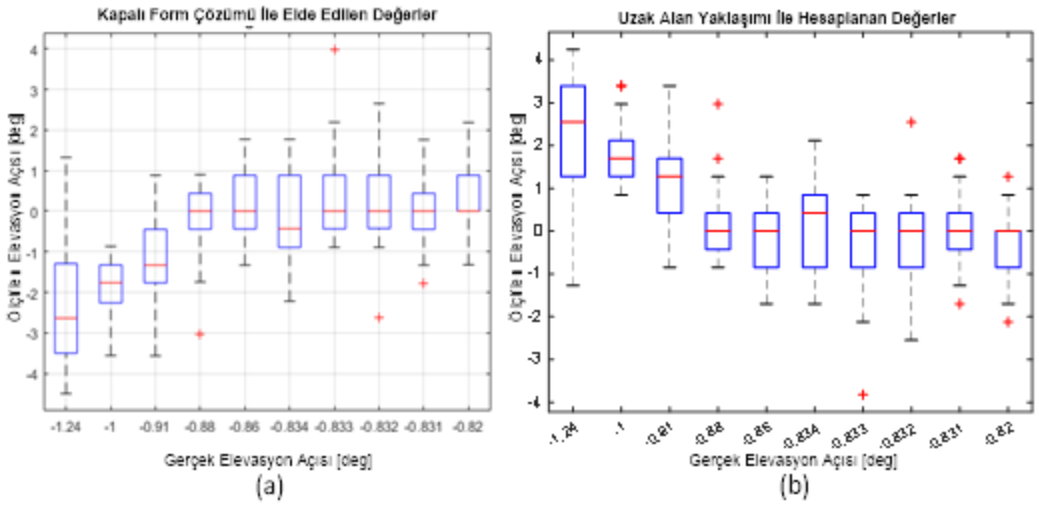
Şekil 6.4'de elde edilen ölçüm değerlerine ait veriler Tablo 6.2 ve Tablo 6.3'de verilmiştir.

Tablo 6.2 Azimuth açısı için Şekil 6.4-a'da gösterilen her mikrofon açısında elde edilen hata değerleri

Gerçek Azimuth	0	5	10	15	20	25	30	40	45	50	60	75	90
Min	-2.14	2.26	6.74	12.96	17.17	21.80	27.55	37.41	43.11	50.14	58.13	73.09	90
Median	0	5.54	9.79	14.35	19.28	23.49	29.12	39.35	45.92	51.44	59.98	75.22	91.72
Max	2.22	7.94	11.57	16.01	21.72	21.80	31.49	41.86	48.48	52.86	62.65	78.18	93.06
RMSE	0.86	1.40	1.19	0.97	1.54	1.83	1.36	1.30	1.34	1.84	1.33	1.35	1.80
µRMSE	1.39												

Tablo 6.3 Azimuth açısı için Şekil 6.4-b'de gösterilen her mikrofon açısında elde edilen hata değerleri

Gerçek Azimuth	0	5	10	15	20	25	30	40	45	50	60	75	90
Min	-0.42	9.10	13.09	15.09	20.35	24.05	28.99	38.11	43.28	49.53	57.29	70.73	82.39
Median	3.49	11.23	14.69	17.95	22.24	25.42	30.38	39.80	45.87	50.97	58.90	72.90	84.65
Max	9.69	12.45	16.00	20.19	23.82	27.17	32.47	42.12	48.15	52.29	61.42	75.66	90.85
RMSE	5.28	6.17	4.68	3.10	2.20	0.98	1.10	1.08	1.24	1.36	1.53	2.58	5.10
µRMSE	2.80												



Şekil 6.5 Deney sahasından ölçülen elevasyon açısının yayılım grafiği (a) kapalı form, (b) uzak alan

Şekil 6.5'de elde edilen ölçüm değerlerine ait veriler Tablo 6.4 ve Tablo 6.5'de verilmiştir.

Tablo 6.4 Elevasyon açısı için Şekil 6.5-a'da gösterilen her mikrofon açısında elde edilen hata değerleri

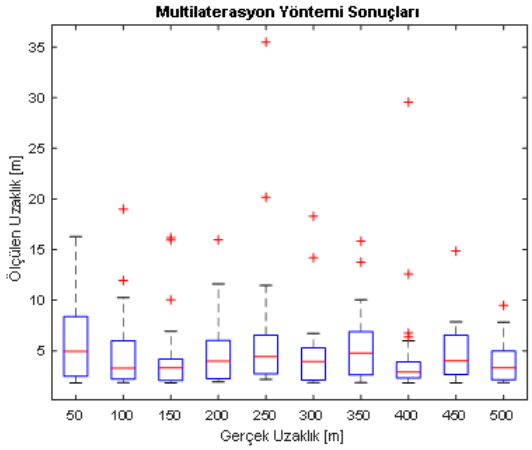
Patlama Noktası	50m	100m	150m	200m	250m	300m	350m	400m	450m	500m
Gerçek Elevation	-1.24	-1	-0.91	-0.88	-0.86	-0.83	-0.83	-0.83	-0.83	-0.82
Min	-4.5	-3.56	-3.57	-3.04	-1.33	-2.22	-0.89	-2.61	-1.79	-1.32
Median	-2.64	-1.76	-1.33	0	0	-0.43	0	0	0	0
Max	1.32	-0.87	0.88	0.90	1.77	1.77	3.98	2.66	1.77	2.19
RMSE	1.75	1.19	1.03	1.04	1.33	1.29	1.58	1.53	1.11	1.42
µRMSE	1.33									

Tablo 6.5 Elevasyon açısı için Şekil 6.5-b'de gösterilen her mikrofon açısında elde edilen hata değerleri

Patlama Noktası	50m	100m	150m	200m	250m	300m	350m	400m	450m	500m
Gerçek Elevasyon	-1.24	-1	-0.91	-0.88	-0.86	-0.83	-0.83	-0.83	-0.83	-0.82
Min	-1.27	0.85	-0.85	-0.85	-1.70	-1.70	-3.82	-2.55	-1.70	-2.12
Median	2.55	1.70	1.27	0	0	0.42	0	0	0	0
Max	4.25	3.40	3.40	2.97	1.27	2.12	0.85	2.55	1.70	1.27
RMSE	3.62	2.95	2.30	1.38	1.06	1.45	1.13	1.21	1.16	0.93
µRMSE	1.72									

6.3 Deneysel Veriler İle Multilaterasyon Çözümü Sonuçları

Beşinci bölümde tek mikrofon dizisi ile multilaterasyon yönteminin konumlama başarısını benzetim çalışması ile elde etmiştik. Mevcut mikrofon dizisi ile elde edilen deneysel TDoA verileri ve mikrofon pozisyon vektörleri kullanılarak denklem (2.37) oluşturulmuş ve LSE çözümü yaptırılmış ve sonuçlar Şekil 6.6'de verilmiştir.



Şekil 6.6 Deneysel TDoA değerleri ile elde edilen multilaterasyon yöntemi sonuçları

Şekil 6.6'de gösterilen grafiğin değerleri Tablo 6.6'de verilmiştir. Tablo 18 incelendiğinde benzetim çalışmasında elde edilen sonuçlarla örtüştüğü görülmektedir. Tek mikrofon dizisi kullanılarak multilaterasyon yöntemi ile konumlama yapılamamaktadır. Sonuç dağılımlarına baktığımızda mesafeden bağımsız olarak median değerinin 3 m ile 5 m arasında olduğu görülmektedir.

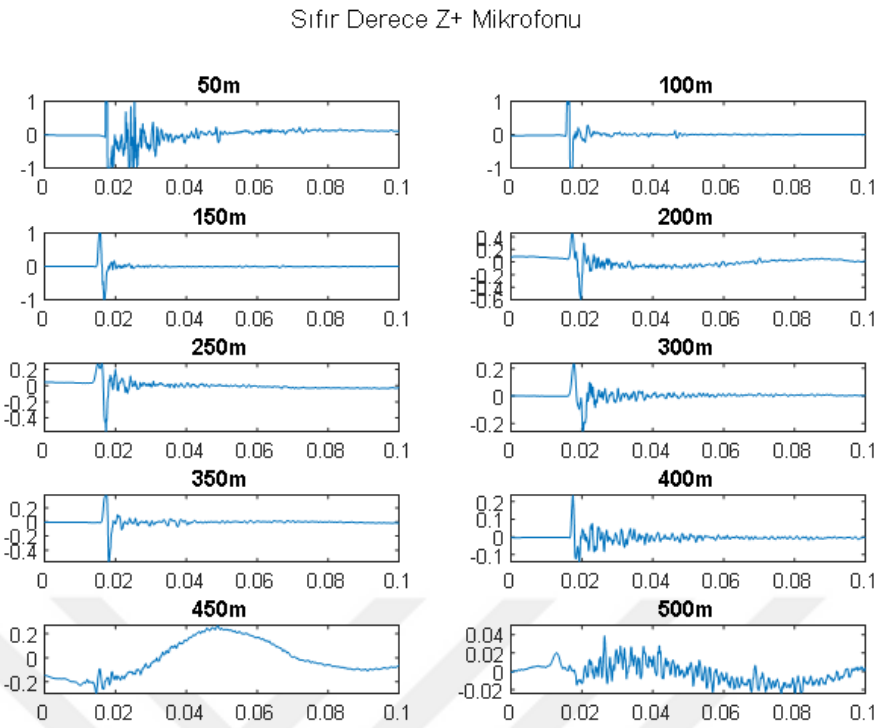
Tablo 6.6 Şekil 6.6'de gösterilen konumlama verileri ve hata yüzdeleri

	50m	100m	150m	200m	250m	300m	350m	400m	450m	500m
Min	1.85	1.85	1.85	1.93	2.19	1.86	1.86	1.85	1.85	1.87
Median	4.97	3.31	3.33	4.00	4.45	3.96	4.79	2.91	4.03	3.33
Max	16.28	18.99	16.19	15.96	35.48	18.29	15.83	29.56	14.85	9.48
RMSE	44.04	95.19	145.65	195.08	243.73	295.34	344.44	395.44	445.31	495.98
%RMSE	88.09	95.19	97.10	97.54	97.50	98.45	98.41	98.86	98.96	99.20

6.4 Derin Öğrenme Verilerinin Oluşturulması

Toplanan kayıtlar Audacity programında 100 ms olarak kırpılmış her patlama 'eksenyönüveadı_mesafesi_patlamanumarası.wav' olarak etiketlenmiştir. Bir patlamaya ait mikrofon kayıtları eş başlangıç süresi ile kırpılarak 6 mikrofon arasındaki TDoA süreleri korunmuştur. Toplamda 1560 adet 100 ms boyutunda ses kaydı elde edilmiştir. Düzenek sıfır derece açısındayken +Z yönündeki mikrofondan kayıt edilmiş farklı mesafelerdeki kırpılmış sesler Şekil 6.7'da görülmektedir.

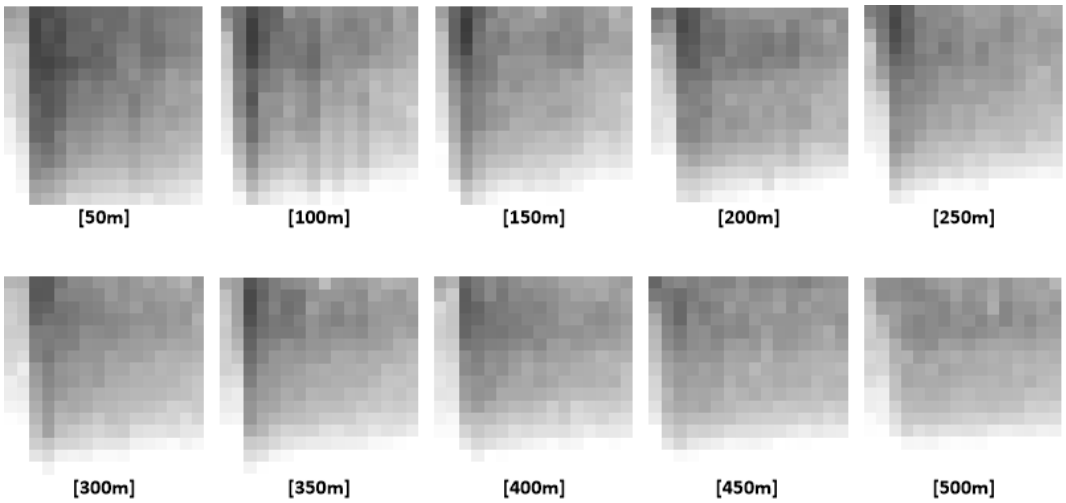
X, Y eksenlerin bulunan mikrofonlardan alınan kayıtlar sadece TDoA hesaplamaları için kullanılmıştır. Z ekseni üzerinde bulunan mikrofonlardan elde edilen kayıtlar ise ayrıca yapay zeka çalışmalarında veri seti oluşturmak için de kullanılmıştır.



Şekil 6.7 Sıfır derece atışında +Z yönündeki mikrofondan kayıt edilen kırpılmış farklı mesafe sesleri

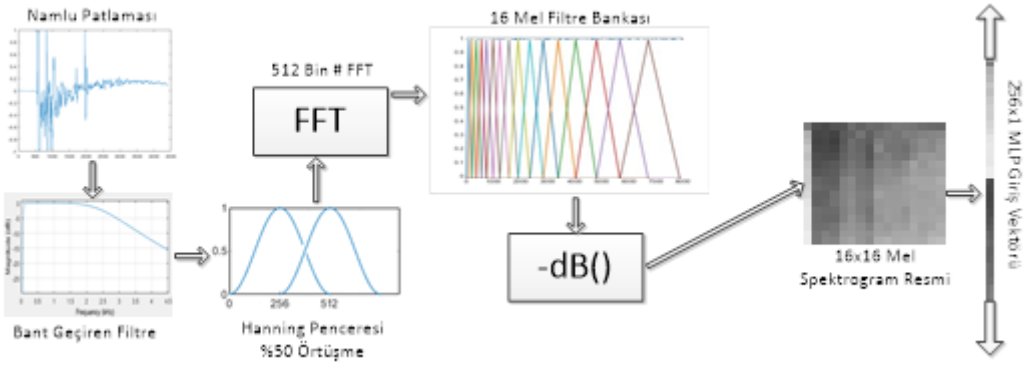
Üçüncü bölümde yer alan ses özelliklerinden Mel spectrogram hem zaman ve frekans özelliğinin yanında duyusal özellik de içermektedir. Derin ağların eğitiminde ihtiyaç duyulan giriş bilgisi, öncelikle Mel sprektrogram özelliğinden elde edilmiştir. Daha sonra MRMR yöntemi ile özellik seçimi yapılarak elde edilen özellik matrisi ile en iyi sonucun eldedildiği CNN ağı eğitilerek özellik seçiminin sonuçlara etkisi gösterilmiştir.

Şekil 6.9 ve 6.12'de veri seti oluşturma adımları görülmektedir. Namlu patlama sesi frekans aralığı 300-1000 Hz'dir. Veri seti için kullanılacak sesler, olası elektriksel ve çevresel gürültülerden arındırmak için 60-2500 Hz kesim frekanslarına sahip band geçiren filtre ile ön filtreleme işlemi yapılmıştır. Filtrelenmiş seslere, 512 boyutlu Hanning çerçeveli, %50 çakışma (overlap) ile 512 noktalı STFT işlemi uygulanmıştır. Elde edilen spektogram görüntüsü 16 adet Mel filtre bandıyla çarpılarak 16×16 boyutunda Mel spektrogram değerleri elde edilmiştir. Mel sprektrogram değerleri desibele çevrildikten sonra 16×16'lık matris, 256×1'lik vektör formuna çevrilerek MLP ağı için veri seti oluşumu tamamlanmıştır. Şekil 6.7'da gösterilen seslerden elde edilen Mel spektrogram görüntüsü Şekil 6.8'de görülmektedir.



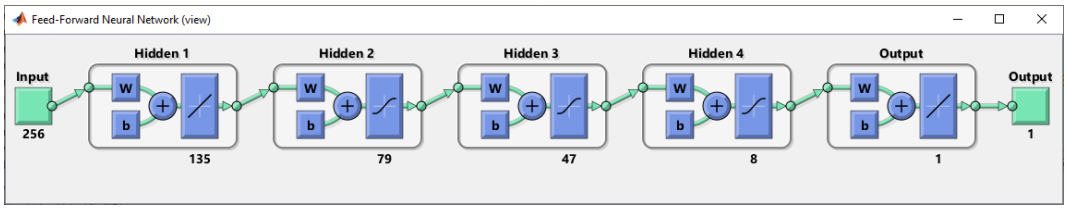
Şekil 6.8 Farklı mesafelerdeki seslerin oluşturduğu Mel sprektrogram resimleri

6.5 Çok Katmanlı Algılayıcı İle Mesafe Tahmini



Şekil 6.9 Ses kayıtlarından veri seti dönüşümü lle MLP ağının giriş vektörü

Yapay sinir ağı olarak Çok Katmanlı Algılayıcı kullanılmıştır ve ağ MATLAB 2022A programında oluşturulmuştur. Tek katmanlı ağ yapısından başlanarak ve farklı hiper parametreler denenerek birçok deneme yapılmıştır. En başarılı tahminleme sonucunun elde edildiği MLP ağı Şekil 6.10'da görülmektedir.



Şekil 6.10 Geliştirilen MLP mimarisi

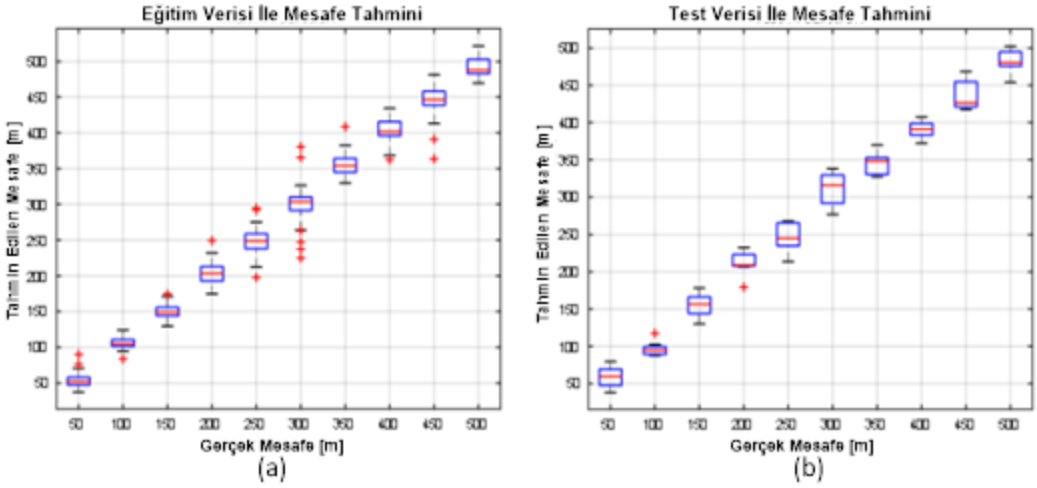
Oluşturulan ağın birinci ve çıkış katmanında doğrusal; ikinci, üçüncü ve dördüncü gizli katmanda tanjant sigmoid aktivasyon fonksiyonları kullanılmıştır. Eğitim fonksiyonu olarak TrainCGP, performans fonksiyonu olarak ise MSEREG kullanılmıştır. Z ekseninde bulunan iki mikrofondan kaydedilen, her mesafe için

52 örnekten oluşan toplam 520 örneğimizin 440 tanesi eğitim 80 tanesi ise test için kullanılmıştır. Diğer hiper parametreler aynı kalarak başarılı sonuçlar elde ettiğimiz bazı ağ yapılarının nöron sayıları ve en iyi tahmin sonuçları kalın yazı tipinde Tablo 6.7'de gösterilmiştir. Eğitim veri setine göre en düşük ortalama hata değeri 13.43 m elde edilse de en iyi genelleme yeteneğine onuncu sıradaki ağ yapısı ile ulaşılmıştır.

Tablo 6.7 Farklı nöron sayılarına göre elde edilen ortalama mesafe hatası

	Katman 1	**Katman 2**	**Katman 3**	**Katman 4**	**Eğitim [m]**	**Test [m]**
1	125	78	51	9	14.05	21.84
2	128	87	49	13	14.14	21.88
3	129	80	44	7	13.43	23.36
4	129	80	46	17	13.78	23.13
5	129	84	52	17	17.81	18.48
6	129	85	43	12	19.94	18.19
7	129	87	48	10	18.24	18.15
8	134	83	51	17	19.66	18.45
9	134	86	52	11	18.72	18.09
10	135	79	47	8	16.84	18.45
11	135	88	45	11	17.44	18.45

Belirtilen ağ ile elde ettiğimiz eğitim ve test tahmin grafikleri sırasıyla Şekil 6.11-a ve Şekil 6.11-b de verilmiştir.

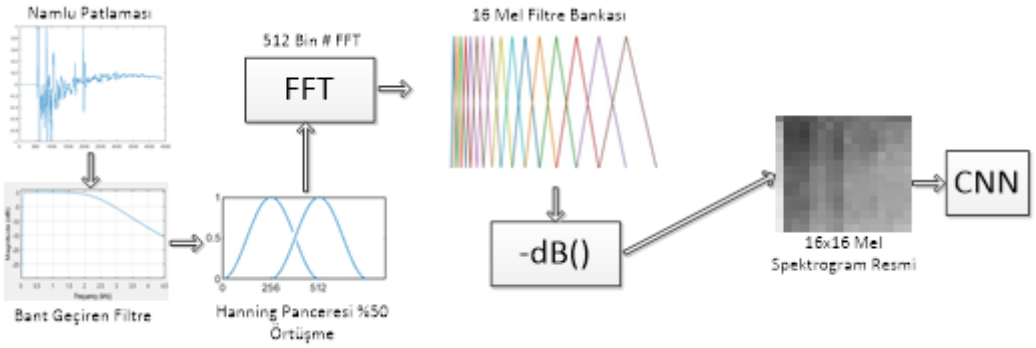


Şekil 6.11 Seçilen ağın tahminleme grafikleri (a) eğitim, (b) test

Bu grafiklerdeki sonuçlara göre elde edilen değerlere göre tahminleme sonuçları ve hata değerleri eğitim ve test veri setleri için Tablo 6.8'de verilmiştir.

Tablo 6.8 En başarılı sonuçları elde ettiğimiz ağın mesafelere göre hata değerleri

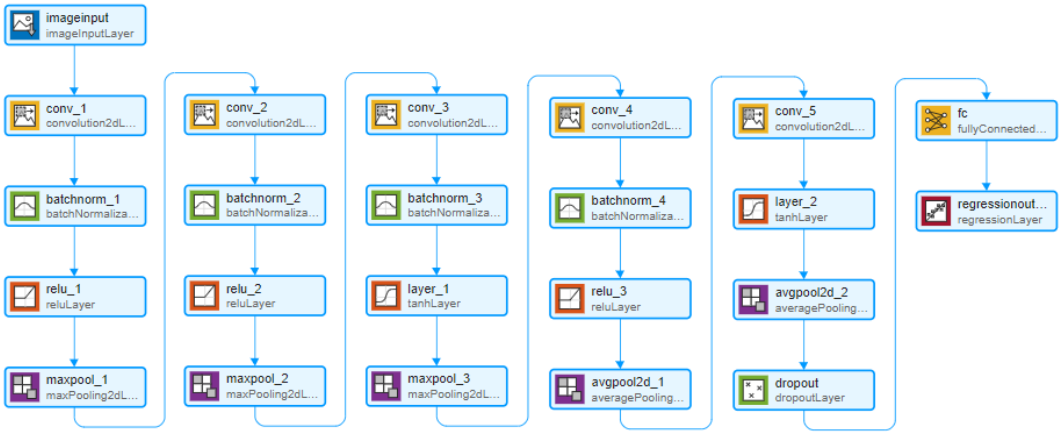
Data	Value	50m	100m	150m	200m	250m	300m	350m	400m	450m	500m
Train	Min	37.44	83.96	129.34	174.48	197.77	225.10	329.61	362.50	363.87	469.59
	Median	52.17	104.80	150.16	203.56	248.76	303.84	354.38	402.03	446.35	488.60
	Max	90.01	124.00	174.39	249.30	294.83	380.43	408.70	434.20	481.36	522.00
	RMSE	10.85	9.45	9.87	15.34	18.82	26.36	16.18	16.51	21.46	15.84
	%RMSE	21.70	9.45	6.58	7.67	7.53	8.79	4.62	4.13	4.77	3.17
Test	Min	38.33	87.45	129.93	179.47	213.77	276.80	327.31	372.47	417.22	453.94
	Median	59.69	94.52	156.73	206.98	245.13	316.12	348.15	391.05	426.68	480.55
	Max	79.42	117.49	178.17	232.39	267.86	338.32	370.04	407.47	468.51	502.51
	RMSE	16.38	9.95	15.47	18.91	18.19	25.00	14.87	14.26	23.17	22.85
	%RMSE	32.75	9.95	10.31	9.45	7.28	8.33	4.25	3.56	5.15	4.57



Şekil 6.12 Ses kayıtlarından veri seti dönüşümü ile CNN ağının giriş matrisi

6.6 Konvolisyonel Ağlar İle Mesafe Tahmini

Konvolüsyonel sinir ağı MATLAB 2022B programında oluşturulmuş ve GeForce RTX3050 GPU kullanılarak ile eğitilmiştir. Bir adet konvolüsyonel katman içeren ağ yapısından başlanarak, farklı ağ modelleri ve eğitim parametreleri denenerek birçok deneme yapılmıştır. En başarılı tahminleme sonucunun elde edildiği 5 konvolüsyonel katman içeren CNN mimarisi Şekil 6.13'de görülmektedir.



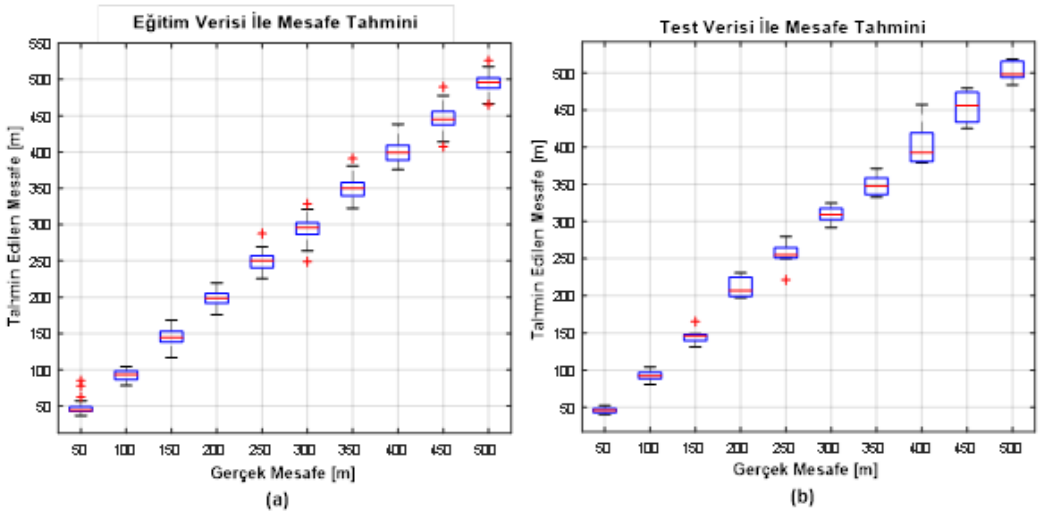
Şekil 6.13 Matlab programında oluşturulan ve eğitilen CNN ağı

Ağı oluşturan katman parametreleri ve eğitim için kullanılan parametreler Tablo 6.9'da verilmiştir.

Tablo 6.9 Ağ ve eğitim parametreleri

Adı	**Parametresi**	**Değeri**
Giriş Katmanı	Veri Boyutu	$16 \times 16 \times 1$
Konvolüsyon Katmanı 1	Filtre Boyutu Filtre Sayısı	11×11 20
Aktivasyon Katmanı 1	Fonksiyon	Rectified Linear Unit
Havuzlama Katmanı 1	Maksimum Havuzlama	2×2
Konvolüsyon Katmanı 2	Filtre Boyutu Filtre Sayısı	9×9 40
Aktivasyon Katmanı 2	Fonksiyon	Rectified Linear Unit
Havuzlama Katmanı 2	Maksimum Havuzlama	2×2
Konvolüsyon Katmanı 3	Filtre Boyutu Filtre Sayısı	5×5 40
Aktivasyon Katmanı 3	Fonksiyon	Hyperbolic tangent
Havuzlama Katmanı 3	Maksimum Havuzlama	2×2
Konvolüsyon Katmanı 4	Filtre Boyutu Filtre Sayısı	3×3 40
Aktivasyon Katmanı 4	Fonksiyon	Rectified Linear Unit
Havuzlama Katmanı 4	Ortalama Havuzlama	2×2
Konvolüsyon Katmanı 5	Filtre Boyutu Filtre Sayısı	7×7 30
Aktivasyon Katmanı 5	Fonksiyon	Hyperbolic Tangent
Havuzlama Katmanı 5	Ortalama Havuzlama	2×2
Eğitim	Optimizasyon Küme Boyutu Öğrenme Oranı Eğitim Sayısı	Root Mean Square Propagation 8 0.0001 30

Belirtilen ağ ile elde ettiğimiz eğitim ve test tahmin grafikleri sırasıyla Şekil 6.14-a ve Şekil 6.14-b de verilmiştir.

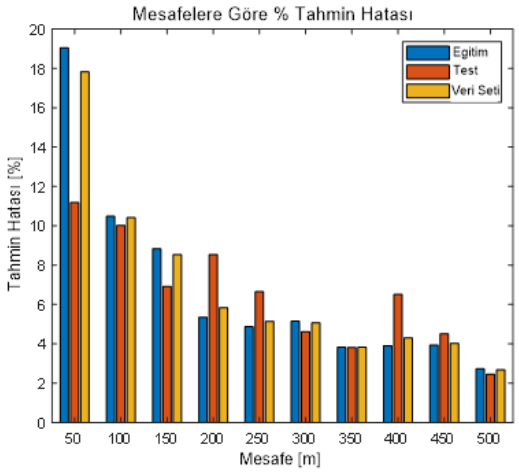


Şekil 6.14 Seçilen ağın tahminleme grafikleri (a) eğitim, (b) test

Şekil 6.14'da gösterilen tahmin sonuçlarının ortalama yüzde hata miktarları eğitim, test ve tüm veri seti için ayrı ayrı Şekil 6.15'de görülmektedir. RMSE hataları denklem (6.1) ile hesaplanmaktadır.

$$RMSE = \sqrt{\frac{1}{N_p}\sum_{0}^{N_p}\left(d_{k_i} - \hat{d}_{k_i}\right)^2} \tag{6.1}$$

Eğitim ve test veri setleri için her mesafe N_p değerleri sırasıyla 44 ve 8'dir



Şekil 6.15 Oluşturulan konvolüsyonel ağın %RMSE tahmin hatasının mesafelere göre değişimi

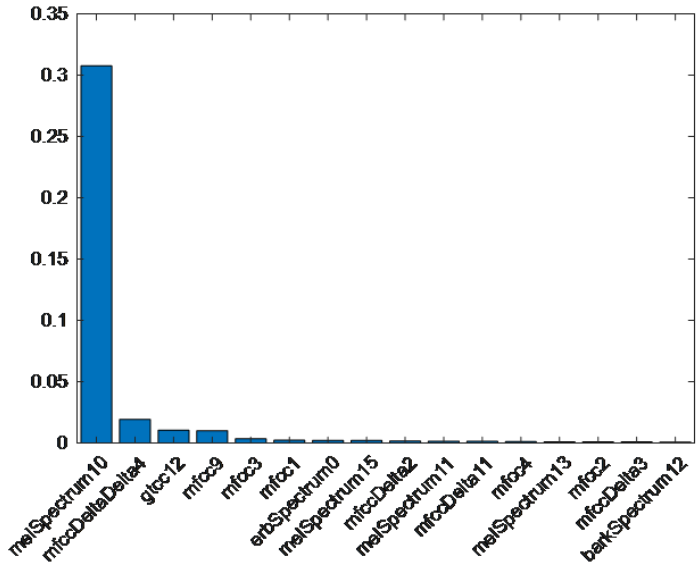
Şekil 6.14'de görülen tahminleme sonuçları ve Şekil 6.15'de görülen hata değerleri eğitim ve test veri setleri için Tablo 6.10'da ayrı ayrı raporlanmıştır.

Tablo 6.10 Oluşturduğumuz konvolüsyonel ağın mesafelere göre hata değerleri

Data	Value	50m	100m	150m	200m	250m	300m	350m	400m	450m	500m
Train	Min	36.85	78.42	116.90	175.67	225.80	248.78	322.56	375.73	407.39	465.26
	Median	44.24	93.14	144.07	198.31	250.38	295.58	349.99	399.12	444.96	495.79
	Max	84.94	104.54	168.64	219.39	287.80	328.55	391.06	438.31	489.55	525.78
	RMSE	9.52	10.48	13.24	10.67	12.17	15.44	13.37	15.60	17.70	13.61
	%RMSE	19.04	10.48	8.83	5.35	4.87	5.15	3.82	3.90	3.93	2.72
Test	Min	40.47	80.84	131.11	197.38	221.35	291.83	332.38	378.67	425.03	483.78
	Median	45.57	92.20	146.39	207.14	254.85	309.23	347.75	392.72	455.82	498.12
	Max	52.43	104.17	165.24	230.89	279.47	324.66	371.30	457.25	479.18	518.22
	RMSE	5.59	10.01	10.38	17.07	16.62	13.84	13.34	26.04	20.29	12.22
	%RMSE	11.17	10.01	6.92	8.53	6.65	4.62	3.81	6.51	4.51	2.45

6.6.1 Özellik Seçiminin Tahminleme Doğruluğuna Etkisi

Veri setini oluşturan seslerin %15'i test verisi olarak ayrıldıktan sonra eğitim versindeki seslere ait üçüncü bölümde belirtilen tüm özellikler elde edilmiştir. Daha sonra MRMR algoritması kullanılarak bu seslere ait özellikler ile çıkış arasındaki



Şekil 6.16 MRMR algoritması ile elde edilen en etkili 16 özelllik

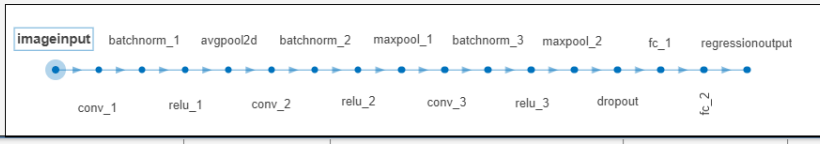
ilişki elde edilmiştir. Bu algoritma ile özelliklerin bir biri ile ilişkisi var ise en yüksek çıkış ilişkisine sahip özellik seçilerek diğer yüksek ilişkiye sahip özellikler elenmiştir. Bu şekilde çıkışla en yüksek ilişkiye, birbirleri ile ise en düşük ilişkiye sahip özellikler seçilmiştir. Oluşturulan 140 özellik içinde en belirleyici onaltı özelliğin sıralanmasıyla oluşan grafik Şekil 6.16'de verilmiştir.

Bu özelliklerden oluşan özellik seçim matrisindeki özellik sütunları kendi içinde normalizasyon işlemlerinden geçirilerek ağın eğitilmesi için gerekli giriş matrisi elde edilmiştir. Ön filtreleme işlemi yapılmadan elde eğitilen giriş matrisleri ile eğitilen ağ Şekil 6.17'da verilmiştir. Bu ağ 102400 öğrenebilir ağırlık ve 17 katman içermektedir.

CNN ağından elde edilen eğitim ve test tahmin sonuçları Şekil 6.18 ve Şekil 6.19'de verilmiştir. Tablo 6.11'de Şekil 6.18 ait değerler sunulmuştur.

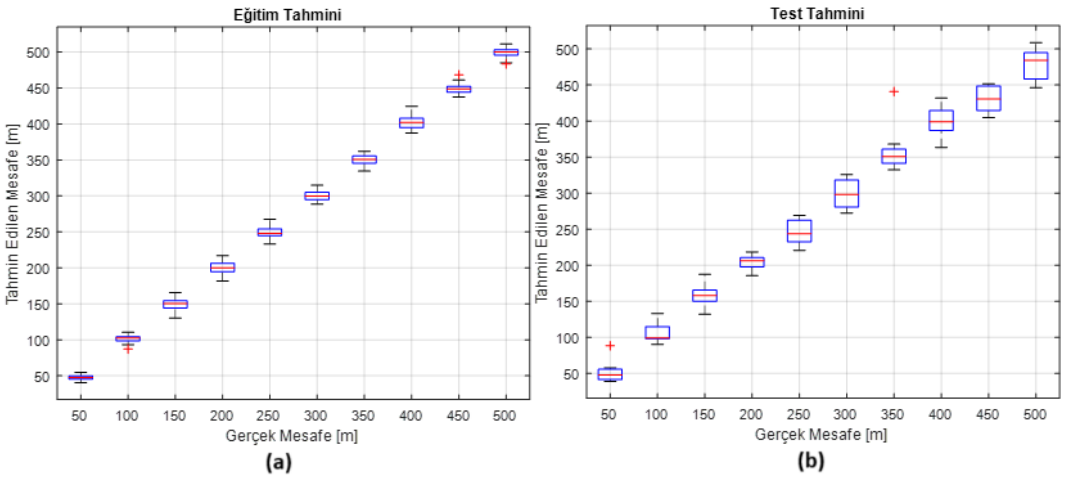
Tablo 6.11 Seçilen özellikler ile eğitilen CNN ağının mesafelere göre hata değerleri

Data	Value	50m	100m	150m	200m	250m	300m	350m	400m	450m	500m
Train	Min	40.53	87.20	130.31	181.55	232.93	288.40	334.40	387.60	437.65	483.40
	Median	47.42	102.33	150.38	199.82	247.78	299.78	350.66	402.13	448.45	500.25
	Max	54.89	110.42	165.80	216.79	267.56	314.84	361.90	424.47	468.34	511.20
	RMSE	3.88	5.06	7.12	7.71	7.40	6.58	5.87	8.72	6.15	5.70
	%RMSE	7.75	5.06	4.75	3.86	2.96	2.20	1.68	2.18	1.37	1.14
Test	Min	39.15	90.23	132.09	185.58	220.68	272.37	332.40	363.43	404.68	446.28
	Median	48.00	99.05	158.00	206.33	243.84	298.08	350.93	399.09	430.39	484.23
	Max	88.43	133.06	187.39	218.46	269.01	325.61	440.82	431.73	451.63	508.71
	RMSE	15.17	15.26	17.12	10.47	16.90	19.07	33.67	20.57	25.93	29.85
	%RMSE	30.34	15.26	11.41	5.24	6.76	6.36	9.62	5.14	5.76	5.97

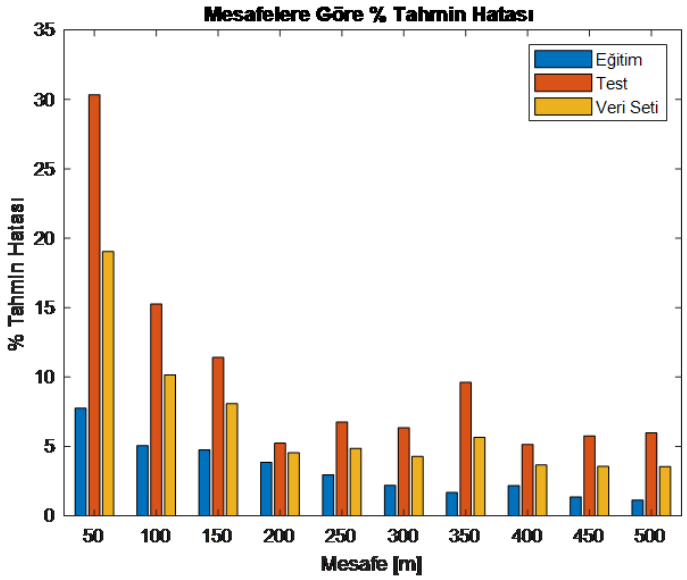


	Name	Type	Activations	Learnable Prope...	States
2	conv_1 20 12×12×1 convolutions with stride [1 1...	2-D Convolution	16(S) × 16(S) × 20(C) × 1(B)	Weig... 12 × 12 × 1... Bias 1 × 1 × 20	-
3	batchnorm_1 Batch normalization with 20 channels	Batch Normalization	16(S) × 16(S) × 20(C) × 1(B)	Offset 1 × 1 × 20 Scale 1 × 1 × 20	TrainedMe... 1 × 1... TrainedVa... 1 × 1...
4	relu_1 ReLU	ReLU	16(S) × 16(S) × 20(C) × 1(B)	-	-
5	avgpool2d 2×2 average pooling with stride [1 1] an...	2-D Average Pooling	15(S) × 15(S) × 20(C) × 1(B)	-	-
6	conv_2 20 10×10×20 convolutions with stride [1 ...	2-D Convolution	15(S) × 15(S) × 20(C) × 1(B)	Wei... 10 × 10 × 2... Bias 1 × 1 × 20	-
7	batchnorm_2 Batch normalization with 20 channels	Batch Normalization	15(S) × 15(S) × 20(C) × 1(B)	Offset 1 × 1 × 20 Scale 1 × 1 × 20	TrainedMe... 1 × 1... TrainedVa... 1 × 1...
8	relu_2 ReLU	ReLU	15(S) × 15(S) × 20(C) × 1(B)	-	-
9	maxpool_1 2×2 max pooling with stride [1 1] and pa...	2-D Max Pooling	14(S) × 14(S) × 20(C) × 1(B)	-	-
10	conv_3 20 8×8×20 convolutions with stride [1 1] ...	2-D Convolution	14(S) × 14(S) × 20(C) × 1(B)	Weig... 8 × 8 × 20... Bias 1 × 1 × 20	-
11	batchnorm_3 Batch normalization with 20 channels	Batch Normalization	14(S) × 14(S) × 20(C) × 1(B)	Offset 1 × 1 × 20 Scale 1 × 1 × 20	TrainedMe... 1 × 1... TrainedVa... 1 × 1...
12	relu_3 ReLU	ReLU	14(S) × 14(S) × 20(C) × 1(B)	-	-
13	maxpool_2 2×2 max pooling with stride [1 1] and pa...	2-D Max Pooling	13(S) × 13(S) × 20(C) × 1(B)	-	-
14	dropout 30% dropout	Dropout	13(S) × 13(S) × 20(C) × 1(B)	-	-
15	fc_1 10 fully connected layer	Fully Connected	1(S) × 1(S) × 10(C) × 1(B)	Weights 10 × 3380 Bias 10 × 1	-
16	fc_2 1 fully connected layer	Fully Connected	1(S) × 1(S) × 1(C) × 1(B)	Weights 1 × 10 Bias 1 × 1	-
17	regressionoutput mean-squared-error with response 'Res...	Regression Output	1(S) × 1(S) × 1(C) × 1(B)	-	-

Şekil 6.17 Özellik seçimi ile eğitilen CNN ağı



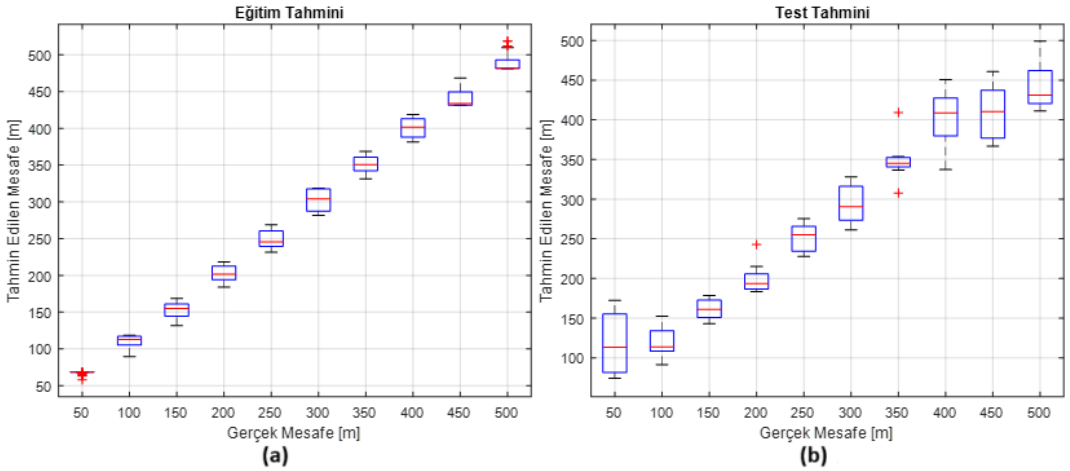
Şekil 6.18 Özellik seçimi ile eğitilen CNN ağının tahminleri (a)- eğitim, (b)- test



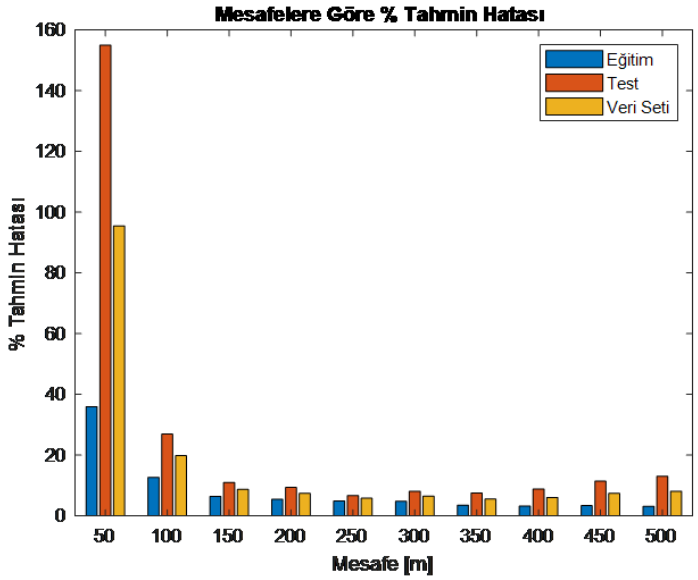
Şekil 6.19 Özellik seçimi ile eğitilen CNN ağının mesafelere göre ortalama tahminleme hataları

6.7 Makine Öğrenmesi Yöntemlerinin Başarısı

Seslere ait giriş matrislerinin her bir sütunu yan yana getirilerek giriş vektörü oluşturulmuştur. Giriş vektörleri Matlab Regression Layer araç kutusunda yer alan 27 farklı makine öğrenmesi yöntemlerinde eğitildikten sonra test verileri ile denenmiştir. En başarılı eğitim verisi tahmini Üstsel GPR algoritması ile, en başarılı test verisi tahmini SVM Kernel algoritması ile elde edilmiştir. SVM Kernel algoritması ile elde edilen eğitim ve test tahmin dağımları Şekil 6.20 ve 6.21'de görülmektedir. Şekil 6.20'de görülen grafikteki değerler Tablo 6.12'de verilmiştir.



Şekil 6.20 En başarılı makine öğrenmesi yöntemi tahmin grafikleri



Şekil 6.21 En başarılı makine öğrenmesi yönteminin tahmin başarısı

Tablo 6.12 Seçilen özellikler ile eğitilen makine öğrenmesi yönteminin sonuçları

Veri	Değer	50m	100m	150m	200m	250m	300m	350m	400m	450m	500m
Train	Min	58.02	89.58	131.49	184.81	231.49	281.47	331.51	381.45	431.44	481.43
	Median	68.51	112.92	154.87	201.59	245.78	303.90	350.84	401.81	433.76	481.80
	Max	68.55	118.56	168.57	218.45	268.68	318.56	368.63	418.57	468.52	518.55
	RMSE	17.98	12.71	9.61	10.93	12.20	14.44	12.35	13.02	15.15	15.88
	%RMSE	35.96	12.71	6.41	5.46	4.88	4.81	3.53	3.25	3.37	3.18
Test	Min	74.19	91.47	142.86	183.47	227.75	261.12	307.83	337.24	366.81	410.97
	Median	113.17	113.49	160.92	193.50	255.18	290.58	344.97	408.85	410.51	431.15
	Max	172.51	152.33	178.43	242.66	275.39	328.17	409.29	450.69	460.73	499.10
	RMSE	77.49	26.94	16.50	18.69	16.78	24.34	26.38	35.43	51.45	65.19
	%RMSE	154.99	26.94	11.00	9.35	6.71	8.11	7.54	8.86	11.43	13.04

7
SONUÇ

Bu tezde literatürde ilk defa altı mikrofondan oluşan küresel tasarımlı tek sensör dizisi ile sadece namlu patlaması sesini algıyarak konumlama yapabilen yapay zeka destekli hibrit bir silah sesi konumlama yöntemi gerçekleştirilmiştir. Elde ettiğimiz kapalı-form çözümün, mevcut çözümlere göre başarısı simulasyon ortamında ortaya konuş, simulasyon çalışmasındaki sonuçların başarısı deneysel çalışma ile doğrulanmıştır. Simulasyon çalışmasındaki grafikler incelendiğinde mikrofonlar arası açıklık çok azaltıldığında sıfır derece hataları oluşmaya başlamaktadır. Fazlaca arttırıldığında ise sonuçlar doğrusallığını korusa da kalıcı bir hata oluşturmaktadır. Örnekleme frekansının artması tahminleme hatasını düşürmekte doğrusal grafiğe yaklaşmaktadır. Kaynak mesafesinin uzaklığının fazlasıyla artmasının sonucu değiştirmediği görülmüştür ve deneysel olarak doğrulanmıştır. Tüm grafiklerde önerdiğimiz yöntemin uzak alan yaklaşımından daha kesin sonuçlar üretmektedir. Ayrıca diğer iki yöntemde açı değeri hesaplandıktan sonra hangi quadrantta olduğunu anlamak için gereken algoritmik yapı önerdiğimiz yöntemde ortadan kalmaktadır. Diğer yöntemlerde yer alan sıfır açı hatası önerdiğimiz yöntemde düşük olmakta, uygun örnekleme frekansı ve mikrofon açıklığı ile 1 derecenin altına inebilmektedir. Uzak alan yaklaşımında 0-30 derece arasında önemli derece ölçüm hatası oluşmasına rağmen önerdiğimiz yöntem tüm açı değerlerinde aynı kararlılığa sahiptir. Eş merkezli ve dik yerleştirilmiş geometriye sahip sensör dizileri için elde ettiğimiz çözüm formülü gerçek uygulamalar için daha başarılı olduğu ortaya konmuştur.

Ses üstü mühimmatlar için kullanılan baristik-akustik model için Tablo 1.3 ve Tablo 1.4 incelendiğinde tek mikrofon dizisi kullanarak sadece namlu patlaması sesi ile konumlama yapılamadığı görülmektedir. Tek mikrofon dizisi ile yapılan konumlamada da sistem parametrelerinin kümilatif hatası oldukça yüksek çıkmaktadır. Multiletarasyon yönteminin doğru sonuçlar üretebilmesi için mikrofonlar arası TDoA sürelerinin yüksek örnek değerlerine sahip olması gerekmektedir. Tablo 5.1 ve Tablo 5.2 birlikte incelendiğinde gerçek zamanlı

uygulamalar için 44.1 kHz gibi düşük örnekleme frekanslarında ortalama %10 hatanın altında sonuçlar almak için 10 m yarıçapa, 500 kHz gibi yüksek örnekleme frekansları için ise 3 m yarıçapa ihtiyaç vardır. Bu yarıçaplarda bile 300 m'den sonra hatanın %10'un üzerine çıktığı görülmektedir. Bu şekliyle karakol gibi sabit sahaları korumak için kullanılabilir bir yöntem olsa da hareketli birlikler tarafından kullanılması mümkün olmamaktadır. Sistemimizin tüm veri seti sonuçlarına göre ortalama mesafe tahmin hatası 13,4 m, tahmin hatası ise %6,76 olmaktadır. Tablo 6.8 incelendiğinde hem eğitim hem de test veri setinde 100 m'den sonra %10 hatanın altına inildiği görülmektedir. 500 m'ye doğru ise hata değerlerimizin %3'ün altına kadar düştüğü görülmektedir.

Tablo 6.2 ve Tablo 6.4 incelendiğinde ise her azimuth ve elevasyon açısı için 2 derecenin altında RMS hata değeri elde edilmiştir. 130 patlama noktasından 260 adet patlama ile elde edilen ortalama Azimuth açısı RMS hatası 1,39 derece, elevasyon açısı RMS hatası 1,33 derece olarak elde edilmiştir. Hesaplanan Azimuth ve elevasyon açılarının hata aralıkları mevcut sistemlerin aralıklarına yakın değerlerde elde edilmiştir. Mevcut sistemler ve çalışmalara göre önerdiğimiz yöntemin mesafe tahmin başarısı Tablo 7.1'de ortaya konmuştur [112].

Tablo 7.1 Mesafe tahmini % başarı oranları

Mesafe Aralığı [m]	**50**	**100**	**150**	**200**	**250**	**300**	**350**	**400**	**450**	**500**
PİLAR [21]	%10 < 200m				%20 < 600m					
Boomerang [18]	%11		%31 @ 300m							
Akman et. Al. [36]	%10	%30 @ 125m								
Önerilen-MLP	%23.36	%9.53	%7.15	%7.94	%7.49	%8.72	%4.56	%4.04	%4.83	%3.39
Önerilen-CNN	%17.83	%10.41	%8.53	%5.84	%5.14	%5.10	%3.82	%4.30	%4.02	%2.68

Bu sonuçlara göre önerdiğimiz yöntemin mesafe tahmin başarısı 50 m mesafe için mevcut çalışmalar göre yüksek hataya sahip olsa da 100 m'den sonraki mesafelerde hata oranı diğer sistemlerden daha düşük olmaktadır. 50-200 m arası önerilen yöntem ortalama %10,65 hataya, 200-500 m arası mesafe için ise %4,18 hata oranına sahiptir. Bu sonuçlara göre önerilen yöntem 50-200 m arasında mevcut sistemlere hemen hemen denk sonuçlar üretse de 200-500 m arasında PILAR sisteminden yaklaşık %79,1 daha başarılı sonuçlar üretmektedir.

Tablo 6.10 ile Tablo 6.11 kıyaslandığında özellik seçimi yapıldığında daha küçük bir ağ yapısıyla daha iyi bir eğitim verisi tahmini elde edilmiştir. Fakat test verisinde beklenen iyileşmenin sağlanamadığı görülmüştür. Tablo 6.12 ile karşılaştırdığımızda ise makine öğrenme yöntemleri ile test verileri için hata oranlarınnın 50 metreden sonra ortalama %11'in üzerinde olduğu görülmektedir.

İleri çalışmalarda sadece derin öğrenme kullanarak açı ve mesafe bilgisini sağlayacak yöntemler geliştirilebilir. Ayrıca farklı ses özelliklerinin ve farklı derin

ağ yapılarının kullanılması, farklı silahlardan oluşan daha zengin veri ile sistemin genelleştirme yeteneği arttırılması ve konumlama ile birlikte silah sınıflamasının birleştirilmesi ileride çalışılması gereken önemli konulardır. Yöntemimiz özellikle kamusal ve askeri alanlarda kendisine kullanım alanı bulacaktır.

KAYNAKÇA

[1] Erişim Tarihi: 29 Eylül 2024. erişim adresi: https://www.iwm.org.uk/collections/item/object/205358152.

[2] J. R. Graves, *Audio gunshot detection and localization systems: history, basic design, and future possibilities*. University of Colorado at Denver, 2012.

[3] R. Costley, "Battlefield acoustics in the First World War: artillery location," *Acoust. Today*, c. 16, ss. 31–39, 2020.

[4] E. A. Page B. Sharkey, "SECURES: system for reporting gunshots in urban environments," *Public Safety/Law Enforcement Technology*, SPIE, c. 2497, 1995, ss. 160–172.

[5] L. G. Mazerolle, *Random gunfire problems and gunshot detection systems*. US Department of Justice, Office of Justice Programs, National Institute of . . ., 1999.

[6] Erişim Tarihi: 29 Eylül 2024. erişim adresi: https://www.soundthinking.com/law-enforcement/leading-gunshot-detection-system/.

[7] R. C. Maher, "Modeling and Signal Processing of Acoustic Gunshot Recordings," *2006 IEEE 12th Digital Signal Processing Workshop & 4th IEEE Signal Processing Education Workshop*, 2006, ss. 257–261. doı: 10.1109/DSPWS.2006.265386.

[8] G. S. Settles, "High-speed Imaging of Shock Waves, Explosions and Gunshots New digital video technology, combined with some classic imaging techniques, reveals shock waves as never before," 2016. erişim adresi: https://api.semanticscholar.org/CorpusID:125657950.

[9] E. M. Carapezza, D. B. Law C. J. Csanadi, "DARPA counter-sniper program: Phase 1 Acoustic Systems Demonstration results," *Command, Control, Communications, and Intelligence Systems for Law Enforcement*, SPIE, c. 2938, 1997, ss. 299–310.

[10] G. L. Duckworth, D. C. Gilbert J. E. Barger, "Acoustic counter-sniper system," *Command, control, communications, and intelligence systems for law enforcement*, SPIE, c. 2938, 1997, ss. 262–275.

[11] R. B. Stoughton, "SAIC SENTINEL acoustic counter-sniper system," *Command, Control, Communications, and Intelligence Systems for Law Enforcement*, SPIE, c. 2938, 1997, ss. 276–284.

[12] G. Prado, H. Dhaliwal P. O. Martel, "Acoustic sniper localization system," *Command, Control, Communications, and Intelligence Systems for Law Enforcement*, SPIE, c. 2938, 1997, ss. 318–325.

[13] S. Magnuson, "IN THE Crosshairs: Technologists take aim at enemy snipers," *National Defense*, c. 92, no. 647, ss. 48–51, 2007.

[14] J. A. Mazurek ve diğ., "Boomerang mobile counter shooter detection system," *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IV*, SPIE, c. 5778, 2005, ss. 264–282.

[15] G. L. Duckworth ve diğ., "Fixed and wearable acoustic counter-sniper systems for law enforcement," *Sensors, C3I, information, and training technologies for law enforcement*, SPIE, c. 3577, 1999, ss. 210–230.

[16] J. Bedard S. Pare, "Ferret: a small arms fire detection system: localization concepts," *Sensors, and command, control, communications, and intelligence (C3I) technologies for homeland defense and law enforcement II*, SPIE, c. 5071, 2003, ss. 497–509.

[17] J. Bédard, "Operating Ferret on a patrol boat," *Unattended Ground, Sea, and Air Sensor Technologies and Applications VIII*, SPIE, c. 6231, 2006, ss. 279–289.

[18] Erişim Tarihi: 29 Eylül 2024. erişim adresi: `https://defense-update.com/20060726_pilar.html`.

[19] A. Donzier S. Cadavid, "Small arm fire acoustic detection and localization systems: gunfire detection system," *Sensors, and Command, Control, Communications, and Intelligence (C3I) Technologies for Homeland Security and Homeland Defense IV*, SPIE, c. 5778, 2005, ss. 245–253.

[20] J. Millet B. Baligand, "Latest achievements in gunfire detection systems," *Proceedings of the of the RTO-MP-SET-107 Battlefield Acoustic Sensing for ISR Applications*, 2006.

[21] Erişim Tarihi: 29 Eylül 2024. erişim adresi: `https://ndiastorage.blob.core.usgovcloudapi.net/ndia/2011/smallarms/WednesdaySI12570Weiland.pdf`.

[22] M. Ertem, R. Pierson, D. Burchick T. Ippolito, "An Acoustic Sensor for the Viper Infrared Sniper Detection System," *Military Sensing Symposia, IRIS Specialty Group on Battlefield Acoustics and Seismics*, 1999.

[23] Erişim Tarihi: 29 Eylül 2024. erişim adresi: `https://www.bu.edu/articles/2005/bu-technology-is-the-brain-behind-new-military-robot-prototype/`.

[24] Erişim Tarihi: 29 Eylül 2024. erişim adresi: `https://www.qinetiq.com/-/media/6a36dd9cf8304e3281709dcaa348ec1f.ashx`.

[25] H. Sébastien, *Sniper Localization using a Helmet Array*, 2006.

[26] P. Volgyesi, G. Balogh, A. Nadas, C. B. Nash A. Ledeczi, "Shooter localization and weapon classification with soldier-wearable networked sensors," *Proceedings of the 5th international conference on Mobile systems, applications and services*, 2007, ss. 113–126.

[27] E. Danicki, “Acoustic sniper localization,” *Archives of acoustics*, c. 30, no. 2, ss. 233–245, 2005.

[28] E. Danicki, “The shock wave-based acoustic sniper localization,” *Nonlinear analysis: theory, methods & applications*, c. 65, no. 5, ss. 956–962, 2006.

[29] T. Mäkinen P. Pertilä, “Shooter localization and bullet trajectory, caliber, and speed estimation based on detected firing sounds,” *Applied acoustics*, c. 71, no. 10, ss. 902–913, 2010.

[30] T. Damarla, L. M. Kaplan G. T. Whipps, “Sniper localization using acoustic asynchronous sensors,” *IEEE Sensors Journal*, c. 10, no. 9, ss. 1469–1478, 2010.

[31] S. Cheinet T. Broglin, “Sensitivity of shot detection and localization to environmental propagation,” *Applied Acoustics*, c. 93, ss. 97–105, 2015.

[32] D. M. P. Calderon J. A. Apolinario Jr, “Shooter localization based on DoA estimation of gunshot signals and digital map information,” *IEEE Latin America Transactions*, c. 13, no. 2, ss. 441–447, 2015.

[33] Ç. Akman, T. Sönmez, Ö. Özuğur, A. B. Başlı M. K. Leblebicioğlu, “Sensor fusion, sensitivity analysis and calibration in shooter localization systems,” *Sensors and Actuators A: Physical*, c. 271, ss. 66–75, 2018.

[34] A. Abiri A. Parsayan, “The bullet shockwave-based real-time sniper sound source localization,” *IEEE Sensors Journal*, c. 20, no. 13, ss. 7253–7264, 2020.

[35] P.-A. Grumiaux, S. Kitić, L. Girin A. Guérin, “A survey of sound source localization with deep learning methods,” *The Journal of the Acoustical Society of America*, c. 152, no. 1, ss. 107–151, 2022.

[36] M. J. Bianco ve diğ., “Machine learning in acoustics: Theory and applications,” *The Journal of the Acoustical Society of America*, c. 146, no. 5, ss. 3590–3628, 2019.

[37] X. Zhang, H. Sun, S. Wang J. Xu, “A new regional localization method for indoor sound source based on convolutional neural networks,” *IEEE Access*, c. 6, ss. 72 073–72 082, 2018.

[38] J. M. Vera-Diaz, D. Pizarro J. Macias-Guarasa, “Towards end-to-end acoustic localization using deep learning: From audio signals to source position coordinates,” *Sensors*, c. 18, no. 10, s. 3418, 2018.

[39] L. Gerosa, G. Valenzise, M. Tagliasacchi, F. Antonacci A. Sarti, “Scream and gunshot detection in noisy environments,” *2007 15th European Signal Processing Conference*, IEEE, 2007, ss. 1216–1220.

[40] A. Pikrakis, T. Giannakopoulos S. Theodoridis, “Gunshot detection in audio streams from movies by means of dynamic programming and bayesian networks,” *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2008, ss. 21–24.

[41] T. Ahmed, M. Uppal A. Muhammad, “Improving efficiency and reliability of gunshot detection systems,” *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, IEEE, 2013, ss. 513–517.

[42] B. Uzkent, B. D. Barkana H. Cevikalp, "Non-speech environmental sound classification using SVMs with a new set of features," *International Journal of Innovative Computing, Information and Control*, c. 8, no. 5, ss. 3511–3524, 2012.

[43] I. L. Freire J. A. Apolinário Jr, "Gunshot detection in noisy environments," *Proceeding of the 7th International Telecommunications Symposium, Manaus, Brazil*, c. 1, 2010.

[44] C.-F. Chan W. Eric, "An abnormal sound detection and classification system for surveillance applications," *2010 18th European Signal Processing Conference*, IEEE, 2010, ss. 1851–1855.

[45] M. Lojka, M. Pleva, E. Kiktová, J. Juhár A. Čižmár, "Efficient acoustic detector of gunshots and glass breaking," *Multimedia Tools and Applications*, c. 75, ss. 10 441–10 469, 2016.

[46] P. Suman, S. Karan, V. Singh R. Maringanti, "Algorithm for gunshot detection using mel-frequency cepstrum coefficients (MFCC)," *Proceedings of Ninth International Conference on Wireless Communication and Sensor Networks: WCSN 2013*, Springer, 2014, ss. 155–166.

[47] K. J. Piczak, "Environmental sound classification with convolutional neural networks," *2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP)*, ss. 1–6, 2015. erişim adresi: https://api.semanticscholar.org/CorpusID:206471429.

[48] F. Colangelo, F. Battisti, M. Carli, A. Neri F. Calabró, "Enhancing audio surveillance with hierarchical recurrent neural networks," *2017 14th IEEE International Conference on Advanced Video and Signal Based Surveillance (AVSS)*, 2017, ss. 1–6. doı: 10.1109/AVSS.2017.8078496.

[49] Ö. Inik, "CNN hyper-parameter optimization for environmental sound classification," *Applied Acoustics*, 2023. erişim adresi: https://api.semanticscholar.org/CorpusID:254781600.

[50] S. M. Khan, A. Divakaran H. S. Sawhney, "Weapon identification using hierarchical classification of acoustic signatures," *Defense + Commercial Sensing*, 2009. erişim adresi: https://api.semanticscholar.org/CorpusID:120126406.

[51] J. Sallai, W. Hedgecock, P. Völgyesi, A. Nádas, G. Balogh Á. Lédeczi, "Weapon classification and shooter localization using distributed multichannel acoustic sensors," *J. Syst. Archit.*, c. 57, ss. 869–885, 2011. erişim adresi: https://api.semanticscholar.org/CorpusID:28237646.

[52] K. Morton, P. Torrione L. M. Collins, "Classification of acoustic gunshot signatures using a nonparametric Bayesian signal model," *Defense + Commercial Sensing*, 2011. erişim adresi: https://api.semanticscholar.org/CorpusID:121180886.

[53] E. Kiktová-Vozarikova, M. Lojka, M. Pleva, J. Juhár A. Cizmár, "Gun type recognition from gunshot audio recordings," *3rd International Workshop on Biometrics and Forensics (IWBF 2015)*, ss. 1–6, 2015. erişim adresi: https://api.semanticscholar.org/CorpusID:20622444.

[54] H. A. Sánchez-Hevia, D. Ayllón, R. Gil-Pita M. Rosa-Zurera, "Maximum Likelihood Decision Fusion for Weapon Classification in Wireless Acoustic Sensor Networks," *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, c. 25, ss. 1172–1182, 2017. erişim adresi: https://api.semanticscholar.org/CorpusID:7230429.

[55] A. Makalesi ve diğ., "AN AUTOMATED GUNSHOT AUDIO CLASSIFICATION METHOD BASED ON FINGER PATTERN FEATURE GENERATOR AND ITERATIVE RELIEFF FEATURE SELECTOR," 2021. erişim adresi: https://api.semanticscholar.org/CorpusID:244427702.

[56] S. A. Qureshi ve diğ., "Gunshots Localization and Classification Model Based on Wind Noise Sensitivity Analysis Using Extreme Learning Machine," *IEEE Access*, c. 10, ss. 87 302–87 321, 2022. doı: 10.1109/ACCESS.2022.3198966.

[57] S. Raponi, G. Oligeri I. M. Ali, "Sound of guns: digital forensics of gun audio samples meets artificial intelligence," *Multimedia Tools and Applications*, c. 81, ss. 30 387–30 412, 2020. erişim adresi: https://api.semanticscholar.org/CorpusID:215814581.

[58] J. Park, Y. Cho, G. Sim, H. Lee J. Choo, "Enemy Spotted: In-game Gun Sound Dataset for Gunshot Classification and Localization," *2022 IEEE Conference on Games (CoG)*, 2022, ss. 56–63. doı: 10.1109/CoG51982.2022.9893670.

[59] N. B. Gaikwad, H. Ugale, A. Keskar N. C. Shivaprakash, "The Internet-of-Battlefield-Things (IoBT)-Based Enemy Localization Using Soldiers Location and Gunshot Direction," *IEEE Internet of Things Journal*, c. 7, ss. 11 725–11 734, 2020. erişim adresi: https://api.semanticscholar.org/CorpusID:226510385.

[60] C. M. J. Galangque S. A. Guirnaldo, "Gunshot Classification and Localization System using Artificial Neural Network (ANN)," *2019 12th International Conference on Information & Communication Technology and System (ICTS)*, ss. 1–5, 2019. erişim adresi: https://api.semanticscholar.org/CorpusID:203606216.

[61] J. Blauert, "Spatial Hearing: The Psychophysics of Human Sound Localization," 1983. erişim adresi: https://api.semanticscholar.org/CorpusID:117099185.

[62] J. C. Middlebrooks D. M. Green, "Sound localization by human listeners.," *Annual review of psychology*, c. 42, ss. 135–59, 1991. erişim adresi: https://api.semanticscholar.org/CorpusID:22801061.

[63] L. Rayleigh, "XII. On our perception of sound direction," *Philosophical Magazine Series 1*, c. 13, ss. 214–232, 1907. erişim adresi: https://api.semanticscholar.org/CorpusID:124589104.

[64] P. Zahorik, D. S. Brungart A. W. Bronkhorst, "Auditory distance perception in humans : A summary of past and present research," *Acta Acustica United With Acustica*, c. 91, ss. 409–420, 2005. erişim adresi: https://api.semanticscholar.org/CorpusID:123368698.

[65] W. A. Yost, "Auditory Perception: A New Analysis and Synthesis," *Trends in Neurosciences*, c. 22, 1999. erişim adresi: https://api.semanticscholar.org/CorpusID:53201309.

[66] J. Ahveninen, N. Kopo I. P. Jääskeläinen, "Psychophysics and neuronal bases of sound localization in humans," *Hearing Research*, c. 307, ss. 86–97, 2014. erişim adresi: https://api.semanticscholar.org/CorpusID:13230729.

[67] S. D. Beck, H. Nakasone K. W. Marr, "Variations in recorded acoustic gunshot waveforms generated by small firearms.," *The Journal of the Acoustical Society of America*, c. 129 4, ss. 1748–59, 2011. erişim adresi: https://api.semanticscholar.org/CorpusID:5296038.

[68] W. E. Baker, "Explosions in air," 1973. erişim adresi: https://api.semanticscholar.org/CorpusID:135758531.

[69] N. Madhu R. Martin, "Acoustic Source Localization with Microphone Arrays," 2008. erişim adresi: https://api.semanticscholar.org/CorpusID:108855157.

[70] M. S. Brandstein H. F. Silverman, "A robust method for speech signal time-delay estimation in reverberant rooms," *1997 IEEE International Conference on Acoustics, Speech, and Signal Processing*, c. 1, 375–378 vol.1, 1997. erişim adresi: https://api.semanticscholar.org/CorpusID:15440482.

[71] D. Zhuo H. Cao, "Fast Sound Source Localization Based on SRP-PHAT Using Density Peaks Clustering," *Applied Sciences*, 2021. erişim adresi: https://api.semanticscholar.org/CorpusID:234214749.

[72] M. Cobos, A. Marti J. J. López, "A Modified SRP-PHAT Functional for Robust Real-Time Sound Source Localization With Scalable Spatial Sampling," *IEEE Signal Processing Letters*, c. 18, ss. 71–74, 2011. erişim adresi: https://api.semanticscholar.org/CorpusID:18207534.

[73] R. O. Schmidt, "Multiple emitter location and signal Parameter estimation," 1986. erişim adresi: https://api.semanticscholar.org/CorpusID:31251598.

[74] A. L. Swindlehurst T. Kailath, "On the sensitivity of the ESPRIT algorithm to non-identical subarrays," *Sadhana*, c. 15, ss. 197–212, 1990. erişim adresi: https://api.semanticscholar.org/CorpusID:16125111.

[75] A. Karbasi A. Sugiyama, "A new DOA estimation method using a circular microphone array," *2007 15th European Signal Processing Conference*, ss. 778–782, 2007. erişim adresi: https://api.semanticscholar.org/CorpusID:8302784.

[76] R. Boora S. K. Dhull, "A TDOA-based multiple source localization using delay density maps," *Sādhanā*, c. 45, 2020. erişim adresi: https://api.semanticscholar.org/CorpusID:255493772.

[77] M. S. Brandstein D. B. Ward, "Microphone Arrays Signal Processing Techniques and Applications," 2001. erişim adresi: https://api.semanticscholar.org/CorpusID:54115066.

[78] Q. Li, X. Zhang H. Li, "Online Direction of Arrival Estimation Based on Deep Learning," *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ss. 2616–2620, 2018. erişim adresi: https://api.semanticscholar.org/CorpusID:52285697.

[79] X. Xiao, S. Zhao, X. Zhong, D. L. Jones, C. E. Siong H. Li, "A learning-based approach to direction of arrival estimation in noisy and reverberant environments," *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, ss. 2814–2818, 2015. erişim adresi: https://api.semanticscholar.org/CorpusID:7672501.

[80] M. Ahmad, M. Muaz M. Adeel, "A Survey of Deep Neural Network in Acoustic Direction Finding," *2021 International Conference on Digital Futures and Transformative Technologies (ICoDT2)*, ss. 1–6, 2021. erişim adresi: https://api.semanticscholar.org/CorpusID:235308345.

[81] M.-A. Chung, H.-C. Chou C.-W. Lin, "Sound Localization Based on Acoustic Source Using Multiple Microphone Array in an Indoor Environment," *Electronics*, 2022. erişim adresi: https://api.semanticscholar.org/CorpusID:247465313.

[82] S. Paulose, E. Sebastian B. Paul, "Acoustic Source Localization," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Energy*, c. 2, ss. 933–939, 2013. erişim adresi: https://api.semanticscholar.org/CorpusID:42147549.

[83] R. B. Calhoun ve diğ., "Precision and accuracy of acoustic gunshot location in an urban environment," *ArXiv*, c. abs/2108.07377, 2021. erişim adresi: https://api.semanticscholar.org/CorpusID:237142402.

[84] H. Agirman, S. Karakutuk E. Gonendik, "MIKES gunshot detection system prototype," *2014 22nd Signal Processing and Communications Applications Conference (SIU)*, ss. 2090–2093, 2014. erişim adresi: https://api.semanticscholar.org/CorpusID:15559222.

[85] C. H. Knapp G. C. Carter, "The generalized correlation method for estimation of time delay," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, c. 24, ss. 320–327, 1976. erişim adresi: https://api.semanticscholar.org/CorpusID:121803300.

[86] P. Song, C. Hao, J. Wu C. Yang, "Acoustic source localization using 10-microphone array based on wireless sensor network," *Sensors and Actuators A-physical*, c. 267, ss. 376–384, 2017. erişim adresi: https://api.semanticscholar.org/CorpusID:115928962.

[87] M. Imran, A. Hussain, N. M. Qazi M. Sadiq, "A methodology for sound source localization and tracking: Development of 3D microphone array for near-field and far-field applications," *2016 13th International Bhurban Conference on Applied Sciences and Technology (IBCAST)*, ss. 586–591, 2016. erişim adresi: https://api.semanticscholar.org/CorpusID:22634024.

[88] T. Kundu, "Acoustic source localization.," *Ultrasonics*, c. 54 1, ss. 25–38, 2019. erişim adresi: https://api.semanticscholar.org/CorpusID:205528441.

[89] C. Ma, Z. Zhou, J. Liu, Z. Cui T. Kundu, "Acoustic source localization using L-shaped sensor clusters: A review.," *Ultrasonics*, c. 132, s. 107 020, 2023. erişim adresi: https://api.semanticscholar.org/CorpusID:258296082.

[90] A. K. Dey S. Saha, "Acoustic Beamforming : Design and Development of Steered Response Power With Phase Transformation (SRP-PHAT).," 2011. erişim adresi: https://api.semanticscholar.org/CorpusID:106751955.

[91] S. Kharel, "Performance Comparison Between Music And Esprit Algorithms For Direction Estimation Of Arrival Signals," 2013.

[92] O. Tuna E. AFACAN, "İşaret Geliş Açısı Kestirim Algoritmaları Estimation of Direction of Arrival Algorithms,"

[93] L. Osman, I. Sfar A. Gharsallah, "Comparative study of high-resolution direction-of-arrival estimation algorithms for array antenna system," *International Journal of Research and Reviews in Wireless Communications (IJRRWC)*, c. 2, no. 1, 2012.

[94] E. Ünal, M. Kayaoğlu, B. Topçu M. U. Doğan, "An acoustic based sniper detection system prototype," *2011 IEEE 19th Signal Processing and Communications Applications Conference (SIU)*, IEEE, 2011, ss. 311–314.

[95] A. Parsayan S. M. Ahadi, "Real time high accuracy 3-D PHAT-based sound source localization using a simple 4-microphone arrangement," *IEEE Systems Journal*, c. 6, no. 3, ss. 455–468, 2011.

[96] K. Sakhnov, E. Verteletskaya B. Simák, "TIME DELAY ESTIMATION ALGORITHMS FOR ECHO CANCELLATION," 2011. erişim adresi: https://api.semanticscholar.org/CorpusID:51859282.

[97] B. G. Ferguson, L. G. Criswick K. W. Lo, "Locating far-field impulsive sound sources in air by triangulation," *The Journal of the Acoustical Society of America*, c. 111, no. 1, ss. 104–116, 2002.

[98] B. G. Ferguson, K. W. Lo R. J. Wyber, "Acoustic sensing of direct and indirect weapon fire," *2007 3rd International Conference on Intelligent Sensors, Sensor Networks and Information*, IEEE, 2007, ss. 167–172.

[99] J. O. Smith J. S. Abel, "Closed-form least-squares source location estimation from range-difference measurements," *IEEE Trans. Acoust. Speech Signal Process.*, c. 35, ss. 1661–1669, 1987. erişim adresi: https://api.semanticscholar.org/CorpusID:36990155.

[100] G. B. Whitham, "The flow pattern of a supersonic projectile," *Communications on pure and applied mathematics*, c. 5, no. 3, ss. 301–348, 1952.

[101] J. Sallai, P. Volgyesi, K. Pence A. Ledeczi, "Fusing distributed muzzle blast and shockwave detections," *14th International Conference on Information Fusion*, IEEE, 2011, ss. 1–8.

[102] T. Zhang C.-C. J. Kuo, "Audio content analysis for online audiovisual data segmentation and classification," *IEEE Trans. Speech Audio Process.*, c. 9, ss. 441–457, 2001. erişim adresi: https://api.semanticscholar.org/CorpusID:13077835.

[103] A. Das, N. Borisov M. C. Caesar, "Fingerprinting Smart Devices Through Embedded Acoustic Components," *ArXiv*, c. abs/1403.3366, 2014. erişim adresi: https://api.semanticscholar.org/CorpusID:14632753.

[104] G. Peeters, "A large set of audio features for sound description (similarity and classification) in the CUIDADO project," *CUIDADO Ist Project Report*, c. 54, no. 0, ss. 1–25, 2004.

[105] S. S. Stevens J. E. Volkmann, "The Relation of Pitch to Frequency: A Revised Scale," *American Journal of Psychology*, c. 53, s. 329, 1940. erişim adresi: https://api.semanticscholar.org/CorpusID:147563163.

[106] S. Umesh, L. Cohen D. J. Nelson, "Fitting the Mel scale," *1999 IEEE International Conference on Acoustics, Speech, and Signal Processing. Proceedings. ICASSP99 (Cat. No.99CH36258)*, c. 1, 217–220 vol.1, 1999. erişim adresi: https://api.semanticscholar.org/CorpusID:39418190.

[107] N. Çalık, L. Durak-Ata, A. Serbes, B. Bolat E. Yavuz, "Performance analysis of feature extraction methods in indoor sound classification," *2015 23nd Signal Processing and Communications Applications Conference (SIU)*, ss. 2025–2028, 2015. erişim adresi: https://api.semanticscholar.org/CorpusID:7844235.

[108] G. Gulgezen, Z. Cataltepe L. Yu, "Stable feature selection using MRMR algorithm," *2009 IEEE 17th Signal Processing and Communications Applications Conference*, ss. 596–599, 2009. erişim adresi: https://api.semanticscholar.org/CorpusID:12834269.

[109] Y. LeCun, L. Bottou, Y. Bengio P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, c. 86, ss. 2278–2324, 1998. erişim adresi: https://api.semanticscholar.org/CorpusID:14542261.

[110] S. Ioffe C. Szegedy, "Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift," *ArXiv*, c. abs/1502.03167, 2015. erişim adresi: https://api.semanticscholar.org/CorpusID:5808102.

[111] K. Zengin A. Yeşildirek, "Closed-Form DoA Solution for Co-Centered Orthogonal Microphone Arrays Based on Multilateration Equations," *Applied Sciences*, c. 13, no. 20, s. 11 297, 2023.

[112] K. Zengin A. Yeşildirek, "A Hybrid Model for 3-D Gunshot Localization Using Muzzle Blast Sound Only," *IEEE Transactions on Instrumentation and Measurement*, c. 73, ss. 1–10, 2024. doı: 10.1109/TIM.2024.3481655.

TEZDEN ÜRETİLMİŞ YAYINLAR

Makale

1. Zengin K, Yeşildirek A. Closed-Form DoA Solution for Co-Centered Orthogonal Microphone Arrays Based on Multilateration Equations. Applied Sciences. 2023; 13(20):11297. https://doi.org/10.3390/app132011297

2. K. Zengin and A. Yeşildirek, A Hybrid Model for 3-D Gunshot Localization Using Muzzle Blast Sound Only, in IEEE Transactions on Instrumentation and Measurement, vol. 73, pp. 1-10, 2024, Art no. 2533810, doi: 10.1109/TIM.2024.3481655.

Konferans Bildirisi

1. Kazım Zengin, Aydın Yeşildirek, "A Noel DoA Estimation Method Based On Multilateration Equations" The 7th International Conference on Computational Mathematics and Engineering Sciences / 20-21 May. 2023, Elazığ – Türkiye