

**BAŐKENT ÜNİVERSİTESİ  
FEN BİLİMLERİ ENSTİTÜSÜ**

**ADLİ UYGULAMALAR İÇİN  
SES İÇERİK ANALİZİ**

**SERCAN SARMAN**

YÜKSEK LİSANS TEZİ

2018



**ADLİ UYGULAMALAR İÇİN  
SES İÇERİK ANALİZİ**

**AUDIO CONTENT ANALYSIS FOR  
APPLICATIONS IN FORENSICS**

**SERCAN SARMAN**

Başkent Üniversitesi  
Lisansüstü Eğitim Öğretim ve Sınav Yönetmeliğinin  
BİLGİSAYAR Mühendisliği Anabilim Dalı İçin Öngördüğü  
YÜKSEK LİSANS TEZİ  
olarak hazırlanmıştır.

2018

“Adli Uygulamalar İçin Ses İçerik Analizi” başlıklı bu çalışma, jürimiz tarafından, 20/03/2018 tarihinde, **BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI'nda YÜKSEK LİSANS TEZİ** olarak kabul edilmiştir.

Başkan : Doç. Dr. Murat KOYUNCU

Üye (Danışman) : Yrd. Doç. Dr. Mustafa SERT

Üye : Yrd. Doç. Dr. Emre SÜMER

**ONAY**

.../03/2018

Prof. Dr. Emin AKATA



**BAŞKENT ÜNİVERSİTESİ FEN BİLİMLERİ ENSTİTÜSÜ**

**YÜKSEK LİSANS / DOKTORA TEZ ÇALIŞMASI ORJİNALLİK RAPOR**

Tarih: 27 / 03 / 2018

Öğrencinin Adı, Soyadı : Sercan SARMAN

Öğrencinin Numarası : 21320094

Anabilim Dalı : Bilgisayar Mühendisliği A.B.D.

Programı : Bilgisayar Mühendisliği Tezli Y.L.

Danışmanın Adı, Soyadı : Mustafa SERT

Tez Başlığı : Adli Uygulamalar İçin Ses İçerik Analizi

Yukarıda başlığı belirtilen Yüksek Lisans tez çalışmamın; Giriş, Ana Bölümler ve Sonuç Bölümünden oluşan, toplam 54 sayfalık kısmına ilişkin, 27 / 03 / 2018 tarihinde tez danışmanım tarafından Turnitin adlı intihal tespit programından aşağıda belirtilen filtrelemeler uygulanarak alınmış olan orijinallik raporuna göre, tezimin benzerlik oranı % 4'tür.

Uygulanan filtrelemeler:

1. Kaynakça hariç
2. Alıntılar hariç
3. Beş (5) kelimedenden daha az örtüşme içeren metin kısımları hariç

“Başkent Üniversitesi Enstitüleri Tez Çalışması Orijinallik Raporu Alınması ve Kullanılması Usul ve Esasları”nı inceledim ve bu uygulama esaslarında belirtilen azami benzerlik oranlarına tez çalışmamın herhangi bir intihal içermediğini; aksinin tespit edileceği muhtemel durumda doğabilecek her türlü hukuki sorumluluğu kabul ettiğimi ve yukarıda vermiş olduğum bilgilerin doğru olduğunu beyan ederim.

Öğrenci İmzası

/ / 2018

## **TEŞEKKÜR**

Yazar, bu çalışmanın gerçekleşmesinde katkılarından dolayı, aşağıda adı geçen kişilere içtenlikle teşekkür eder.

Tez danışmanım Sayın Yrd. Doç. Dr. Mustafa SERT'e, çalışmanın sonuca ulaştırılmasında ve karşılaşılan güçlüklerin aşılmasında her zaman yardımcı ve yol gösterici olduğu için...

Eşime ve anneme her zaman destek oldukları için...

Sayın Araş. Gör. Tunç AŞUROĞLU ve Araş. Gör. Koray AÇICI'ya yanımda oldukları için...

## ÖZ

### ADLİ UYGULAMALAR İÇİN SES İÇERİK ANALİZİ

Sercan Sarman

Başkent Üniversitesi Fen Bilimleri Enstitüsü

Bilgisayar Mühendisliği Anabilim Dalı

Günümüzde artan şiddet olayları, adli incelemelerin de önemini artırmıştır. Şiddet olaylarının ardından gerçekleştirilecek olan adli incelemeler esnasında, erişilebilir durumda olan bütün işitsel ve görsel veriler oldukça kıymetlidir. Olayın gerçekleştiği konumun tespit edilmesi, şiddetin türünün belirlenmesi ve benzeri süreçler, adli ses analizi kapsamında yer almaktadır. Günümüzde çevrimiçi içeriğe erişimin akıllı cihazlar aracılığıyla konum bağımsız olarak gerçekleştirilebiliyor olması ve sunulan içeriğin miktarının hızlı bir şekilde artmasıyla; içeriğin otomatik olarak sınıflandırılmasının önemini artırmıştır. Özellikle çocuk ve gençleri olumsuz olarak etkileyebilecek içeriğin otomatik olarak tespit edilmesi, içerik miktarının hızlı artışıyla birlikte önem kazanmıştır. Buna karşılık, sinyal işleme alanında, özellikle de adli ses analizi kapsamında gerçekleştirilen çalışmaların başarımı, diğer alanlarda kullanılan makine öğrenmesi yöntemlerinin şiddet sahnesi sınıflandırması alanına uygulanabileceğini göstermiştir. Bu tez çalışması kapsamında, silah seslerinin ve video verilerinin şiddet içeren sahnelerinin ses tabanlı sınıflandırılması problemleri ele alınmıştır. Bu amaçla, makine öğrenmesi metotlarının ve topluluk öğrenmesi yaklaşımları probleme uygulanmıştır. Yöntemler, performans veri kümeleri üzerinde karşılaştırmalı olarak incelenmiş ve silah sesleri sınıflandırılması alanında %66, şiddet sahnesi sınıflandırması alanında %62'ye varan sınıflandırma başarımları elde edilmiştir.

**ANAHTAR SÖZCÜKLER:** adli ses analizi, destek vektör makineleri, topluluk öğrenme, şiddet sahnesi sınıflandırması

**Danışman:** Yrd. Doç. Dr. Mustafa Sert, Başkent Üniversitesi, Bilgisayar Mühendisliği Bölümü

## **ABSTRACT**

### **AUDIO CONTENT ANALYSIS FOR APPLICATIONS IN FORENSICS**

Sercan Sarman

Başkent University Institute of Science and Engineering

Department of Computer Engineering

Nowadays, the increase in violent events has enhanced the importance of forensic investigations. All accessible auditory and visual data are highly valuable during the examination to be performed after violent events. Audio forensics analysis contains determination of location in which violent incident occur and determination of type of violence. Recently, the location-free and easier access to online content via smart devices and the increase of content have enhanced the importance of automatical classification of content. With the rapid growth in the amount of content, it has become crucial to automatically determine the content that can adversely affect children and youth. On the other hand, the success of the studies carried out in the field of signal processing, especially in the context of audio forensic analysis, shows that the methods of machine learning used in other areas can be applied to the field of violent scene classification. In this study, we study the problem of gunshot sounds and violent scene classification. For this purpose, machine learning and ensemble learning approaches applied to this problem. We examine classification rates of various machine learning and ensemble learning approaches comparatively and we achieve classification accuracies of 66% and 62% in audio gunshot classification and violent scene classification, respectively.

**KEYWORDS:** audio forensic analysis, support vector machines, ensemble learning, violent scene classification

**Advisor:** Asst. Prof. Mustafa Sert, Ph.D., Başkent University, Department of Computer Engineering



# İÇİNDEKİLER LİSTESİ

	<u>Sayfa</u>
1. GİRİŞ.....	1
1.1. Problemin Tanımı .....	1
1.2. Kapsam .....	1
1.3. Amaç .....	2
1.4. Hedefler.....	2
1.5. Literatüre Katkısı .....	3
2. İLGİLİ ÇALIŞMALAR.....	4
3. GENEL KAVRAMLAR VE TANIMLAR .....	9
3.1. Öznitelikler.....	9
3.1.1. Mel-Frequency Cepstral Coefficients (MFCC).....	9
3.1.2. Perceptual Linear Prediction (PLP) .....	12
3.1.3. Rasta Filtered Perceptual Linear Prediction (Rasta-PLP).....	13
3.1.4. Linear Prediction Coding (LPC).....	13
3.1.5. Zero Crossing Rate (ZCR).....	14
3.1.6. Sound Pressure Level (SPL) .....	14
3.2. Sınıflandırıcılar .....	14
3.2.1. Destek vektör makineleri .....	15
3.2.2. Torbalama .....	17
3.2.3. Rastgele ormanlar .....	18
4. SİLAH SESLERİNİN SINIFLANDIRILMASI.....	20
4.1. Veri kümesi.....	22
4.2. Destek vektör makinesi deneyleri.....	23
4.3. Torbalama Deneyleri .....	31
4.4. Rastgele ormanlar .....	35
5. ŞİDDET SAHNESİ SINIFLANDIRMASI .....	37
5.1. Destek Vektör Makineleri.....	41
5.2. Torbalama .....	46
5.3. Rastgele Ormanlar .....	50
5.4. Önerilen Yöntemler .....	51
6. SONUÇLAR VE TARTIŞMA.....	53
KAYNAKLAR LİSTESİ .....	55

<b>ŞEKİLLER LİSTESİ</b>	<u>Sayfa</u>
Şekil 3.1 MFCC özniteliğinin çıkarımı aşamaları.....	9
Şekil 3.2 PLP özniteliğinin çıkarımı aşamaları.....	12
Şekil 4.1 Silah seslerinin sınıflandırılması işleminin gerçekleştirilme aşamaları.....	21
Şekil 5.1 Şiddet sahnesi sınıflandırma işleminin aşamaları.....	37
Şekil 5.2 Şiddet sahnesi sınıflandırması veri kümesinde MFCC özniteliğinin dağılımı.....	39
Şekil 5.3 Şiddet sahnesi sınıflandırması veri kümesinde ZCR özniteliğinin dağılımı.....	40
Şekil 5.4 Şiddet sahnesi sınıflandırması, deneysel sonuçların literatür [20] ile karşılaştırılması grafiği.....	52

## ÇİZELGELER LİSTESİ

	<u>Sayfa</u>
Çizelge 4.1 Silah seslerinin sınıflandırılması veri kümesi dağılımı.....	22
Çizelge 4.2 Silah seslerinin sınıflandırılması, c-destek vektör makinesi doğrusal çekirdek sınıflandırması sonuçları.....	24
Çizelge 4.3 Silah seslerinin sınıflandırılması, c-destek vektör makinesi polinomal çekirdek sınıflandırması sonuçları.....	25
Çizelge 4.4 Silah seslerinin sınıflandırılması, c-destek vektör makinesi radyal çekirdek sınıflandırması sonuçları.....	26
Çizelge 4.5 Silah seslerinin sınıflandırılması, c-destek vektör makinesi sigmoid çekirdek sınıflandırması sonuçları.....	27
Çizelge 4.6 Silah seslerinin sınıflandırılması, nu-destek vektör makinesi doğrusal çekirdek sınıflandırması sonuçları.....	28
Çizelge 4.7 Silah seslerinin sınıflandırılması, nu -destek vektör makinesi polinomal çekirdek sınıflandırması sonuçları.....	29
Çizelge 4.8 Silah seslerinin sınıflandırılması, nu-destek vektör makinesi radyal çekirdek sınıflandırması sonuçları.....	30
Çizelge 4.9 Silah seslerinin sınıflandırılması, nu -destek vektör makinesi sigmoid çekirdek sınıflandırması sonuçları.....	31
Çizelge 4.10 Silah seslerinin sınıflandırılması, Torbalama yöntemi ile elde edilen sonuçlar, 10 torba c-destek vektör makinesi sonuçları...	30
Çizelge 4.11 Silah seslerinin sınıflandırılması, Torbalama yöntemi ile elde edilen sonuçlar, 20 torba c-destek vektör makinesi sonuçları...	33
Çizelge 4.12 Silah seslerinin sınıflandırılması, Torbalama yöntemi ile elde edilen sonuçlar, 10 torba nu-destek vektör makinesi sonuçları.	33
Çizelge 4.13 Silah seslerinin sınıflandırılması, Torbalama yöntemi ile elde edilen sonuçlar, 20 torba nu-destek vektör makinesi sonuçları.	34
Çizelge 4.14 Silah seslerinin sınıflandırılması, Rastgele Ormanlar yöntemi ile elde edilen sonuçlar, 10 ağaç sonuçları.....	35
Çizelge 4.15 Silah seslerinin sınıflandırılması, Rastgele Ormanlar yöntemi ile elde edilen sonuçlar, 20 ağaç sonuçları.....	36
Çizelge 5.1 Şiddet sahnesi sınıflandırması için kullanılan MediaEval Violent Scene Detection Task veri kümesinde yer alan filmlerin eğitim ve test kümelerine dağılımı, şiddet içeren ve içermeyen sahnelerin toplam sahnelere oranı.....	41
Çizelge 5.2 Şiddet sahnesi sınıflandırması, c-destek vektör doğrusal çekirdek sınıflandırması sonuçları.....	41
Çizelge 5.3 Şiddet sahnesi sınıflandırması, c-destek vektör polinomal çekirdek sınıflandırması sonuçları.....	42
Çizelge 5.4 Şiddet sahnesi sınıflandırması, c-destek vektör radyal	

	çekirdek sınıflandırması sonuçları.....	42
Çizelge 5.5	Şiddet sahnesi sınıflandırması, c-destek vektör sigmoid çekirdek sınıflandırması sonuçları.....	43
Çizelge 5.6	Şiddet sahnesi sınıflandırması, nu-destek vektör doğrusal çekirdek sınıflandırması sonuçları.....	44
Çizelge 5.7	Şiddet sahnesi sınıflandırması, nu-destek vektör polinomal çekirdek sınıflandırması sonuçları.....	44
Çizelge 5.8	Şiddet sahnesi sınıflandırması, nu-destek vektör radyal çekirdek sınıflandırması sonuçları.....	45
Çizelge 5.9	Şiddet sahnesi sınıflandırması, nu-destek vektör sigmoid çekirdek sınıflandırması sonuçları.....	45
Çizelge 5.10	Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, c-destek vektör makineleri doğrusal çekirdek sonuçları.....	46
Çizelge 5.11	Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, c-destek vektör makineleri polinomal çekirdek sonuçları.....	47
Çizelge 5.12	Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, c-destek vektör makineleri radyal çekirdek sonuçları.....	47
Çizelge 5.13	Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, c-destek vektör makineleri sigmoid çekirdek sonuçları.....	48
Çizelge 5.14	Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, nu-destek vektör makineleri doğrusal çekirdek sonuçları.....	48
Çizelge 5.15	Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, nu-destek vektör makineleri polinomal çekirdek sonuçları.....	49
Çizelge 5.16	Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, nu-destek vektör makineleri radyal çekirdek sonuçları.....	49
Çizelge 5.17	Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, nu-destek vektör makineleri sigmoid çekirdek sonuçları.....	50
Çizelge 5.18	Şiddet sahnesi sınıflandırması, Rastgele Ormanlar yöntemi ile elde edilen sonuçları.....	51
Çizelge 5.19	Şiddet sahnesi sınıflandırması, deneysel sonuçların literatür [20] ile karşılaştırılması.....	52

## SİMGELER VE KISALTMALAR LİSTESİ

Bagging	Torbalama (Bootstrap Aggregation)
BoAW	Ses Kelimeleri Torbası (Bag of Audio Words)
CC	Cepstral Katsayılar (Cepstral Coefficients)
CV	Katlamalı Çapraz Doğrulama (Cross-Validation)
DFT	Ayrık Fourier Dönüşümü (Discrete Fourier Transform)
FBI	Federal Araştırma Bürosu (Federal Bureau of Investigation)
FFT	Hızlı Fourier Dönüşümü (Fast Fourier Transform)
GMM	Gaussian Karıştırma Modeli (Gaussian Mixture Model)
IEEE	Elektrik Elektronik Mühendisleri Enstitüsü (Institute of Electrical and Electronics Engineering)
LOO	Birini Dışarıda Bırak (Leave one out)
LPC	Doğrusal Tahmin Katsayıları (Linear Prediction Coefficients)
MFCC	Mel Frekansı Cepstral Katsayıları (Mel Frequency Cepstral Coefficients)
OvA	Bire Karşı Hepsi (One Versus All)
PLP	Algısal Doğrusal Tahmin (Perceptual Linear Prediction)
RF	Rastgele Ormanlar (Random Forests)
SPL	Ses Basınç Seviyesi (Sound Pressure Level)
SVC	Destek Vektör Sınıflandırması (Support Vector Classification)
SVM	Destek Vektör Makineleri (Support Vector Machines)
ZCR	Sıfırı Geçiş Oranı (Zero Crossing Rate)

# 1. GİRİŞ

## 1.1. Problemin Tanımı

Günümüzde artan şiddet olayları adli incelemelerin de önemini artırmıştır. Bu istenmeyen olaylar zaman zaman gözlem verilerinin erişilebilirliğini de etkilemektedir. Video, ses ya da benzeri ham verilere, adli olaydan sonra bütün olarak erişmek mümkün olmayabilmektedir. Bu noktada erişilebilir durumda olan ham verinin değeri daha da artmaktadır.

Gerçekleşmiş olan herhangi bir şiddet olayının ardından görüntü kayıtlarına ulaşmak her zaman mümkün olmamaktadır. Özellikle de patlama ve benzeri olayların ardından cihazların zarar görmesi sebebiyle, ses ve görüntü gibi hassas verilerin bütünlüğünün korunamadığı görülmektedir. Örneğin patlamanın ardından oluşabilecek toz bulutu görüntünün sürekliliğini bozarken, ses kayıtları etkilenmemektedir. Benzer şekilde, görüntü kayıtları sis gibi iklim koşullarından etkilenirken ses kayıtları etkilenmemektedir. Bu noktada ses içeriği ön plana çıkmaktadır. Zira ses içeriğinden; olayın gerçekleştiği konum, şiddetin türü ve benzeri içerik çıkarılabilmektedir. Bu veriler tek başına kullanılabildiği gibi, görsel içerikten elde edilen bilgiye ek olarak da kullanılabilmektedir. Ses içeriği kullanılarak şiddetin tespit edilmesi, şiddetin kaynağının sınıflandırılması, şiddetin kaynağının belirlenmesi ve benzeri işlemler gerçekleştirilebilmektedir. Buna bağlı olarak, gerçek zamanlı müdahale gerektiren şiddet olaylarında, müdahalenin şekli belirlenirken yukarıda bahsedilen yöntemlerden faydalanılabilir. Örneğin bir patlama sonucunda ortaya bir yangın çıkması durumunda, yangın söndürme faaliyetleri patlamanın kaynağına göre yönlendirilebilir.

## 1.2. Kapsam

Bu tez çalışmasında; silah sesleri sınıflandırması ve şiddet sahnesi sınıflandırması amacıyla ses sinyallerinin normalizasyonu, öznitelik çıkarımı, sınıflandırıcılar için model öğrenme ve sınıflandırma işlemleri gerçekleştirilmiştir. Öncelikli olarak öznitelik çıkarım yönteminin sınıflandırma işleminin başarımına etkisi, ardından

sınıflandırma işleminde yer alan çeşitli parametrelerin sınıflandırma işleminin başarımına etkisi, son olarak topluluk öğrenme metotlarının uygulanmasının sistemin bütününe olan etkisi karşılaştırmalı olarak ele alınmıştır.

### **1.3. Amaç**

Bu tez kapsamında yapılan çalışmalar, gerçekleşmiş bir şiddet olayının ardından şiddeti tespit etme ve şiddetin kaynağını sınıflandırmayı amaçlamaktadır. Yapılan çalışmalar mevcut ses kayıtlarının çeşitli makine öğrenmesi yöntemleri yardımıyla sınıflandırılması ve tespit edilmesi sürecini gerçekleştirmektedir. Çalışmada, bahsi geçen işlemler için kullanılan yöntemlerin performansları karşılaştırmalı olarak sunulmuştur.

Bu tez çalışmasında, silah sesleri sınıflandırması kapsamında çeşitli silah seslerinin sınıflandırılması amaçlanmaktadır. Bu işlem için sınıflandırma algoritmalarından faydalanılmış olup, bunun için modeller oluşturulmuş ve bu modeller üzerinden sınıflandırma süreci gerçekleştirilmiştir. Uygulanan yöntemlerin sınıflandırma başarımı ise bu çalışma kapsamında karşılaştırmalı olarak paylaşılmıştır.

Bu tez kapsamında gerçekleştirilen bir diğer işlem ise, popüler filmler içerisinde yer alan şiddet içerikli sahnelerin tespit edilmesidir. Şiddet kaynaklarının sınıflandırılması konusunda öğrenilen yöntemler, popüler sinema filmlerinde yer alan şiddet içerikli sahnelerin tespit edilmesinde kullanılmış ve performansları karşılaştırmalı olarak verilmiştir. Bu noktada amaç, tüm filmler için gerçekleştirilen şiddet sınıflandırması işlemini gerçekleştirmektir.

### **1.4. Hedefler**

Bu tez kapsamında yapılan çalışmalar neticesinde, silah sesleri üzerinde adli analizler yapılarak silah seslerinin sınıflandırılması amaçlanmıştır. Bu bağlamda, elde edilen bilgi birikimi kullanılarak oluşturulan bir diğer hedef ise, şiddet sahnesi sınıflandırmasıdır.

## **1.5. Literatüre Katkısı**

Bu tez çalışması; adli ses analizi konusunda silah seslerinin sınıflandırılması ve şiddet sahnesi sınıflandırması başlıklarını incelemektedir. Geleneksel öznitelik çıkarımı ve sınıflandırıcı başarımlarını karşılaştırarak, topluluk yöntemlerinin belirtilmiş olan alan içerisindeki başarımlarını karşılaştırmalı olarak sunmaktadır. Ayrıca bu çalışmada, sınıflandırıcı parametrelerinin başarıma olan etkisine detaylı olarak yer verilmektedir. Son olarak, destek vektör makinelerinin, torbalama yönteminin ve rastgele ormanlar yönteminin silah seslerinin sınıflandırılması ve şiddet sahnesi sınıflandırması problemlerindeki sınıflandırma başarımlarını karşılaştırmalı olarak ortaya konmaktadır.



## 2. İLGİLİ ÇALIŞMALAR

Başkan Nixon'a ait teyplerin tüm transkriptlerini içeren bu kitap [1], adli ses analizinin başlangıç noktası olarak gösterilebilir. Birleşik Devletlerde, doğrudan ses kayıtları ile ilgilenen ilk dava 1958 yılında gerçekleşmiştir. İlk kez, yargıç McKeever davasında kanıtın gerçekliği ve doğruluğu sorgulanmıştır. Yargıç, konuşmanın transkriptinin mahkemeye sunulmasını istemiş; fakat sunulan kaydın geçerli olabilmesi için 6 spesifik gereklilikten bahsetmiştir: ses kayıtları, kayıt operatörü, doğrulanmış kayıt, kayıt üzerinde oynama olmaması, kaydın korunması ve konuşmacıların tanımlanması [2]. Koenig, Federal Araştırma Bürosu'nun (FBI) herhangi bir adli kaydın orijinal olup olmadığını araştırırken uyguladığı aşamaları detaylı olarak tanımlamıştır. Bu sürecin amacı, kaydın orijinal olup olmadığını araştırmak, kaydın kopya olup olmadığını anlamak, kayıta herhangi bir değişiklik yapıp yapılmadığını ortaya çıkarmaktır. Bu bilgiler, iştirakçi tarafından sunulan materyal, kayıtların oynatılması, genel bir analiz, kritik dinleme, fiziksel inceleme, manyetik iyileştirme, enstrümantal analiz (dalga boyu, dar bant spektrumu, spektrografi), çeşitli teknikler, inceleme örnekleri ve şahitlikten elde edilen veriler bütünleştirilerek ortaya çıkarılmaktadır [3].

Ses Mühendisliği Topluluğu'nun S43-2000 standardı, ses kayıtlarının doğruluğunun tespiti için bilimsel bir prosedür tanımlamıştır. Bu standardın amacı, kanıt olarak ya da kamusal olarak kullanıma sunulan kayıtların doğruluğunu bilimsel bir yaklaşımla formülize etmektir [4]. Ses Mühendisliği Topluluğu'na ait bir diğer standartta ise adli soruşturmalar içerisinde sunulmuş olan ses kayıtlarının saklanması, incelenmesi ve etiketlenmesi gibi süreçler tanımlanmıştır. Bu standart, diğerinden farklı olarak doğrulama süreciyle ilgilenmemektedir [5].

Koenig et al., geçtiğimiz yarım yüzyıllık süre içerisinde, adli ses incelemeleri tanınan bir iş kolu olma sürecinde köklü değişiklikler yaşandığını belirtmiştir. Sayısal sinyal işleme araçlarından önce adli ses konusunda çalışanlar, analog manyetik teypler ve basit kayıt stüdyosu ekipmanları ile çalışmaktaydılar. Bu çalışmalarda kullanılan ekipmanlar; analog filtreler, çeşitli hızlarda oynatma yeteneğine sahip cihazlar ile osiloskop, mikroskop ve ses spektrografları gibi ekipmanlardır [6]. Zakariah et al. ise

adli ses analizi sürecinin üç başlıktan bir ya da daha fazlasına yanıt aradığını belirtmiştir:

1) Bütünlük: Ses kaydı, oluşturulmasından inceleme aşamasına kadar değişikliğe uğramış mıdır?

2) İyileştirme ve anlaşılabilirlik: Kayıt daha anlaşılabilir hale getirilebilir mi?

3) Tanımlama: Kaydın içerisinde yer alan konuşmacılar tanımlanabilir mi?

Aynı çalışmada, sayısal sesin bütünlüğünün doğrulanma metotlarının iki ana başlıkta incelenebileceği belirtilmiştir. Bu başlıklar, içerik tabanlı doğrulama ve kapsayıcı tabanlı doğrulamadır. Ses kaydına ait dosya yapısı ile metadatayı ilgilendiren çalışmalar kapsayıcı tabanlı doğrulamanın, ses kaydına ait bit ve byte düzeyindeki işlemler ise içerik tabanlı doğrulamanın temelini oluşturmaktadır [7]. Zhao and Malik, akustik sahne imzasına dayanan bir sahte ses tespiti sunmuştur. Buna göre, test ses kayıtlarından elde edilen akustik sahne imzaları, sahte ses tespitinde kullanılmıştır. Bilgi edinme sistemi, nihai kaydın içerisinde izlerini bırakmaktadır ve bu bilgi kanıtın doğruluğu tespit etmek için kullanılmıştır. Bu veri kaynaktan kaynağa değişeceği için, gürültü sinyalinde yer alan değişiklikler doğruluğu tespit etmek için kullanılabilir [8]. Maher, kendisinden önce gelenlerin önerilerini ve yaklaşımlarını bir araya getirerek bütünlük bir yol haritası çizmiştir. Maher'e göre yapılması gerekenler şu şekilde sıralanmıştır:

1) Ses kaydının ya da ses kanıtının gerçekliğini, doğruluğunu, güvenilirliğini tespit etmek

2) Konuşma anlaşılabilirliğini ve düşük düzeydeki seslerin işitilebilirliğini sağlamak için ses kayıtlarında iyileştirmeleri gerçekleştirmek

3) Kanıtı yorumlamak ve dokümanete etmek: Konuşmacıları belirlemek, diyalogları tanımlamak, olayı yeniden canlandırmak ve zaman çizelgelerini ortaya koymak [9].

Maher, keskin nişancıların tespiti için gerçek zamanlı silah atışı tespiti ve atışın gerçekleştirildiği konumun tespiti konularında çalışma gerçekleştirmiştir. Bu çalışmada ortaya çıkan sese ait akustik bilgiler kullanılarak akustik modeller oluşturulmuştur ve bu bilgilerin ışığında konum bilgisinin yanı sıra mermi hızı ve yörüngesi bilgileri ile birlikte silahın ve mühimmatın türüne dair bilgilerin de elde edilebileceğini öne sürmüştür [10]. Martin, çalışmasında silah seslerinin destek vektör makineleri aracılığıyla sınıflandırılması konusunu ele almıştır. Bu bağlamda, farklı öznitelik türlerinin ve çekirdek parametrelerinin başarıma olan etkisini incelemiştir. Genlik tabanlı özniteliklerin çevresel koşullara karşı hassasiyeti bulgusunun yanı sıra, silah seslerinin akustik olarak anlık bir ses ortaya çıkmasının öznitelik oluşturma aşamasında ele alınmasının başarıyı arttırabileceğini belirtmiştir [11]. Navratil et al., yapay sinir ağları kullanarak çeşitli silah seslerini sınıflandırma konusunda çalışmıştır. Bu çalışmada deneysel sonuçları elde edebilmek için kendi veri kümelerini kontrollü olarak oluşturmuştur. Kullanılan tabancaları türlerine göre sınıflandırma işlemini gerçekleştirdiklerinde %59 başarımları elde etmiştir. Aynı yöntemi 3 ana silah türü tabanca, büyük çaplı silah ve patlama olarak uyguladıklarında ise %96 başarımları elde etmiştir [12]. Djeddou and Touhami, silah sesinin gürültülü çevrede elde edilmiş imzasını kullanarak silah seslerini sınıflandırma problemine odaklanmıştır. Bu bağlamda GMM algoritmasından ve hiyerarşik sınıflandırmadan faydalanmıştır. Çalışmada LPC, PLP, CC, MFCC öznitelikleri yer almaktadır [13]. Kiktova et al., akıllı akustik etkinlik tespit sistemlerini, sadece silah seslerini tespit eden sistemlerden silahın türünü tespit eden sistemlere geliştiren bir çalışma gerçekleştirmiştir. Bunun için Saklı Markov Modeli kullanmıştır ve %80 dolayında başarımları elde etmiştir. Bunun için hem zaman hem de frekans düzleminde öznitelikler kullanmıştır. Çalışmada belirtilen en büyük handikap ise veri kümesinin sınırlı olması olmuştur [14].

Şiddet sahnesi sınıflandırması söz konusu olduğunda çalışmalar 3 ana başlık altında incelenebilmektedir. Bu başlıklar, çalışmanın dayanağını oluşturan içeriğin niteliğinden gelmektedir ve bunlar görsel, işitsel ve görsel-işitsel olarak ele alınmaktadır. İlk grupta yer alan Mittal et al., çarpışmaya kalan zaman, sahne geçişi, hareketli şablonlar gibi algısal seviyede hareket özniteliklerinden oluşan bir öznitelik

konsepti geliřtirmişlerdir. Video görüntülerinin bu seviyede sunuluyor olmasının řiddet tespiti gibi içerik tabanlı çalışmalarda oldukça güçlü kıldığını ortaya koymuşlardır [15]. Datta et al., hareketin yönüne ve büyüklüğüne odaklanan bir hızlanma ölçüm vektörü tanımlamıştır ve bu vektör yardımıyla yumruk yumruğa kavga ve tekme gibi insan kaynaklı řiddet unsurlarını video görüntüleri içerisinde tespit etmeye çalışmıştır [16]. Kim et al., otonom olarak iş birliği yapan kamera sistemlerini destekleyen akıllı bir gözetleme sistemi önermiştir. Çok katmanlı ve birbirinden bağımsız olarak çalışan güvenlik seviyeleri tasarlamıştır ve bunun için ilgi bölgeleri oluşturmuştur. Bölgelerin güvenlik seviyelerini belirlemek için olağan dışı durumları kullanmıştır [17]. Souza et al. çalışmasında doğrusal destek vektör makineleri kullanan görsel kod çizelgeleri konseptine dayanan bir řiddet tespiti gerçekleştirmiştir. Mevcut řiddet tespitlerinden farklı olarak uzay-zamansal öznitelikleri görsel kelime torbaları ile birlikte kullanmıştır [18].

İkinci grupta ise işitsel çalışmalar yer almaktadır. Giannakopoulos et al., Bayesian Networks temelli çok sınıflı sınıflandırma algoritması ortaya koymuştur. Bu çalışmada filmlerden elde edilmiş ses segmentlerine yer vermiştir ve çalışmanın amacı olarak çocuklar gibi hassas yaş gruplarını korumayı ele almıştır. 12 ses özniteliği kullanmıştır ve ses segmentlerini üçü řiddet içeren üçü içermeyen olacak şekilde altı sınıfa ayırmıştır. Sınıflandırma aşamasında OvA (One-Versus-All, birine karşı hepsi) mimarisini kullanmıştır. Çalışmasının sonuç aşamasında farklı türlere ait 30 filmi ve 5000 ses segmentini kullanmıştır [19]. Acar et al., MFCC özniteliklerini kullanan bir BoAW metodu řiddet sahnesi sınıflandırması gerçekleştirmiştir. BoAW için Vector Quantization ve Sparse Coding yöntemlerini kullanmıştır. Sahnelerin řiddet içeren ve içermeyen olmak üzere ikili sınıflandırma işlemi için destek vektör makinelerinden yararlanmıştır. Sparse Coding içeren yöntemin Vector Quantization içeren yöntemden daha yüksek başarımla elde ettiğini gözlemlemiştir [20].

Son grupta ise görsel ve işitsel yaklaşımları bir arada kullanan çalışmalar yer almaktadır. Mu et al., çalışmasında işitsel özniteliklerin řiddetli sahne tespitindeki başarımlarını incelemiştir. Bu amaçla konvolüsyonel sinir ağlarından faydalanmıştır.

Bununla birlikte görsel ve işitsel özniteliklerin bir arada kullanımının başarımını arttırdığını da göstermiştir [21]. Sjöberg et al., şiddet sahnesi sınıflandırması probleminde yapay sinir ağlarının başarımını incelemiştir. Bu çalışmada genel amaçlı öznitelik çıkarımı yöntemleri ve genel amaçlı sınıflandırıcıları kullanarak şiddet sahnesi sınıflandırması gerçekleştirebileceklerini göstermiştir [22]. Giannakopoulos et al., video paylaşım sitelerinde şiddetli sahnelerin tespiti için ses, hareketli görüntü ve metin bilgilerini bütünleştiren bir çalışma gerçekleştirmiştir. Bu çalışmada metin bilgisi, kullanıcılar tarafından eklenen yorumlara karşılık gelmektedir. İkili sınıflandırma yapılan bu modelde, 7'si ses düzlemine ait olmak üzere 9 öznitelik kullanılmıştır. 210 YouTube video paylaşımı üzerinden gerçekleştirilen deneyler sonucunda %82 sınıflandırma başarımı elde edilmiştir [23].

İncelenen bu çalışmalar, adli ses analizi alanında yer alan silah seslerinin sınıflandırılması ve şiddet sahnesi sınıflandırması konularında özniteliklerin ve sınıflandırıcıların başarıma katkılarını sunmaktadır. Bu tez çalışmasında, literatürde gerçekleştirilmiş olan deneylerin yanı sıra topluluk öğrenmesinin adli ses analizine olan katkısı incelenmiştir.

### 3. GENEL KAVRAMLAR VE TANIMLAR

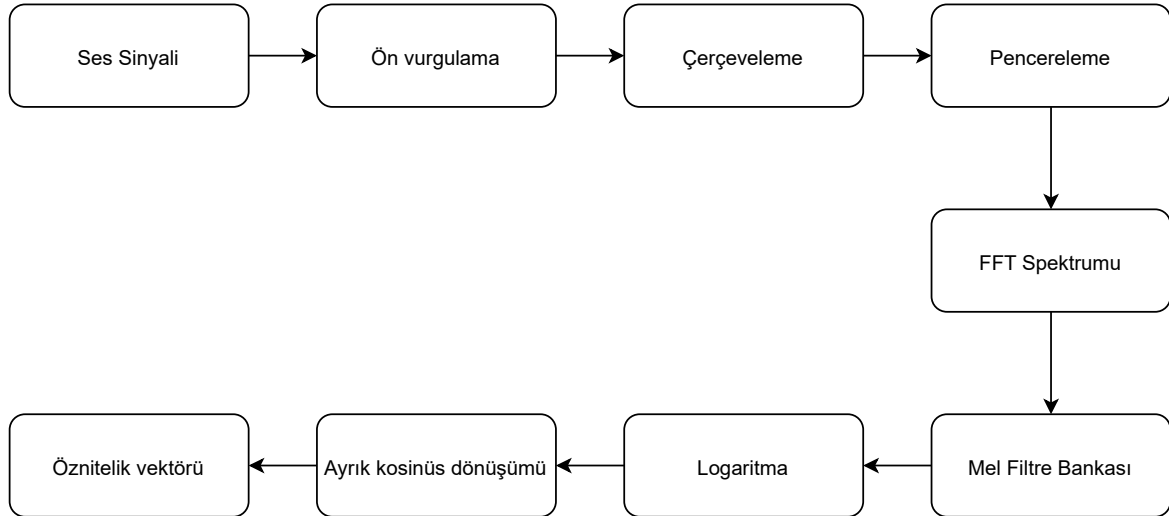
Bu bölümde, tez çalışmasında kullanılan temel kavramlar ve tanımlar yer almaktadır.

#### 3.1. Öznitelikler

Bu tez çalışmasında MFCC özniteliğine, insanın işitme özelliğini baz alarak tasarlandığı için, PLP ve Rasta-PLP özniteliklerine gürültü içeren sinyallerdeki başarımları için, LPC'nin yaygın olarak kullanılması sebebiyle, ZCR ve SPL özniteliklerine şiddet içeren ve içermeyen ses kayıtları baz alındığında ayırt edilebilirliği için yer verilmiştir.

##### 3.1.1. Mel-Frequency Cepstral Coefficients (MFCC)

Algı temelli sesi temsil eden özniteliklerdir. MFCC, frekans düzleminde yer alan bir özniteliktir. Bu özniteliğin oluşturulması esnasında gerçekleştirilen aşamalar aşağıda detaylı olarak verilmiştir:



Şekil 3.1 MFCC özniteliğinin çıkarımı aşamaları [29]

Şekil 3.1’de görüldüğü üzere öncelikli olarak ses sinyali, Ön Vurgulama aşamasından geçmektedir. Bu aşamanın amacı, yüksek frekanslardaki enerji miktarını artırmaktır. Frekansa bağlı olarak gerçekleşen enerji düşüşü, Spektral Eğim olarak adlandırılmaktadır. Yüksek frekanslı enerjiyi arttırmak, bu frekanslardan gelen bilgiyi akustik modele daha uygun hale getirmektedir.

Ön Vurgulama işlemi FIR (Finite Impulse Response-Sonlu Darbe Cevaplı) filtre aracılığıyla gerçekleşmektedir. Bunlar geri beslemesi olmayan filtreler olup, sinyal zaman içerisinde sifıra ulaştığı için sonlu sayıda terim içermektedir. Filtrenin çıkış değeri sadece şu andaki giriş değerine ve geçmiş çıkış değerlerine bağlıdır.

Ön Vurgulama işlemine ait denklem ise eşitlik (3.1)’de sunulmaktadır.

$$\text{Giriş sinyali } x[n] \text{ ve } 0.9 \leq \alpha \leq 1.0 \text{ olmak üzere, } y[n] = x[n] - \alpha x[n - 1] \quad (3.1)$$

Sinyalin zamana bağlı olarak istatistiksel özelliklerinin ortak olduğu alt bütünleri oluşturabilmek (çıkarmak) amacıyla, çerçeveleme işlemi gerçekleştirilmektedir.

Bu aşama, sinyalin ardışık N tane örnekten oluşan alt parçalar halinde ele alınabilmesini sağlamaktadır. Her çerçevenin zaman düzleminde bir uzunluğu bulunmakta olup, sinyalin örnekleme sıklığı ile birlikte ele alındığında çerçevenin kaç tane örnek içerdiği bulunmaktadır. Ardışık çerçeveler birbirinden M örnek uzaklıkta bulunmaktadır ve M değerinin N değerinden küçük olması durumunda çerçeveler birbiriyle örtüşmektedir.

Çerçeveleme aşaması ile bütünleşik bir şekilde ele alınabilecek olan bu aşamada; sinyalin, çerçeveleme aşamasında belirlenen ve tanımlanan alt bütünlerinin oluşturulması gerçekleşmektedir. İki çeşit pencereleme yöntemi bulunmaktadır: Dikdörtgen ve Hamming. En basit çerçeveleme yöntemi Dikdörtgen olmakla birlikte, sinyalin sınır değerlerinde oluşturduğu kesintilerden dolayı Fourier Analizi aşamasında süreksizlik nedeniyle sorunlar oluşabileceğinden MFCC öznitelik çıkarımı işleminde genellikle Hamming yöntemi kullanılmaktadır. Hamming, süreksizlikleri önleyerek pencere sınırlarında sinyal değerlerini sifıra doğru çekmektedir.

Pencereleme yönteminin denklemleri (3.2) ve (3.3)'de sunulmaktadır.

$$\text{Dikdörtgen} \quad w[n] = \begin{cases} 1, & 0 \leq n \leq L - 1 \\ 0, & \text{diğer durumlarda} \end{cases} \quad (3.2)$$

$$\text{Hamming} \quad w[n] = \begin{cases} 0.54 - 0.46\cos\left(\frac{2\pi n}{L}\right), & 0 \leq n \leq L - 1 \\ 0, & \text{diğer durumlarda} \end{cases} \quad (3.3)$$

Bir sonraki aşamada, pencerelenmiş sinyal içerisinde spektral bilginin çıkarılması yer alır. Burada amaç, sinyalin çeşitli frekans bantları içerisinde ne kadar enerji barındırdığını bulmaktır. Ayırık zamanda (örneklerde) yer alan ayırık frekans bantlarındaki sinyalin içerisindeki spektral bilgiye ulaşmamızı sağlayan araç, Ayırık Fourier Dönüşümüdür (DFT).

Ayırık Fourier Dönüşümü için girdi, pencerelenmiş sinyaldir  $x[n] \dots x[m]$ . Çıktı ise her bir N ayırık frekans bandı için  $X[k]$  kompleks değeridir. Bu değer, orijinal sinyalde yer alan frekans bileşeninin büyüklüğünü ve fazını temsil etmektedir. Büyüklüğün frekansa bağlı grafiği ortaya konulduğunda ise spektrum gözlenmektedir. Ayırık Fourier Dönüşümü için genellikle Hızlı Fourier Dönüşümü (FFT) kullanılmaktadır.

Fourier Dönüşümünün çıktısı, her bir frekans bandı için enerji miktarıdır. İnsan algısı her bir frekans bandı için eşit hassasiyette olmayıp, yüksek frekanslar için daha az hassasiyet barındırmaktadır. İnsanın işitmesinin bu özelliğinin öznelik çıkarımı aşamasında kullanılmasının, başarımı artırdığı gözlenmiştir. MFCC'lerde model kullanım şekli, bir önceki aşamada üretilen frekansları mel ölçeğine (skalasına) çevirmektir. Mel, perdede algısal olarak eşit mesafedeki ses çiftlerinin eşit sayıda mele ayrılması için tanımlanan bir perde birimidir. Hertz ile mel ölçeği arasındaki eşleme, 1000 Hz altında doğrusal ve 1000 Hz üzerinde ise logaritmiktir. Ham akustik frekanstan mel frekansı hesaplama yöntemi (3.4)'te sunulmaktadır:

$$mel(f) = 1127 \ln\left(1 + \frac{f}{700}\right) \quad (3.4)$$



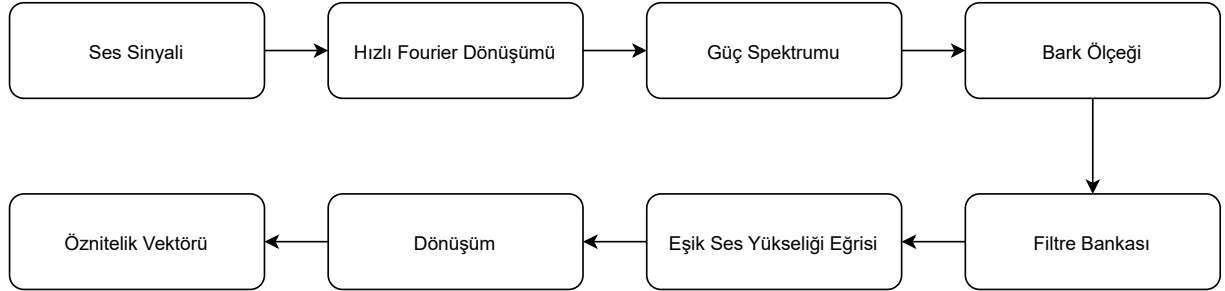
Bu sezgiselliği, 1000 Hz'den aşağı doğrusal aralıklı 10 filtre ile her bir frekans bandından enerji toplayan bir filtre grubu yaratarak ve geriye kalan filtreleri logaritmik olarak 1000 Hz'nin üzerinde yayarak oluşturmaktadır.

Genel olarak insan işitmesinin sinyal seviyesine tepkisi, logaritmik olarak değişmektedir. Genlik düzeyinde bakıldığında insanlar, yüksek genlikte büyük değişikliklere daha az duyarlıdır. Buna ek olarak logaritma fonksiyonu kullanmak, öznelikleri girdideki değişimlere daha az duyarlı hale getirmektedir.

Bu aşamada amaç, Mel spektrumunun log değerini zaman düzlemine çevirmektir. Bu aşamanın sonucunda MFCC öznelikleri oluşmaktadır. Oluşan bu değere akustik vektörler de denmektedir.

### 3.1.2. Perceptual Linear Prediction (PLP)

İşitsel spektrum tahminini türetmek için, işitme psikofiziğine ilişkin üç kavram kullanılır: kritik bant spektral çözünürlüğü, eşit ses yüksekliği eğrisi ve yoğunluk-ses şiddeti güç yasası. PLP, frekans düzleminde yer alan bir özneliktir. Bu özneliğin oluşturulması esnasında gerçekleştirilen aşamalar Şekil 3.2'de detaylı olarak verilmektedir.



Şekil 3.2 PLP özneliğinin çıkarımı aşamaları [30]

Özneliğin ortaya çıkış amacı, insan işitmesinin psikofiziğini öznelik çıkarımı yöntemi olarak kullanmaktır. Bu özneliğinin çıkarımının ilk, pencerelenmiş olan sinyalden Hızlı Fourier Dönüşümü kullanılarak güç spektrumunun hesaplanmasıdır. İkinci aşama, güç spektrumuna Bark ölçeğinin uygulanmasıdır. Bark ölçeği işitsel olarak anlamlı olan ölçeğin bir diğer türüdür. Bu aşamada filtre bankası çıktıları

oluşmaktadır. Üçüncü aşamada Eşik Ses Yüksekliği eğrisi oluşturulmaktadır. Bu eğri, işitme duyarlılığını simüle etmek için filtre bankası çıktıları kullanmaktadır. Sonrasında ise eşitlenmiş değerler, Stevens'ın Güç Yasasına göre 0.33 kuvveti ile dönüştürülmektedir.

### 3.1.3. Rasta Filtered Perceptual Linear Prediction (Rasta-PLP)

Rasta-PLP, PLP özneliğini doğrusal spektral bozulmalara karşı daha güçlü hale getiren bir metodolojidir [31].

Özneliğin çıkarım aşamalarının ilki, PLP'de olduğu gibi kritik bandın hesaplanması ve logaritmasının alınmasıdır. İkinci olarak, logaritmik kritik bant spektrumunun regresyon hattı kullanılarak 5 ardışık spektral değerler aracılığıyla zamansal türevi hesaplanmaktadır. Üçüncü olarak, eşik değer veya medyan filtreleme ile doğrusal olmayan işleme yapılmaktadır. Dördüncü aşamada Logaritmik kritik bandın türevi, birinci derece sonsuz darbe yanıtı (Infinite Impulse Response) aracılığıyla yeniden entegre edilmektedir. Sistemin başlangıç pozisyonu, efektif pencere boyutunu ayarlamak için ayarlanabilmektedir. Beşinci aşamada, göreceli logaritmik spektrumun ters logaritması alınarak, göreceli işitsel spektrum oluşturulmaktadır.

Son aşamada geleneksel PLP tekniği izlenerek bu spektrumun tüm kutuplu modeli hesaplanmaktadır.

### 3.1.4. Linear Prediction Coding (LPC)

LPC, çoğunlukla ses sinyali işleme ve konuşma işlemede kullanılan, doğrusal prediktif model bilgisi kullanılarak sayısal bir konuşma sinyalinin sıkıştırılmış formdaki spektral zarfını temsil eden özneliktir [32]. Bir sonraki noktayı, geçmiş noktaların doğrusal kombinasyonu olarak tahmin etmeye çalışmaktadır.

$$s[n] = \sum_{k=1}^p a_k s[n-k] + e[n] \quad (3.5)$$

(3.5) LPC denklemine göre göre,  $a_k$ ,  $p$  dereceden doğrusal tahmin katsayısıdır.  $e[n]$  ise tahmin hatasıdır. LPC, en küçük kareler anlamında tahmin hatasını en aza indirgeyerek ileri bir doğrusal öngörücünün katsayılarını belirlemektedir.

### 3.1.5. Zero Crossing Rate (ZCR)

ZCR, sinyalin zamana bağlı olarak işaret değiştirme hızını temsil etmektedir [33]. Bir diğer deyişle, sinyalin X-eksenini geçme miktarını temsil etmektedir. ZCR özniteliğinin denklemi (3.6) ve (3.7)'de sunulmaktadır.

$$\sum |sign[x(m)] - sign[x(m - 1)]| \quad (3.6)$$

Öyle ki

$$\begin{cases} sign[x(n)] = 1, x(n) \geq 0 \\ sign[x(n)] = -1, x < 0 \end{cases} \quad (3.7)$$

### 3.1.6. Sound Pressure Level (SPL)

Akustik basınç seviyesi olarak da tanımlanan bu öznitelik, sesin bir referans değere göre efektif basıncının logaritmik ölçümünü temsil etmektedir [34]. SPL özniteliğinin denklemi (3.8)'de sunulmaktadır:

$$SPL = 10 \log_{10} \left( \frac{p^2}{p_0^2} \right) \quad (3.8)$$

$p$ : sesin Pascal cinsinden basınç seviyesi.  $p_0$ : referans sesin Pascal cinsinden basınç seviyesi

## 3.2. Sınıflandırıcılar

Ses sinyalleriyle oluşturulan öznitelikler kullanarak modeller oluşturmak ve bu modeller yardımıyla sınıflandırma işlemlerini gerçekleştirmek amacıyla çeşitli sınıflandırıcılar kullanılmıştır. Bu sınıflandırıcılar, Destek Vektör Makineleri, Rastgele Ormanlar ve Bootstrap Aggregation (Bagging)'dir. Sınıflandırıcılar hakkında detaylı bilgiler aşağıda verilmektedir:

### 3.2.1. Destek vektör makineleri

Destek Vektör Makineleri, karar verme düzlemlerini belirleme konsepti üzerine kurulmuş bir yol gösterilmiş makine öğrenmesi metodudur. Karar verme düzlemi, farklı sınıflara ait objeleri birbirinden ayırmak için var olan bir düzlemdir. Bu düzlemlere destek vektörleri denmektedir [24].

Bir sınıflandırma çalışmasına, eğitim ve test kümelerinin oluşturulması ile başlanmaktadır. Eğitim kümesinde yer alan her örnek bir hedef değer ya da sınıf etiketi ile işaretlenmektedir ve örnekler çeşitli özellikler taşıyabilmektedir. Destek vektör makinelerinin amacı, eğitim kümesinde yer alan örneklerin bu özelliklerini ve etiketlerini kullanarak bir model oluşturmak; ardından bu modeli kullanarak test kümesinde yer alan örnekleri, sadece ilgili örneklerin özelliklerini kullanarak sınıf etiketlerini tahmin etmeye çalışmaktır.

Verilmiş olan bir eğitim setindeki örnek-sınıf etiketi çiftleri  $(x_i, y_i), i = 1, \dots, l$  öyle ki  $x_i \in R^n$  ve  $y \in \{1, -1\}^l$ , Destek Vektör Makineleri [24;25], (3.9)'da yer alan problemi çözmeye çalışmaktadır.

$$\xi_i \geq 0, i = 1 \dots > N, y_i(W^T \varphi(X_i) + b) \geq 1 - \xi_i \text{ olmak koşuluyla} \quad (3.9)$$

$$\frac{1}{2} W^T W + C \sum_{i=1}^l \xi_i$$

Burada yer alan eğitim kümesi vektörleri  $X_i$ ,  $\varphi$  fonksiyonu aracılığıyla daha yüksek düzlem uzayına taşınmıştır. Destek vektör makineleri, bu yeni ve daha yüksek seviyeli düzlem uzayında en yüksek marjin değerine sahip doğrusal bir ayrıştırıcı bulmaktadır.  $C > 0$  ise ceza parametresidir.  $K(X_i, X_j) \equiv \varphi(X_i)^T \varphi(X_j)$  ise çekirdek fonksiyonu olarak adlandırılmaktadır. Bu çalışma kapsamında da incelenecek olan, 4 temel çekirdek fonksiyonu bulunmaktadır. Bunlar; doğrusal, radyal, polinomal, radyal ve sigmoid çekirdek fonksiyonları olup, bunların detayları ve farkları ilerleyen bölümlerde yer almaktadır.

Tüm makine öğrenmesi yöntemlerinin başarımları, eğitim ve test sınıflarındaki seçilime ve dağılıma bağlıdır. Bu çalışmada yer verilen yöntemlerin başarımlarını eğitim ve test kümelerinin seçiminden bağımsız hale getirebilmek amacıyla, Katlamalı Çapraz Doğrulama (Cross-Validation) kullanılmaktadır. Bu yöntemde, veri kümesi  $k$  eşit parçaya ayrılmakta,  $k-1$  parça eğitim için kullanılmakta ve geriye kalan parça test için kullanılmaktadır. Bu işlem  $k$  kez tekrarlanarak her bir alt kümenin hem test hem de eğitim kümesinde en az bir kez yer alması sağlanmaktadır. Bu işlem,  $k$ -fold ya da  $k$ -katlamalı çapraz doğrulama olarak adlandırılmaktadır. Başarımları hesaplamak için,  $k$  işlemin ortalaması alınmaktadır. *Birini Dışarıda Bırak (Leave-one-out)* yöntemi ise,  $N$  elemanı olan bir veri kümesinde  $n$ -fold katlamalı çapraz doğrulama işlemine eşdeğerdir. Buna göre, kümedeki her eleman için geriye kalan elemanlar ile eğitim aşaması gerçekleştirilmekte olup, son eleman için test işlemi yapılmaktadır. Bu işlem eleman sayısı kadar tekrarlanmakta ve başarımları hesaplayabilmek için  $N$  adet iterasyonun ortalama değeri alınmaktadır.

Bu çalışma kapsamında Destek vektör makinelerinin iki farklı sınıflandırma metodu kullanılmıştır: c-svc ve nu-svc. Takip eden bölümde bu iki sınıflandırma metodunun detayları paylaşılmaktadır. Ardından çekirdek fonksiyonlarının türleri yer almaktadır.

Destek vektör makinelerinin 4 farklı çekirdek özelliği, bu çalışmada karşılaştırma amacıyla kullanılmıştır. Çekirdek fonksiyonlarında kullanılan  $\gamma, r, d$  çekirdek fonksiyonlarının parametreleridir.

**C-destek vektör sınıflandırması:** Bu sınıflandırma metodunda, eğitim aşamasının amacı (3.10)'da yer alan hata fonksiyonunu minimize etmektir:

$$\xi_i \geq 0, i = 1 \dots N, y_i(W^T \phi(X_i) + b) \geq 1 - \xi_i \text{ olmak koşuluyla} \quad (3.10)$$

$$\frac{1}{2} W^T W + C \sum_{i=1}^N \xi_i$$

Burada yer alan  $C$  kapasite sabiti,  $w$  katsayı (coefficient) vektörü,  $\xi_i$  ayrıştırılmayan girdileri ele alan parametredir. İmleç  $i$ ,  $N$  tane eğitim durumu göstermektedir.  $y_i \in \{1, -1\}$  sınıf etiketlerini ve  $X_i$  bağımsız değişkenleri temsil etmektedir.  $\Phi$  çekirdek

fonksiyonu ise girdi verisini bağımsız öznelilik uzayına dönüştürmektedir. Daha yüksek C değerleri, hatanın daha fazla cezalandırılması demektir.

**Nu-destek vektör sınıflandırması:** C-destek vektör sınıflandırmasından farklı olarak nu-destek vektör sınıflandırması metodunun amacı (3.11)'de yer alan hata fonksiyonunu minimize etmektir:

$$\xi_i \geq 0, i = 1 \dots N, \rho \geq 0, y_i(W^T \varphi(X_i) + b) \geq \rho - \xi_i \text{ olmak koşuluyla} \quad (3.11)$$

$$\frac{1}{2}w^T w - \nu\rho + C \sum_{i=1}^N \xi_i$$

**Çekirdek fonksiyonları :** 4 temel çekirdek türüne ait denklemler takip eden bölümde yer almaktadır.

Doğrusal çekirdek denklemi (3.12)'de yer almaktadır.

$$K(X_i, X_j) = X_i^T X_j \quad (3.12)$$

Polinomal çekirdek denklemi (3.13)'de yer almaktadır.

$$\gamma > 0 \text{ olmak üzere, } K(X_i, X_j) = (\gamma X_i^T X_j + r)^d \quad (3.13)$$

Radyal çekirdek denklemi (3.14)'de yer almaktadır.

$$\gamma > 0 \text{ olmak üzere, } K(X_i, X_j) = \exp(-\gamma |X_i - X_j|^2) \quad (3.14)$$

Sigmoid çekirdek denklemi (3.15)'te yer almaktadır.

$$K(X_i, X_j) = \tanh(\gamma X_i^T X_j + r) \quad (3.15)$$

### 3.2.2. Torbalama

Bootstrap Aggregation kelimelerinin kısaltması olarak kullanılan Bagging yöntemi, Torbalama olarak da adlandırılmaktadır. Sınırlı sayıda örneklemin bulunduğu sınıflandırma problemlerinde eğitim kümesini güçlendirmek amacıyla sıklıkla

kullanılan bir topluluk öğrenme metodu olan Torbalama, birden fazla makine öğrenmesi modelinin (örneğin birden fazla destek vektör makineleri) bir araya gelerek topluluğun ortak karar vermesi sonucunda başarımı artırmayı amaçlamakta olup, basit ve güçlü bir yöntemdir [35].  $N$  tane örneklemden tekrarlı  $M$  tane alt örnek oluşturarak makine öğrenme metotlarında bu alt örnek kümelerini kullanmakta, bunu yaparken de test kümesini sabit tutmaktadır. Nihai karar oluşturulurken de topluluğun ortak kararı baz alınmaktadır.

İlk olarak eğitim kümesinden örnekleme işlemi yapılır. Her bir örnek eğitim kümesi kullanılarak modeller oluşturulur. Ardından her bir örnek eğitim kümesi kullanılarak, nihai test kümesi için tahminler oluşturulur. Tüm tahminler bir araya getirilerek oylama yapılır ve en yüksek oya sahip öge, nihai tahmin olarak kabul edilir. Topluluğun karar vermesi esası, bu metodu güçlü kılan noktadır.

### **3.2.3. Rastgele ormanlar**

Topluluk öğrenme yöntemlerinden Torbalama yönteminin gelişmiş hali olarak kabul edilen Rastgele Ormanlar yöntemi, ağaç tipi bir sınıflandırıcıdır. Breiman'ın önermiş olduğu Rastgele Ormanlar yönteminde, Torbalama ve Rastgele Özellik seçimi birlikte ele alınmaktadır [36]. Torbalama yönteminde olduğu gibi tekrarlı eğitim kümeleri meydana getirilmekte ve bu alt eğitim setlerinden ağaçlar oluşturulmaktadır. Oluşturulan ağaçlar budanmamaktadır.

Diğer makine öğrenmesi yöntemleri gibi Rastgele Ormanlar da eğitim ve test aşamalarından oluşmaktadır. Rastgele Ormanlar metodunun sözde kodu aşağıdaki gibidir:

Önkoşul: Eğitim kümesi  $S: (x_1, y_1), \dots, (x_n, y_n)$ , özellikler  $F$  ve ağaç sayısı  $B$  olmak üzere

```
1    function RastgeleOrmanlar(S,F)
2         $H \leftarrow \emptyset$ 
3        for  $i \in 1, \dots, B$ 
4             $S^{(i)} \leftarrow A$     S'den torbalama örneği
5             $h_i \leftarrow \text{RastgeleAğaçÖğrenmesi}(S^{(i)}, F)$ 
6             $H \leftarrow H \cup \{h_i\}$ 
7        end
8        return H
9    end function
10   function RastgeleAğaçÖğrenmesi (S,F)
11       Her düğümde:
12            $f \leftarrow F'$  e ait çok küçük bir alt küme
13            $f'$  e ait en iyi özelliğe göre ayır
14       return öğrenilmiş ağaç
15   end function
```



#### 4. SİLAH SESLERİNİN SINIFLANDIRILMASI

Deneyler Matlab<sup>1</sup> platformunda gerçekleştirilmiştir. Deneyler 8 çekirdekli, 2.4 GHz frekansında çalışan 64-bitlik i7 işlemciye ve 8 GB rastgele erişimli belleğe sahip bir bilgisayarda gerçekleştirilmiştir.

Silah seslerinin sınıflandırılması konusu ile ilgili olarak daha önce gerçekleştirilen çalışmalara dair bilgiler, İlgili Çalışmalar başlığında sunulmuştur. Maher, silah atışı ve atışın gerçekleştiği konum ile ilgili modelleme çalışması gerçekleştirmiştir [10]. Martis, destek vektör makineleri ile sınıflandırma problemine odaklanmıştır [11]. Navratil, silah seslerini sınıflandırma konusunda yapay sinir ağlarından faydalanmıştır [12]. Djeddou, GMM ve hiyerarşik sınıflandırma kullanmıştır [13]. Kiktova, Saklı Markov Modelinden faydalanmıştır [14].

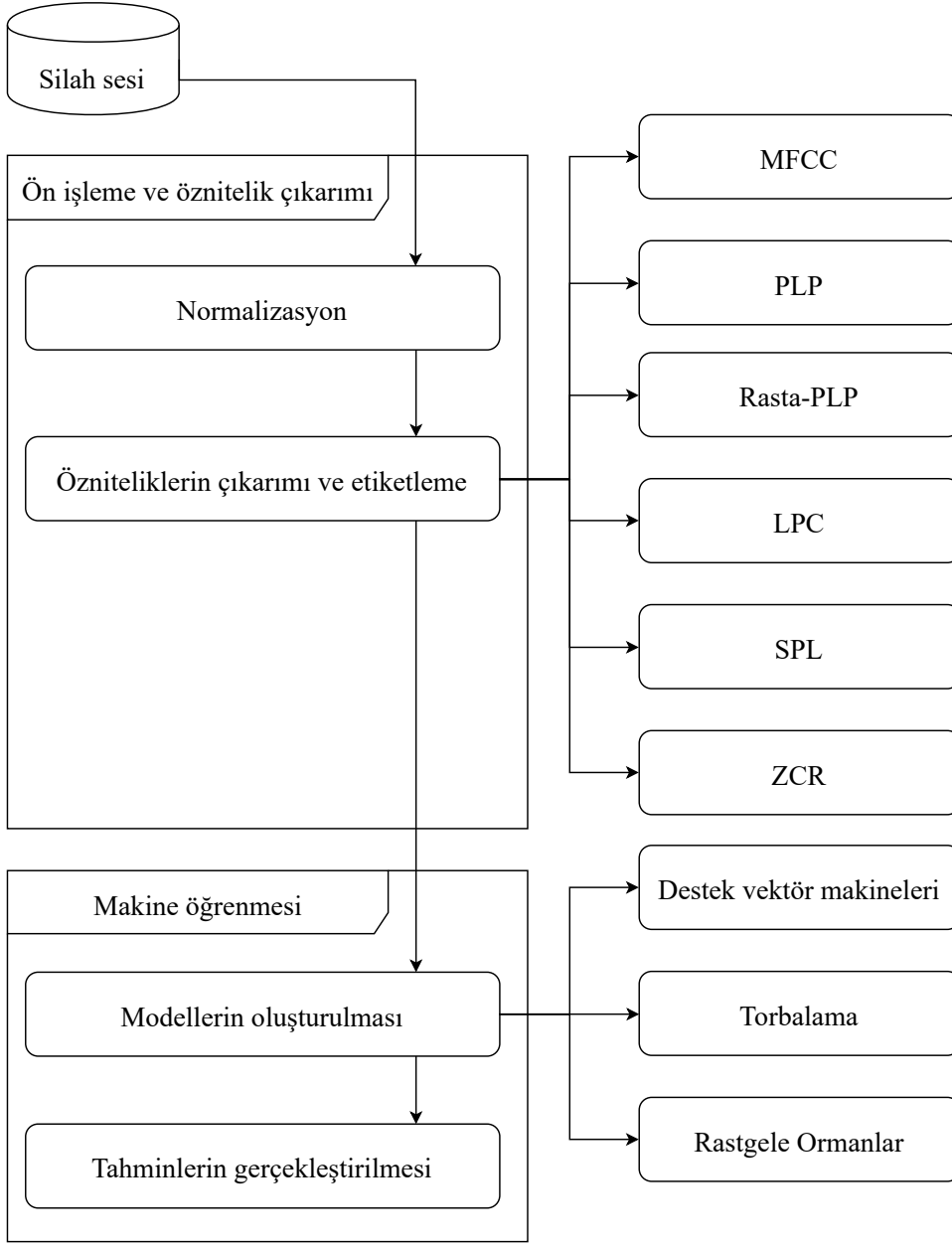
Tez çalışmasının bu bölümünde, Maher'in [2; 9; 10] veri kümesi kullanılarak silah seslerinin sınıflandırılması gerçekleştirilmiştir. Öznitelikler bölümünde bahsedilen yöntemler kullanılarak ses sinyallerinden elde edilen öznitelikler, Sınıflandırıcılar bölümünde bahsedilen sınıflandırıcılar yardımıyla ayrıştırılmaya çalışılmaktadır.

İlgili çalışmada Maher kayıtları Çizelge 4.1'de yer aldığı şekilde etiketlemiş ve çalışmamıza da aynı etiketler ile konu olmuştur. Çalışmanın bu aşamasında takip edilen adımlar, Şekil 4.1'de sunulmaktadır:

---

<sup>1</sup> Matlab, Mathworks firmasının bir ürünüdür.

<http://www.mathworks.com/products/matlab/>



Şekil 4.1 Silah seslerinin sınıflandırılması işleminin gerçekleştirilme aşamaları

Öncelikli olarak normalizasyon gerçekleştirilmektedir. Bu bağlamda sinyal birden fazla kanaldan oluşmaktaysa tek bir kanala indirgenmektedir. Bu işlem için, sinyalin tüm kanallarının matematiksel ortalaması alınmaktadır.

Bu işlemin ardından, her bir ses sinyali için öznelik çıkarımı yöntemleri kullanılarak öznelikler çıkarılmakta ve ses sinyalinin ait olduğu sınıfa ait etiket ile

ilişkilendirilmektedir. Yukarıda bahsedildiği üzere 5 farklı sınıfa ait ses sinyali bulunmaktadır. Her bir ses sinyali için MFCC, PLP, Rasta-PLP, LPC, ZCR, SPL öznitelikleri çıkarılmaktadır.

Her bir ses sinyaline ait öznitelikler etiketlendikten sonra, sınıflandırma aşamasına geçilmektedir. Bu aşamada, tüm ses sinyallerine ait aynı türdeki öznitelikler ayrı ayrı ele alınmaktadır. Modeller oluşturulduktan sonra sınıflandırma aşamasına geçilmektedir.

Tez çalışmasının bu bölümünde, destek vektör makineleri ile çok sınıflı sınıflandırma problemi ele alındığından birebir (OvO) sınıflandırma işlemi gerçekleştirilmektedir. Topluluk öğrenme yöntemlerinden olan Torbalama yönteminde ise destek vektör makinelerinden faydalanılmıştır. Son olarak da Rastgele Ormanlar yöntemi uygulanmıştır.

Silah seslerinin sınıflandırılması için tasarlanmış olduğumuz makine öğrenmesi metodunun başarımını ölçmek amacıyla öncelikli olarak farklı destek vektör makineleri metotlarının sınıflandırma başarımına etkisi incelenmiştir. Bu çalışma kapsamında Silah seslerinin sınıflandırılması işlemi gerçekleştirilebilmek amacıyla Destek vektör makineleri, Rastgele Ormanlar ve Torbalama yöntemleri kullanılmıştır.

Destek vektör makinelerinin tipleri, sınıflandırma işleminin metoduna göre isimlendirilmektedir. Bu çalışma kapsamında Silah seslerinin sınıflandırılması işlemi için c-destek vektör sınıflandırması ve nu-destek vektör sınıflandırması metotları kullanılmıştır. Karşılaştırma işlemlerinde eğitim ve test kümeleri 2-fold, 5-fold, 10-fold ve Leave-one-out (LOO) olarak sunulmuştur.

#### **4.1. Veri kümesi**

Tez çalışmasının bu bölümünde, Maher'in [2; 9; 10] veri kümesi kullanılmıştır. Maher, kayıtları gerçekleştirirken çok yönlü iki mikrofon kullanmıştır ve bu mikrofonlar 25 cm ara ile yerleştirmiştir. Kayıtlar 2 ses kanalına sahip olup, 48 kHz olarak örneklenmiş

ve bu örneklerin her biri 16-bit ile temsil edilmiştir. Çizelge 4.1, deneyde yer alan silah seslerine ait etiket bilgilerini sunmaktadır.

Çizelge 4.1 Silah seslerinin sınıflandırılması veri kümesi dağılımı

1-6. Atışlar	308 Winchester, NATO FNM 83-23
7-12. Atışlar	223 Remington, Ruger mini 14, NATO 5.56x45mm Ball M1A4
13-20. Atışlar	9mm 115 grain full metal jacket (Askeri olmayan) "Sellier & Bellot" (Çek Cumhuriyeti)
21-29. Atışlar	HK USP 230 grain full metal jacket, 45ACP, Remington
30-37. Atışlar	HK USP compact, 40 Smith and Wesson, Federal Hydroshock

#### 4.2. Destek vektör makinesi deneyleri

Destek vektör makinesi ile elde edilen deneysel sonuçlar, takip eden bölümde sunulmaktadır. Bu başlıkta öncelikle c-destek vektör sınıflandırması ve farklı çekirdek fonksiyonları ile elde edilen başarımlar yer almaktadır. Ardından nu-destek vektör sınıflandırması sonuçları ve farklı çekirdek fonksiyonları ile elde edilen başarımlar yer almaktadır.

Çalışma kapsamında bulunan tüm öznitelik çıkarımı yöntemlerinin sonucunda elde edilen özniteliklerin c-destek vektör makineleri sınıflandırma metodu ve doğrusal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 4.2'deki gibidir.

Çizelge 4.2 Silah seslerinin sınıflandırılması, c-destek vektör makinesi doğrusal çekirdek sınıflandırması sonuçları

Öznitelik Adı	2 fold (Sınıflandırma başarımı %)	5 fold (Sınıflandırma başarımı %)	10 fold (Sınıflandırma başarımı %)	LOO (Sınıflandırma başarımı %)
MFCC	31.25	31.13	31.27	31.21
PLP	28.24	20.93	28.31	28.28
Rasta-PLP	22.39	22.39	22.39	22.39
LPC	21.87	21.87	20.83	19.09
ZCR	33.33	33.33	30.56	30.56
SPL	22.22	25	27.78	25

Çizelge 4.2'ye göre, destek vektör makinelerinin c-destek vektör makinesi metodu doğrusal çekirdek fonksiyonu kullanıldığında elde edilen en yüksek başarımlar ZCR öznitelik çıkarımı yöntemi ile, 2-fold ve 5-fold CV aracılığıyla ortaya çıkmaktadır. MFCC özniteliğinin başarımları da %2 dolaylarında düşük olmakla birlikte, ZCR özniteliğinin değerlerine oldukça yakındır.

Çalışma kapsamında bulunan tüm öznitelik çıkarımı yöntemlerinin sonucunda elde edilen özniteliklerin c-destek vektör makineleri sınıflandırma metodu ve polinomal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar Çizelge 4.3'deki gibidir.

Çizelge 4.3 Silah seslerinin sınıflandırılması, c-destek vektör makinesi polinomal çekirdek sınıflandırması sonuçları

Öznitelik Adı	2 fold (Sınıflandırma başarımı %)	5 fold (Sınıflandırma başarımı %)	10 fold (Sınıflandırma başarımı %)	LOO (Sınıflandırma başarımı %)
MFCC	29.68	29.68	29.7	29.35
PLP	22.39	22.39	22.39	22.39
Rasta-PLP	22.39	22.39	22.39	22.39
LPC	21.7	21.87	22.04	22.22
ZCR	30.55	36.11	36.11	36.11
SPL	33.33	33.33	33.33	22.22

Çizelge 4.3'e göre, destek vektör makinelerinin c-destek vektör makinesi metodu polinomal çekirdek fonksiyonu kullanıldığında elde edilen en yüksek başarımlar ZCR öznitelik çıkarımı yöntemi ile, 5-fold, 10-fold ve LOO yöntemi aracılığıyla ortaya çıkmaktadır.

Çalışma kapsamında bulunan tüm öznitelik çıkarımı yöntemlerinin sonucunda elde edilen özniteliklerin c-destek vektör makineleri sınıflandırma metodu ve radyal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 4.4'deki gibidir.

Çizelge 4.4 Silah seslerinin sınıflandırılması, c-destek vektör makinesi radyal çekirdek sınıflandırması sonuçları

Öznitelik Adı	2 fold (Sınıflandırma başarımı %)	5 fold (Sınıflandırma başarımı %)	10 fold (Sınıflandırma başarımı %)	LOO (Sınıflandırma başarımı %)
MFCC	31.47	31.4	31.66	31.71
PLP	28.27	28.45	28.35	28.34
Rasta-PLP	22.39	22.39	22.39	22.39
LPC	21.52	21.18	20.13	20.83
ZCR	33.33	33.33	30.56	30.56
SPL	36.11	36.11	36.11	33.33

Çizelge 4.4'e göre, destek vektör makinelerinin c-destek vektör makinesi metodu radyal çekirdek fonksiyonu kullanıldığında elde edilen en yüksek başarımlar SPL öznitelik çıkarımı yöntemi ile, 2-fold ve 10-fold CV yöntemi aracılığıyla ortaya çıkmaktadır. Aynı Destek Vektör Makinesi türündeki diğer çekirdek fonksiyonlarından farklı olarak SPL öznitelik çıkarım yöntemi ile daha yüksek başarımlar elde edilmiş olsa da, ZCR özneliğinin %3 daha düşük başarımlar ile geriye kalan öznitelik çıkarım yöntemlerinden daha başarılı olmuştur.

Çalışma kapsamında bulunan tüm öznitelik çıkarım yöntemlerinin sonucunda elde edilen özneliklerin c-destek vektör makineleri sınıflandırma metodu ve sigmoid çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 4.5'deki gibidir.

Çizelge 4.5 Silah seslerinin sınıflandırılması, c-destek vektör makinesi sigmoid çekirdek sınıflandırması sonuçları

Öznitelik Adı	2 fold (Sınıflandırma başarımı %)	5 fold (Sınıflandırma başarımı %)	10 fold (Sınıflandırma başarımı %)	LOO (Sınıflandırma başarımı %)
MFCC	13.95	14.92	13.41	13.38
PLP	28.25	28.29	28.31	28.27
Rasta-PLP	22.39	22.39	22.39	22.39
LPC	23.43	22.22	22.56	21
ZCR	33.33	33.33	30.56	30.56
SPL	22.22	22.22	22.22	22.22

Çizelge 4.5'e göre, destek vektör makinelerinin c-destek vektör makinesi metodu sigmoid çekirdek fonksiyonu kullanıldığında elde edilen en yüksek başarımlar ZCR öznitelik çıkarımı yöntemi ile, 2-fold ve 5-fold CV yöntemi aracılığıyla ortaya çıkmaktadır. Bir diğer dikkat çekici nokta ise SPL özniteliğinin c-destek vektör makinelerinin, sigmoid çekirdek fonksiyonu ile birlikte kullanıldığı zaman, CV değerine bağlı olarak başarımların değişmemiş olmasıdır.

Çalışma kapsamında bulunan tüm öznitelik çıkarımı yöntemlerinin sonucunda elde edilen özniteliklerin nu-destek vektör makineleri sınıflandırma metodu ve doğrusal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 4.6'daki gibidir.



Çizelge 4.6: Silah seslerinin sınıflandırılması, nu-destek vektör makinesi doğrusal çekirdek sınıflandırması sonuçları

Öznitelik Adı	2 fold (Sınıflandırma başarımı %)	5 fold (Sınıflandırma başarımı %)	10 fold (Sınıflandırma başarımı %)	LOO (Sınıflandırma başarımı %)
MFCC	23.01	25.39	23.58	24.41
PLP	22.47	21.96	22.95	22.68
Rasta-PLP	21.93	19.91	18.77	18.35
LPC	19.79	22.74	20.65	19.09
ZCR	36.11	33.33	33.33	36.11
SPL	22.22	25	27.78	25

Çizelge 4.6'ya göre, destek vektör makinelerinin nu-destek vektör makinesi metodu doğrusal çekirdek fonksiyonu kullanıldığında elde edilen en yüksek başarımlar ZCR öznitelik çıkarımı yöntemi ile, 2-fold ve LOO CV yöntemi aracılığıyla ortaya çıkmaktadır. Yine ZCR yönteminin, 10-fold hariç tüm CV değerleri için başarımlarının diğer öznitelik çıkarımı yöntemlerinden yüksek olduğunu gözlemlemekteyiz.

Çalışma kapsamında bulunan tüm öznitelik çıkarımı yöntemlerinin sonucunda elde edilen özniteliklerin nu-destek vektör makineleri sınıflandırma metodu ve polinomal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar Çizelge 4.7'deki gibidir.

Çizelge 4.7 Silah seslerinin sınıflandırılması, nu-destek vektör makinesi polinomal çekirdek sınıflandırması sonuçları

Öznitelik Adı	2 fold (Sınıflandırma başarımı %)	5 fold (Sınıflandırma başarımı %)	10 fold (Sınıflandırma başarımı %)	LOO (Sınıflandırma başarımı %)
MFCC	22.57	22.69	22.11	23.55
PLP	23.85	24	23.24	21.12
Rasta-PLP	20.64	18.55	18.09	19.03
LPC	23.43	22.22	18.75	19.61
ZCR	30.55	44.44	36.11	36.11
SPL	22.22	22.22	22.22	25

Çizelge 4.7'ye göre, destek vektör makinelerinin c-destek vektör makinesi metodu polinomal çekirdek fonksiyonu kullanıldığında elde edilen en yüksek başarımlar ZCR öznitelik çıkarımı yöntemi ve 2 fold CV aracılığıyla ortaya çıkmaktadır. Bu değer aynı zamanda bundan önce gerçekleştirilen deneyler arasındaki en yüksek değerdir.

Çalışma kapsamında bulunan tüm öznitelik çıkarımı yöntemlerinin sonucunda elde edilen özniteliklerin nu-destek vektör makineleri sınıflandırma metodu ve radyal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 4.8'deki gibidir.

Çizelge 4.8 Silah seslerinin sınıflandırılması, nu-destek vektör makinesi radyal çekirdek sınıflandırması sonuçları

Öznitelik Adı	2 fold (Sınıflandırma başarımı %)	5 fold (Sınıflandırma başarımı %)	10 fold (Sınıflandırma başarımı %)	LOO (Sınıflandırma başarımı %)
MFCC	21.34	22.7	23.66	23.25
PLP	26.14	25.02	24.92	24.19
Rasta-PLP	21.32	18.15	18.34	18.7
LPC	19.09	19.79	22.56	18.22
ZCR	30.55	36.11	44.44	38.88
SPL	38.88	33.33	36.11	36.11

Çizelge 4.8'e göre, destek vektör makinelerinin nu-destek vektör makinesi metodu radyal çekirdek fonksiyonu kullanıldığında elde edilen en yüksek başarımlar ZCR öznitelik çıkarımı yöntemi ve 10 fold CV aracılığıyla ortaya çıkmaktadır. Bu değer, nu-svc ve polinomal çekirdek bir arada kullanıldığında ZCR özniteliği ile elde edilen en yüksek başarımla eşittir. Bunun yanı sıra, SPL özniteliği ile elde edilen başarımlar ZCR özniteliğine oldukça yakındır.

Çalışma kapsamında bulunan tüm öznitelik çıkarımı yöntemlerinin sonucunda elde edilen özniteliklerin nu-destek vektör makineleri sınıflandırma metodu ve sigmoid çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 4.9'daki gibidir.

Çizelge 4.9 Silah seslerinin sınıflandırılması, nu-destek vektör makinesi sigmoid çekirdek sınıflandırması sonuçları

Öznitelik Adı	2 fold (Sınıflandırma başarımı %)	5 fold (Sınıflandırma başarımı %)	10 fold (Sınıflandırma başarımı %)	LOO (Sınıflandırma başarımı %)
MFCC	26.23	26.17	26.19	26.07
PLP	25.53	26.22	26.44	25.7
Rasta-PLP	21.86	19.47	19.85	20.03
LPC	19.44	17.36	16.32	16.32
ZCR	30.55	36.11	44.44	44.44
SPL	22.22	22.22	22.22	19.44

Çizelge 4.9'a göre, destek vektör makinelerinin c-destek vektör makinesi metodu sigmoid çekirdek fonksiyonu kullanıldığında elde edilen en yüksek başarımlar ZCR öznitelik çıkarımı yöntemi 5 ve 10 fold CV aracılığıyla ortaya çıkmaktadır. Bu değer, nu-svc ve polinomal çekirdek bir arada kullanıldığında ZCR öznitelik ile elde edilen en yüksek başarımla eşittir.

### 4.3. Torbalama Deneyleri

Çalışmanın bu bölümünde, mevcut veri kümesi üzerinde Torbalama metodu uygulanmış ve elde edilen sonuçlar yorumlanmıştır. Takip eden bölümde farklı destek vektör metodlarına göre ayrıştırılmış olup, 10 ve 20 torba parametreleri için sonuçlar paylaşılmıştır. Bu kısımda, ZCR ve SPL özniteliklerine skalar öznitelik olmaları sebebiyle yer verilmemiştir.

Öncelikle Torbalama yöntemi ve c-destek vektör makineleri kullanılarak elde edilen sonuçlar yer almaktadır. Ardından nu-destek vektör makineleri kullanılarak elde edilen sonuçlar yer almaktadır. Sonuçlar, kullanılan torba sayısına göre 10 torba ve 20 torba olarak iki grupta ele alınmıştır.

C-destek vektör makineleri ve 10 torba kullanarak elde edilen sonuçlar takip eden çizelgede sunulmaktadır. Mevcut veri kümesi üzerinde eğitim ve test sınıfları belirlenmiş olmadığından, veri kümesi belirli oranlarda bölünerek eğitim ve test amacıyla ayrılmıştır. Bu oranlar sırasıyla eğitim ve test kümeleri olmak üzere, 50-50, 70-30 ve 90-10 olarak seçilmiştir.

Çizelge 4.10 Silah seslerinin sınıflandırılması, Torbalama yöntemi ile elde edilen sonuçlar, 10 torba c-destek vektör makinesi sonuçları

Eğitim ve Test Kümeleri oranı	50-50	70-30	90-10
MFCC (Sınıflandırma başarımı %)	31.09	32.36	32.35
PLP (Sınıflandırma başarımı %)	28.03	28.07	27.97
Rasta-PLP (Sınıflandırma başarımı %)	22.38	22.38	22.36
LPC (Sınıflandırma başarımı %)	24.65	25.22	26.31

Çizelge 4.10'da Torbalama Yöntemi, 10 Torba parametresi ile uygulanmış destek vektör makineleri, c-destek vektör makinesi sınıflandırması sonuçları yer almaktadır. C-destek vektör makinesi sınıflandırması ile en iyi sonuçların elde edildiği radyal çekirdek fonksiyonu kullanılmıştır. Çizelgenin ilk satırı eğitim ve test kümelerinin dağılımını sırasıyla ve oransal olarak göstermektedir. Buna göre, LPC öznitelik çıkarım yöntemi hariç, geriye kalan yöntemler için torbalama metodunda eğitim ve test kümelerinin dağılım oranının başarımı dominant bir şekilde etkilemediği görülmüştür. Bunun yanı sıra, destek vektör makinelerinin aynı sınıflandırma tipi ve aynı çekirdek metodunda olduğu gibi MFCC diğer öznitelik çıkarımı yöntemlerinden daha yüksek sonuçlar elde etmemizi sağlamıştır. C-destek vektör makineleri ve 20 torba kullanarak elde edilen sonuçlar takip eden çizelgede sunulmaktadır.

Çizelge 4.11 Silah seslerinin sınıflandırılması, Torbalama yöntemi ile elde edilen sonuçlar, 20 torba c-destek vektör makinesi sonuçları

Eğitim ve Test Kümeleri oranı	50-50	70-30	90-10
MFCC (Sınıflandırma başarımı %)	31.12	32.36	32.55
PLP (Sınıflandırma başarımı %)	28.12	28.14	28.37
Rasta-PLP (Sınıflandırma başarımı %)	22.37	22.38	22.36
LPC (Sınıflandırma başarımı %)	23.61	21.51	24.56

Çizelge 4.11'de Torbalama yöntemi, 20 Torba parametresi ile uygulanmış destek vektör makineleri, c-destek vektör makinesi sınıflandırması sonuçları yer almaktadır. C-destek vektör makinesi sınıflandırması ile en iyi sonuçların elde edildiği radyal çekirdek fonksiyonu kullanılmıştır. Çizelgenin ilk satırı eğitim ve test kümelerinin dağılımını sırasıyla ve oransal olarak göstermektedir. Bunun yanı sıra, destek vektör makinelerinin aynı sınıflandırma tipi ve aynı çekirdek metodunda olduğu gibi MFCC diğer öznitelik çıkarımı yöntemlerinden daha yüksek sonuçlar elde etmemizi sağlamıştır.

Nu-destek vektör makineleri ve 10 torba kullanarak elde edilen sonuçlar takip eden çizelgede sunulmaktadır.

Çizelge 4.12 Silah seslerinin sınıflandırılması, Torbalama yöntemi ile elde edilen sonuçlar, 10 torba nu-destek vektör makinesi sonuçları

Eğitim ve Test Kümeleri oranı	50-50	70-30	90-10
MFCC (Sınıflandırma başarımı %)	25.01	24.65	24.41
PLP (Sınıflandırma başarımı %)	23.75	23.53	22.69
Rasta-PLP (Sınıflandırma başarımı %)	17.86	18.22	18.49
LPC (Sınıflandırma başarımı %)	22.91	23.75	29.82

Çizelge 4.12'de Torbalama Yöntemi, 10 Torba olarak uygulanmış destek vektör makineleri, nu-destek vektör makinesi sınıflandırması sonuçları yer almaktadır. Nu-

destek vektör makinesi sınıflandırması ile en iyi sonuçların elde edildiği radyal çekirdek fonksiyonu kullanılmıştır. Çizelgenin ilk satırı eğitim ve test kümelerinin dağılımını sırasıyla ve oransal olarak göstermektedir. Buna göre, LPC öznitelik çıkarım yöntemi hariç geriye kalan yöntemler için Torbalama metodunda eğitim ve test kümelerinin dağılım oranının başarıyı dominant bir şekilde etkilemediği görülmüştür. LPC özniteliği, eğitim ve test kümelerinin 90-10 dağılım gösterdiği deneylerde daha önceki deneylere oranla önemli bir yükseliş göstererek bu bölümde en yüksek başarıyı elde eden öznitelik olmuştur.

Nu-destek vektör makineleri ve 20 torba kullanarak elde edilen sonuçlar takip eden çizelgede sunulmaktadır.

Çizelge 4.13 Silah seslerinin sınıflandırılması, Torbalama yöntemi ile elde edilen sonuçlar, 20 torba nu-destek vektör makinesi sonuçları

Eğitim ve Test Kümeleri oranı	50-50	70-30	90-10
MFCC (Sınıflandırma başarımları %)	25.35	24.67	26.01
PLP (Sınıflandırma başarımları %)	22.04	26.89	27.83
Rasta-PLP (Sınıflandırma başarımları %)	18.95	17.22	18.02
LPC (Sınıflandırma başarımları %)	22.91	22.09	24.56

Çizelge 4.13'de Torbalama yöntemi, 20 torba olarak uygulanmış destek vektör makineleri, nu-destek vektör makinesi sınıflandırması sonuçları yer almaktadır. nu-destek vektör makinesi sınıflandırması ile en iyi sonuçların elde edildiği radyal çekirdek fonksiyonu kullanılmıştır. Çizelgenin ilk satırı eğitim ve test kümelerinin dağılımını sırasıyla ve oransal olarak göstermektedir. Torbalama yöntemi ile gerçekleştirilen deneyler içerisinde PLP özniteliğini ilk kez MFCC özniteliğinden baskın gelerek, daha yüksek başarımları elde etmiştir.

#### 4.4. Rastgele ormanlar

Çalışmanın bu bölümünde mevcut veri kümesi üzerinde Rastgele Ormanlar metodu uygulanmış, elde edilen sonuçlar yorumlanmıştır. Bu bölümde ağaç sayısına göre ayrıştırılmış bir şekilde sonuçlar paylaşılmıştır.

Bu bölümde 10 ağaç parametresi ile elde edilen Rastgele Ormanlar metodunun sınıflandırma sonuçları paylaşılmıştır.

Çizelge 4.14 Silah seslerinin sınıflandırılması, Rastgele Ormanlar yöntemi ile elde edilen sonuçlar, 10 ağaç sonuçları

Öznitelik Adı	50-50	70-30	90-10
MFCC	24.24	24.65	24.81
PLP	23.98	24.29	25.03
Rasta-PLP	21.36	21.51	21.42
LPC	20.13	25.58	28.07
ZCR	44.44	50	66.66
SPL	33.33	40	33.33

Çizelge 4.14'te Rastgele Ormanlar yöntemi ve 10 ağaç parametresi ile elde edilen sonuçlar yer almaktadır. Çizelgenin ilk satırı eğitim ve test kümelerinin dağılımını sırasıyla ve oransal olarak göstermektedir. MFCC, PLP ve Rasta-PLP öznitelikleri bu deneyde eğitim ve test kümelerinin dağılım oranlarına bağlı olarak dikkate değer değişiklikler göstermese de; LPC, ZCR ve SPL öznitelikleri eğitim kümesinde yer alan örnek sayısının artışına bağlı olarak artma eğilimi göstermiştir. Bunun yanı sıra ZCR özneliği, tüm eğitim-test kümesi dağılımlarında en yüksek sınıflandırma başarımları gösteren öznelik olmuştur.

Bu bölümde 20 ağaç parametresi ile elde edilen Rastgele Ormanlar metodunun sınıflandırma sonuçları paylaşılmıştır.



Çizelge 4.15 Silah seslerinin sınıflandırılması, Rastgele Ormanlar yöntemi ile elde edilen sonuçlar, 20 ağaç sonuçları

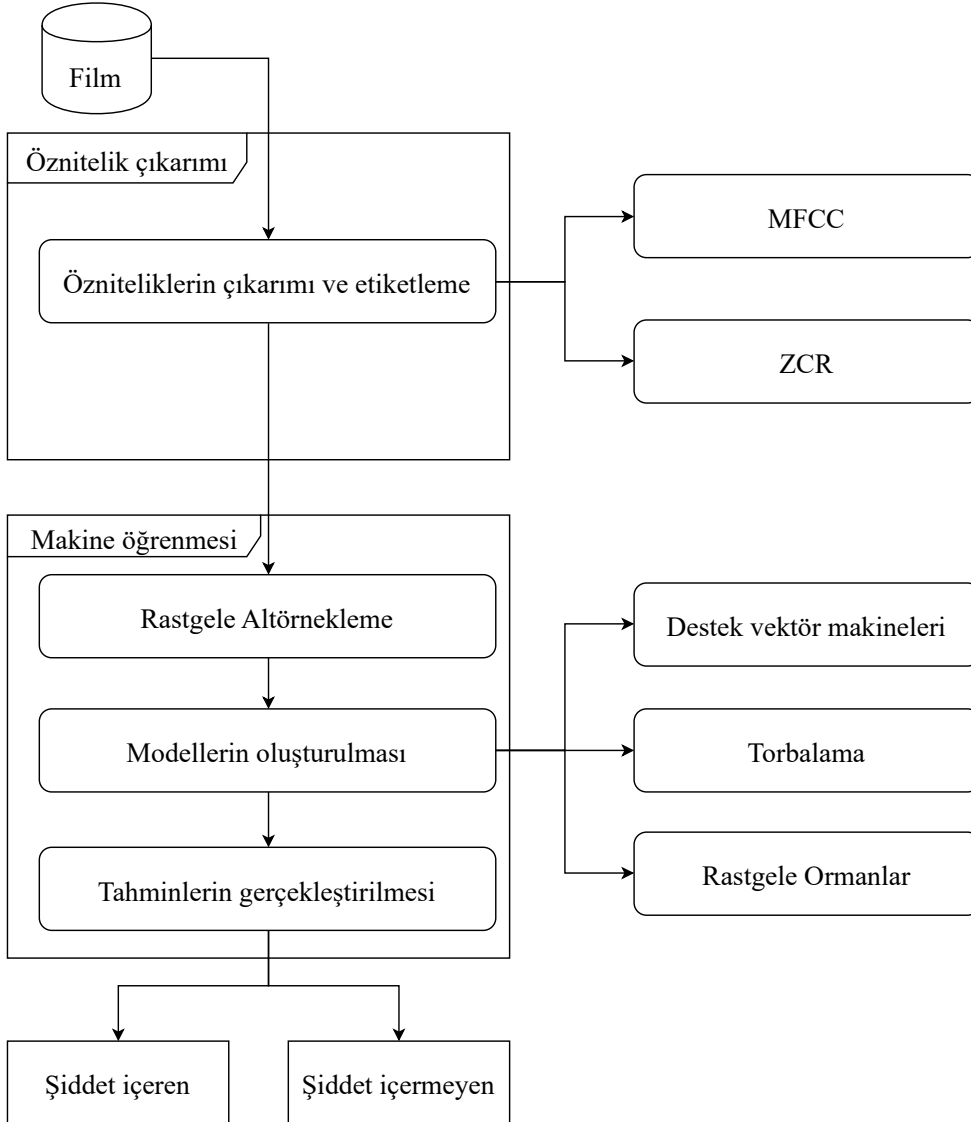
Öznitelik Adı	50-50	70-30	90-10
MFCC	24.77	24.87	24.28
PLP	24.27	24.8	24.16
Rasta-PLP	20.99	21.49	21.96
LPC	20.13	22.09	28.07
ZCR	44.44	50	66.66
SPL	33.33	50	66.66

Çizelge 4.15'te Rastgele Ormanlar yöntemi ve 20 ağaç parametresi ile elde edilen sonuçlar yer almaktadır. Çizelgenin ilk satırı eğitim ve test kümelerinin dağılımını sırasıyla ve oransal olarak göstermektedir. MFCC, PLP ve Rasta-PLP öznitelikleri bu deneyde eğitim ve test kümelerinin dağılım oranlarına bağlı olarak dikkate değer değişiklikler göstermese de; LPC, ZCR ve SPL öznitelikleri eğitim kümesinde yer alan örnek sayısının artışına bağlı olarak artma eğilimi göstermiştir. Bunun yanı sıra ZCR ve SPL öznitelikleri, 50-50 ve 90-10 eğitim-test kümesi dağılımlarında en yüksek sınıflandırma başarımı elde eden öznitelikler olmuşlardır.

## 5. ŞİDDET SAHNESİ SINIFLANDIRMASI

Çalışmanın bu bölümünde, filmler içerisindeki sahnelerin şiddet içerip içermemesine göre sınıflandırılmaya çalışılmıştır. Bu amaçla, filmler içerisindeki sahnelerin ses açısından öznitelikleri çıkarılmış, bu öznitelikler Sınıflandırıcı bölümünde bahsedilen sınıflandırıcılar yardımıyla ayrıştırılmaya çalışılmıştır.

Çalışmanın bu aşamasında takip edilen adımlar Şekil 5.1'de verilmektedir:



Şekil 5.1 Şiddet sahnesi sınıflandırma işleminin aşamaları

Çalışmanın bu bölümünde, ana akım filmler kullanılmış ve bu filmlerde yer alan şiddet içerikli sahneler sınıflandırılmaya çalışılmıştır. Şiddet tanımı “Fiziksel şiddet içerdiği için 8 yaşında bir çocuğun izlemesine izin verilmeyecek olan” şeklinde sübjektif olarak yapılmıştır. Objektif tanım ise “İnsan yaralanması ya da acı ile sonuçlanan fiziksel şiddet ya da kaza” olup, işitsel olarak filmlerin içerisindeki sahnelerde silah ateşlenmesi, patlama ve çığlık içeriklerden en az bir tanesi mevcutsa o sahnenin şiddet içerdiği varsayılmıştır [26].

Bu veri kümesinde Léon, Reservoir Dogs, Armageddon, I am Legend, Saving Private Ryan, Eragon, Harry Potter and the Order of the Phoenix, Billy Elliot, Pirates of the Caribbean – the Curse of the Black Pearl, The Sixth Sense, The Wicker Man, Midnight Express, Kill Bill 1, The Wizard of Oz, The Bourne Identity, Dead Poets Society, Independence Day, Fight Club, Fantastic Four 1, Fargo, Forrest Gump, Legally Blond, Pulp Fiction, The God Father 1, The Pianist, 8 Mile, Braveheart, Desperado, Ghost in the Shell, Jumanji, Terminator 2, V for Vendetta filmleri yer almaktadır. Bu filmler farklı türlerde olup, her biri farklı düzeyde şiddet içermektedir ve bu kapsamda toplam 24 film bulunmaktadır. Bu filmlerin 17 tanesi eğitim için, 7 tanesi ise test için tanımlanmıştır.

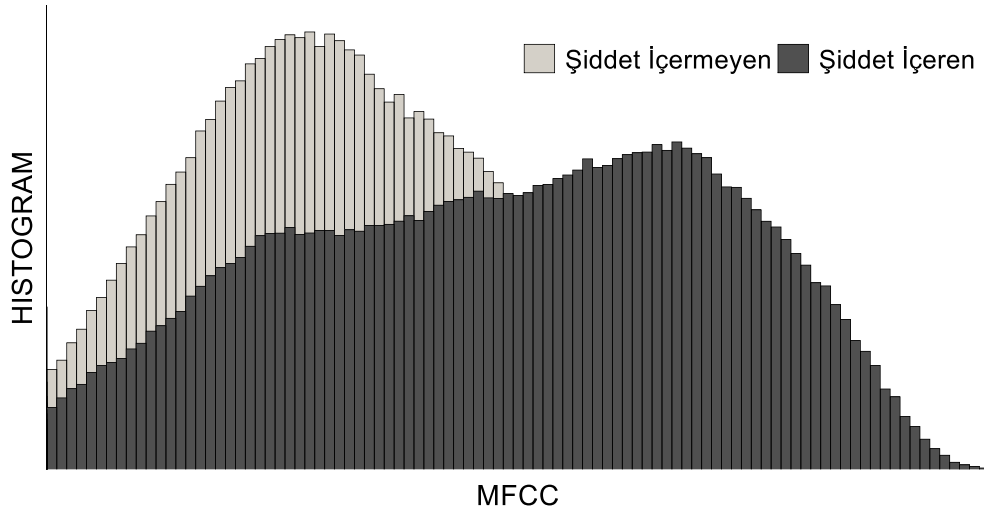
Veri kümesinde 17 filme ait MFCC ve ZCR öznitelikleri, bu filmlere ait şiddet içeren sahnelerin başlangıç ve bitiş bilgileri ile birlikte paylaşılmaktadır.

Ses düzleminde 3 farklı konsept şiddetli olarak varsayılmıştır. Bu konseptler silah ateşlenmesi, patlama ve çığlıktır. Silah ateşlenmesi ile ilgili olarak, herhangi bir sahnede bir ya da birden fazla silah ateşlenmesi mevcut ise şiddetli olarak etiketlenmiştir. Top mermisi ateşlenmesi de silah ateşlenmesi etiketi altında yer almıştır. Patlama ile ilgili olarak, aynı şekilde bir ya da birden fazla patlamanın bulunduğu sahneler şiddetli olarak tanımlanmıştır, bunlara sihirli patlamalar da dahildir. Top mermisi atışı, silah atışı olarak ele alınsa da, patlamanın gerçekleştiği sahne Patlama olarak etiketlenmiştir. Bir ya da birden fazla çığlığın yer aldığı sahneler de şiddetli olarak tanımlanmıştır. Hayvanların çığlıkları ise göz önünde bulundurulmamıştır. Çığlık bilgisine sadece 9 filmde erişilebilmektedir.

Çizelge 5.1'de görüldüğü üzere Şiddet İçermeyen Sahne oranı, Şiddet içeren sahnelere oranına göre oldukça yüksektir. Bu nedenle, makine öğrenmesi yöntemleri sahneleri Şiddet İçermeyen şeklinde tahmin etmektedir. Bunun önüne geçilebilmesi için birçok yöntem bulunmaktadır [27;28;29]. Bu çalışma kapsamında bizler Rastgele Alt örnekleme yöntemi tercih edilmiş [7] ve bu sayede, içeriklerin tahminleme sonuçlarının daha sağlıklı bir şekilde gerçekleştirildiği gözlenmiştir.

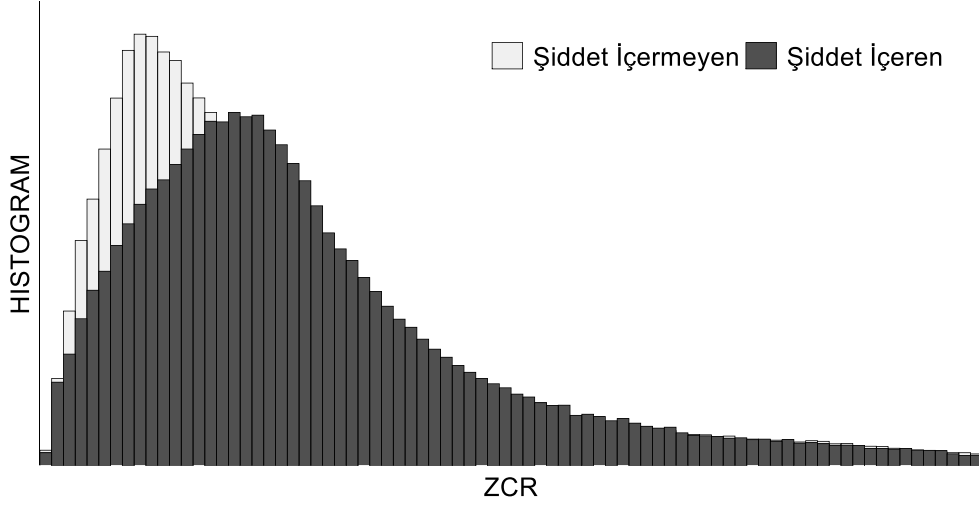
Rastgele alt örnekleme aşamasından sonra modellerin eğitilmesi aşamasında her bir öznitelik için ayrı ayrı modeller oluşturulmuştur. Bu modeller destek vektör makineleri, torbalama ve rastgele ormanlar metotları aracılığıyla oluşturulmuş ve test kümesinde yer alan sahneler kullanılarak başarımları hesaplanmıştır.

Destek vektör makinelerinde her iki sınıflandırma metoduna ait en iyi başarımın elde edildiği çekirdek fonksiyonlar kullanılmıştır. Deneysel sonuçlar ve bu sonuçlara ait değerlendirmeler takip eden bölümde detaylı olarak yer almaktadır.



Şekil 5.2 Şiddet sahnesi sınıflandırması veri kümesinde MFCC özniteliğinin dağılımı

Şekil 5.2'de Şiddet sahnesi sınıflandırması veri kümesinde, MFCC özniteliğinin dağılım grafiği yer almaktadır. Bu grafik MFCC değerlerinin veri kümesindeki dağılımını göstermektedir. Buna göre şiddet içeren ve içermeyen MFCC değerleri dağılımsal olarak farklılık gösterse de, aynı MFCC değerine sahip sahnelerin şiddetli ve şiddetsiz olarak etiketlendiği görülebilmektedir ve bu sınıflandırma problemini daha da güçleştirmektedir.



Şekil 5.3 Şiddet sahnesi sınıflandırması veri kümesinde ZCR özniteliğinin dağılımı

Şekil 5.3'de Şiddet sahnesi sınıflandırması veri kümesinde ZCR özniteliğinin dağılım grafiği yer almaktadır. Bu grafik, ZCR değerlerinin veri kümesindeki dağılımını göstermektedir. Buna göre şiddet içeren ve içermeyen ZCR değerleri dağılımsal olarak çok yakınlık göstermektedir. Yine de, bu veri kümesi üzerinde ZCR özniteliğinin sınıflandırma problemindeki başarımı, MFCC özniteliğinden çok düşük değildir.

Şiddet sahnesi sınıflandırması ve şiddetli sahnelerin tespiti için tasarlanan makine öğrenmesi metodunun başarımını ölçmek amacıyla öncelikli olarak farklı destek vektör makineleri metodlarının sınıflandırma başarımına etkisi incelenmiştir. Bu çalışma kapsamında şiddet sahnesi sınıflandırması işlemini gerçekleştirebilmek

amacıyla Destek vektör makineleri, Rastgele Ormanlar ve Bagging yöntemleri kullanılmıştır.

Çalışma kapsamında sunulan 17 filme ait özet bilgiler Çizelge 4.16'da yer almaktadır.

Çizelge 5.1 Şiddet sahnesi sınıflandırması için kullanılan MediaEval Violent Scene Detection Task veri kümesinde yer alan filmlerin eğitim ve test kümelerine dağılımı, şiddet içeren ve içermeyen sahnelerin toplam sahnelere oranı

<i>Veri kümesi</i>	<i>Film Sayısı</i>	<i>Şiddet İçeren Sahne Oranı</i>	<i>Şiddet İçermeyen Sahne Oranı</i>
Eğitim	14	%9.2	%90.8
Test	3	%5.5	%94.5
Toplam	17	%8.45	%91.55

### 5.1. Destek Vektör Makineleri

Destek vektör makinesi ile elde edilen deneysel sonuçlar, takip eden bölümde sunulmaktadır. Bu başlıkta öncelikle c-destek vektör sınıflandırması ve farklı çekirdek fonksiyonları ile elde edilen başarımlar yer almaktadır. Ardından nu-destek vektör sınıflandırması sonuçları ve farklı çekirdek fonksiyonları ile elde edilen başarımlar yer almaktadır.

Çalışma kapsamında bulunan öznitelik çıkarımı yöntemlerini sonucunda elde edilen özniteliklerin c-destek vektör makineleri sınıflandırma metodu ve doğrusal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar Çizelge 5.2'deki gibidir.

Çizelge 5.2 Şiddet sahnesi sınıflandırması, c-destek vektör doğrusal çekirdek sınıflandırması sonuçları

Öznitelik Adı	Başarım (%)
MFCC	62.36
ZCR	50

Çizelge 5.2'ye göre, Şiddet sahnesi sınıflandırması probleminde c-destek vektör makinesi sınıflandırması doğrusal çekirdek fonksiyonu ile birlikte uygulandığında MFCC özniteliğinin başarımı, ZCR özniteliğinin başarımına oranla %12 daha yüksek görünmektedir.

Çalışma kapsamında bulunan öznitelik çıkarımı yöntemlerini sonucunda elde edilen özniteliklerin c-destek vektör makineleri sınıflandırma metodu ve polinomal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 5.3'deki gibidir.

Çizelge 5.3 Şiddet sahnesi sınıflandırması, c-destek vektör polinomal çekirdek sınıflandırması sonuçları

Öznitelik Adı	Başarım (%)
MFCC	61.47
ZCR	49.6

Çizelge 5.3'e göre, Şiddet sahnesi sınıflandırması probleminde c-destek vektör makinesi sınıflandırması polinomal çekirdek fonksiyonu ile birlikte uygulandığında MFCC özniteliğinin başarımı doğrusal çekirdek uygulandığında olduğu gibi ZCR özniteliğine göre yaklaşık %12 daha yüksek çıkmaktadır.

Çalışma kapsamında bulunan öznitelik çıkarımı yöntemlerini sonucunda elde edilen özniteliklerin c-destek vektör makineleri sınıflandırma metodu ve radyal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 5.4'deki gibidir.

Çizelge 5.4 Şiddet sahnesi sınıflandırması, c-destek vektör radyal çekirdek sınıflandırması sonuçları

Öznitelik Adı	Başarım (%)
MFCC	63.46
ZCR	55.44

Çizelge 5.4'e göre, Şiddet sahnesi sınıflandırması probleminde c-destek vektör makinesi sınıflandırması radyal çekirdek fonksiyonu ile birlikte uygulandığında MFCC özniteliğinin başarımı, ZCR özniteliğinin başarımına oranla %8 daha yüksek görünmektedir.

Çalışma kapsamında bulunan öznitelik çıkarımı yöntemlerini sonucunda elde edilen özniteliklerin c-destek vektör makineleri sınıflandırma metodu ve doğrusal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 5.5'deki gibidir.

Çizelge 5.5 Şiddet sahnesi sınıflandırması, c-destek vektör sigmoid çekirdek sınıflandırması sonuçları

Öznitelik Adı	Başarım (%)
MFCC	38.55
ZCR	50

Çizelge 5.5'e göre, Şiddet sahnesi sınıflandırması probleminde c-destek vektör makinesi sınıflandırması sigmoid çekirdek fonksiyonu ile birlikte uygulandığında ZCR özniteliğinin başarımı, MFCC özniteliğinden yaklaşık %12 daha yüksek çıkmaktadır. Şiddet sahnesi sınıflandırması probleminde c-destek vektör makinesi ile gerçekleştirilen deneylerde sadece sigmoid çekirdekte ZCR özniteliğinin başarımı MFCC özniteliğinden daha yüksek olmaktadır.

Bunun yanı sıra, MFCC özniteliği ile elde edilen başarımlar, C-destek vektör makinesi sınıflandırmasında MFCC özniteliği ile elde edilen en düşük başarımdır.

Nu-Destek Vektör Sınıflandırması kullanılarak elde edilen sonuçlar takip eden bölümde, çekirdek fonksiyonlarına göre ayrılmış şekilde sunulmaktadır.

Çalışma kapsamında bulunan öznitelik çıkarımı yöntemlerini sonucunda elde edilen özniteliklerin nu-destek vektör makineleri sınıflandırma metodu ve doğrusal çekirdek



kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 5.6'daki gibidir.

Çizelge 5.6 Şiddet sahnesi sınıflandırması, nu-destek vektör doğrusal çekirdek sınıflandırması sonuçları

Öznitelik Adı	Başarım (%)
MFCC	53.07
ZCR	49.82

Çizelge 5.6'ya göre, Şiddet sahnesi sınıflandırması probleminde nu-destek vektör makinesi sınıflandırması doğrusal çekirdek fonksiyonu ile birlikte uygulandığında MFCC özniteliğinin başarımları, ZCR özniteliğinin başarımlarına oranla %4 daha yüksek görünmektedir. C-destek vektör makinesi ile gerçekleştirilen deneyler ile karşılaştırıldığında, MFCC özniteliğinin başarımları %10 daha düşük olmakla birlikte, ZCR özniteliğine ait başarımlar oldukça benzerdir.

Çalışma kapsamında bulunan öznitelik çıkarımı yöntemlerini sonucunda elde edilen özniteliklerin c-destek vektör makineleri sınıflandırma metodu ve polinomal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 5.7'deki gibidir.

Çizelge 5.7: Şiddet sahnesi sınıflandırması, nu-destek vektör polinomal çekirdek sınıflandırması sonuçları

Öznitelik Adı	Başarım (%)
MFCC	59.04
ZCR	50

Çizelge 5.7'ye göre, Şiddet sahnesi sınıflandırması probleminde nu-destek vektör makinesi sınıflandırması polinomal çekirdek fonksiyonu ile birlikte uygulandığında MFCC özniteliğinin başarımları doğrusal çekirdek uygulandığında olduğu gibi ZCR özniteliğine göre yaklaşık %9 daha yüksek çıkmaktadır. ZCR özniteliğinin başarımları

doğrusal çekirdeğe oranla değişiklik göstermemiş, MFCC özneliğinin başarımı artış eğiliminde bulunmuştur.

Çalışma kapsamında bulunan öznelik çıkarımı yöntemlerini sonucunda elde edilen özneliklerin c-destek vektör makineleri sınıflandırma metodu ve radyal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 5.8'deki gibidir.

Çizelge 5.8 Şiddet sahnesi sınıflandırması, nu-destek vektör radyal çekirdek sınıflandırması sonuçları

Öznelik Adı	Başarım (%)
MFCC	52.56
ZCR	52.26

Çizelge 5.8'e göre, Şiddet sahnesi sınıflandırması probleminde nu-destek vektör makinesi sınıflandırması radyal çekirdek fonksiyonu ile birlikte uygulandığında MFCC ve ZCR özneliklerine ait başarımların birbirine çok yakın oldukları gözlenmiştir. Her iki özneliğin de c-destek vektör makinelerine oranla başarımları düşmüş olsa da, ZCR özneliğinin başarımı nu-destek vektör makinelerindeki en yüksek ZCR özneliği başarımı olarak dikkat çekmektedir.

Çalışma kapsamında bulunan öznelik çıkarımı yöntemlerini sonucunda elde edilen özneliklerin nu-destek vektör makineleri sınıflandırma metodu ve doğrusal çekirdek kullanılarak destek vektör makineleri yardımıyla sınıflandırılması sonucunda elde edilen başarımlar, Çizelge 5.9'daki gibidir.

Çizelge 5.9 Şiddet sahnesi sınıflandırması, nu-destek vektör sigmoid çekirdek sınıflandırması sonuçları

Öznelik Adı	Başarım (%)
MFCC	61.47
ZCR	50

Çizelge 5.9'a göre, Şiddet sahnesi sınıflandırması probleminde nu-destek vektör makinesi sınıflandırması sigmoid çekirdek fonksiyonu ile birlikte uygulandığında ZCR özniteliğinin başarımı, MFCC özniteliğinden yaklaşık %11 daha yüksek çıkmaktadır.

Bunun yanı sıra, MFCC özniteliği ile elde edilen başarım, nu-destek vektör makinesi sınıflandırmasında MFCC özniteliği ile elde edilen en yüksek başarımıdır.

## 5.2. Torbalama

Çalışmanın bu bölümünde mevcut veri kümesi üzerinde Torbalama metodu uygulanmış, elde edilen sonuçlar yorumlanmıştır. Takip eden bölümde farklı destek vektör metotlarına ve çekirdek türlerine göre ayrıştırılmış, 10 ve 20 torba parametreleri için sonuçlar paylaşılmıştır.

Bu bölümde torbalama yöntemi, c-destek vektör makineleri ile birlikte doğrusal çekirdek fonksiyonu kullanılarak elde edilen deneysel sonuçlar, 10 ve 20 torba parametreleri için paylaşılmıştır.

Çizelge 5.10 Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, c-destek vektör makineleri doğrusal çekirdek sonuçları

Öznitelik Adı	10 Torba	20 Torba
MFCC (Sınıflandırma başarımı %)	62.33	62.33
ZCR (Sınıflandırma başarımı %)	50.79	50.04

Çizelge 5.10'a göre, Şiddet sahnesi sınıflandırması probleminde torbalama yöntemi c-destek vektör makinesi doğrusal çekirdek uygulandığında MFCC özniteliği için 10 torba ve 20 torba parametreleri sınıflandırma başarımını değiştirmemektedir. Referans değer olarak torbalama yöntemi olmaksızın c-destek vektör makinesi doğrusal çekirdek başarımı baz alındığında, torbalama yönteminin pozitif yönlü bir etkisi görünmemektedir. ZCR özniteliğinin başarımı, 10 torba parametrelili deneyde %1'e yakın bir artış göstermiştir. ZCR özniteliğinin başarımı 10 ve 20 torba parametrelerine bağlı olarak kayda değer bir değişim göstermemiştir.

Bu bölümde torbalama yöntemi, c-destek vektör makineleri ile birlikte polinomal çekirdek fonksiyonu kullanılarak elde edilen deneysel sonuçlar, 10 ve 20 torba parametreleri için paylaşılmıştır.

Çizelge 5.11 Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, c-destek vektör makineleri polinomal çekirdek sonuçları

Öznitelik Adı	10 Torba	20 Torba
MFCC (Sınıflandırma başarımı %)	48.74	57.29
ZCR (Sınıflandırma başarımı %)	50.12	50.04

Çizelge 5.11'e göre, Şiddet sahnesi sınıflandırması probleminde torbalama yöntemi c-destek vektör makinesi polinomal çekirdek uygulanıp referans değeri olarak torbalama yöntemi olmaksızın c-destek vektör makineleri polinomal çekirdek başarımı baz alındığında MFCC özneliği ile 10 torba parametresi için başarımlar yaklaşık %12 azalmaktadır. 20 torba parametresinde ise bu değer yaklaşık %4'e düşmektedir. ZCR özneliği için ise 10 torba ve 20 torba parametrelerindeki başarımlar değeri, referans değere göre yaklaşık %0.5 artmaktadır.

Bu bölümde torbalama yöntemi, c-destek vektör makineleri ile birlikte radyal çekirdek fonksiyonu kullanılarak elde edilen deneysel sonuçlar, 10 ve 20 torba parametreleri için paylaşılmıştır.

Çizelge 5.12 Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, c-destek vektör makineleri radyal çekirdek sonuçları

Öznitelik Adı	10 Torba	20 Torba
MFCC (Sınıflandırma başarımı %)	63.59	63.66
ZCR (Sınıflandırma başarımı %)	55.48	55.16

Çizelge 5.12'ye göre, Şiddet sahnesi sınıflandırması probleminde torbalama yöntemi c-destek vektör makinesi radyal çekirdek uygulanıp referans değeri olarak torbalama yöntemi olmaksızın c-destek vektör makineleri radyal çekirdek başarımı baz

alındığında hem MFCC hem de ZCR özneliğinin başarımlar değeri, referans değere oldukça yakındır.

Bu bölümde torbalama yöntemi, c-destek vektör makineleri ile birlikte sigmoid çekirdek fonksiyonu kullanılarak elde edilen deneysel sonuçlar, 10 ve 20 torba parametreleri için paylaşılmıştır.

Çizelge 5.13 Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, c-destek vektör makineleri sigmoid çekirdek sonuçları

Öznitelik Adı	10 Torba	20 Torba
MFCC (Sınıflandırma başarımları %)	43.16	42.06
ZCR (Sınıflandırma başarımları %)	50	50

Çizelge 5.13'e göre, Şiddet sahnesi sınıflandırması probleminde torbalama yöntemi c-destek vektör makinesi radyal çekirdek uygulanıp referans değeri olarak torbalama yöntemi olmaksızın c-destek vektör makineleri radyal çekirdek başarımları baz alındığında hem MFCC özneliğinin başarımları 10 torba parametresi ile yaklaşık %4.5, 20 torba parametresi ile %3.5 artmaktadır. ZCR özneliğinin 10 torba ve 20 torba parametreleri ile başarımları ise, referans değere göre değişmemektedir.

Bu bölümde torbalama yöntemi, nu-destek vektör makineleri ile birlikte doğrusal çekirdek fonksiyonu kullanılarak elde edilen deneysel sonuçlar, 10 ve 20 torba parametreleri için paylaşılmıştır.

Çizelge 5.14 Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, nu-destek vektör makineleri doğrusal çekirdek sonuçları

Öznitelik Adı	10 Torba	20 Torba
MFCC (Sınıflandırma başarımları %)	52.13	59.57
ZCR (Sınıflandırma başarımları %)	49.99	49.99

Çizelge 5.14'e göre, Şiddet sahnesi sınıflandırması probleminde torbalama yöntemi nu-destek vektör makinesi doğrusal çekirdek uygulanıp referans değeri olarak

torbalama yöntemi olmaksızın nu-destek vektör makinesi doğrusal çekirdek başarımı baz alındığında, MFCC özniteliğinin başarımı 10 torba parametresi için %1 azalmakta fakat 20 torba parametresi için %6 artış göstermektedir. ZCR özniteliğinin başarımı ise referans değer ile aynı kalmaktadır.

Bu bölümde torbalama yöntemi, nu-destek vektör makineleri ile birlikte polinomal çekirdek fonksiyonu kullanılarak elde edilen deneysel sonuçlar, 10 ve 20 torba parametreleri için paylaşılmıştır.

Çizelge 5.15 Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, nu-destek vektör makineleri polinomal çekirdek sonuçları

Öznitelik Adı	10 Torba	20 Torba
MFCC (Sınıflandırma başarımı %)	59.25	56.21
ZCR (Sınıflandırma başarımı %)	49.96	49.75

Çizelge 5.15'e göre, Şiddet sahnesi sınıflandırması probleminde torbalama yöntemi nu-destek vektör makinesi polinomal çekirdek uygulanıp referans değer olarak torbalama yöntemi olmaksızın nu-destek vektör makinesi polinomal çekirdek başarımı baz alındığında, MFCC özniteliğinin başarımı 10 torba parametresi için benzer kalmakta; fakat 20 torba parametresi ile yaklaşık %3 azalmaktadır. Buna karşılık, ZCR özniteliğinin başarımı referans değer ile benzerlik göstermektedir.

Bu bölümde torbalama yöntemi, nu-destek vektör makineleri ile birlikte radyal çekirdek fonksiyonu kullanılarak elde edilen deneysel sonuçlar, 10 ve 20 torba parametreleri için paylaşılmıştır.

Çizelge 5.16 Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, nu-destek vektör makineleri radyal çekirdek sonuçları

Öznitelik Adı	10 Torba	20 Torba
MFCC (Sınıflandırma başarımı %)	51.88	51.84
ZCR (Sınıflandırma başarımı %)	50.90	50.48

Çizelge 5.16'ya göre, Şiddet sahnesi sınıflandırması probleminde torbalama yöntemi nu-destek vektör makinesi radyal çekirdek uygulanıp referans değer olarak torbalama yöntemi olmaksızın nu-destek vektör makinesi radyal çekirdek başarıımı baz alındığında, MFCC özniteliğinin başarıımı 10 torba ve 20 torba parametreleri için birbirine oldukça yakındır ve referans değerden yaklaşık %0.5 azdır. ZCR özniteliğinde ise hem 10 torba hem de 20 torba parametresinde başarımda referans değere göre başarıım yaklaşık %2 daha azdır.

Bu bölümde torbalama yöntemi, nu-destek vektör makineleri ile birlikte Sigmoid çekirdek fonksiyonu kullanılarak elde edilen deneysel sonuçlar, 10 ve 20 torba parametreleri için paylaşılmıştır.

Çizelge 5.17 Şiddet sahnesi sınıflandırması, Torbalama yöntemi ile elde edilen sonuçlar, nu-destek vektör makineleri sigmoid çekirdek sonuçları

Öznitelik Adı	10 Torba	20 Torba
MFCC (Sınıflandırma başarıımı %)	61.48	61.47
ZCR (Sınıflandırma başarıımı %)	50	50

Çizelge 5.17'ye göre, Şiddet sahnesi sınıflandırması probleminde torbalama yöntemi nu-destek vektör makinesi sigmoid çekirdek uygulanıp referans değer olarak torbalama yöntemi olmaksızın nu-destek vektör makinesi sigmoid çekirdek başarıımı baz alındığında, hem MFCC hem de ZCR öznitelikleri için 10 torba ve 20 torba parametrelerinde elde edilen başarımlar referans değer ile özdeştir. Ayrıca MFCC özniteliği ile elde edilen bu başarıım, torbalama yönteminde nu-destek vektör makineleri ile elde edilen en yüksek başarımdır.

### 5.3. Rastgele Ormanlar

Çalışmanın bu bölümünde, Şiddet sahnesi sınıflandırması problemi veri kümesinde Rastgele Ormanlar metodu uygulandığında elde edilen deneysel sonuçlar yer almaktadır.

Çizelge 5.18 Şiddet sahnesi sınıflandırması, Rastgele Ormanlar yöntemi ile elde edilen sonuçları

Öznitelik Adı	10 Ağaç	20 Ağaç
MFCC	57.33	57.16
ZCR	55.46	55.50

Çizelge 5.18'e göre, Şiddet sahnesi sınıflandırması probleminde Rastgele ormanlar metodu kullanıldığında elde edilen başarımlar, hem MFCC hem de ZCR özneliği için destek vektör makineleri ya da torbalama yöntemi ile elde edilen en yüksek sonuçlardan daha yüksek değildir. MFCC özneliğinde ağaç sayısı arttığında başarımlar %0.2 dolaylarında düşüş gösterse de, ZCR özneliğinde %0.04 artış göstermiştir.

#### 5.4. Önerilen Yöntemler

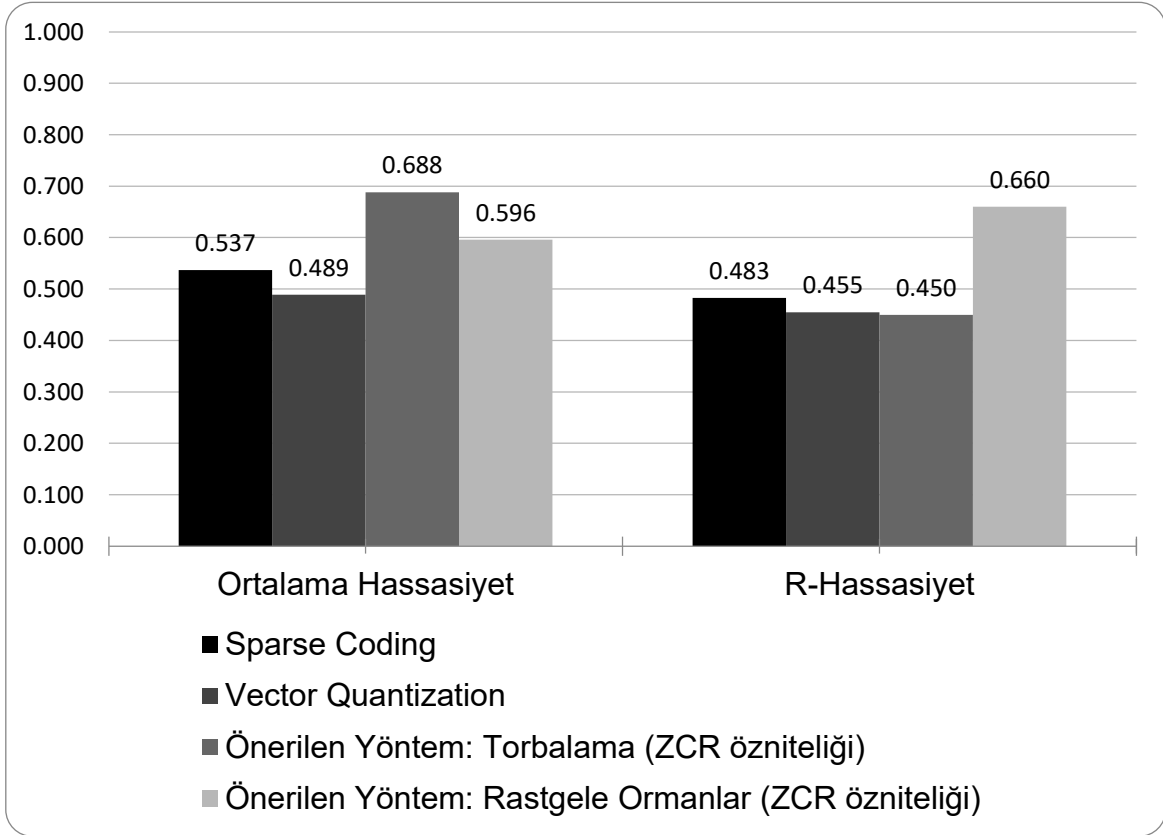
Şiddet sahnesi sınıflandırması probleminde ise literatürdeki çalışmalar ile karşılaştırma yapabilmek için sınıflandırma başarımları değeri yeterli olmamaktadır. Acar'ın çalışmasını ele aldığımızda [20], ortalama hassasiyet ve R-hassasiyet değerlerinin mevcut olduğunu görmekteyiz. Deneysel sonuçlar sonucunda elde ettiğimiz en yüksek ortalama hassasiyet ve R-hassasiyet sonuçlarının, [20] ile karşılaştırılması sonucunda elde edilen değerler aşağıda sunulmuştur. Literatürde kullanılan Ortalama Hassasiyet@20, ilk 20 değer içerisindeki verilerin hassasiyet değerlerinin ortalaması olarak hesaplanmaktadır. R-Hassasiyet@20 ise, ilk 20 değer içerisinde yer alan ilgili verilerin oranıdır. İlgili karşılaştırma Çizelge 5.19'da sunulmuştur. Karşılaştırmanın grafik hali ise Şekil 5.4'de yer almaktadır.



Çizelge 5.19 Şiddet sahnesi sınıflandırması, deneysel sonuçların literatür [20] ile karşılaştırılması

Yöntem	Ortalama Hassasiyet@20	R-Hassasiyet@20
Sparse Coding [20]	0.537	0.483
Vector Quantization [20]	0.489	0.455
Önerilen Yöntem: Torbalama (ZCR özneliği)	0.596	0.660
Önerilen Yöntem: Rastgele Ormanlar (ZCR özneliği)	0.688	0.450

Şekil 5.4: Şiddet sahnesi sınıflandırması, deneysel sonuçların literatür [20] ile karşılaştırılması grafiği



Çizelge 5.19 ve Şekil 5.4'e göre, çalışmamızda sunmuş olduğumuz yöntemlerin Ortalama Hassasiyet ve R-Hassasiyet metrikleri kullanıldığında da daha başarılı olduğunu görmekteyiz.

## 6. SONUÇLAR VE TARTIŞMA

Bu tez çalışması kapsamında, silah seslerinin ve video verilerinin şiddet içeren sahnelerinin ses tabanlı sınıflandırılması problemleri ele alınmıştır. Çalışmada sınıflandırıcı parametrelerinin destek vektör makinesi sınıflandırmasına etkisi incelenmiş, topluluk öğrenme yöntemleri olan Torbalama ve Rastgele Ormanlar karşılaştırmalı olarak sunulmuştur.

Silah seslerinin sınıflandırılması ve şiddet sahnesi sınıflandırılması problemlerinde öznitelik çıkarım yöntemlerinin sınıflandırma başarımına etkisi, sınıflandırıcıların ve sınıflandırıcılara ait alt parametrelerin başarıma etkisi incelenmiştir. Bu bağlamda, diğer parametrelerin sabit olduğu durumda en yüksek başarımın genel olarak ZCR özniteliği ile elde edildiği gözlenmiştir. ZCR özniteliğinin seste gerçekleşen değişikliklere olan hassasiyetinin bu başarımı oluşturduğu düşünülmektedir. Aynı özniteliklerin farklı sınıflandırıcılardaki performansı göz önünde bulundurulduğunda ise en yüksek başarımın Rastgele Ormanlar metodu ile elde edildiği gözlenmiştir.

Silah seslerinin sınıflandırılmasını ve şiddetli sahneler sınıflandırılarak tespit edilmesini sağlayan bu tez çalışmasının daha yüksek sınıflandırma başarımları elde edebilmesi için var olan veri miktarının artırılması gerekmektedir. Silah sesleri sınıflandırılması için açık bir veri kümesine ulaşılamamıştır. Bu amaçla, Montana Üniversitesinden Robert C. Maher'in [2;9;10] çalışmalarında kullandığı veri kümesi kullanılmıştır. Bu veri kümesi üzerinde, daha önce sınıflandırma çalışması yapılmadığından, bu tezde sunulan sınıflandırma sonuçları bilgimiz dahilinde veri kümesi için bir ilktir.

Silah sesleri sınıflandırması problemi için daha fazla veri kümesine ait seslerin eğitim amacıyla kullanılması, daha fazla sayıda sınıfa ait silah sesinin tanımlanabilmesini sağlayacaktır. Bununla birlikte mevcut sınıflara ait daha fazla silah sesinin eğitim amacıyla kullanılması da mevcut silah türlerinin daha yüksek başarımla tanımlanabilmesini sağlayacaktır. Şiddet sahnesi sınıflandırması problemi ele

alındığında ise daha çok sayıda filmin etiketlenerek eğitim kümesi olarak sağlanması, şiddet içeren sahnelerin daha yüksek başarımla tespit edilmesini sağlayacaktır.

Bu çalışmada önerilen yöntemler işitsel özniteliklere dayanmakta olup, farklı konulara kolaylıkla uyarlanabilir yöntemlerdir. Örneğin, gözetleme sistemlerine entegre edilmesi durumunda şiddet içeren olaylar hızla tespit edilebilir ve silah içeren olaylarda silahın türü belirlenebilir. İşitsel veriden elde edilebilecek şiddet durumları sınırlıdır. Görsel – işitsel öznitelikler bir arada kullanılarak daha fazla sayıda şiddet durumunun tespit edilebileceği ve daha yüksek başarımla elde edilebileceği öngörülmektedir. Gelecek çalışmalar için, mevcut özniteliklere ek olarak daha fazla sayıda özniteliğe yer verilebilir, öznitelik seçimi gerçekleştirilebilir. İşitsel özniteliklerin yanı sıra, görsel özniteliklerin de kullanımı sağlanabilir. Daha geniş kapsamlı sınıflandırıcıların uygulanması ile daha yüksek sınıflandırma başarımla elde edilebilir. Eğitim sınıfında yer alan örnek sayısının artırılması ve farklı sınıflara ait örneklerin kullanılması, hem sınıflandırma başarımlarını artıracak hem de sistemlerin daha çeşitli kapsamdaki şiddet olaylarını sınıflandırabilmesini sağlayacaktır.

## KAYNAKLAR LİSTESİ

- [1] Washington Post, The Presidential Transcripts, 3<sup>rd</sup> edition, 1974
- [2] Maher, R., C., Audio Forensic Examination, IEEE Signal Processing Magazine, Vol. 26, Sayı: 2, s. 84-94, Mart 2009
- [3] Koenig, B., E., Authentication of forensic audio recordings, J. Audio Engineer Soc., vol. 38, no. 1/2, s. 3–33, 1990
- [4] AES Standard for Forensic Purposes, Criteria for the Authentication of Analog Audio Tape Recordings, AES Standard 43-2000
- [5] AES Recommended Practice for Forensic Purposes, Managing Recorded Audio Materials Intended for Examination, AES Standard 27-1996
- [6] Koenig, B., E., Lacey, D., S., Killion, S., A., Forensic enhancement of digital audio recordings, J. Audio Eng. Soc., vol. 55, no. 5, s. 252–371, Mayıs 2007
- [7] Zakariah, M., Khan, M., K., Malik, H., Digital multimedia audio forensics: past, present and future, Multimedia tools and applications, vol. 77, s. 1009-1040, Ocak 2018
- [8] Zhao, H., Malik, H., "Audio recording location identification using acoustic environment signature, Information Forensics and Security, s.1746–1759, 2013
- [9] Maher, R., C., Audio Forensic Examination, IEEE Signal Processing Magazine, vol. 26, sayı: 2, s. 84-94, Mart 2009
- [10] Maher, R., C., Modeling and Signal Processing of Acoustic Gunshot Recordings, 2006 IEEE 12nd Digital Signal Processing Workshop & 4th IEEE Signal Processing Education Workshop, Teton National Park, s. 257-261, 2006
- [11] Martin, J., Support Vector Machine Classification of Gunshots, Aralık 2007
- [12] Navratil, M., Kresalek, V., Dostalek, P., Neural network classification of gunshots using spectral characteristics, Ocak 2011
- [13] Djeddou, M., Touhami, T., Classification and Modeling of Acoustic Gunshot Signatures, Arabian Journal for Science and Engineering, vol. 38, sayı 12, s. 3399–3406, Aralık 2013
- [14] Kiktova, E., Lojka, M., Pleva, M., Juhar, J., Cizmar, A., Gun type recognition from gunshot audio recordings, 3rd International Workshop on Biometrics and Forensics (IWBF 2015), Gjovik, s. 1-6., 2015
- [15] Mittal, A., Cheong, L., F., Sing, L., T., Dynamic Bayesian framework for extracting temporal structure in video, Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001, Vol. 2, 2001
- [16] Datta, A., Shah, M., Lobo, N., D., V., Person-on-Person Violence Detection in Video Data, Pattern Recognition, 16th International Conference, 2002
- [17] Kim, S., Nam, Y., Kim, J., Cho., W., D., ISS: Intelligent Surveillance System using Autonomous Multiple Cameras, Proceedings of the 4th Int. Conf. on Ubiquitous Information Technologies & Applications, Fukuoka, Japan, s. 1-6, 2009

- [18] Souza, F., Chavez, G., C., Valle, E., Araujo, A., Violence Detection in Video Using Spatio-Temporal Features., In SBC Conference on Graphics, Patterns and Images, s. 224–230, 2010
- [19] Giannakopoulos, T., Pikrakis, A., Theodoridis, S., A Multi-Class Audio Classification Method With Respect To Violent Content In Movies Using Bayesian Networks, Multimedia Signal Processing, MMSP 2007. IEEE 9th Workshop, 2007
- [20] Acar, E., Hopfgartner, F., Albayrak, S., Detecting violent content in Hollywood movies by mid-level audio representations, Content-Based Multimedia Indexing (CBMI), 11th International Workshop, 2013
- [21] Mu, G., Cao, H., Jin, Q., Violent Scene Detection Using Convolutional Neural Networks and Deep Audio Features, Communications in Computer and Information Science, vol 663., 2016
- [22] Sjöberg, M., Schlüter, J., Ionescu, B., Schedl, M., FAR at MediaEval 2013 Violent Scenes Detection: Concept-based Violent Scenes Detection in Movies, Working Notes Proc., 2013
- [23] Giannakopoulos, T., Pikrakis, A., Theodoridis, S., A multimodal approach to violence detection in video sharing sites, In Pattern Recognition (ICPR), 2010 20th International Conference, s. 3244-3247, 2010
- [24] Cortes, C., Vapnik, V., Support-Vector Networks., Mach. Learn. vol. 20, sayı 3, s. 273-297, Eylül 1995
- [25] Boser, B., E., Guyon, I., M., Vapnik, V., N., A training algorithm for optimal margin classifiers, In Proceedings of the fifth annual workshop on Computational learning theory COLT '92. ACM, New York, NY, USA, s. 144-152, 1992
- [26] Demarty, C.,H., Penet, C., Soleymani, M., Gravier, G., VSD, a public dataset for the detection of violent scenes in movies: design, annotation, analysis and evaluation., Multimedia Tools and Applications, Mayıs 2014
- [27] He, H., Garcia, E., Learning from imbalanced data. Knowledge and Data Engineering, IEEE Transactions, s.1263–1284, 2009
- [28] Akbani, R., Kwek, S, Japkowicz, N., Applying support vector machines to imbalanced datasets, ECML, s. 39–50, 2004
- [29] Magre, S., Janse, P., Deshmukh, R., A Review on Feature Extraction and Noise Reduction Technique, 2014
- [30] Hermansky, H., Perceptual linear predictive (PLP) analysis of speech, The Journal of the Acoustical Society of America, 1990
- [31] Hermansky, H., Morgan, N., Bayya, A., Kohn, P., RASTA-PLP speech analysis technique, ICASSP-92: 1992 IEEE International Conference on Acoustics, Speech, and Signal Processing, Vol. 1, s.121-124, 1992
- [32] Makhoul, J., Linear prediction: A tutorial review," in Proceedings of the IEEE, vol. 63, no. 4, s. 561-580, 1975
- [33] Chen, C. H., Signal processing handbook, Dekker, New York, 1988
- [34] International Electrotechnical Commission, Letter symbols to be used in electrical technology – Part 3: Logarithmic and related quantities, and their units, 3. Baskı, 2002

- [35] Breiman, L., Bagging predictors. Machine Learning, vol. 24, 1996, s. 123-140
- [36] Breiman, L., Random Forests. Machine Learning, vol. 45, 2001, s. 5-32