

**KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

**DUYMA ENGELLİLER İÇİN GİYİLEBİLİR TİTREŞİM TABANLI SES TANIMA
UYGULAMALARI**

DOKTORA TEZİ

Bilgisayar Yük. Müh. Mete YAĞANOĞLU

**MAYIS 2018
TRABZON**



KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ

BİLGİSAYAR MÜHENDİSLİĞİ ANABİLİM DALI

DUYMA ENGELLİLER İÇİN GİYİLEBİLİR TİTREŞİM TABANLI SES TANIMA
UYGULAMALARI

Bil. Yük. Müh. Mete YAĞANOĞLU

Karadeniz Teknik Üniversitesi Fen Bilimleri Enstitüsünde
"DOKTOR (BİLGİSAYAR MÜHENDİSLİĞİ)"
Unvanı Verilmesi İçin Kabul Edilen Tezdir.

Tezin Enstitüye Verildiği Tarih : 09 / 04 / 2018

Tezin Savunma Tarihi : 04 / 05 / 2018

Tez Danışmanı : Prof. Dr. Cemal KÖSE

Trabzon 2018

**KARADENİZ TEKNİK ÜNİVERSİTESİ
FEN BİLİMLERİ ENSTİTÜSÜ**

**Bilgisayar Mühendisliği Anabilim Dalında
Mete YAĞANOĞLU Tarafından hazırlanan**

**DUYMA ENGELLİLER İÇİN GİYİLEBİLİR TİTREŞİM TABANLI SES TANIMA
UYGULAMALARI**

**başlıklı bu çalışma, Enstitü Yönetim Kurulunun 10 /04/2018 gün ve 1748 sayılı
kararıyla oluşturulan jüri tarafından yapılan sınavda
DOKTORA TEZİ
olarak kabul edilmiştir.**

Jüri Üyeleri

Başkan : Prof. Dr. Recep Alp KUT

Üye : Prof. Dr. Cemal KÖSE

Üye : Doç. Dr. Ümit Deniz ULUŞAR

Üye : Doç. Dr. Bekir DİZDAROĞLU

Üye : Dr. Öğr. Üyesi Önder AYDEMİR



**Prof. Dr. Sadettin KORKMAZ
Enstitü Müdürü**

ÖNSÖZ

Teknolojinin gelişmesi ile birlikte giyilebilir sistemler hayatımızın her alanında bizlere kolaylık sağlamaktadır. Son zamanlarda özellikle ses tanıma kullanılarak yapılan çalışmalar popüler olmaktadır. Bu tez çalışmasında duyma engelliler için giyilebilir titreşim tabanlı bir sistem önerilmektedir. Duyma engelliler hayat kalitesini artırmak, daha refah bir yaşam sürmesini sağlamak için geliştirilen giyilebilir sistemimizde ses yön bulma, ses tanıma, yanında yüksek sesli konuşma, ebeveyn sesi tanıma gibi önemli olabilecek sorunlara çözüm olunmuştur. Bu çalışmada çıkan ürünler duyma engellilerin daha iyi bir hayat standardına sahip olmasına katkı sağlamaktadır.

Çalışmalarım süresince bilgi, görüş ve önerileriyle bana yardımcı olan çok değerli danışman hocam Sayın Prof. Dr. Cemal KÖSE'ye teşekkürlerimi bir borç bilir ve şükranlarımı sunarım.

Çalışmalarım süresince değerli görüş ve katkılarını esirgemeyen Sayın Doç. Dr. Bekir DİZDAROĞLU ve Sayın Dr. Öğr. Üyesi Önder AYDEMİR'e teşekkürlerimi sunarım. Çalışmalarım süresince bana destek olan değerli mesai arkadaşım Arş. Gör. Ferhat BOZKURT'a teşekkürlerimi sunarım.

Öğrenim hayatım boyunca buralara gelmemi sağlayan sevgili babama, anneme ve abime çok teşekkür ederim. Hayatıma anlam kazandıran ve tez çalışması süresince her zaman bana destek olan çok sevdiğim eşim Sevda YAĞANOĞLU'na sonsuz teşekkürlerimi sunarım.

Mete YAĞANOĞLU

Trabzon 2018

TEZ ETİK BEYANNAMESİ

Doktora Tezi olarak sunduđum “Duyma Engelliler İin Giyilebilir Titreřim Tabanlı Ses Tanıma Uygulamaları” bařlıklı bu alıřmayı bařtan sona kadar danıřmanım Prof. Dr. Cemal KÖSE’nin sorumluluđunda tamamladıđımı, verileri / örnekleri kendim topladıđımı, deneyleri / analizleri ilgili laboratuarlarda yaptıđımı / yaptırdıđımı, bařka kaynaklardan aldıđım bilgileri metinde ve kaynakada eksiksiz olarak gösterdiđimi, alıřma sürecinde bilimsel arařtırma ve etik kurallara uygun olarak davrandıđımı ve aksinin ortaya ıkması durumunda her türlü yasal sonucu kabul ettiđimi beyan ederim. 04/05/2018

Mete YAĐANOĐLU

İÇİNDEKİLER

	<u>Sayfa No</u>
ÖNSÖZ.....	III
TEZ ETİK BEYANNAMESİ.....	IV
İÇİNDEKİLER.....	V
ÖZET.....	VIII
SUMMARY	IX
ŞEKİLLER DİZİNİ.....	X
TABLolar DİZİNİ.....	XII
SEMBOLLER DİZİNİ.....	XIII
1. GENEL BİLGİLER.....	1
1.1. Giriş.....	1
1.2. Ses Parmak İzi	3
1.2.1. Ses Parmak İzi Özellikleri	6
1.3. Giyilebilir Teknoloji	8
1.4. Tezin Amacı ve Kapsamı.....	10
1.5. Literatür Çalışması.....	14
1.5.1. Sağır Kişiler ile İlgili Çalışmalar	14
1.5.2. Ses Parmak İzi Çalışmaları	20
1.5.3. Ses Tanıma Çalışmaları	23
2. YAPILAN ÇALIŞMALAR.....	27
2.1. Ses Tanıma Çalışması.....	27
2.1.1. Sistemin Yapısı.....	27
2.1.2. Ses Parmak İzi Yöntemi	29
2.1.2.1. Ses Parmak İzi Genel Prensipler.....	31
2.1.2.2. Ses Parmak İzi Çalışma Mekanizması.....	34
2.1.2.2.1. Örnekleme.....	34
2.1.2.2.2. Spektrogram.....	35
2.1.2.2.3. Zirve Bulma.....	39

2.1.2.2.4.	Ses Parmak İzi Özetleme	43
2.1.2.2.5.	Ses Parmak İzi Oluşturmak İçin Kullanılan Yapılar	44
2.1.2.2.6.	Kullanılan Kütüphaneler.....	45
2.1.3.	Diğer Metotlar.....	46
2.1.3.1.	Sıfır Geçiş Oranı	46
2.1.3.2.	Mel Frekansı Kepstrum Katsayıları.....	46
2.1.3.3.	Spektral Altbant Merkezleri	49
2.1.3.4.	Çizgisel Spektral Frekanslar	49
2.1.3.5.	Doğrusal Kestirim Katsayıları	50
2.1.3.6.	Doğrusal Öngörülü Kepstral Katsayıları	51
2.1.3.7.	Log Alanı Oranı	52
2.1.3.8.	Sınıflandırma	52
2.1.3.8.1.	K-En Yakın Komşu	52
2.1.3.8.2.	Destek Vektör Makineleri.....	53
2.1.3.8.2.1.	Doğrusal Ayrılabilir Veriler İçin DVM.....	54
2.1.3.8.2.2.	Doğrusal Ayrılamayan Veriler İçin DVM.....	55
2.2.	Yön Bulma Çalışması.....	56
2.2.1.	Sistemin Yapısı.....	56
2.2.2.	Eğitim Verisi	59
2.2.3.	Ön İşleme	59
2.2.4.	Kullanılan Öznitelik Yöntemleri	59
2.2.4.1.	Çarpıklık	60
2.2.4.2.	Basıklık	60
2.2.4.3.	Kare Ortalamalarının Karekökü	60
2.2.4.4.	Varyans ve Standart Sapma	61
2.2.4.5.	Yerel Maksimum ve Yerel Minimum.....	61
2.2.5.	Öznitelik Seçme	61
2.2.6.	Mikrofon Sınıfı Kullanan Lokalizasyon Algoritmaları	62
2.3.	Ebeveyn Sesi Tahmini Çalışması	63
2.4.	Diğer Uygulamalar	64
2.4.1.	Horlama Sesi Tahmini Çalışması	64
3.	BULGULAR.....	66

3.1.	Değerlendirme Ölçütleri	66
3.2.	Deney Ortamı	67
3.3.	Ses Tanıma Çalışması.....	71
3.4.	Yön Bulma Çalışması.....	80
3.5.	Ebeveyn Sesi Tahmini Çalışması	88
3.6.	Horlama Sesi Tahmini Çalışması	92
4.	SONUÇLAR.....	95
5.	ÖNERİLER VE TARTIŞMA.....	97
6.	KAYNAKLAR	102

ÖZGEÇMİŞ

Doktora Tezi

ÖZET

DUYMA ENGELLİLER İÇİN GİYİLEBİLİR TİTREŞİM TABANLI SES TANIMA
UYGULAMALARI

Mete YAĞANOĞLU

Karadeniz Teknik Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalı
Danışman: Prof. Dr. Cemal KÖSE
2018, 116 Sayfa

Duyma engelli kişilerin günlük yaşantısında ses algılaması hayat kalitesi açısından çok önemlidir. Duyma engelli kişiler dışarıda yürürken ya da ev içerisindeyken birçok önemli sesi duyamazlar. Geliştirdiğimiz giyilebilir sistem ile duyma engelliler sesin yönünü ve çevresinde oluşan önemli sesleri ayırt edebilecektir. Mikrofon ve titreşim motorundan oluşan giyilebilir cihazımız kullanıcının sırtına monte edilmiştir. Ses tanıma çalışmasında gerçek-zamanlı uygulamalarda ses parmak izi yöntemi kullanılmıştır. Geliştirdiğimiz giyilebilir cihazda yönü algılamak için ise sınıflandırma tekniği kullanılmıştır. Bu tez çalışmasında yanında konuşan kişinin bağırp bağırmadığı da tespit edilmiştir. Duyma engelli geliştirmeyi hedeflediğimiz sistem sayesinde kişinin bağırp bağırmadığını anlayabilecektir ve önemli bir durum varsa bunu algılaması hızlanmaktadır. Bu tez çalışmasında ebeveyn sesi tahmini de yapılmıştır. Kullanan kişinin kendi anne ve baba sesini tanıyacak bir sistem tasarlanmıştır. Bu çalışmaların hepsinde giyilebilir ve titreşim tabanlı bir sistem tasarlanmıştır. Bu çalışma sağır insanların yararına yeni bir fikir sunmaktadır. Burada sunulan bilgilerin, duyma engelliler için sistem geliştirirken ihtiyaç duyulan giyilebilir hesaplama ve insan-bilgisayar etkileşimi alanındaki araştırmalarda bilgilere katkı sağlayacağına inanmaktayız.

Anahtar Kelimeler: Giyilebilir hesaplama, İnsan bilgisayar etkileşimi, Ses tanıma, Ses yön tanıma, Sağırılar için titreşim tabanlı ses algılama, Ses parmak izi.

PhD. Thesis

SUMMARY

WEARABLE VIBRATION BASED SOUND RECOGNITION APPLICATIONS
FOR DEAF

Mete YAĞANOĞLU

Karadeniz Technical University
The Graduate School of Natural and Applied Sciences
Computer Engineering Graduate Program
Supervisor: Prof. Dr. Cemal KÖSE
2018, 116 Pages

Sensing of sound is very important in daily life of deaf people in term of quality of life. When they walk outside or are at home, they are unable to hear many important sounds. Using the wearable vibration-based system we have designed and developed, deaf people can distinguish the direction and importance of sounds occurring around them. Our wearable device consists of a microphone and vibration motor mounted on the user's back. In the study of sound recognition, audio fingerprint method is used in real-time applications. Classification technique is used to sense direction in the developed wearable device. In this thesis, it is determined whether who talks by the one shouts. Thanks to the system we aim at developing, a deaf person can understand whether or not a person shouting and if there is a critical situation, it alerts them to it. In this thesis, a system that enables users to recognize their own parents' voices is designed. In all these studies, a wearable and vibration-based system is designed. This study presents a new idea that will benefit deaf people. The information presented here will contribute to the research in the fields of wearable processing and human-computer interaction, which are required for the development of systems for deaf people.

Key Words: Wearable processing, Human computer interaction, Speech recognition, Sound direction recognition, Vibration based sound detection for deaf, Audio fingerprint.

ŞEKİLLER DİZİNİ

	<u>Sayfa No</u>
Şekil 1.1. Ses tanıma işlemi için ses parmak izi kullanımı.....	6
Şekil 1.2. Ses parmak izi genel sistemi [38].....	8
Şekil 1.3. Giyilebilir teknoloji genel yapısı.....	10
Şekil 1.4. Tezin genel akış şeması.....	13
Şekil 2.1. Ses tanıma çalışması için geliştirdiğimiz giyilebilir sistem	28
Şekil 2.2. Giyilebilir sistem ekipmanları	29
Şekil 2.3. Ses parmak izi yapısı [143]	30
Şekil 2.4. Puanlama süreci [143].....	31
Şekil 2.5. Ses parmak izi çıkarma genel yapısı [38].....	32
Şekil 2.6. Ses parmak izi (a) öznitelik çıkarma, (b) eşleştirme adımları [143]	33
Şekil 2.7. Giyilebilir sistemimizin ses parmak izi kullanımı.....	34
Şekil 2.8. Ses örnekleri ve spektrogramları	36
Şekil 2.9. Ses verisi spektrogramı takımyıldızı haritası	41
Şekil 2.10. Çapa noktası gösterimi	42
Şekil 2.11. Ses parmak izi akış şeması [19]	45
Şekil 2.12. MFKK adımları	47
Şekil 2.13. Yeni gelen örneğin sınıfının K-EYK ile belirlenmesi.....	53
Şekil 2.14. Doğrusal ayrılabilen veriler için hiper-düzlem seçimi [163]	54
Şekil 2.15. Doğrusal ayrılamayan veriler için hiper-düzlem seçimi [163].....	56
Şekil 2.16. Yön tanıma çalışması için geliştirdiğimiz giyilebilir cihaz.....	58
Şekil 2.17. Ebeveyn sesi tahmini için geliştirdiğimiz giyilebilir cihaz	63
Şekil 2.18. Horlama sesi tahmini için geliştirdiğimiz giyilebilir cihaz	65
Şekil 3.1. Ses tanıma çalışması için kullanılan giyilebilir cihaz test ortamı	68
Şekil 3.2. Yön bulma çalışması için kullanılan giyilebilir cihaz test ortamı	70
Şekil 3.3. Giyilebilir test sistemi	70
Şekil 3.4. İç ortam seslerinin karşılaştırılması.....	71
Şekil 3.5. İç ortam sesleri ROC eğrileri	72
Şekil 3.6. Dış ortam seslerinin karşılaştırılması	73
Şekil 3.7. Dış ortam sesleri ROC eğrileri	73
Şekil 3.8. Farklı test ortamında SPİ yöntemi karşılaştırılması	74

Şekil 3.9.	Duyuma engelliler ile normal duyan kişilerin karşılaştırılması	77
Şekil 3.10.	Ses tanıma çalışması gerçek zamanlı test sonuçları	79
Şekil 3.11.	Öznitelik yöntemlerini karşılaştırılması	80
Şekil 3.12.	Yön bulma çalışması bilgisayar ortamında elde edilen sonuçlar	81
Şekil 3.13.	Sol, sağ, ön ve arka yönlere ait ROC eğrileri	82
Şekil 3.14.	Yön bulma çalışması gerçek zamanlı elde edilen sonuçlar	83
Şekil 3.15.	Farklı bireylerin ses algılama başarı oranları	86
Şekil 3.16.	Sağır birey yön bulma test sonuçları	86
Şekil 3.17.	Normal ses ve bağırma sesi algılama başarı oranları	87
Şekil 3.18.	Ev ortamında sağır ve normal kişilerin ebeveyn sesi tahmin oranları	88
Şekil 3.19.	Ev ortamında MFKK ve SPİ yöntemleri karşılaştırılması	89
Şekil 3.20.	Uyurken sağır ve normal kişilerin ebeveyn sesi tahmin oranları	90
Şekil 3.21.	Uyurken MFKK ve SPİ yöntemleri karşılaştırılması	90
Şekil 3.22.	Sağır kişilerin ebeveyn sesi tahmin oranları	91
Şekil 3.23.	Sağır bireylerin ev ortamında ebeveyn sesi tahmini ROC eğrisi	91
Şekil 3.24.	Horlama sesi tahmin etme sistemi.....	92
Şekil 3.25.	Horlama sesleri ROC eğrisi.....	93

TABLolar DİZİNİ

	<u>Sayfa No</u>
Tablo 1.1. Ses parmak izi yönteminde kullanılan öznelikler.....	22
Tablo 3.1. Hata matrisi gösterimi	66
Tablo 3.2. Çevresel ses tanıma çalışmaları karşılaştırılması	75
Tablo 3.3. Ses tanıma çalışması bilgisayar ortamı sonuçları hata matrisi.....	78
Tablo 3.4. İç ortam sesleri gerçek zamanlı test sonuçları hata matrisi	78
Tablo 3.5. Dış ortam sesleri gerçek zamanlı test sonuçları hata matrisi	78
Tablo 3.6. Ses tanıma çalışması gerçek zamanlı test kappa değerleri.....	79
Tablo 3.7. Yön bulma çalışması gerçek zamanlı ve bilgisayar ortamı başarı sonuçları..	81
Tablo 3.8. Yön bulma çalışması bilgisayar ortamı sonuçları hata matrisi	81
Tablo 3.9. Bilgisayar ortamında elde edilen kappa değerleri	82
Tablo 3.10. Yön bulma çalışması gerçek zamanlı elde edilen hata matrisi.....	83
Tablo 3.11. Yön bulma çalışması gerçek zamanlı elde edilen kappa değerleri.....	84
Tablo 3.12. Farklı ortamlarda elde edilen sonuçlar	84
Tablo 3.13. Mesafeye göre yön bulma başarı oranları	84
Tablo 3.14. 1 metre mesafe test sonuçları	85
Tablo 3.15. 4 metre mesafe test sonuçları	85
Tablo 3.16. Dört ve sekiz yön başarı oranları.....	87
Tablo 3.17. Sistemimizin mikrofon sınıfı kullanan algoritma ile karşılaştırılması.....	88
Tablo 3.18. Horlama sesi çalışmaları karşılaştırması	94

SEMBOLLER DİZİNİ

SGO	Sıfır Geçiş Oranı (Zero Crossing Rate)
MFKK	Mel Frekansı Kepstrum Katsayıları (Mel Frequency Cepstral Coefficients)
SAM	Spektral Altbant Merkezleri (Spectral Subband Centroids)
DVM	Destek Vektör Makineleri (Support Vector Machine)
K-EYK	K En Yakın Komşuluk (K-Nearest Neighbor Algorithm)
ÇSF	Çizgisel Spektral Frekanslar (Line Spectral Frequencies)
DÖKK	Doğrusal Öngörülü Kepstral Katsayıları (Linear Prediction cepstral coefficients)
DKK	Doğrusal Kestirim Katsayıları (Linear prediction coefficient)
LAO	Log alanı oranı (Log Area Ratios)
KOK	Kare Ortalamalarının Karekökü (Root Mean Square)
GMM	Gizli Markov Model (Hidden Markov Model)
YSA	Yapay Sinir Ağları (Artificial Neural Network)
FFT	Hızlı Fourier Dönüşümü (Fast Fourier Transform)
SPİ	Ses Parmak İzi (Audio Fingerprint)
TP	Doğru Pozitif (True Positive)
TN	Doğru Negatif (True Negative)
FP	Yanlış Pozitif (False Positive)
FN	Yanlış Negatif (False Negative)
SE	Duyarlılık (Sensitivity)
SP	Özgüllük (Specificity)
AUC	Eğri Altındaki Alan (Area Under Curve)
SNR	Sinyal Gürültü Oranı (Signal/Noise Ratio)
TBA	Temel Bileşenler Analizi (Principal Component Analysis)
ç	Çarpıklık
b	Basıklık
Ymak	Yerel Maksimum
Ymin	Yerel Minimum
KOK	Kare Ortalamalarının Karekökü

1. GENEL BİLGİLER

1.1. Giriş

Teknolojinin gelişmesiyle birlikte, bilgi teknolojisi hayatımızın birçok yönünde etkili bir şekilde kullanılmaktadır. İnsan ve bilgi arasında iletişim sorunları giyilebilir teknolojileri giderek daha önemli hale getirmiştir. Giyilebilir teknolojiler gerçek-zamanlı olarak hesaplama yaparak insan-bilgisayar iletişimini hızlandırmaktadır. Bu sayede, giyilebilir hesaplamalar, ses tanıma sistemlerinde daha yaygın kullanıma başlamıştır.

Ses tanıma için yardımcı cihazların geliştirilmesi önemli bir araştırma alanıdır ve artan bir ilgi görmektedir. Bununla birlikte, çevresel sesleri tanımlamak için özel olarak geliştirilmiş çok az yöntem vardır. Ev ortamlarındaki çevresel sesin sınıflandırılmasını kullanmak, ev otomasyonu veya akıllı evler için yeni bir araştırma konusu haline gelmiştir [1-6]. Ses sınıflarını tanımlamak, ev ortamının izlenmesine önemli derecede yardımcı olabilir. Bir evdeki ortam seslerinin içinde insanın konuşma sesi ve konuşma dışı büyük bir dizi ses sınıfı vardır. Örneğin, bir sistem köpek sesini, alarm sesini ve kapı ziline çaldığını saptayabilir. Aynı zamanda sistem aile fertlerinden gelen ağlama, bağırma ve çığlık gibi özellikli sesleri tanımlayabilir [7].

Günümüz teknolojisi, dünya nüfusunun büyük bir yüzdesini, onları kaydeden, işleyen ve internet üzerinde onlara ilgi duyan herhangi bir ses parçasını yayınlama olanağı sağlamıştır. Bu yeni ses içeriğiyle ilgili araştırmaların çoğunluğu önemli konuşma ve müzik kategorilerine odaklanmış olsa da, işitsel dünyamızın çoğunu oluşturan doğal ve çevresel sesler genellikle göz ardı edilmektedir. Mevcut yaklaşımların çoğunluğu, konuşma tanıma tekniklerini genellikle yüksek hesaplama karmaşıklığına maruz bırakan görev için uyarlamaya çalışmaktadır.

Dünyanın dört bir yanındaki sağır insanlar birçok iletişim problemiyle karşı karşıya kalmaktadır. İnsanları umumi yerlerde normal insanlara kendilerini ifade etmede zayıflatan konuşma ve işitmeye dair yaşanmış olan çeşitli zorluklar mevcuttur. Dinleme, günlük yaşamımızda hayati bir rol oynayan önemli bir insan becerisidir. Buna ek olarak, insanların reaksiyonları duydukları seslere oldukça fazla bağlıdır. Nitekim duyma engelli insanlar yol boyunca yürürlerken tehditleri tanımlayabilmede birçok zorluğa maruz kalmaktadırlar. Duyma engelli birçok yaya, yol boyunca yürürlerken ya da karşıya geçerlerken hayatlarını kaybetmişlerdir. Dünyada duyma engelli yayalara yardımcı olan birçok cihaz kullanılıyor

olmasına rağmen, bu cihazların çoğu oldukça pahalıdır ve kullanıcıya detaylı bir bildirim sağlamamaktadır. Bundan dolayı, gelişmiş bir duyma engelli veya duyma engelli cihaz çokça tercih edilmektedir.

İşitme, insanlar için çok önemli bir duyumsal işlevdir. İşitme engelli kişilerin yaşam kalitesini iyileştirmek için evde oluşan çeşitli sesleri algılayabilen ve sınıflandırabilen bir uygulama geliştirmek yardıma yönelik temel bir gereklilik olarak kabul edilmektedir. Bir yangın alarmı veya tehlikeye karşı uyarıda bulunan bir telefon sesi tespit edildiğinde bunları algılayan bir sistem duyma engelliler için fayda sağlayacaktır. İşitme, konuşma, okuma ve yazma günlük aktivitelerimiz için önemlidir ve işitme herkesin yaşamında hayati bir rol oynamaktadır. İnsanlar duydukları sese göre tepki verir. Yol boyunca yürürken ya da karşı yola geçerken, duyma engelli insanlar doğuştan engellerinden kaynaklanan birçok zorlukla karşı karşıya gelmektedir. Bundan dolayı, bu tez duyma engellilerin yanlarından ya da arkalarından gelen tehditleri anlamalarında ve çevresinde olabilecek önemli sesleri ayırmalarında yardımcı olabilecek bir cihaz ortaya koymaktadır.

İşitme, insanlar için çok önemli bir duyumsal işlevdir. Bununla birlikte, tüm insanlar işitme duyusuna sahip değildir. Dünya Sağlık Federasyonu tarafından yayınlanan bir rapora göre, sağır insanlar dünya nüfusunun 72 milyonunu oluşturuyorlar [8]. Küresel olarak, Ulusal Sağlık İstatistikleri Merkezi, Amerika Birleşik Devletleri'nde otuz yedi milyondan fazla kişide işitme kaybının bir şekli olduğu, bu oranın nüfusun yaklaşık yüzde on yedisi olduğunu bildirmektedir [9]. Her binde bir çocuk sağır doğuyor veya üç yaşına kadar sağır halde oluyor [10]. Bu insanlar iletişim kurmak için işaret dili öğrenirler. Maalesef, çoğu insan ortalama hareketlerini anlamıyor ve ne söylemeye çalıştığını tespit edemiyor. İşitme kaybı olan kişilerin büyük oranının yaşı altmış beş ya da daha büyük olsa da, birçok insan da hayatı boyunca çoğu sesi duymadan yaşıyor.

Sağır insanların iletişim güçlerini geliştirmelerine, çevrelerine uyum sağlamasına ve toplumda daha etkili bir şekilde işlev görmesine yardımcı olmak için birçok cihaz ve sistem mevcuttur. Uyarı sistemleri veya basitçe sinyal vericiler, telefon sesi, kapı zili, bebeğin ağlaması, hareketi, hava durumu uyarıları veya duman alarmları gibi farklı olayları bildirmeye yardımcı olmak için tasarlanmıştır. Genellikle sağır insanların ev ziyaretçisinin farkında olmaması ve ayrıca ileri yaştaki insanlar yürümekte veya hareket etmede ve kapıda kimin olduğunu görmekte güçlük çekerler. Bu tez çalışmasından çıkan giyilebilir ürün ile sağır insanların ev ortamında ve dış ortamda daha refah bir yaşam sağlaması hedeflenmektedir.

Son yıllarda hem akademik hem de endüstriyel anlamda giyilebilir teknoloji alanında birçok gelişme yaşanmaktadır [11]. Günümüzde sağlık hizmetlerinin pahalı oluşu, mikro denetleyici teknolojisinin gelişmesi, algılayıcı boyutlarının, teknolojinin gelişimi ile ilişkili olarak küçülmesi ve giyilebilir teknoloji dünyasında önemli gelişmeler yaşanması gibi etkenlerden dolayı sağlık alanındaki giyilebilir teknoloji ürünlerine uygulamalara da sık sık rastlanılmaktadır.

Giyilebilen teknolojiler ya da cihazlar, giyilebilecek, kullanıcının aksesuarlarının, elbiselerinin ya da vücudunun içine, üzerine ya da altına yerleştirilebilecek cihazlardır. Bu cihazların hesaba dayalı ve algılayıcı aygıtlar vasıtasıyla geliştirilmesine dair araştırma alanı giyilebilir hesaplama olarak adlandırılmaktadır. 21. yüzyılın başında, giyilebilir cihazlar tüketici tarafından kullanılmaya başlanmıştır. İnsanlar, cep telefonlarından Bluetooth kulaklıklarından kablosuz bir şekilde haberleşmeye başlamışlardır.

1.2. Ses Parmak İzi

Ses parmak izi (SPI), içerik tabanlı bir ses alma tekniğidir. Genellikle büyük bir ses dosyalarından oluşan bir sorgu ses içeriğinin kaynağının belirlenmesinde kullanılır. Ses formatı olarak bilinen, kompakt akustik özelliklerin çıkartılmasıyla, bu teknik, çok sayıda ses dosyasının yalnızca yeni dijital baskı verilerini depolayan bir veritabanı oluşturur. Daha sonra, bilinmeyen bir ses parçası sunulduğunda, özellikleri aynı şekilde hesaplanır ve veritabanında saklanan bu özelliklere karşı eşleştirmek için kullanılır. Sorgu ses içeriğinin parmak izi veritabanında başarıyla eşleşirse, bunlar aynı ses içeriği olarak tanımlanır ve bu ses parçasının meta verileri geri gönderilir.

SPI, ses içeriğinin tam olarak tanımlanması için kullanılan bir teknolojidir. Tipik kullanımı, bir sorgunun potansiyel olarak bozulmuş veya değiştirilmiş olduğu bir ses alıntısı olduğu zaman, kısa bir sorgu göz önüne alındığında, büyük bir koleksiyonun bir parçasını tam olarak tanımlamaktır. Ses parmak izleri endüstrinin ilgisini çeker ve günlük yaşamda iyi bilinen bir araçtır ve milyonlarca kullanıcı bu teknolojiyi kullanmaktadır. Örneğin cep telefonları üzerinden, şu anda oynatılan bir şarkı hakkında bilgi edinmek için sıklıkla kullanılmaktadır. Medya endüstrisi, dünyanın herhangi bir yerindeki çeşitli TV ve radyo yayın kanalları için hangi zaman diliminde hangi içeriğin oynandığının bilgisini toplamak için otomatik SPI alma sistemleri kullanmaktadır.

Ses parmak izleri, etiketlenmemiş sesi içeriğe dayalı tanımlama için yoğun şekilde kullanılmıştır [12- 19]. Ses parmak izleri, genellikle ses sinyalinin hesaplanan özelliklere dayanır. Bu alanda yapılan önceki çalışmalara göre, ideal bir ses parmak izi sistemi birkaç gerekliliği karşılamalıdır [20]. Öncelikle, ses parmak izlerinin dayanıklı olması gerekir. Gerçek zamanlı uygulamalarda gürültüden ve diğer sinyal bozulmalarından etkilenmemelidir. İkincisi, ölçeklenebilir olmalıdır. Veri tabanı, büyüyen büyük boyutlu bir dijital ses kataloğu içermelidir. Üçüncü olarak, parmak izlerinin hesaplanması, veritabanının boyutunu ve uzak servisler için iletim gecikmesini en aza indirmek için kompakt ve efektif olmalıdır. Dördüncü olarak, kısa sorgu parçasının milyonlarca sesli dosyalardan oluşan bir veritabanındaki karşılık gelen belgeyle eşleşmesi için oldukça özellikli olmalıdır. Ve son olarak, veritabanı araştırmalarını gerçekleştirme stratejisi çok etkili olmalıdır. Güvenilir büyük ölçekli ses parmak izi uygulamaları geliştirirken, tüm bu gereksinimlerin ciddiye alınması gerekir.

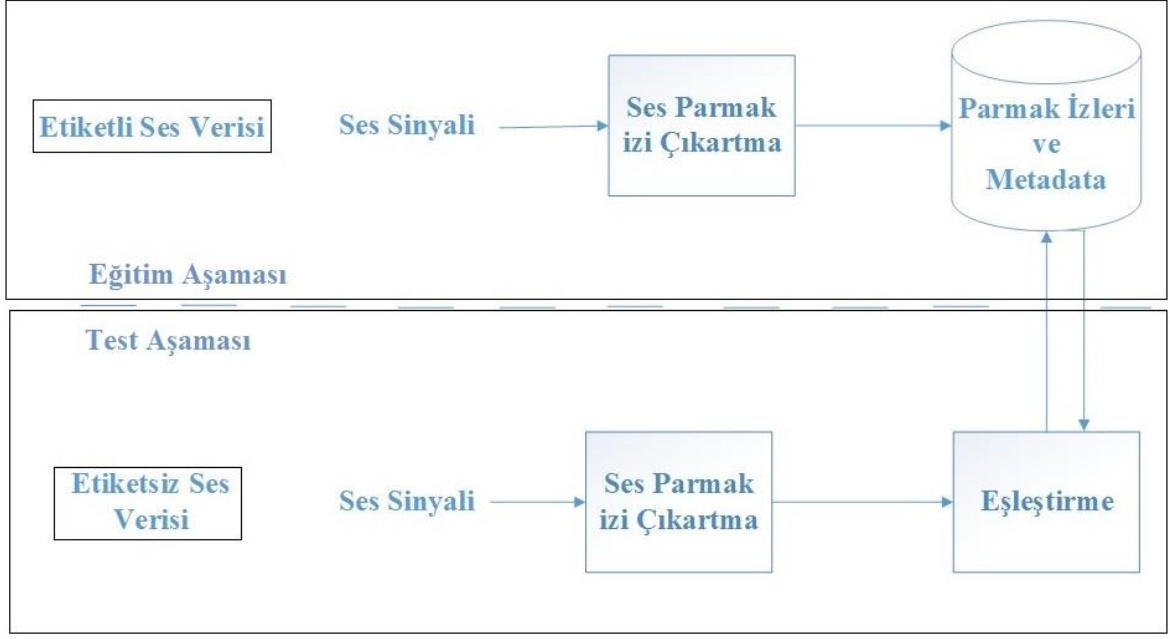
Günümüzde, ses parmak izine dayanan birçok pratik uygulama vardır. Bunlar üç kategoriye sınıflandırılabilirler [20]. Birinci olarak çoğu ülkede, radyo istasyonları bir müzik yayınlamadan önce telif ücreti ödemelidir. Telif haklarının düzgün bir şekilde ödenip ödenmediğinden endişelenen bazı hak sahipleri, müziklerinin izinsiz kullanıp kullanmadığını ses parmak izi yöntemi ile öğrenebilirler. İkinci olarak mobil ses tanıma yapılmaktadır. Buna iyi bir örnek, akıllı telefonlar gibi mobil cihazlarda müzik tanımanın yapılmasıdır. Bir restoran veya evde olduğunuzu ve aniden güzel bir şarkı duyduğunuzu, ancak adını bilmediğini düşünün. Bu ses parmak izi işlemi o şarkıyla ilgili daha fazla bilgi edinmenize yardımcı olabilir. Akıllı telefonlarda Shazam [21] ve SoundHound [22] gibi popüler müzik tanıma uygulamaları kullanılmaktadır. Üçüncü olarak veri bütünlüğü sistemleri olarak söylenebilir. Bazı senaryolarda, ses dosyalarının bütünlüğü, gerçekte kullanılmadan önce doğrulanmalıdır. Dürüstlük, ses dosyalarının değiştirilmediği veya çok fazla bozulma olmadığı anlamına gelir. Bir başka olası uygulama, şirketlerin reklamlarının gerekli uzunluk ve hızda yayınlanıp yayınlanmadığını kontrol etmek istemeleridir.

Bu tez çalışmasında ses tanıma aşamasında SPİ yöntemi kullanılmıştır. SPİ, ses veritabanında karşılık gelenlere veya benzer öğelere hızlı bir şekilde ve ses sinyalinin kısa ve etiketlenmemiş parçasını tanıtmak için kullanılabilir hızlı bir yöntemdir ve ses nesnesinin kısa bir dijital özetidir. Bu, müzik arama, gürültülü ortamlarda sorgulama, sanatçı tanımlama, tür sınıflandırılması, müzik önerisi vs. gibi tüketici cihazları için birçok uygulama veya hizmetler için kullanılabilir [23-28]. SPİ, ses sinyalinin kompakt, düşük

seviyeli bir temsilidir [29]. SPI uygulamaları cep telefonu kullanarak müzik tanımlama, radyo, televizyon ve internet ortamında şarkı belirlenmesi gibi dijital müzik kütüphanesi organizasyonunu içerir. Parmak izi içerik bağımlı işaretleri elde etmek, eşleme sorunları çözmek ve içeriği değiştirilmiş olup olmadığını kontrol etmek için kullanılabilir. SPI, bilinmeyen seslerin içerik tabanlı tahmini için yaygın olarak kullanılmaktadır [29-37].

Çoğu ses çıkarma algoritması temelde ses dosyalarındaki belirgin özelliklerin çıkarılmasına dayanmaktadır ki bu akustik parmak izi olarak adlandırılan bir tekniktir. Ses parmak izi sesin frekans özelliklerinden oluşmaktadır. Bu özellikler, seste gürültü olsa dahi değişmez ya da minimal bir değişim geçirir. İyi bir parmak izi algoritması, sesin diğer seslerden ayırt edebilen ve sesin algısal özelliklerini göz önünde bulundurmalıdır. Bu algısal özelliklerden bazıları, spektrum düzlük, ortalama spektrum, bir dizi frekans bandındaki önemli tonlar, sıfırdan geçiş sayısı ve bant genişliğidir. Frekans bantlarındaki önemli tonlar, ses parmak izlerini tanımladığı düşünülen özelliktir. Parmak izleri daha güçlü, ölçeklenebilir ve sinyali önemli ölçüde tanımlayan özellikleri muhafaza ederler. Frekans analizi de bu bağlamda oldukça ilgi odağındadır. Kısacası, frekans özellikleri bir sinyalin benzersizliğini gösterir.

Teknolojilerin gelişmesiyle birlikte düşük maliyetli, taşınabilir, uzaktan erişimli hasta izleme sistemlerinin geliştirilmesi mümkün olmuştur. Kişilerin vücuduna yerleştirilen ve giyilebilen bu sistemler sayesinde gerçek zamanlı veri analizi yapılabilmektedir. Teknolojide yaşanan gelişmelere paralel olarak insanlar ile bilgisayar arasındaki iletişim için çeşitli yöntemler kullanılmaya başlanmıştır. İnsan ile bilgisayar arasında iletişim kurmak için geliştirilen yöntemlerden biri de ses tanımadır. Şekil 1.1’de ses tanıma işlemi için kullanılan ses parmak izi sistemlerinin genel bir yapısı görülmektedir. Öncelikle etiketli ses verisi eğitim aşamasında ses parmak izi çıkarma öznitelikleri kullanılarak veritabanına kayıt edilir. Bu aşamada sesle ilgili özellikler de tutulur. Test aşamasında ise bilinmeyen ses verisi gelir ve ses parmak izi çıkarma öznitelikleri kullanılarak veritabanında daha önce kayıt edilen veriler ile eşleştirilir. Eğer eşleşme sonunda benzer özellikli ses verisi bulunursa, veritabanından ses ile ilgili özellikler çekilir ve kullanıcıya aktarılır.



Şekil 1.1. Ses tanıma işlemi için ses parmak izi kullanımı

1.2.1. Ses Parmak İzi Özellikleri

Ses parmak izi özellikleri aşağıdaki gibi sıralanabilir: [38]

- Doğruluk: Doğru tanımlamalarının sayısının yüksek olması beklenir. Ses parmak izi özellikle telif hakları ve ses tanıma için önemli bir özelliktir. Yüksek doğruluk oranı ile tespit etmesi beklenir.
- Güvenilirlik: Ses dosyasının yanlış olarak tanınmaması gerekir. Tanımlanacak parça havuzunda olan ya da olmayan bir sorgunun telif hakkını uygulayan organizasyonlar için çalma listesi oluşturmada büyük bir öneme sahip olduğunu değerlendirme metotları güvenilirlik olarak tanımlanabilir. Sistemin güvenilir olması için Yanlış Pozitif (FP) değeri küçük olmalıdır.
- Sağlamlık: Sıkıştırma ve bozukluk seviyesine ya da iletim kanalına müdahaleye bakılmaksızın bir parçayı net bir şekilde tanımlayabilme kabiliyetidir. Yani sağlamlık özelliği, bir ses dosyasının bozulmalarda bile tanımlanmasıdır. Sistemin sağlam olabilmesi için Yanlış Negatif (FN) değeri küçük olmalıdır.
- Taneciklilik: Birkaç saniye uzunluğundaki ses bölümlerinden bütün bir sesi tanımlayabilme kabiliyetidir.

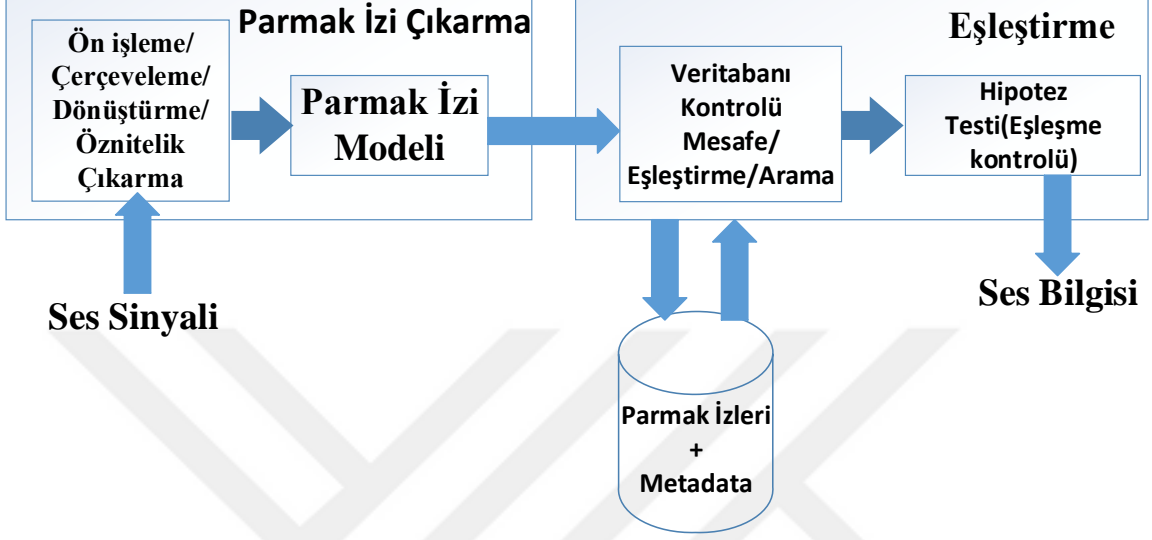
- Güvenlik zafiyeti: Çözümün parçalanabilirlik ya da tahrifata karşı güvenlik zafiyetidir.
- Çok Yönlülük: Ses formatına bakılmaksızın sesi tanımlayabilme ve farklı uygulamalar için aynı veritabanını kullanabilme kabiliyetidir.
- Ölçeklenebilirlik: Oldukça büyük parça veritabanları ve birçok eş zamanlı tanımlama ile gösterdiği performanstır. Bu, sistemin hassasiyetini ve karmaşıklığını etkiler.
- Karmaşıklık: Bu parmak izi çıkarmanın sayısal maliyetleri, parmak izinin büyüklüğü, aramanın karmaşıklığı, parmak izi karşılaştırmasının karmaşıklığı, veritabanına yeni parçalar ekleme maliyeti vb. ile ilgilidir.
- Kırılganlık: Anlam bütünlüğünü doğrulama sistemleri gibi bazı uygulamalar, anlamdaki değişikliklerin saptanmasını gerekli kılabilir. Dayanıklılık gereksinimine karşın, parmak izi anlamı/içeriği koruyan dönüşümlere karşı dayanıklı olmalıdır.

Genellikle parmak izi;

- Kaydın algısal bir özeti olmalıdır. Parmak izi, akustiğe ilişkin bilginin maksimumunu muhafaza etmelidir. Bu özet, büyük bir sayıdaki parmak izinde ayırım yapmaya olanak sağlamalıdır. Bu, karmaşıklık ve dayanıklılık gibi diğer gereksinimlerle uyuşmayabilir.
- Bozulmalarla değişmemelidir. Bu dayanıklılık gereksiniminden ortaya çıkmaktadır. Gelen ses verisi gürültülü ortamda bile olsa yüksek doğrulukla tanımlanmalıdır.
- Kompakt olmalıdır. Büyük bir sayıdaki parmak izinin depolanması ve karşılaştırılması gerektiği için küçük boyutlu bir gösterim karmaşıklığa sebep olur. Ancak, fazlaca kısa bir gösterim kayıtlar arasında ayırım yapabilmek için yeterli olmayabilir ki bu da doğruluğu, güvenilirliği ve dayanıklılığı etkiler.
- Kolay bir şekilde hesaplanabilir olmalıdır. Karmaşıklığın sebeplerinden ötürü parmak izi özeti fazla vakit harcamamalıdır.

Şekil 1.2 'de görüldüğü gibi ses parmak izi sistemlerinde temel iki süreç vardır: parmak izi çıkarma ve eşleme algoritması. Parmak izi çıkartması, bir kaydın alakalı algılamaya özelliklerini kısaca ve sağlam bir biçimde türetir. Parmak izi çok sayıda diğer parmak izi üzerinde ayırmacılık gücü, bozulmalara karşı da çalışabilmesi ve hesaplama basitliği özelliklerine sahip olmalıdır. Eşleme algoritması ise veritabanından benzerliklere bakarak yeni gelen ses verisinin özelliklerini veritabanından çeker. Eşleme algoritması hızlı olmalı

ve ses verisinin az bir kısmından tahmin edebilmelidir [38]. Bir ses parmak izi sisteminde öncelikli olarak parmak izleri oluşturularak yeni sesler öğrenilir. Daha sonra ise bilinmeyen sesleri veritabanında arayarak tanıma ve eşleştirme işlemi yapar.



Şekil 1.2. Ses parmak izi genel sistemi [38]

1.3. Giyilebilir Teknoloji

Giyilebilir teknoloji, bireylerin biyolojik yapısını irdelemek, kullanıcı ortamını kayıt altına almak, kullanıcıların formunu ve konumunu takip etmek ve kullanıcıya bilgi vermek üzere birçok fonksiyona sahiptir. Bütün bu fonksiyonlar kullanıcılara faydalar sağlamaktadır. Giyilebilir teknoloji, sağlık, askeriye, mühendislik ve çevre gibi farklı birkaç endüstri için önemli bir etki ve faydalara sahiptir. Örneğin, sağlık endüstrisinde, Google Glass'ın sağlık sektöründe kullanılmasıyla birlikte giyilebilir teknolojinin kullanımında bir artış görülmüştür. Google Glass, modifiye bir çift gözlük vasıtasıyla kullanıcının video veya resim kaydedebilmesine, elindeki veriyi paylaşmasına ve internet üzerinden bilgi aramasına olanak sağlayan giyilebilir bir cihazdır.

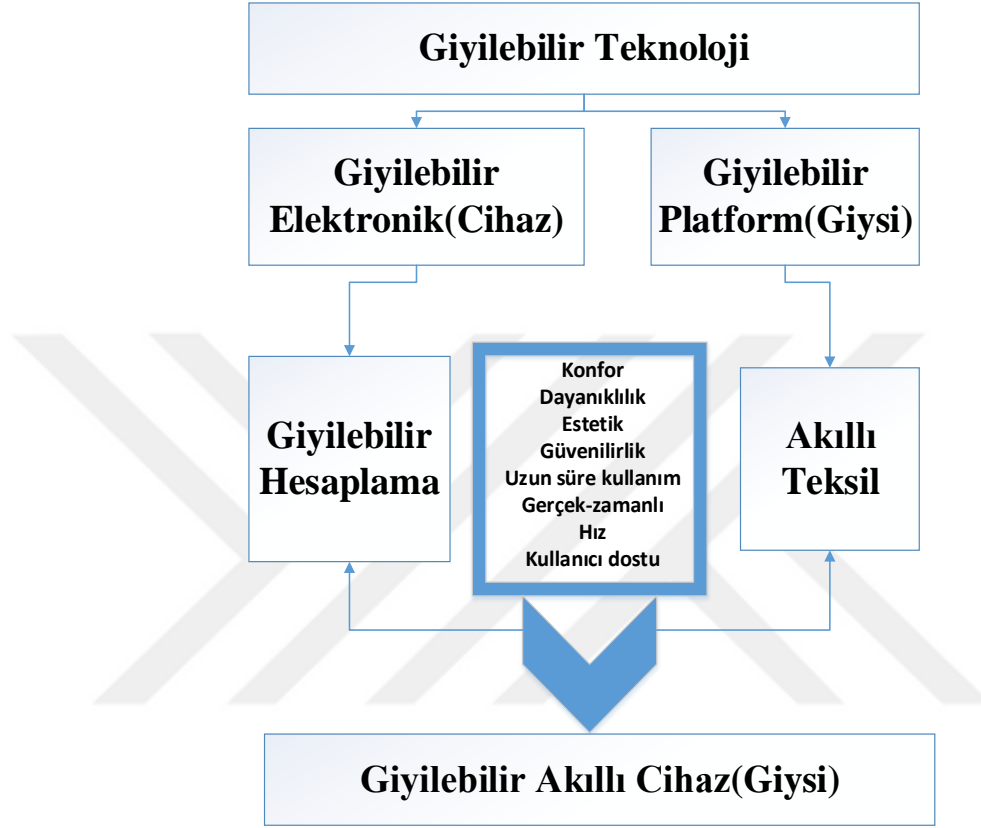
Giyilebilir teknoloji, farkındalığı artırma hususunda eşi benzeri olmayan bir kabiliyet sunar. Bu cihazlar özellikle her bireyi yeni bir kişisel yönetim, kabiliyeti algılama ve durumsal anlayış düzeyine çıkararak, bu cihazları kullananlar hakkında bilgi toplamak için tasarlanmıştır. Bu teknoloji eşsiz bir dizi sensör kullanmakta ve ya mevcut cihaz üzerinden ya da kablosuz/kablolu bir şekilde senkronize bir cihaz veya uygulama vasıtasıyla bilgi

göstermektedir. Bu cihazlar kullanıcıların verimliliklerini en üst noktaya çıkarmalarını sağlayan yeni bir yazılımla güncellenmek için tasarlanmıştır. Bu cihazların algıladığı veri miktarı şaşırtıcı seviyededir; bunlar bir kullanıcının kalp atış hızını veya tansiyonuna odaklanma, bir kullanıcıya ideal kardiyo düzeyine ulaştığında bilgi verme veya bir sonraki kavşakta sağa dönmesini söyleme kabiliyetlerine sahiptir.

Son yıllarda, giyilebilir teknoloji, bilgi ve iletişim teknolojisinin hızla gelişmesi ile yeni bir tür insan-bilgisayar etkileşimi konusudur. Giyilebilir teknoloji, hareketlilik ve bağlantı özelliklerini kullanıcılara bırakır, böylece kullanıcılar çevrimiçi bilgiye kolaylıkla erişebilir ve hareket halindeyken hemen başkalarıyla iletişim kurabilirler [39]. Fiziksel beden bilgilerini izlemek için giyilebilir cihazlar ve sensörler geliştirilmiştir. Algılama teknolojilerinin geliştirilmesi ile çeşitli amaçlara ulaşmak için pazarda çeşitli giyilebilir cihazlar piyasaya sürülmüştür. Tipik bir örnek, kullanıcıların hareket mesafesini ve hareket koşullarını, zaman içinde, bilek tarafından takılan giyilebilir cihazlara dahil edilen hareket sensörleri aracılığıyla sürekli olarak ölçen bir uygulama ve ölçülen sonuçları görüntüleyen giyilebilir cihazlar bulunmaktadır. Giyilebilir alanın en temel uygulamaları, kalori alımını ve tüketiminin ölçülmesi, uyku takibi, kan basıncı ve kalp atış hızıdır. Bu nedenle, giyilebilir uygulamalar çeşitli insan faaliyetlerini günlük yaşamda sayısal olarak niceleştirerek başlamıştır. Son birkaç yıldır, artan performansla amaçlarına göre daha fazla giyilebilir cihazlar geliştirilmiştir. Sonuç olarak, bireysel günlük yaşamlarını kendi başarılarına ölçmek için bu cihazların talebi artmıştır. Bu talebin yanı sıra, gerçek yaşama uygulanması için bireysel koşulları analiz ederek yaşam kalitesini iyileştirmeye yönelik yöntemlerle ilgili daha fazla çalışma yapılmıştır. Hareketleri izlenen hedefler, gerçekleştirilen fiziksel aktiviteler ve çevresel bilgiler gibi çeşitli kişisel bilgileri analiz eden giyilebilir teknolojiler bulunmaktadır [40-42].

Giyilebilir bilgisayarlar insan üzerine takıldığı veya bir giysi üzerine dikildiği için öncelikli olarak konforlu olmalıdır. Kullanan kişinin hareket etme kabiliyetini engellememelidir. Kullanan kişi giyilebilir sistemi kullandığında herhangi bir rahatsızlık hissetmemelidir. Giyilebilir bilgisayarların diğer önemli özelliği dayanıklılığıdır. Herhangi bir çarpma, düşme v.b. durumlarda giyilebilir sistem hasar görmemelidir. Giyilebilir cihazlar estetik olmalı ve görünümü rahatsız etmemelidir. Giyilebilir cihazların en önemli özelliklerinden birisi de güvenilir olmasıdır. Giyilebilir hesaplama sonunda hesaplanan veriler kullanıcıya aktarılması en önemli özelliğidir. Bu sayede özellikle fayda sağlanacak konularda bilgi sahibi olunur ve yaşam kalitesi artabilir. Örneğin duyma engelliler için yön

tespiti hesaplandıktan sonra doğru bir şekilde duyma engelliye aktarılırsa, duyma engelli kişi bu yöne yönelebilir ve yaşam kalitesi artabilir. Giyilebilir cihazlar uzun süreli kullanıma da uygun olmalıdır. Şekil 1.3’de giyilebilir sistemlerin genel yapısı görülmektedir.



Şekil 1.3. Giyilebilir teknoloji genel yapısı

Giyilebilir bilgisayar, belirli görev veya vücut pozisyonuna rağmen çıkışın sürekli algılanabileceği bir veya daha fazla çıkış cihazı ile vücuda bağlanan bir veri işleme sistemidir. Kullanıcıları ile iletişime geçme ve günlük yaşamlarının gerçek anlamda yaygınlaşması potansiyeline sahiptir. Akıllı telefonlar gibi mobil cihazların aksine, giyilebilir bilgisayarlar daima açık, hazır ve her zaman erişilebilir [43].

1.4. Tezin Amacı ve Kapsamı

Bu tez çalışmasında, duyma engellilerin konuşmanın nereden geldiğini ve hangi ses geldiğini algılamasını sağlayan uygulama yapılmıştır. Duyma engelli kişilerde sesin yönünü

bulmak büyük bir problemdir. Duyma engelli bir kişi ses varsa bunu algılamalı ve ani bir durum için yönünü de anlamalıdır. Bu çalışma ile sesin yönü tespit edilmiştir ve bu şekilde duyma engellilerin sesin nereden geldiğini algılaması sağlanmıştır. Örneğin arkasından konuşan bir kişiyi fark edebilir, ani fren sesi varsa bunun nereden geldiğini anlayabilir ve daha güvenli bir şekilde seyahat edebilmektedir. Bu çalışma ile aynı zamanda duyma engelli birisinin yanında bağıarak konuşulup, konuşulmadığı da tespit edilmiştir. Bu sayede panik durumunda olan bir durumu daha erken fark etmesi hedeflenmiştir. Bu çalışma ile kişiler için önem arz eden seslerin; telefon sesi, alarm sesi, kapı zil sesi, korna sesi, fren sesi, köpek havlaması, insan sesi tespiti de hedeflenmektedir. Örneğin duyma engelli kişi uyku halindeyken telefon alarmı çalarsa bunu duyması sağlanacak ve bu sayede daha güvenli bir yaşam standardına sahip olacaktır. Bu çalışmanın en temel amacı ise titreşimler sayesinde ses tanıma sistemi geliştirmektir.

Yön bulma çalışması için geliştirilen giyilebilir sistemde dört mikrofona girişi, iki adet titreşim motoru, dört adet LED bulunmaktadır. Sağ ve sol parmak uçlarına yerleştirilen titreşim motorları sayesinde, sesin hangi yönden geldiği belli titreşim sıklıkları ile gösterilecektir. Ses algılandığı zaman titreşim hangi yönden geldiğini algılasa o çıkışa ait olan LED de yanacaktır.

Gerçek zamanlı çalışacak sistem aynı zamanda elde ettiğimiz ses verileri ön işleme, öznitelik çıkarma ve sınıflandırma adımlarından geçirilecektir. Kullanılacak farklı öznitelikler ve sınıflandırma yöntemleri sayesinde bu alanda araştırma yapacak kişilere katkı sağlayacağını düşünmekteyiz. Sonuç olarak, bu çalışma duyma engelli kişilere faydalı olacağını düşünmekteyiz.

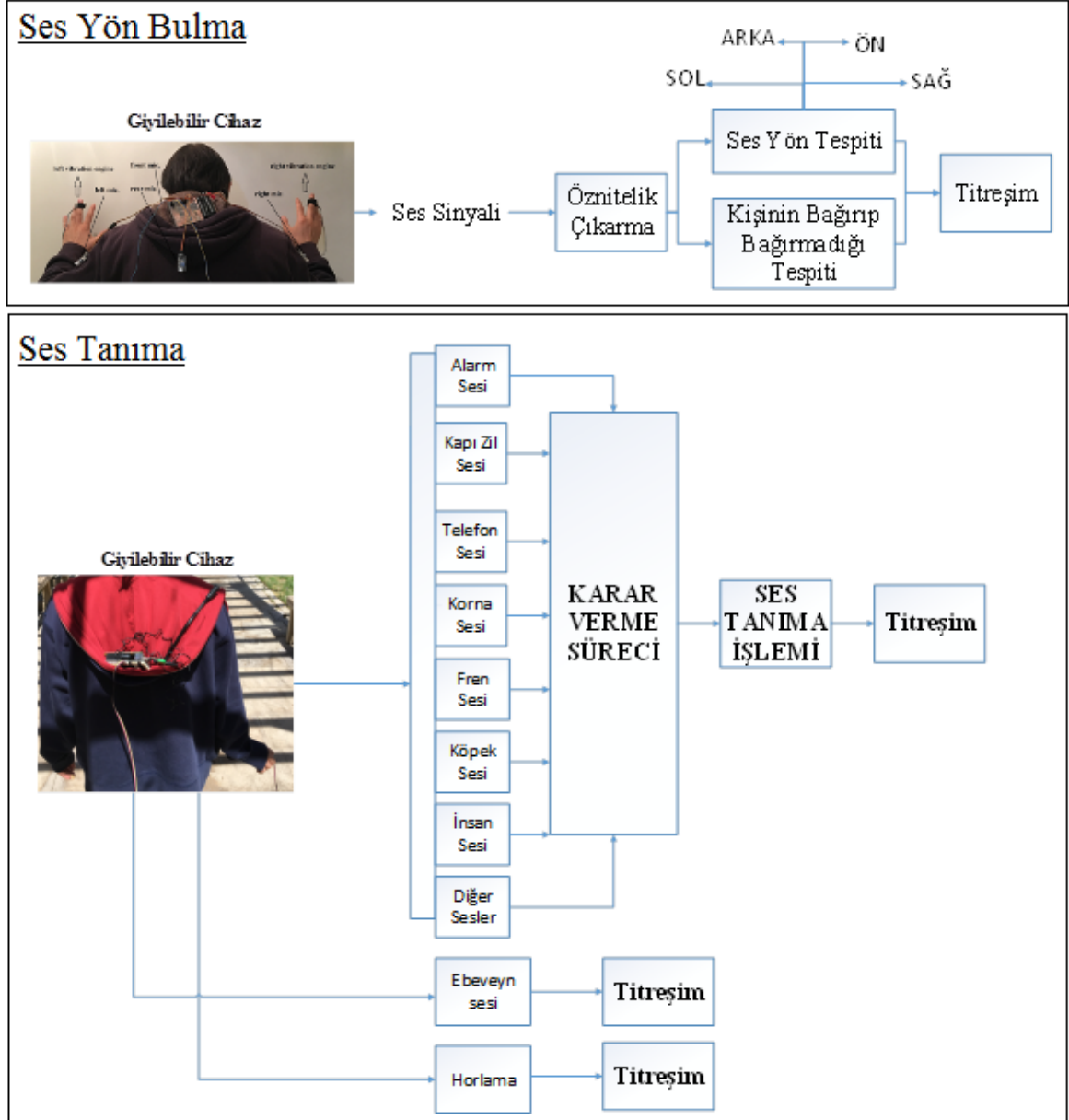
Bu tez çalışmasında giyilebilir bir ürün elde edilmiştir. Yapılan ürün ile ses tanıma, sesin yönü tespiti, konuşan kişinin bağırap bağırmadığı tespiti, ebeveyn sesi tespiti ve önemli seslerin tespiti yapılmıştır. Geliştireceğimiz ürün sayesinde duyma engelliler toplum içinde daha rahat ve güvenilir bir hayat yaşamaları sağlanacaktır.

Bu çalışmanın amaçlarını maddeler halinde özetleyecek olursak;

1. Sesin nereden geldiği tespit edilecektir. Sesin yönü tespit edilerek, o yöne yönelmesi hedeflenmektedir.
2. Telefon sesi, alarm sesi, kapı zil sesi, korna sesi, fren sesi, köpek havlaması, insan sesi seslerinin gerçek zamanlı olarak tespiti hedeflenmektedir.

3. Ses tanıma sistemi yapılması hedeflenmektedir. Sesin yönü tespit edildiği zaman aynı zamanda ebeveynlerden hangisinin konuştuğu tespit edilerek titreşimlerle kişinin algılaması hedeflenmektedir.
4. Bireyin düşük sesle mi yoksa yüksek sesle mi konuştuğunu algılayarak duyma engelli kişinin dikkatini daha da belirginleştirmesini sağlamaktır. Örneğin yanında yüksek sesle konuşan biri olursa bunu algılayabilecek ve olaylara tepki süresi hızlanacaktır. Kişinin bağırıp bağırmadığını belli bir eşik değeri (70 dB(A)) belirlendikten sonra ayırt edebiliyoruz. Bunu yapmamızın temel amacı duyma engelli birisinin olaylara tepki süresini hızlandırmaktır. Yanında bağırarak konuşan birinin kendisine bir sorundan bahsetmiş olması ve buna daha dikkatli davranması gerekebilir.
5. Horlama sesi tespiti yapılacaktır. Horlayan kişilerin horlamadan çıkması için titreşimler verilmiştir.
6. Gerçek zamanlı olarak çalışan bir sistem tasarlanması hedeflenmektedir. Duyma engelli kişilerin geliştirdiğimiz ürünü kullanarak bundan fayda sağlanması hedeflenmektedir.
7. Giyilebilir bir ürün tasarlanması hedeflenmiştir. Giyilebilir ürünü kullanan bireye titreşimler verilerek yönü algılaması ve hangi sesin geldiğini algılaması sağlanacaktır.

Duyma engellilerin hayatını kolaylaştırmaya yönelik bilgisayar ve internet teknolojisi tabanlı birçok önemli çalışma yapılmaktadır. Bu çalışmalar sayesinde duyma engellilerin yaşam kalitesi artırmaya çalışılmaktadır. Duyma engelli kişilerin en büyük problemi sesin nereden geldiğini algılayamamasıdır. Bu çalışmada öncelikli hedefimiz duyma engellilerin sesin yönünü algılaması ve o yöne yönelmesini sağlamaktır. Aynı zamanda bir diğer hedefimiz konuşan kişinin düşük sesle mi yoksa yüksek sesle mi konuştuğunu algılayarak duyma engelli kişinin dikkatini daha da belirginleştirmesini sağlamaktır. Diğer önemli hedefimiz ise duyma engelliler için önemli olabilecek sesleri tahmin ederek titreşim tabanlı giyilebilir sistem sayesinde duyma engelli kişinin algılamasını sağlamaktır. Tezin genel akışı Şekil 1.4'te görüldüğü gibidir.



Şekil 1.4. Tezin genel akış şeması

Tezin birinci bölümünde, ses tanıma, ses parmak izi ve giyilebilir teknolojiler ile ilgili genel bilgiler verilmiştir. Ayrıca, sağır kişiler, ses parmak izi çalışmaları ve ses tanıma çalışmaları ile ilgili literatürde yapılan çalışmalardan bahsedilmiştir.

İkinci bölümde, ses tanıma çalışmasında kullanılan giyilebilir sistemin yapısından bahsedilmiştir. Geliştirdiğimiz titreşim tabanlı giyilebilir cihazımızın özelliklerinden ve kullanılan metotlardan bahsedilmiştir. Gerçek zamanlı cihazımız için en iyi sonuçları aldığımız ses parmak izi yöntemi detaylıca açıklanmıştır. Aynı zamanda kullandığımız diğer yöntemlerde anlatılarak yön bulma çalışması ve yanında konuşan kişinin bağırıp

bağırmediği tespiti çalışmasından bahsedilmiştir. Yön bulma çalışmasında öncelikli olarak sistemin yapısından bahsedilip kullanılan ön işleme, öznelik yöntemi ve sınıflandırma yönteminden bahsedilmiştir. Daha sonra ise ebeveyn sesi tahmini çalışmasından bahsedilmiş ve son olarak horlama sesi tahmini çalışması sistem yapısına yer verilmiştir.

Tezin üçüncü bölümünde, deney ortamından bahsedilmiş ve ses tanıma çalışması, ses yön bulma çalışması, yanında konuşan kişinin bağırıp bağırmediği tespiti çalışması, ebeveyn sesi tahmini çalışması ve horlama sesi tahmini çalışması sonuçları değerlendirilmiş ve literatürde kullanılan değerlendirme ölçütleri ile başarı oranları gösterilmiştir.

Son olarak tezin dördüncü ve beşinci bölümlerinde bu çalışmanın sonuçlarından, tartışma ve önerilerden ve gelecekteki çalışmalardan bahsedilmiştir.

1.5. Literatür Çalışması

Son on yıl süresince, konuşma tanıma araştırması için yapılmış çok miktarda çalışma mevcuttur. Konuşma ses dalga formlarının analizi üzerine birçok teknik ve ilerleme sağlanmıştır. Ayrıca, kaydedilen konuşmadaki çeşitli modelleri tanımlayabilmek amacıyla yapılan sınıflandırma için birçok farklı istatistiksel yöntem kullanılmıştır. Konuşma tanıma uygulamalarında gelen sesi tanımlamak için birçok farklı ses özellikleri önerilmiştir [30, 44-47]. Araştırmacılar sesleri tanımlamak için kullanılacak özellikleri sınıflandırmaya çalışmışlardır. Ses özellikleri ise zamansal, enerji, spektral, harmonik ve algılama özellikleri gibi temel gruplara ayrılabilir.

1.5.1. Sağır Kişiler ile İlgili Çalışmalar

Mobil araçlar, bireyselliği, işbirliğini ve sürekliliği koruyarak bilgiye her yerde ve her zaman erişimi sağlar. Mobil araçlarda gelişmeler sadece öğrenme için değil, aynı zamanda özel ihtiyaçları olan öğrencilerin değerlendirmeleri için de bunları uygun hale getirir. Sağır ve duyma zorluğu olan ve öğrenme bozukluğu olan öğrenciler, zengin görsel içerik, titreşimli uyarı ve yanıp sönen ekran gibi mobil araçların öne çıkan özelliklerinden faydalanabilirler. Vassis ve diğerleri [48] çalışmasında, mobil değerlendirme prosedürünü koruma altına almak için standart kablosuz teknolojiler ve açık anahtarlı kriptografi temelinde kablosuz bir sistem önerilmiştir. Ayrıca, öğrencilerin ihtiyaçlarına göre ekran

görünümünü ve test sonuçlarını adapte edebilmek için kişiselleştirme teknikleri uygulanmıştır.

Shivakumar ve Rajasenathipathi [49] çalışmalarında, sağır ve kör insanların titreşimli eldiven vasıtasıyla, son bilgisayar teknolojilerinden yararlanmalarına yardımcı olabilmek için, donanım kontrol prosedürleri ve bir ekran girdi programı kullanarak iletişim amacıyla bilgisayara bağlanmıştır. Eldivende altı farklı pozisyondaki titreşimler Braille koduyla eşleşmiştir. Böylelikle görme engelli kişi İngilizce karakterleri anlayabilmektedir. Bu çalışmada, İngilizce metnin Braille koduna dönüştürülmesindeki yazılım dizaynı, Braille eldiveninin donanım dizaynı, donanım kontrol prosedürleri rakamları kullanarak Braille eldivenindeki altı farklı pozisyonun Braille eşdeğer titreşimlerinin neler olduğu ve Ki-kare testi kullanarak bunun nasıl test edildiği anlatılmıştır. Braille görme engellilerin okuma ve yazmada kullandıkları önemli bir dildir. İletişim ve eğitim amaçlı kullanımları oldukça yaygındır.

Sağır-kör insanların dünyaya açılan pencereleri küçüktür. Yeni teknolojinin bu duruma yardımcı olabilir fakat bu pahalıdır. Arato ve diğerleri [50] çalışmalarında, titreşimli dahili motoru olan Android akıllı telefon kullanarak Macar bir sağır-kör kişi ile SMS mesajları yazmak ve okumak için çok uygun fiyatlı bir yöntem geliştirip bunu test ettiler. Kelimeler ve karakterler titreşimli Braille kodlarına ve Mors kelimelerine dönüştürülmüştür. Mors, kod olarak karakterleri algılamak ve dil olarak kelimeleri algılamak için öğretilmiştir.

Nanayakkara ve diğerleri [51] çalışmasında, dokunsal ve görsel bilgi kombinasyonunun duyma engelliler açısından müzik deneyimini artırmak için kullanılıp kullanılamayacağı, kullanılacaksa nasıl kullanılacağı sorusuna yanıt aramaktadır. Bu çalışmada sunulan kavramların diğer çevresel ses türlerinin görsel gösterim veya dokunsal girdi aletlerine dönüştürülmesinde faydalı olabileceği ve böylece, örneğin, sağır bir insanın zil sesini, arkasından yaklaşan ayak seslerini, ona seslenen birisini duymasını ve konuşmaları anlamasını sağlayacaktır. Ayrıca, prototip sistem bir müzik enstrümanı çalmayı ve akortlu şarkı söylemeyi öğrenmede yardımcı olarak kullanılabilir. Bu araştırma, sağır topluluğun müziği deneyimleme biçimi önemli ölçüde değiştirmede var olan teknolojiyi kullanmadaki önemli potansiyelini göstermektedir.

Gollner ve diğerleri [52] çalışmalarında, sağır-kör insanların iletişimi destekleyecek ve böylece özgürlüklerini pekiştirecek yeni bir iletişim sistemi tanıtmaktadır. Mobil Eldiven Lorm, sağır ve körler için mobil bir iletişim ve çeviri aracı sağlamaktadır. Eldiven, duyma ve görme zorluğu yaşayan insanların kullandığı yaygın bir iletişim formu olan dokunmatik

alfabe Lorm'u metne çevirir ya da bunun tam tersini yapar. Geliştirilen donanım, sağır-kör kullanıcının eldivenin avuç kısmına alıcının elde taşınırına SMS olarak gönderilmek üzere tekstil basınçlı sensörler aracılığıyla mesaj oluşturmasını sağlar.

Schmitz ve Ertl [53] çalışmalarında, sağır-kör kişilerin elektronik haritaları kullanma ve keşfetmelerini sağlamak için, standart gürültülü oyun tablası kullanarak haritaları dokunsal yolla gösteren bir sistem geliştirmiştir. Bu sistem hem alan içi hem de alan dışı kullanım için amaçlanmıştır ve bu yüzden küçük alanların keşfinin yanı sıra daha geniş bölgelerin geniş taslağını almak için mekanizmalar içerir. Bu sayede titreşimler kullanarak dijital haritaları erişilir kılmak hedeflenmiştir.

Ketabdar ve Polzehl [54] çalışmalarında, sesli bir olay olması durumunda sesli içeriği, dokunsal konuyu ve görsel uyarıları analiz edebilen mobil telefonlar için bir uygulama geliştirmiştir. Bu uygulama özellikle çevrelerinde olan sesli olaylar hakkında uyarılmaları adına sağır veya duyma bozukluğu olan insanlar için kullanışlıdır. Sesli içerik analiz algoritması mobil telefonun mikrofonunu kullanarak veri yakalar ve kullanıcının etrafındaki sesli aktivitelerdeki değişimi kontrol eder. Eğer bir değişiklik olursa ve başka diğer koşullarla karşılaşılırsa, uygulama sesli içeriğin değişimiyle orantılı olarak görsel veya titreşimsel-dokunsal uyarılar verir. Bu kullanıcıyı olay hakkında bilgilendirir.

Nanayakkara ve diğerleri [55] çalışmasında kulaktan ses alımı yerine kanallar aracılığıyla bilginin duyuşal girdisini geliştirerek sağırılar için müzik deneyimini zenginleştirmek için tasarlanan bir sistem önerilmiştir. Titreşimli Haptik Sandalye ve müzik özelliklerine uygun olan bilgilendirici görsel efektlerin bilgisayar gösterimi olmak üzere sistemin iki temel bileşeni bulunmaktadır. Haptik Sandalye dokunma yoluyla titreşimlerin duyuşal girdisini sağlar.

Düşük frekanslı titreşimler hem dokunma hem de işitsel sistemler tarafından tespit edilebilir. Caetano ve Jousmaki [56] çalışmalarının amacı tüm kafa derisi manyetoensefalografisi(MEG) aracılığıyla titreşim uyarısının tek başına kişi işitsel alanlarını aktive edip etmeyeceğini tespit etmektir. 11 normal duyan yetişkinden 200 Hz titreşime kadar MEG sinyalleri kaydedip, sağ elin parmak uçlarına iletmişlerdir. Tüm denekler titreşimli tüpe dokunduklarında bir ses algıladıklarını ve tüpe dokunmadıklarında hiçbir şey algılamadıklarını rapor ettiler. Bu çalışmanın sonucunda, doğuştan sağır yetişkinlerde daha önce tespit edilen sonuçlarla uyumlu olacak şekilde, normal duyuşu olan yetişkinlerde işitsel kortekse titreşim uyarılı girdinin yaklaşmasını önermektedir.

Damper ve Evans [57] çalışmalarında çeşitli evsel seslerin (kapı zili, telefon veya yangın alarmı gibi), oluşması durumunda sağır-kör bir kişiyi uyarmak ve bu seslerin kaynağı hakkında bilgi vermek amacıyla geliştirilen bir elektronik sistemi açıklamaktadırlar. Sistem etkili güvenilir, nispeten ucuz, evsel alet-edevata direk bağlantı gereksinimini ortadan kaldıracak ve diğer potansiyel kullanıcı gereksinimlerine uyarlanabilir şekilde tasarlanmıştır. Sistem belli bir çevresel sesi tespit etme ve tanıma durumunda sağır-kör insan tarafından takılan titreşimli taşınabilir üniteye bu bilgiyi aktaran sabit pozisyonlu bir ses analizi ve tanımlama biriminden oluşmaktadır. Titreşimin detaylı örüntüsü sesin kaynağını kodlar ve bu şekilde kullanıcıya bunun kimliğini belirtir.

Heracleous ve diğerleri [58] çalışmasında, normal duyan ve sağır deneklerde GMM kullanarak Fransızca işaretli konuşmanın otomatik tanınmasını yapılmıştır. İşaretli konuşma elleri farklı pozisyonda ve dudak kalıplarıyla kombinasyon halinde kullanarak sağır kişiler tarafından rahatlıkla anlaşılabilir sesler üreterek konuşulan dili aktarmaya yarayan görsel bir moddur. İşaretli konuşmanın amacı dudak okuma problemlerinin üstesinden gelmek ve böylece sağır çocukların ve yetişkinlerin konuşulan dili tamamen anlamalarını sağlamaktır

Jeyalakshmi ve diğerleri [59] çalışmalarında, normal çocuklar ve işitme engelli çocuklara yönelik olarak konuşma tanıma sisteminin performansı sunulmaktadır. MFKK öznitelik çıkarma ve GMM kullanılarak bir sistem geliştirilmiştir. Sistemin doğruluğu işitme engelli konuşmalar için % 92.4 ve normal konuşmalar için % 98.4'tür.

Lozano ve diğerleri [60] çalışmasında, işitme özürü kişiler için ciddi sorunlarla karşılaşmasına sebep olabilecek günlük durumları çözmek amacıyla uygulama geliştirmişlerdir. Bu çalışma, yüksek doğruluk ve güvenilirlik ile kapı zili, çalar saat, bebek ağlaması gibi seslerin sınıflandırılmasına ilişkin bir dizi parametre elde etmek için yürütülen çalışmanın ön sonuçlarını sunmaktadır.

Beskow ve diğerleri [61] çalışmalarında, evde işitme projesi geliştirmişlerdir. Proje, işitme engelli kişilerin ev ortamlarındaki ihtiyaçlarına odaklanmaktadır. İşitme desteği için ses yükseltme, ses azaltma, ses sınıflandırma ve olay algılama gibi konuşma ve ses algılamayı destekleyen çeşitli entegre özelliklerle ve gerçek zamanlı konuşma okuma desteği sağlayan yenilikçi bir medya merkezi çözümü geliştirmişlerdir. Sınıflandırıcının ortalama doğruluğu %82'dir.

Carrera ve diğerleri [62] çalışmalarında, sağır-kör bir kullanıcı ile görüşülen kişinin vibrotaktik cihaz vasıtasıyla iletişiminin canlılığını değerlendirmek üzere tasarlanmış bir eldiven sunulmaktadır.

Sakajiri ve diğerkleri [63] çalıřmalarında, iřitme engelli kiřiler için, dokunmatik bir ekran aracılıđıyla řarkılarına yardımcı olmak amacıyla bir sesli adım kontrol sistemi geliřtirmiřlerdir.

Shull ve Damian [64] çalıřmasında giyilebilir haptik cihazların rehabilitasyon, vestibüler kayıp, görme kaybı ve iřitme kaybı gibi çeřitli klinik uygulamalar için fonksiyon geliřtirdiđini belirtmiřlerdir.

Mirri ve diğerkleri [65] çalıřmalarında, Malossi alfabesi kullanılarak duyma-görme engelli insanları iletiřimde destekleyebilmek için giyilebilir bir cihazdan faydalanan düşük maliyetli ve açık kaynaklı bir destek sisteminin tasarımı ve uygulanmasını ortaya koymaktadırlar. GlovePi olarak adlandırılan sistem, bir eldiven, bir Raspberry Pi ve geniřletme kartıyla birlikte bir MPR121 kapasitif dokunmatik sensor modülü gibi üç düşük maliyetli temel malzemeden meydana gelmektedir. MPR121 modülü, veri transferine olanak sađlayarak eldivendeki sensörler ve Raspberry Pi arasında bir köprü gibi çalıřır. Buna karřılık, Raspberry Pi bir Android uygulamasına veri transfer eden bir sunucu gibi çalıřır. Bu řekilde, duyma-görme engelli kullanıcı Malossi alfabesini kullanarak diđer kullanıcılara mesaj göndermek için bu eldiveni kullanabilir. GlovePi, duyma-görme engelli insanlar için destekleyici bir teknoloji iřlevi gören giyilebilir bir cihazdır. GlovePi, bir asistan ya da tercümana ihtiyaç duyulmaksızın kullanıcıların bařka insanlarla doğrudan etkileřime girebilmelerini sađlayan ve dünyanın geri kalanıyla otonom olarak iletiřim kurabilmelerini sađlayan ve kullanıcı dostu düşük maliyetli bir çözümdür. GlovePi, temel olarak bir Raspberry Pi'ye bađlanmış sensörlerle donatılmış bir eldiven ve mobil cihazlar için bir uygulamadan oluşur.

Jadhav ve Joshi [66] çalıřmalarında normal insanlar ile konuřması zayıf insanlar arasındaki iletiřim kopukluđunu azaltmak için, sistem tasarlamıř ve uygulamıřlardır. Gömülü sistem, parmakların hareketini algılayabilmesi için kullanılan sensörlerin yanı sıra giyilebilir ve algılayabilen eldivenlerden oluşmaktadır. Kelimeleri saptayabilmek için iřaret dili kullanılmaktadır.

Koklear implant (CI) kullanıcıları da normal iřiten insanların yaptıđı gibi ek dokunsal yardımcılardan yararlanabilir. Zhong ve diğerkleri [67] çalıřmasında dokunsal titreřim kaynađı olarak iki kemiđe demirlemiř (birleřtirilmiř) iřitme (bone-anchored hearing aids - BAHA) kullanılır. Belirli bir mesafe ve açığı korumak için özel bir aygıt ile birbirine bađlanmış iki BAHA'nın her ikisinin yönlü mikrofonları vardır ve biri ön sol tarafa diđer i ise ön sađ tarafa programlanmıřtır.

Kalıcı işitmesi olmayan, ancak düşük frekans bilgilerinde mevcut ipuçlarından faydalana bilecek çok sayıda CI kullanıcıları vardır. İnsanın haptik duyusunun frekans hassasiyeti düşük frekanslarda insan akustik işitmesinin frekans hassasiyetine benzer olduğu için, Wang ve diğerleri [68] çalışmalarında düşük frekanslı ipuçlarını iletmek için dokunsal yardımcıların becerisini incelemişlerdir. 5 CI kullanıcısı ve 10 normal işitme katılımcısı, sözcük dağılımının düşük öngörülebilirliği için tasarlanmış uyarlamaları sağlar ve heceleme frekansıyla dengelenmiş ampirik ifadeleri kullanarak kelime segmentasyonundaki doğru ve yanlış kelimelerin oranını hesaplar. BAHA kullanılarak yapılan sonuçlar gösteriyor ki, dokunsal yardımcı ile doğru kelimelerin oranı üzerinde küçük ancak fakat anlamlı bir iyileşme vardır ve kelime segmentasyon hataları azalmıştır. Bu bulgular sözcük segmentasyonunun algısal görevinde dokunsal bilgilerin yararını desteklemektedir.

Ching ve diğerleri [69] çalışmalarında sağır tarafından ortam ve uyarı bilgi seslerinin izlenmesine izin veren görsel bir ekranın tasarlamış ve değerlendirmişlerdir. Sağır insanlara çevre sesi hakkında bilgi vermek için görsel ekran tasarlamışlardır. Tasarımları, zil sesleri gibi konuşma dışı ses üzerine odaklanmaktadır.

Sağır ve işitme engelli insanlar, kapı zili, çocuk ağlaması veya telefon sesi gibi ev içi seslerin varlığını tanımlamada günlük zorluklarla yüzleşirler. Kumari ve diğerleri [70] çalışmalarında, evlerinde yalnız yaşayan işitme engelli insanlar için kapı zili çaldığını bildiren düşük maliyetli bağımsız bir cihaz tasarlamış ve uygulamışlardır. Sistem kamera, vibratör, kablosuz GSM ve bluetooth'u içeren Raspberry Pi'yi temel almaktadır. Ziyaretçi kapı ziline bastığında, kapıda doğru kişinin ya da hırsız olduğunu tanımaya yardımcı olarak çekilen görüntü, giyilebilir cihaza aktarılır. Resim aktarıldıktan sonra, giyilebilir cihaz titreşerek bildirir. Ayrıca, mesaj ev sahibine GSM üzerinden gönderilir. Daha sonra bilgi almak için ziyaretçinin görüntüsü tarih ve saat ile birlikte sunucuya gönderilir. Sistem, güvenilir, etkili ve kullanımı kolaydır ve aynı zamanda kullanıcının güvenliğini de artırmaktadır.

Matthews ve diğerleri [71] çalışmalarında, çevresel seslerin görsel görüntülerinin tekrarlamalı bir şekilde incelenmesini sunmuşlardır. Katkılarından biri sağır olan insanlar için ortam seslerinin ne olduğu konusunda zengin bir anlayış sağlamalarıdır.

Gao ve diğerleri [72] çalışmalarında sağırın başkalarıyla iletişim kurmasına yardımcı olmak için tasarlanmış bir sistem sunulmuştur. Tasarımda ve uygulamada bazı yararlı yeni fikirler önermektedir. Değişmeyen özelliğin imzalayıcı konuma çıkarılması amacıyla geometrik analize dayalı bir algoritma sunulmuştur. İşaret sinyallerinin veri

akışından alt sözcükleri otomatik olarak bölümlenmek için kombine bir yaklaşımı kullanılmıştır. Deneyle, bu çalışmada önerilen teknikler, tanıma hızı veya tanıma performansı üzerinde etkili olduğunu göstermektedir.

1.5.2. Ses Parmak İzi Çalışmaları

Son on yılda, birkaç parmak izi sistemleri geliştirilmiştir. Cano ve diğerele [38] parmak izi algoritmaları üzerinde yapılmış çalışmasında sağlamlık, çarpışma direnci ve veritabanı arama verimliliği üzerinde durulmuştur. Sağlamlık, müziğin bozuk parçasının parmak izi, bozulmamış kayıttın parmak izine yeterince yakın olmalıdır. Çarpışma direnci, müziğin iki farklı parçalarının parmak izi yeterince farklı olmalıdır. Veritabanı arama verimliliği, ölçeklenebilir veritabanı muhafaza etmek için, parmak izi gösterimi verimli veritabanı aramasına izin vermelidir.

Ses parmak izi bulma çalışmalarında son zamanlarda popüler olmaktadır. Özellikle kurumsal büyük firmalar bu alanda algoritmalar geliştirmektedir. En iyi bilinen herkesin kullanımına açık ses parmak izi algoritması Shazam'dır [19]. Shazam lokal ses parmak izlerine dayanmaktadır. Shazam ile insanların akıllı telefonlar kullanarak, aradıkları şarkıyı bulabilirler. Shazam ses sinyalinin spektrogramında gözlemlenen dorukları şarkının yerel özellik noktaları olarak kullanır. Çıkarılan parmak izlerinin ses sıkıştırma ve gürültü durumlarda diğerele türlerine karşı son derece sağlam olduğu gösterilmiştir [73].

Ses parmak izi algoritmalarında spektrogramlar, sinyal ve görüntü işleme yöntemleri sıklıkla kullanılmıştır. Genel parmak izlerine dayalı ses sinyalinin spektrogramından çıkarılmasına dayalı algoritma Haitsma ve Kalker [13] tarafından sunulmuştur. Baluja ve diğerele [74] tarafından dalga tabanlı ses parmak izi algoritması Waveprint önerilmiştir. Son zamanlarda, Google ses arama sistemi Waveprint parmak izi yöntemi kullanmıştır. Waveprint anahtar algoritma dalgalara dayanır ve bu yöntem sesin bir bit hızı değişimi konusunda sağlamdır.

Zhu ve diğerele [75] değişmez ölçek özelliği dönüşümü tabanlı algoritma sunmuşlardır. Chunk ve Ko [76] çalışmalarında parmak izlerinin sayısını azaltabilir yeni bir yöntem önerdiler. Önerilen yöntem genellikle görüntü sinyali işleme sırasında özellik çıkarımı olarak kullanılan Gauss farkını kullanır.

Cano ve diğerleri [38] doğruluk, güvenilirlik, sağlamlık, ayrıntı düzeyi, güvenlik, çok yönlülük, ölçeklenebilirlik, karmaşıklık ve hassaslık gibi ses parmak izi özelliklerine sahip olan bir ses tanımlama sistemi tanımlamıştır.

Rein ve Reisslein [77] dik dalga dispersiyon vektörü ve sinir ağları kullanılması yoluyla tanımlanamayan bir klasik müzik kompozisyonunun tanımlanmasını sağlayan ses parmak izlerinden kullanılması için bir yöntem önermiştir.

Ellis ve diğerleri [78] çalışmasında spektrogram noktaları kullanılarak parmak izi üretilmiştir. Serra ve diğerleri [79] ses sinyalinin bir iki boyutlu temsiline dayanan ses özelliği çıkarma algoritması önermiştir. Chung ve Ko [80] çalışmalarında görüntü sinyali işleme sırasında özellik çıkarımı olarak kullanılan Gauss farkını kullanmışlardır. Çalışmalarında önerilen yöntem ve dinamik zaman eğilmesini kullanmışlardır ve deneysel aramalarda aynı sesi %90 başarı oranı ile belirlemişlerdir. Pan ve diğerleri [81] çalışmalarında ses parmak izi üretmek için yerel enerji ağırlık merkezini kullanmışlardır. Jiang ve diğerleri [82] çalışmalarında ses parmak izi için gerçek zamanlı doruk keşfedilmesi yöntemi sunmuştur. Wang [19] çalışmasında, her ses kaydı frekansındaki yerel belirgin doruklarına bulmak için kısa süreli Fourier dönüşümü kullanılarak analiz etmiştir. Ouali ve diğerleri [83] çalışmalarında veritabanında büyük bir müzik koleksiyonundan sesin dönüştürülmüş kopyasını tespit eden ses parmak izi sistemi önermişlerdir.

Badii ve Thiemert [84] ses, görsel, metinsel, meta-veri ve çevre modülü dahil çok türlü taşıma parmak izi için bir çerçeve sunmuşlardır. Saracoglu ve diğerleri [85] içerik tabanlı kopya algılama için görsel-işitsel parmak izini geliştirmiştir. Ayrı ayrı ses ve görüntü parmak izi eşleştirmesinden elde edilen güvene göre en uygun sonucu seçmektedirler. Cheung ve Zakhor [86] çalışmalarında karşılık gelen çerçeveler elde edilen parlaklık histogramları arasındaki mesafeyi ölçen bir yaklaşım geliştirdiler. Oostveen ve diğerleri [87] diferansiyel blok parlaklığına dayalı temel algısal özelliklerini çıkarmak için bir yöntem belirttiler. Hampapur ve Bolle [88] video parmak izi gibi baskın renk ile gradyan büyüklükleri boyunca ağırlık merkezini önerdi. Seo ve diğerleri [89] ses parmak izi araştırmasına normalize spektral bant momentlerini uyguladı. Bu önerilen altbant özellikleri arasında birinci dereceden normalize edilmiş momentin parmak izi için en iyi performans olduğunu gösterdi. Yuan ve diğerleri [90] tarafından spektral altbant merkezi ses parmak izi olarak seçilmiştir. Park ve diğerleri [91] frekans filtreleme teknikleri ile Philips ses parmak izi düzeni yöntemini genişletmiştir.

Kim ve diğeri [92] çalışmalarında, akustik ses tabanlı içerik tanımlaması için, işitsel parmak izi sistemi sundular. Önerilen yöntemde, işitsel parmak izi belirgin doruk tespiti için adaptif eşikleme yöntemi ve dönüşüm tabanlı ses çıkarma kullanılarak oluşturulmuştur. Etkili eşleştirme gürültü, yankı, yeniden örnekleme, denkleştirme veya sıkıştırmaya karşı korumak için tekrarlanmayan ön sesin sağlam bir doruk eşleşmeye dayalı özetleme işlevi kullanılarak yapılmıştır.

Fenet ve diğeri [93] yükseklik değişmesinde sağlam olmayı içeren sabit Q dönüşümleri tabanlı parmak izi ile birleştirilmiş özetleme tekniği önerdi. Anguera [94] tarafından spektral noktaları etrafında spektrogram maskeleyen kullanılan yaklaşım önerilmiştir. Kim ve diğeri [95] tarafından ses işitsel parmak izi için modüle edilmiş karmaşık gömülmüş dönüşüme dayalı doruk algılama sunulmuştur.

Vaia ve Nikolaos [96] sistem işleyişini hızlı ve dayanıklı kılabilmek için bir ses parmak izi sistemini kullanmıştır. Bu parmak izlerinin, çeşitli yaygın bozulmalara karşı oldukça dayanıklıdır ve her şarkı segmentini test etmek için yaklaşık 10 saniye harcamaktadır. Veri kümesini sınıflandırmak için aşırı uyumdan dolayı büyük bir veri kümesi için Gauss Karışım Modellerini (GMM) kullanmışlardır.

Ses parmak izi algoritmalarında birçok öznelik yöntemi kullanılmıştır. Tablo 1.1'de ses parmak izi yöntemlerinde kullanılan öznelikler görülmektedir.

Tablo 1.1. Ses parmak izi yönteminde kullanılan öznelikler

Yapılan Çalışma	Öznelik
Saracoglu ve diğeri [85]	Düşük düzey öznelik çıkarma yöntemleri, altbant enerji farkları
Haitsma ve Kalker [13]	Altbant enerji
Lebosse ve diğeri [97]	Yerel maksimum renk enerjisi, enerji zirveleri etrafında altbant enerji
Park ve diğeri [91]	Altbant enerji farkları
Younessian ve diğeri [98]	Altbant enerji değişiklikleri
Mukai ve diğeri [99]	Vektör nicemleme (VQ) altbant enerjisi
Nagano ve diğeri [100]	VQ normalleştirilmiş spektrum
Ke ve diğeri [101]	Viola-Jones filtreleri
Ramalingam ve Krishnan [102]	Kısa zamanlı Fourier dönüşümü özellikleri, Gauss karışım modelleri, MFKK, Shannon entropisi, Renyi entropisi, spektral merkez, spektral bant genişliği, spektral düzlük ölçüsü
Allamanche ve diğeri [103]	MPEG-7 düşük seviye tanımlayıcıları
Cotton ve Ellis [104]	Gabor sözlüğü elemanları çiftleri
Ibarrola ve Chaves [105]	Entropi delta
Baluja ve Covell [106]	Haar dalgacık katsayısı momenti

Tablo 1.1'in devamı

Yapılan Çalışma	Öznitelik
Baluja ve Covell [74]	Haar dalgacık katsayısı maksimum
Kim ve Yoo[107]	Güçlendirilmiş spektral altbant momentleri
Jang ve diğerleri [108]	Arttırılmış filtreler
Liu ve diğerleri [73]	Modülasyon frekansı özelliklerinde meydana gelen değişiklikler
Sukittanon ve Atlas [32]	Modülasyon frekansı özellikleri
Burges ve diğerleri [12]	Konvolüsyonel sinir ağı
Son ve diğerleri [23]	Haitsma parmak izi maskesi
Uchida ve diğerleri [109]	Filtre bankası enerji farklılıkları
Jegou ve diğerleri [110]	Niceleşmiş mel spektrografik enerji değerleri
Khemiri ve diğerleri [111]	GMM, MFKK
Ngo ve diğerleri [112]	MFKK VE SIFT tabanlı görsel söz dizimi sözcükleri
Seo ve diğerleri [113]	Spektral altbant merkezleri
Seo ve diğerleri [89]	Spektral altbant momentleri
Six ve Leman [114]	Spektral tepe üçlüsü
Sonnleitner ve Widmer [115]	Spektral tepe dörtlüsü
George ve Jhunjhunwala [116]	Spektral tepeler
Fenet ve diğerleri [93]	Spektral tepe çiftleri
Shi ve diğerleri [117]	Spektral parlaklık maksimumu
Anguera ve diğerleri [94]	Spektral zirve, yerel enerji farklılıkları
Dupraz ve Richard [118]	Spektral tepeler
Ouali ve diğerleri [119]	İkili spektrogram

1.5.3. Ses Tanıma Çalışmaları

Wang ve diğerleri [120] çalışmalarında ev otomasyonu uygulamaları için hızlı gürültü bastırma yaklaşımını kullanan bir çevresel ses tanıma sistemi önermektedir. Ses sınıflandırma modülünde, dalgacık alt uzayına dayanan ve bir sinyalin dalgacık paket dönüşümünde on yedi kritik banttan türetilen bir öznitelik çıkarma yöntemini kullanmışlardır. Gürültüsüz ortamda %90.63 lük bir başarı elde etmişlerdir, SNR 15 dB olduğunda başarı oranı %88.51, SNR 5 dB olduğunda ise %85.10'luk başarı elde edilmiştir.

Salomon ve Bella [121] çalışmalarında çevresel seslerin sınıflandırılması için derin konvolüsyonel sinir ağları mimarisi önermişlerdir. Veri kıtlığı sorununun üstesinden gelmek için ses veri genişletme yöntemi kullanılmışlardır.

Kugler ve diğerleri [122] çalışmalarında, gömülü uygulamalara odaklanarak tasarlanmış, çevresel seslere ayrılmış bir ses tanıma yöntemi önermektedirler.

Kingsbury ve diğeri [123] modülasyon spektrogram kullanarak gürültülerden arındırılmış güçlü bir konuşma tanıma sistemi üzerinde çalışma yapmışlardır. GMM kullanarak modülasyon spektrogram ile temizledikleri ses örneklerini sınıflandırmışlardır.

Reynolds ve diğeri [124] kimlik doğrulama yapan bir konuşmacı tanıma sistemi geliştirmişlerdir. Sistemde tanıma işlemi için Gauss yöntemi kullanılmıştır.

Oberle ve Kaelin [125] çalışmalarında, GMM kullanılarak dokunmatik işitme cihazları için sinyal tanıma düzeni sunmuştur. Özellikle, her alarm sınıfının gözlem olasılığı yoğunluk fonksiyonunun dört durumlu bir GMM tarafından modellendiği bir maksimum olasılık sınıflandırıcısı önerilmektedir. Performans, dört tipik alarm sınıfından 205 alarm sinyali içeren bir veritabanı kullanılarak değerlendirilir ve geleneksel bir minimum mesafe sınıflandırıcısı ve bir sinir ağı yaklaşımı ile karşılaştırılmıştır. Belirtilen alternatifler ile karşılaştırıldığında GMM tabanlı sınıflandırıcının üstün tanıma performansı göstermiştir. Sunulan tanıma programı, düşük hesaplama maliyetleri nedeniyle gerçek zamanlı uygulama için uygundur.

Seoane ve diğeri [126] çalışmalarında farklı yürütme evreleri, giyilebilir biyomedikal ölçüm sistemleri, deney düzeneği, biyomedikal sinyal analizi ve gerçekleştirilen konuşma işlemi yapabilen proje geliştirmişlerdir.

Reed ve Delhome [127] çalışmalarında, çevresel seslerin giyilebilir bir dokunsal yardımıyla tanımlanabileceğini araştırmışlardır.

Alias ve diğeri [128] çalışmasında, en olağan ses sinyallerini analiz etmek için geliştirilmiş konuşma, müzik ve çevresel seslerin özneliklerini çıkarma teknikleriyle ilgili güncel bir inceleme sunmaktadırlar.

Eronen [129] çalışmasında, Mel-frekansı, kepsralin lineer tahmini ve delta kepsral katsayıları gibi müzik enstrümanının tanınmasında hesap için birkaç özellik kullanmıştır. Çalışmalarında hem enstrüman ailesini hem de tekil olarak enstrümanı test etmiştir. Akustik ve sentetik solo tonların tanınmasında en iyi performans, tekil enstrümanlar için %32 ve enstrüman aileleri için %57 olmuştur. Bu sonuçlar, kararlı ve başlangıç özellikleriyle iki MFKK dizisinden oluşan bir özellik dizisiyle elde edilmiştir.

Barbedo ve Tzanetakis [130], spektral ve temporal içeriklerindeki çakışma sorununa bir çözüm sunmuştur. İzole kısımları tanımlayarak enstrümanlar arasındaki spektral uyumsuzluğu irdelemişlerdir. Onların teknikleri, herhangi bir enstrümanın içeriğinin izole gözüktüğü zaman frekans alanlarındaki bölgeleri aramaktır. Deneylerini 25 enstrümanla yürütmüşlerdir.

Lin ve diğeri [131] çalışmalarında Gabor dalgacık özelliklerine dayalı olarak bir ses sınıflandırma ve bölütleme yöntemi önermişlerdir. Bu yöntem, önce bir ses akışını her biri bir saniyelik ses bilgisi içeren kliplere bölecektir. Ardından, her bir klip iki sınıftan veya beş sınıftan biri olarak sınıflandırılır. Önerilen yöntemin çeşitli ses kaynakları için sağlam olmasını sağlamak için, en yüksek ayırmacılığa sahip özellikleri elde etmek için Fisher Lineer Ayırmacı kullanmışlardır.

Chen ve Soong [132] çalışmasında ses tanınması için 5000 çok kullanılan Çin kelimesini GMM kullanarak tahmin etmişlerdir. El-Ramyl ve diğeri [133] Arapça ifadeler için ses tanıma işlemini YSA kullanarak yapmışlardır. 2 farklı ifade için %91 ve %82 lik başarı oranları elde etmişlerdir. Yaniv ve Burshtein [134] çalışmalarında ses tanıma için dinamik zaman çözümlü modeli için GMM ve k-ortalamlar kullanmışlardır. Lin ve Wei [135] çalışmalarında konuşma duygu tanıma sistemi için DVM, GMM, K-EYK, MFKK kullanmışlardır. DVM kullanarak %88.9 luk başarı elde etmişlerdir. Debyeche ve diğeri [136] çalışmalarında ses tanıma için YSA kullanarak %89 luk başarı elde etmişlerdir. Kruger ve diğeri [137] çalışmalarında GMM ve DVM tabanlı ses tanıma sisteminde %92.23 lük başarı elde etmişlerdir. Lam ve Leung [138] çalışmalarında ses tanıma için YSA kullanarak %90 lık başarı elde etmişlerdir. Wachter ve diğeri [139] çalışmalarında Bayes yaklaşımı ses tanıma işlemi yapmışlardır. Al-Haddad ve diğeri [140] çalışmalarında ses tanıma için hibrit model kullanarak %94 başarı elde etmişlerdir.

Ma ve diğeri [141] çalışmalarında 12 çeşit akustik çevre, 10 farklı akustik ortam için GMM ve MFKK kullanarak %92 başarı elde etmişlerdir. İnsanların, daha önce açıklanan deneylerle aynı test veri setini kullanarak, sesleri sınıflandırma konusunda önceden herhangi bir eğitim almayan kısa ses kayıtlarına dayalı ortamları ne kadar iyi tanıyabildiğini bulmak için yapılan deneylerde ise genel başarı oranı %35 olarak hesaplanmıştır.

Bu tez çalışmasında ise yönü algılamak için sınıflandırma tekniği kullanılmıştır. Literatürde yön bulmak için genellikle lokalizasyon algoritmaları kullanılmaktadır. Yön bulma çalışmasında giyilebilir sistemimiz lokalizasyon tabanlı algoritmalara göre daha hızlı (0.68 sn) sonuç vermektedir. Aynı zamanda daha düşük maliyetle daha başarılı sonuçlar üretilmiştir. Yönü algılamak sadece titreşimler kullanılmıştır ve giyilebilir sistemimizin anlaşılabilirliği kolaydır. Bu çalışmada aynı zamanda yanında konuşan kişinin bağırıp bağırmadığı da tespit edilmiştir. Literatürde bu çalışmaya benzer çalışma bulunmamaktadır. Bu tez çalışmasında önemli seslerin tespiti titreşimler ile ayırt edilmiştir. Ses tanıma

alışması için kullandığımız yöntemde öznitelik olarak spektrogram zirvelerinin karmaları kullanılmıştır. Spektrogramlar oluşturulduktan sonra, doruk noktaları, parmak izi karmaları, takımyıldızı haritasındaki zirve noktalarının ikili kombinasyonlar şeklinde eşleştirilmesiyle elde edilmektedir ve özetleme aşamasında ikili alan kullanılmaktadır. Aynı zamanda giyilebilir sistemimizin hızını artırmak için gelen sesin en ayırt edici olabilecek kısmına kadar alınmıştır. Farklı sesler için ise farklı modlar da önerilmiştir. Ses tanıma çalışmasında yüksek başarı oranı, gerçek zamanlı çalışması ve düşük maliyete sahip giyilebilir bir cihaz tasarlanmıştır. Bu tez çalışmasında ebeveyn sesi tahmini de yapılmıştır. Kullanan kişinin anne ve babasını ayırt ederken evde bulunan diğer erkek ve bayanları sınıflandırmayacak bir sistem tasarlanmıştır. Literatürde giyilebilir olarak tasarlanan ebeveyn sesi tahmini çalışması bulunmamaktadır. Son olarak ise horlama sesi tahmini yapılmıştır. Horlayan bir kişi tespit edildikten sonra horlamasını engellemesi için titreşimler gönderilmiştir. Bu çalışmaların hepsinde giyilebilir ve titreşim tabanlı bir sistem tasarlanmıştır. Gelen ses ya da yön tespit edilerek titreşimlerle hangi ses ya da yön olduğu tespit edilerek kullanıcıya aktarılmaktadır. Duyma engelliler için geliştirilen cihazlar ile duyma engelli kişilerin yaşam kalitesini artırması düşünülmektedir.

2. YAPILAN ÇALIŞMALAR

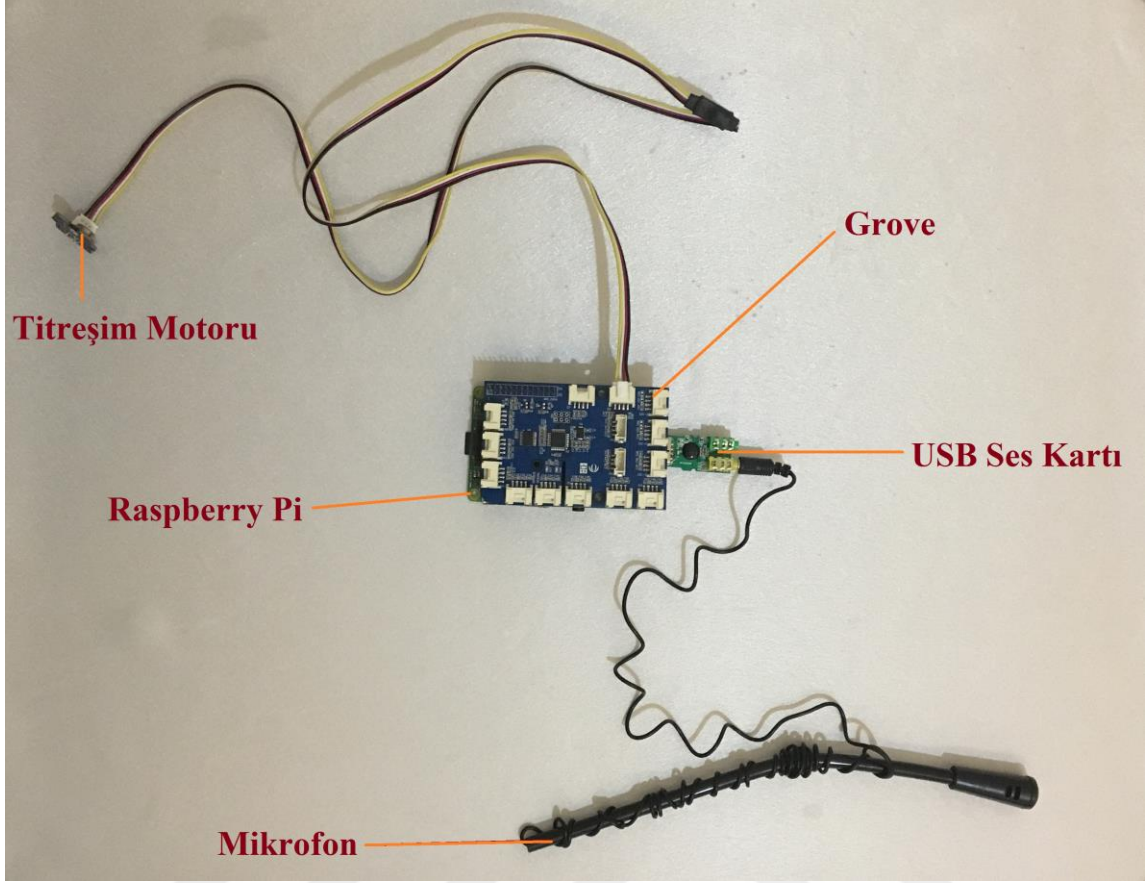
2.1. Ses Tanıma Çalışması

2.1.1. Sistemin Yapısı

Teknoloji devriminde oyun konsolları, bilgisayarlar, akıllı telefonlar, tabletler yer alırken günümüzde sırayı akıllı ve giyilebilir cihazlar almaktadır. Giyilebilir cihazlar özellikle görme ve duyma engelli bireyler için önemli bir yere sahiptir. Giyilebilir cihazlar sayesinde görme ve duyma engelli bireylerin yaşam kalitesi artırılmaya çalışılmıştır. Bu çalışmada, duyma engelli kişilerin ev ortamında ve dış ortamda kullanımını kolaylaştırmak için giyilebilir bir cihaz geliştirilmiştir. Cihazımız kullanıcının sırt bölgesine dikilerek yerleştirilmiştir.

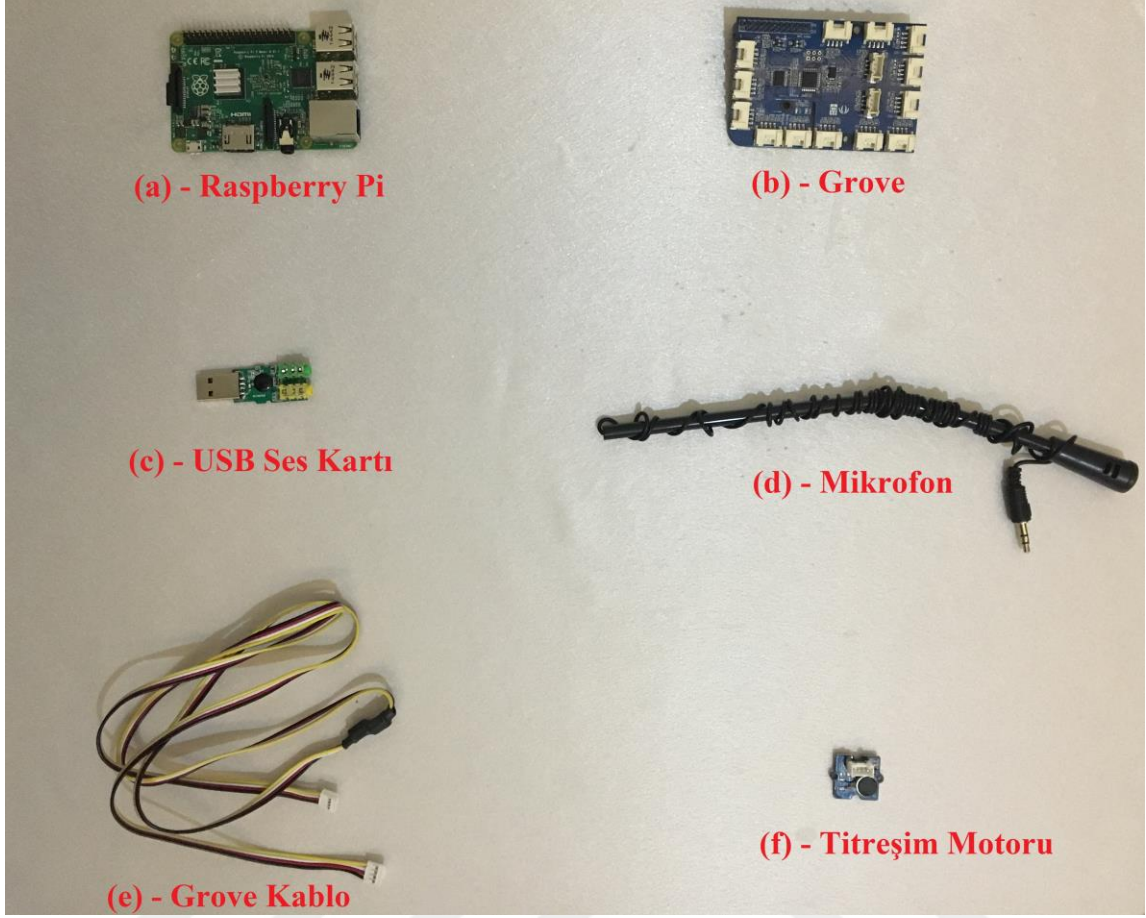
Şekil 2.1'de görüldüğü gibi cihazımız 5 parçadan oluşmaktadır. Ses tanıma çalışması için tasarladığımız giyilebilir cihazda Raspberry Pi, Grove, mikrofon, USB ses kartı ve titreşim motoru bulunmaktadır.

Şekil 2.2 (a)'da görüldüğü gibi bir kredi kartı büyüklüğünde olan Raspberry Pi İngiltere'de geliştirilmiştir. Raspberry Pi, 1.2 GHz 4-çekirdekli 64-bit işlemci, 1 GB RAM bellek, bluetooth, ethernet kartı, HDMI girişi, kamera arabirimi, ekran arabirimi, 4 USB girişi ve 40 genel giriş/çıkış ünitesine sahiptir. Raspberry Pi, okullarda bilgisayar temelli eğitimi geliştirmek ve çocuklara bilgisayarı öğretmeyi amaçlamak için geliştirilmiştir. Kredi kartları ebatlarında, içerisine Linux, Android ve Windows dağıtımları kurulabilmektedir. Raspberry Pi, ARM mimarisine sahip mini bir bilgisayar olarak ifade edilmektedir. Düşük maliyet, güç tüketimi ve küçük boyuta sahip olmasından dolayı Raspberry Pi'nin son zamanlarda popülerliğini artırmaktadır. Bu çalışmada Raspberry Pi'ye, Rasbian işletim sistemi kurularak çalıştırılmıştır.



Şekil 2.1. Ses tanıma çalışması için geliştirdiğimiz giyilebilir sistem

Şekil 2.2 (b)'de görüldüğü gibi Grove, Raspberry Pi'yi Grove sensörlerine kullanmamızı sağlayan ve üzerinde 15 adet sensör bağlantısı olan karttır. Grove, Raspberry Pi üstüne takılarak kullanılır ve Raspberry Pi ile iletişim kurar. Şekil 2.2 (d)'de görüldüğü gibi mikrofon, ortamdaki sesi almamız ve işlememiz için kullanılır. Mikrofon, Raspberry Pi cihazına USB ses kartı sayesinde bağlanmıştır. Raspberry Pi içinde ses kartı olmadığı için sesi almamız için USB ses kartı kullanılmıştır. Şekil 2.2 (f)'de görüldüğü gibi titreşim motoru ise kullanıcıya dokunma duyusunda algılaması için uyarı göndermektedir. Titreşim motoru, Grove modülüne bağlı olarak çalışmaktadır. Her ses için farklı titreşim uyararı verilerek duyma engelli kişinin hangi ses olduğunu anlaması sağlanmıştır. Güç kaynağı olarak şarj edilebilir pil kullanılabileceği gibi USB güç kaynağı da kullanılabilir. Mikrofon, sesi aldıktan sonra Raspberry Pi ortamına örneklendirilerek aktarılır. Mikrofondan ses alındıktan sonra bu sesin hangi ses olup olmadığı SPI yöntemi ile tespit edilmiştir.



Şekil 2.2. Giyilebilir cihaz ekipmanları (a) Raspberry Pi, (b) Grove, (c) USB ses kartı, (d) Mikrofon, (e) Grove kablo, (f) Titreşim motoru

2.1.2. Ses Parmak İzi Yöntemi

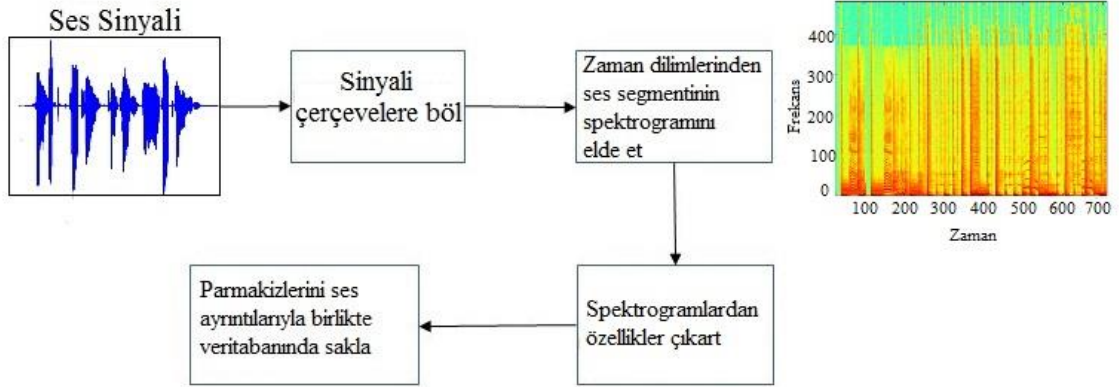
Ses parmak izi (SPİ) , ses dosyasının özetidir ve ses kaydını özetleyen kompakt içerik temelli imzadır. Ses izi, ses veritabanında karşılık gelenlere veya benzer öğelere hızlı bir şekilde ulaşan ve ses sinyalinin kısa ve etiketlenmemiş parçasını tanıtmak için kullanılabilir ses nesnesinin kısa bir dijital özetidir. Son zamanlarda, Google başta olmakla birçok ses arama siteleri aynı sesi aramak için ses parmak izi teknolojisini kullanmıştır. Ses parmak izi genellikle benzer ses dosyaları ve aynıısının içerik tabanlı döndürmesi için kullanılır.

SPİ, en çok ses formatına bakılmaksızın etiketsiz sesi ona karşılık gelen tanımlayıcı veri ile ilişkilendirme kabiliyetiyle bilinir. SPİ ya da içerik temelli ses tanımlama sistemleri, parmak izi gibi ses içeriğinin bir parçasından algısal bir özet çıkarır ve veritabanında bunu depolar. Etiketsiz bir sesle sunulduğunda, parmak izi hesaplanır ve veritabanında

depolananlarla eşleştirilir. Parmak izleri kullanılarak ve algoritmalar eşleştirilerek, bir kaydın karmaşık sürümleri aynı ses içeriği olarak tanımlanabilir.

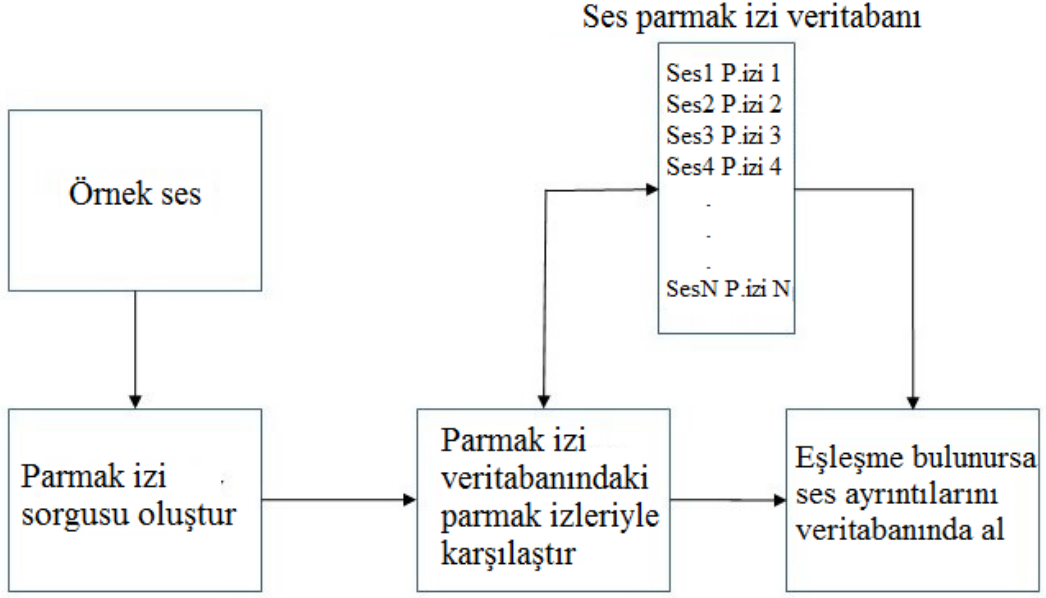
SPI, müzik bilgisini toplama/elde etme ve video tanımlama gibi başka alanlara desteğin yanı sıra ses tanımlama, indeksleme, arama, navigasyon, izleme ve diğer parasal amaçlar için yaygın bir şekilde kullanılmaktadır. Kasıtlı ya da kasıtsız bir ses sinyalini bozmanın kolaylığından dolayı, dayanıklılık ve doğruluk ses parmak izinde hataları önleyebilmek için oldukça önemli özelliklerdir [142].

Parmak izi insanlar için eşi benzeri olmayan bir kimliktir ve aynı teori ses dosyalarına da uygulanabilir. Ses parmak izini çıkarmak, ses dosyaları için ayırt edici imzaların üretilmesi anlamına gelmektedir. Şekil 2.3'de ses parmak izi oluşturma adımları gösterilmektedir. Şekilde görüldüğü gibi öncelikle ses sinyali çerçevelere bölünür. Daha sonra spektrogram elde edilir. Bu spektrogramlardan çeşitli özellikler çıkartılır. En son olarak da parmak izleri ses ayrıntılarıyla birlikte saklanır [143].



Şekil 2.3. Ses parmak izi yapısı [143]

Şekil 2.4'de ses parmak izi puanlama süreci görülmektedir. Yeni bir ses geldiğinde öncelikle bu sesin parmak izi sorgusu oluşturulur. Ses parmak izi veritabanındaki parmak izleriyle karşılaştırılır. Eğer eşleşme bulunursa ses ayrıntılarını veritabanından alarak bu sesin tahmini yapılır [143].

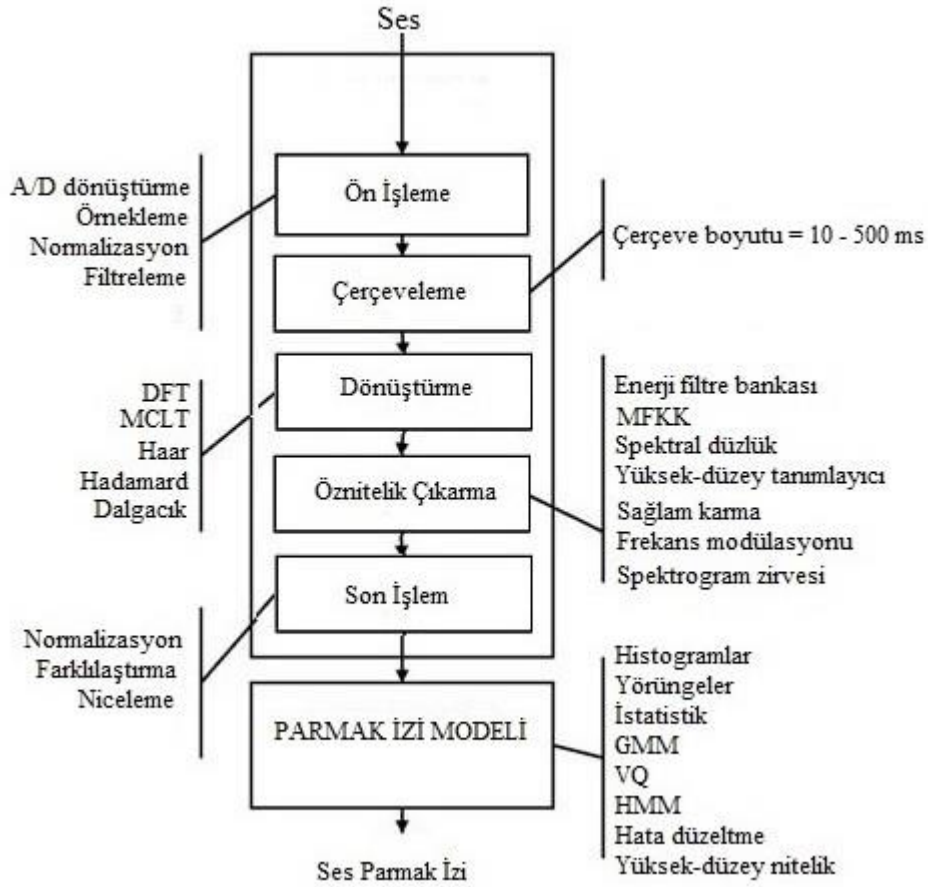


Şekil 2.4. Puanlama süreci [143]

İdeal bir parmak izi sistemi birkaç gereksinimi karşılamalıdır. Sıkıştırma seviyesine ya da iletim kanalındaki müdahaleye bakılmaksızın bir parçayı tam olarak tanımlayabilmelidir. Uygulanmasına bağlı olarak, sadece bir kaç saniyelik kısımlardan sesi tanımlayabilmelidir. Parmak izi sistemi sayısal olarak da verimli olmalıdır. Verimlilik, hem tanımlı olmayan sesin parmak izinin hesaplanmasında hem de daha çok parmak izi havuzunda en iyi eşleşmenin yapılmasında gerçek bir uygulamada çok önemlidir. Bu sayısal maliyet, parmak izlerinin büyüklüğüne, arama algoritmasının ve parmak izi derlemesinin karmaşıklığına bağlıdır [38].

2.1.2.1. Ses Parmak İzi Genel Prensipler

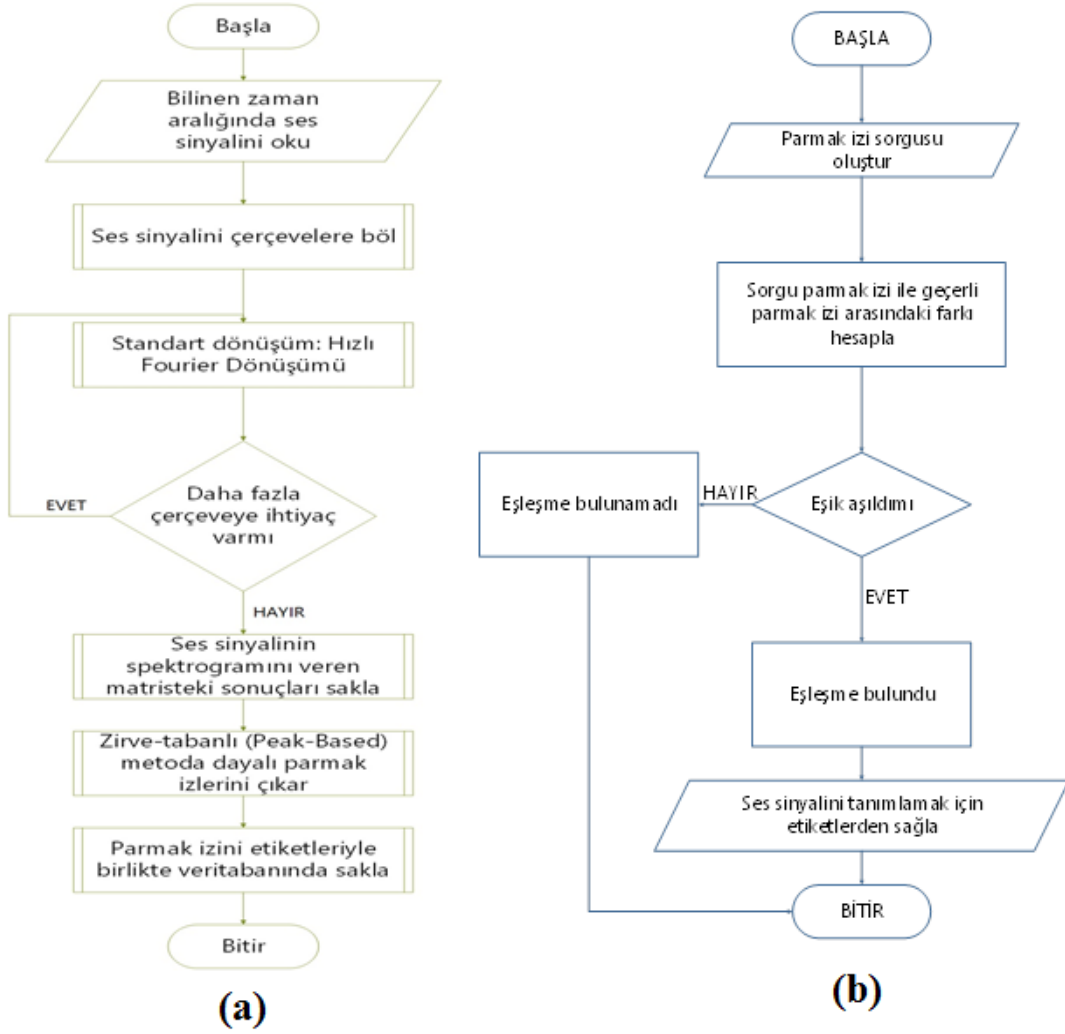
Ses dosyasının kısa bir özeti olan SPI, ses içeriğinin bir parçasının algısal özetini çıkarır ve veri tabanında saklar. Öncelikle ses verisinden spektrogramlar elde edilir. Daha sonra doruk noktaları bulunup, özetleme parmak izi oluşturulur. Daha sonra sistemimiz bu sesin hangi ses olduğunu belirler. Ses parmak izi çıkarma sistemleri genel yapısı Şekil 2.5’de görüldüğü gibidir [38].



Şekil 2.5. Ses parmak izi çıkarma genel yapısı [38]

Ses parmak izi sistemi bilinmeyen bir ses parçasının bilinen milyonlarca ses parçasıyla etkin bir şekilde karşılaştırmasını yapabilmelidir. Ses parmak izi, ses içeriğinin bir parçasının algısal özetini çıkarır ve veri tabanında saklar. Etiketsiz ses ile sunulduğunda, onun parmak izi hesaplanır ve veritabanında saklanan kişilerle eşleştirilir. Parmak izi ve eşleşme algoritmalarını kullanarak, kaydın farklı sürümleri aynı ses içeriği olarak tespit edilebilir. Parmak izi sistemi de hesaplama açısından verimli olmalıdır. Verimlilik hem bilinmeyen sesin parmak izi hesaplanmasında hem gerçek bir uygulamada hem de büyük parmak izi deposunda en iyi eşleşme için aramada kritik öneme sahiptir. Bu hesaplama maliyeti parmak izi boyutu, arama algoritması karmaşıklığı ve parmak izi çıkarma karmaşıklığı ile ilişkilidir.

Şekil 2.6 (a)'da ses parmak izi öznitelik çıkarma adımları görülmektedir. Bu çalışmada zirve-tabanlı metoda dayalı parmak izi özneliği kullanılmıştır. Şekil 2.6 (b)'de ise ses parmak izi eşleştirme adımları görülmektedir. Eğer yeni gelen ses verisi veri tabanında eşleşirse, bu ses verisinin tüm özellikleri veritabanından çekilerek kullanıcıya aktarılır [143].

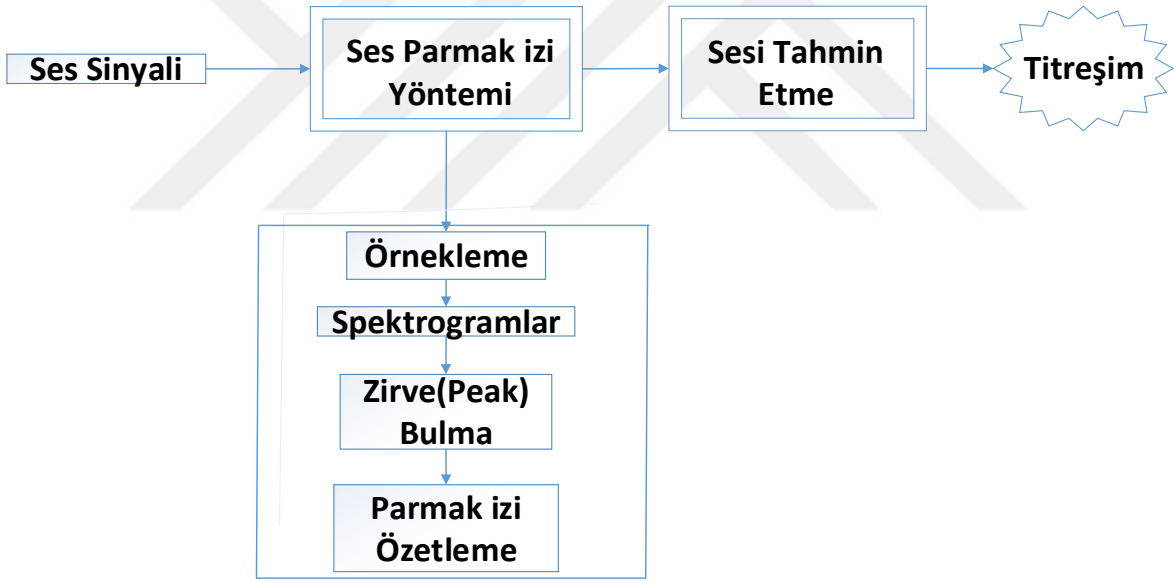


Şekil 2.6. Ses parmak izi (a) öznitelik çıkarma, (b) eşleştirme adımları [143]

Ses parmak izi teknikleri ses sinyalleri hafif veya ciddi derecede bozuk olduğunda bile içerik tabanlı ses tanımlaması yaparken başarılı sonuç iletmeyi amaçlar. Ses parmak izinin etkili olabilmesi için ilk olarak, girdi ses sinyalinin boyutunun önemli ölçüde azaltılmış olması gerekir. İkincisi, ortaya çıkan özelliklerin girdinin muhtemel bozulmalarına karşı sağlam olması gerekir. Örneğin, radyoda çalan şarkıların tespiti yapılacaksa, sistem doğrusal olmayan bozulmaların her türüne karşı sağlam olmalıdır. Üçüncü olarak, ortaya çıkan özellikler bilgilendirici olmalıdır. Ses tanımlama için, farklı ses dosyalarını bazı uygun metriklerde, uzak olan özelliklere eşlemek gerekir. Dördüncü olarak, özellik çıkarma işleminde hesaplamaların verimli olması gerekir.

2.1.2.2. Ses Parmak İzi Çalışma Mekanizması

Şekil 2.7’de görüldüğü gibi sistemimiz dört aşamadan oluşmaktadır. Öncelikli olarak ses algılanır. Daha sonra SPİ yöntemi kullanılarak sesin hangi ses olduğu tahmin edilir. En son olarak da her ses için farklı titreşimler verilerek, kullanıcının sesi algılaması sağlanır. SPİ yöntemi aşamasında öncelikli olarak sesin örneklenmesi yapılmaktadır. Alınan örnekler sinyal olduğundan, sesin spektrogramını oluşturmak için sesin örneklerinden küçük zaman pencerelerinde defalarca FFT kullanılır. Ses sinyalinin spektrogramından genlik doruk noktalarını bulmalıyız. Doruk nokta çevresindeki komşuluklardan en büyüğü olan genlik değerine karşılık gelen zaman, frekans çiftidir. Son olarak ise parmak izi özetleme yapılır. Doruk noktaları parmak izleri ile birleştirerek o ses için uygun izi buluruz. Bunun için özetleme fonksiyonu kullanılmıştır.



Şekil 2.7. Giyilebilir sistemimizin ses parmak izi kullanımı

2.1.2.2.1. Örnekleme

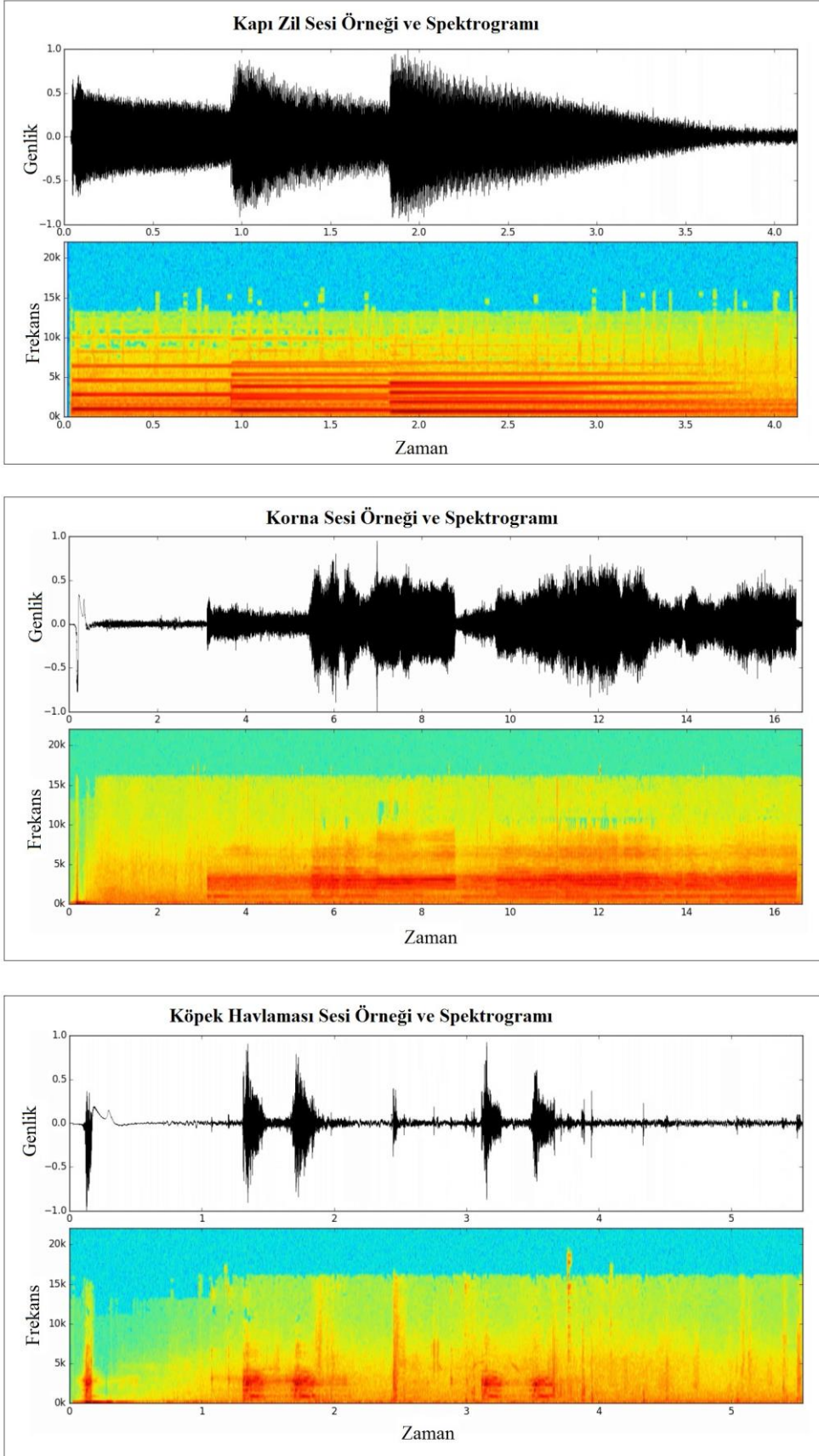
Öncelikle ses dijital hale dönüştürülür ve belirlenmiş bir formata getirilir. Ses kaydedilmesi durumunda kabul edilen kural, insan kulakları 22050 Hz üzerindeki frekansları kaçırmaz. Nyquist-Shannon Örnekleme Teoremine göre, örnekleme oranı en yüksek frekans

bileşeninin iki katı olmalıdır. Bu nedenle Nyquist Teoremine göre saniyede 44100 örnekleme yapılmaktadır.

2.1.2.2.2. Spektrogram

Alınan örnekler bir tür sinyal olduğundan, sesin spektrogramını oluşturmak için sesin örneklerinden küçük zaman pencerelerinde defalarca FFT kullanılmıştır. Şekil 2.8’de kapı zil sesi, korna sesi ve köpek havlaması ses örnekleri ve spektrogramları görülmektedir. Görüldüğü gibi spektrogram zaman ve frekans fonksiyonu olarak bir dalga genliğinde 2 boyutlu dizidir. Uygulanan FFT sonucunda bize bir sütun vererek, özel frekansta sinyal gücünü gösterir. 2 boyutlu dizi spektrogramlarını elde etmek için pencereleme işlemi yeterli ölçüde yapmamız gerekmektedir. Spektrogram, bir sesin benzersizliğini tanımlamak için kullanılmıştır. Dış ortam ve iç ortamda çeşitli sesler gürültüye sebep olmaktadır. Buradaki amacımız ses sinyalinin en ayırt edici parmak izini yakalamaktır.

Bir ses sinyali spektrogramı, bu sinyalin zaman-frekans analizidir. Başka bir deyişle, spektrogram zaman-alan sinyalinin frekans spektrumunun görsel bir gösterimini ortaya koyar. Spektrogram, sinyaldeki farklı tanımlanamayan spektrum özelliklerinin daha net bir şekilde anlaşılmasını sağlar. Bu, araştırmacıların her ses sinyalini birbirinden ayıran özellikleri çıkartmak için spektrogram analizini kullanmalarının temel nedenidir. Fourier dönüşümleri kullanılarak çeşitli zaman karelerinde bir spektrogram oluşturulabilir. Ses parmak izleri doğrudan Fourier dönüşümlerinden hesap edilebilirken, Fourier dönüşümlerinin sonuçları bir filtre bankasından geçirildikten sonra da hesaplanabilir. Filtre bankası, ses sinyalinin zaman üzerindeki çeşitli frekans bantlarındaki spektrum bileşenlerini çıkartan bir bant geçiren filtre kümesidir. Sayısal sinyal işlemenin gelmesiyle birlikte, Fourier dönüşümleri kullanılarak spektrogram işleme için alternatif bir yaklaşım tanımlanmıştır. Öncelikle, giriş sinyali örneklenir ve zaman karelerine bölünür. Sonra, sinyalin her zaman karesinin FFT’si hesaplanır ve giriş sinyalinin spektrogramını ortaya çıkarmak için bir araya getirilir.



Şekil 2.8. Ses örnekleri ve spektrogramları

Öncelikle olarak giriş sinyali, sayısallaştırılır [144].

$$y(n) = x(n) + w(n), 0 < n \leq N \quad (2.1)$$

Burada $x(n)$ ses sinyali, $w(n)$ ise gürültüdür. Ayrık Fourier dönüşümünden sonra eşitlik (2.2) elde edilir [144].

$$Y(k) = X(k) + W(k), 0 < k \leq N \quad (2.2)$$

Burada eşitlik (2.3), (2.4), (2.5), (2.6) kullanılır [144].

$$y(n) \leftrightarrow Y(k), x(n) \leftrightarrow X(k), w(n) \leftrightarrow W(K) \quad (2.3)$$

$$Y(k) = \sum_{n=0}^{N-1} y(n) e^{-2\pi i k n / N} \quad (2.4)$$

$$X(k) = \sum_{n=0}^{N-1} x(n) e^{-2\pi i k n / N} \quad (2.5)$$

$$W(k) = \sum_{n=0}^{N-1} w(n) e^{-2\pi i k n / N} \quad (2.6)$$

Böylece spektral çıkarma Eşitlik (2.7) ile tanımlanır:

$$H(k) = 1 - \frac{\mu(k)}{|X(k)|} \quad (2.7)$$

$\mu(k)$, konuşma çerçevelerindeki spektrumun ortalama değeridir. $H(k)$, spektral çıkarma filtresidir. Bu şekilde, gürültü spektrumunu giriş sinyalinden çıkarılır ve nispeten temiz bir sinyal alırız.

FFT boyutu N_{FFT} ye eşit olan ve N_{hop} boyutunda atlama boyutu olan iki önemli parametre olan N_{win} pencere boyutu vardır. N_{FFT} , spektrumdaki iki frekans bileşeni arasındaki mesafe olan spektrogramın frekans çözünürlüğüne karar verir. Eşitlik (2.8) ile hesaplanır [144].

$$f_{res} = \frac{f_{hedef}}{N_{FFT}} \quad (2.8)$$

Hop boyutu, çerçeveler arasında çakışma oluşturmak için pencere boyutunda farklıdır. Örtüşme gereklidir çünkü Kısa Süreli Fourier Dönüşümde pencere işlevi genellikle pencere sınırlarına yakın olarak çok küçük veya sıfırdır. Hop boyutu N_{hop} , pencere fonksiyonunun seçimine bağlıdır. Hanning penceresi için, değeri tipik olarak pencere boyutunun yarısıdır. Hanning penceresi Eşitlik (2.9) ile hesaplanır [144].

$$w_n = 0.5 \left(1 - \cos\left(2\pi \frac{n}{N}\right)\right), 0 \leq n \leq N \quad (2.9)$$

Hesaplamadan sonra, spektrogram iki boyutlu bir dizi ile temsil edilebilir. Herhangi bir çakışma yoksa, sinyalin büyük bir kısmı bastırılır.

$$S(f, t), 0 \leq t < N_{\text{çerçeve}}, 0 \leq f < N_{bin} \quad (2.10)$$

$N_{\text{çerçeve}}$ çerçeve numarasıdır ve Eşitlik (2.11) ile hesaplanır. N_{bin} ise Eşitlik (2.12) ile ifade edilir [144].

$$N_{\text{çerçeve}} = \left\lceil \frac{L}{N_{hop}} \right\rceil \quad (2.11)$$

$$N_{bin} = \frac{N_{FFT}}{2} \quad (2.12)$$

Zirve noktalar elde edilmeden önce spektrogramda büyüklük bilgisi hesaplanır ve faz bilgisi önemsenmez [144].

$$S(f, t) = |S(f, t)|, 0 \leq t < N_{\text{cerçeve}}, 0 \leq f < N_{\text{bin}} \quad (2.13)$$

Daha sonra logaritmik hesaplama yapılır.

$$S(f, t) = \log(S(f, t)), 0 \leq t < N_{\text{cerçeve}}, 0 \leq f < N_{\text{bin}} \quad (2.14)$$

Başlangıç geçişlerini en aza indirmek için spektrogramın sıfır-ortalaması hesaplanır.

$$S(f, t) = S(f, t) - E(S), 0 \leq t < N_{\text{cerçeve}}, 0 \leq f < N_{\text{bin}} \quad (2.15)$$

Daha sonra ise Eşitlik (2.16)'deki gibi yüksek geçişli filtre uygulanır.

$$y(n) = x(n) - x(n - 1) + py(n - 1) \quad (2.16)$$

Bu eşitlikte p değeri yavaşça değişen bileşenleri kaldırmak ve hızla değişen bileşenleri vurgulamak için kullanılmıştır.

2.1.2.2.3. Zirve Bulma

Ses parmak izi sistemini kurabilecek birkaç yaklaşım mevcuttur. Bu yaklaşımlardan birçoğu temel bir kavram olan spektrogramdan türemiştir. Bir ses sinyali spektrogramından ses parmak izini oluşturabilmek için birkaç yol mevcuttur. Tepeye bağlı yöntem, ses imzalarını oluşturabilmek için spektrogramdan faydalanma yaklaşımıdır. Bu yaklaşımda, spektrogram filtreden geçirildikten sonra, ses dosyaları için imzalar oluşturmak için nihai verideki bantların her birisindeki tepe değerlerinden faydalanılır.

Gauss yumuşatma, spektrogramdaki bir sütuna karşılık gelen bir vektörde belirgin olmayan maksimumu bastırmak için kullanılır. Giriş vektöründen öncelikli olarak yerel maksimumlar Eşitlik (2.17) ile elde edilir [144].

$$(a_i, I_i), 0 \leq i < I \quad (2.17)$$

Bu denklemde a_i ve I_i , maksimum olanın genliği ve zamanı gösterir, I ise vektördeki zirvelerin sayısıdır. Daha sonra her bir yerel maksimumda Gauss üst üste konulur. Gauss maksimum (a_i, I_i) için Eşitlik (2.18)'deki gibi hesaplanır [144].

$$G_i(n) = a_i e^{-\frac{(n-I_i)^2}{2\sigma^2}}, 0 \leq n < N \quad (2.18)$$

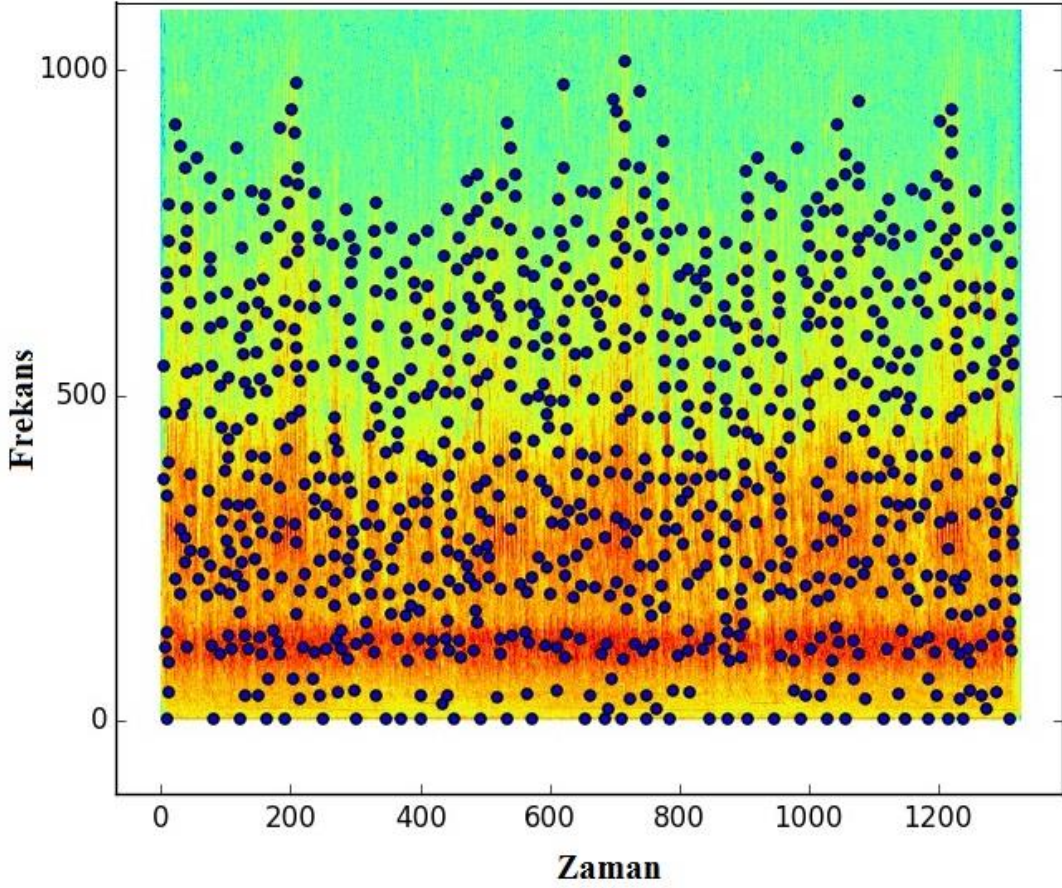
Sesin bozulması durumunda zirve noktalarının korunması ihtimali yüksek olduğu ve yüksek bir enerji düzeyine sahip olduğu için bu çalışmada zirve noktaları tercih edilmiştir. Yüksek oranda gürültü ve sinyal bozulmasına karşı dayanıklı olacağından dolayı spektrogram zirvelerini parmak izi özneteliği olarak kullanılmıştır.

Öncelikle ses sinyalinin spektrogramından genlik doruk noktalarını bulmalıyız. Doruk nokta çevresindeki komşuluklardan en büyüğü olan genlik değerine karşılık gelen zaman, frekans çiftidir. Çevresindeki diğer çiftlerden genlikleri düşüktür ve bu gürültünün azalmasını sağlar. Gürültüye dayanıklı dorukları çıkardıktan sonra, onu tanımlayan bir sesin ilgi çekici noktaları bulunmuştur. Zirve noktalarının gösterimi, takımyıldızı haritası (Constellation map) olarak isimlendirilmiştir. Şekil 2.9'da örnek bir ses verisinin takımyıldızı haritası gözükmemektedir. Ses parmak izi, takımyıldızı haritasındaki yıldız çiftlerinin karmasından oluşur ve bu çiftlerden birisi sabit noktadır, diğeri de herhangi bir sabit nokta zamanı için zaman farkıdır. İki frekansın karmasını oluşturan bu sabit için hedef noktadaki esneme faktörüyle saptanan nokta kümesi sayılarından birisidir. [19].

Zirve noktalar Eşitlik (2.19) deki gibi hesaplanır:

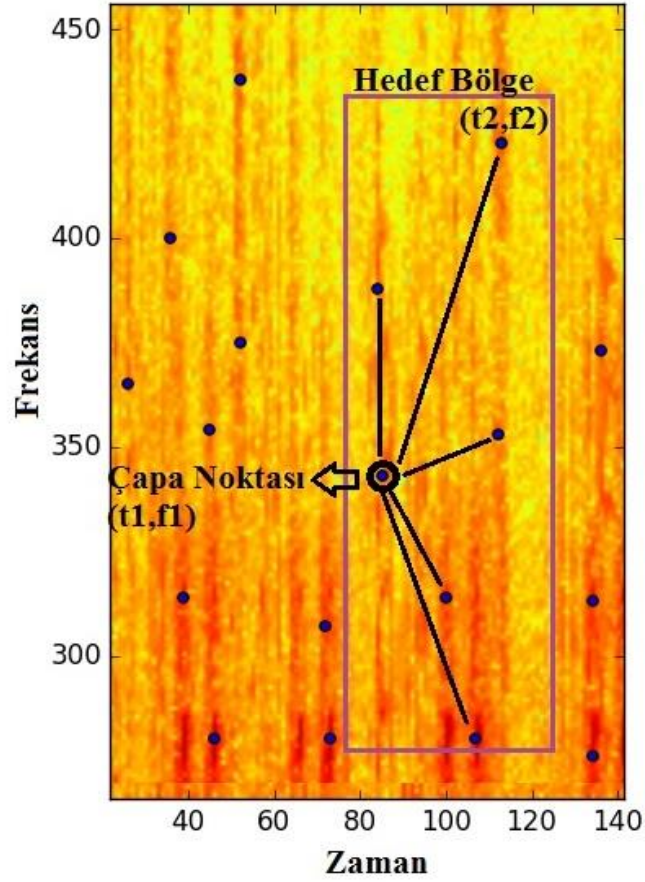
$$S(f, t), t < N_{\text{çerçeve}}, 0 \leq N_{\text{bin}} \rightarrow \{(f_n, t_n)\}, 0 \leq n < N_{\text{zirve}} \quad (2.19)$$

Burada (f_n, t_n) spektrogramdaki tepe koordinatı ve N_{zirve} , bir ses parçasındaki zirvelerin sayısıdır. Spektrogramdaki zirveler gökyüzündeki birçok yıldız benzediğinden, takımyıldızı haritası olarak adlandırılır [144].



Şekil 2.9. Ses verisi spektrogramı takımyıldızı haritası

Tek başına frekans bileşeninin gerekli özgünlüğe sahip olmadığı için tepe değerlerini doğrudan kullanmak mümkün değildir. Doruk çiftleri göz önüne alındığında, parmak izi karmaları, takımyıldızı haritasındaki zirve noktalarının ikili kombinasyonlar şeklinde eşleştirilmesiyle elde edilmektedirler. Öncelikle Şekil 2.10'da görüldüğü çapa noktaları (anchor point) olarak kullanılmak üzere bir zirve tespit edilir ve daha sonra bir hedef bölgeye atanır. Ardından, çapa nokta ve hedef bölgedeki her zirve noktası oluşur ve her iki frekans değerinin ve tepeler arasındaki zaman farkının bir kombinasyonu olarak her zirve çifti için bir karma değer elde edilir.



Şekil 2.10. Çapa noktası gösterimi

Doruk noktası olarak her zirveden yararlanılarak, endekslenecek öğelerin sayısı, hedef bölgedeki zirvelere bağlı olarak bir etken artar. Eğer nokta etrafında merkezlenmiş bir bölgedeki bütün komşularından daha yüksek bir enerji içeriğine sahipse zaman-frekans noktası da aday bir tepedir. Aday tepeleri, yoğunluk kriterine göre seçilir. Yoğunluk kriteri, komşularından daha yüksek frekansa sahip şarkı parçalarıdır. Her zaman frekans yerindeki tepeler, en yüksek genlik tepelerinin ses bozulmalarına karşı en dirençli tepeler olduğundan genliğe göre seçilir. Parmak izi karmaları, takımyıldızı haritasından meydana gelir. Bu karma işaretler, frekanslara ilişkin zaman-frekans noktası çiftleridir. Bazı noktalar, sabit nokta olarak seçilir ki bir hedef bölgeye sahip her sabit nokta bununla bağlantılıdır. Her sabit nokta sıralı olarak hedef bölgesindeki noktalarla eşleşir. Her çift iki frekans bileşenini ve noktalar arası zaman frekansını verir. Her bir zirve çifti, çapa noktası ve hedef bölgedeki nokta arasındaki zaman ve frekans farklılığı ile temsil edilir. Örneğin, $(t_1; f_1)$ ve $(t_2; f_2)$ çifti ile Eşitlik (2.20) ile temsil edilebilir.

$$t_1: (f_1, f_2, \Delta t); \Delta t = (t_2 - t_1) \quad (2.20)$$

Bu kombinasyonel karma stratejisinin üç avantajı vardır. Öncelikle ortaya çıkan parmak izleri, tekil zirvelerden daha yüksek bir özgüllük gösterir ve alımın hızlanmasına yol açar. İkincisi, parmak izleri, mutlak zaman bilgisi yakalamadığı için değişmezdir. Üçüncü olarak, tanıtılan parmak izi sayısının kombinasyonel çarpımı sinyal bozulmalarına karşı sağlamlığı artırmaktadır.

2.1.2.2.4. Ses Parmak İzi Özetleme

Dorukları bulduğumuzda benzer doruk noktaları olabilir, doruk noktaları parmak izleri ile birleştirilerek o ses için uygun izi buluruz. Bunun için özetleme fonksiyonu kullanılmıştır. Özetleme fonksiyonu girdi olarak tamsayı alır ve çıktı olarak başka bir tamsayıya dönüştürür. Spektrogramların dorukları ve doruk frekans birleşmelerine aralarındaki zaman farkı ile bakarak, sesleri ayırt edecek bir özetleme oluşturabiliriz. Bu çalışmada birden fazla doruğun dikkatini alarak, daha ayrıntılı entropisi olan bir başka deyişle daha fazla bilgi içeren parmak izi oluşturulmuştur.

Parmak izi modelleme bloğu, ön işleme, çerçeveleme, dönüştürme ve öznitelik çıkarma aşamalarından sonra çıkarılan özellik vektörlerinin dizisine dayanan son parmak izini hesaplar. Her çerçeve bir özellik vektörü üretir, bu nedenle özellik vektörlerinin başlangıç dizisi doğrudan parmak izi modeline bağlıdır.

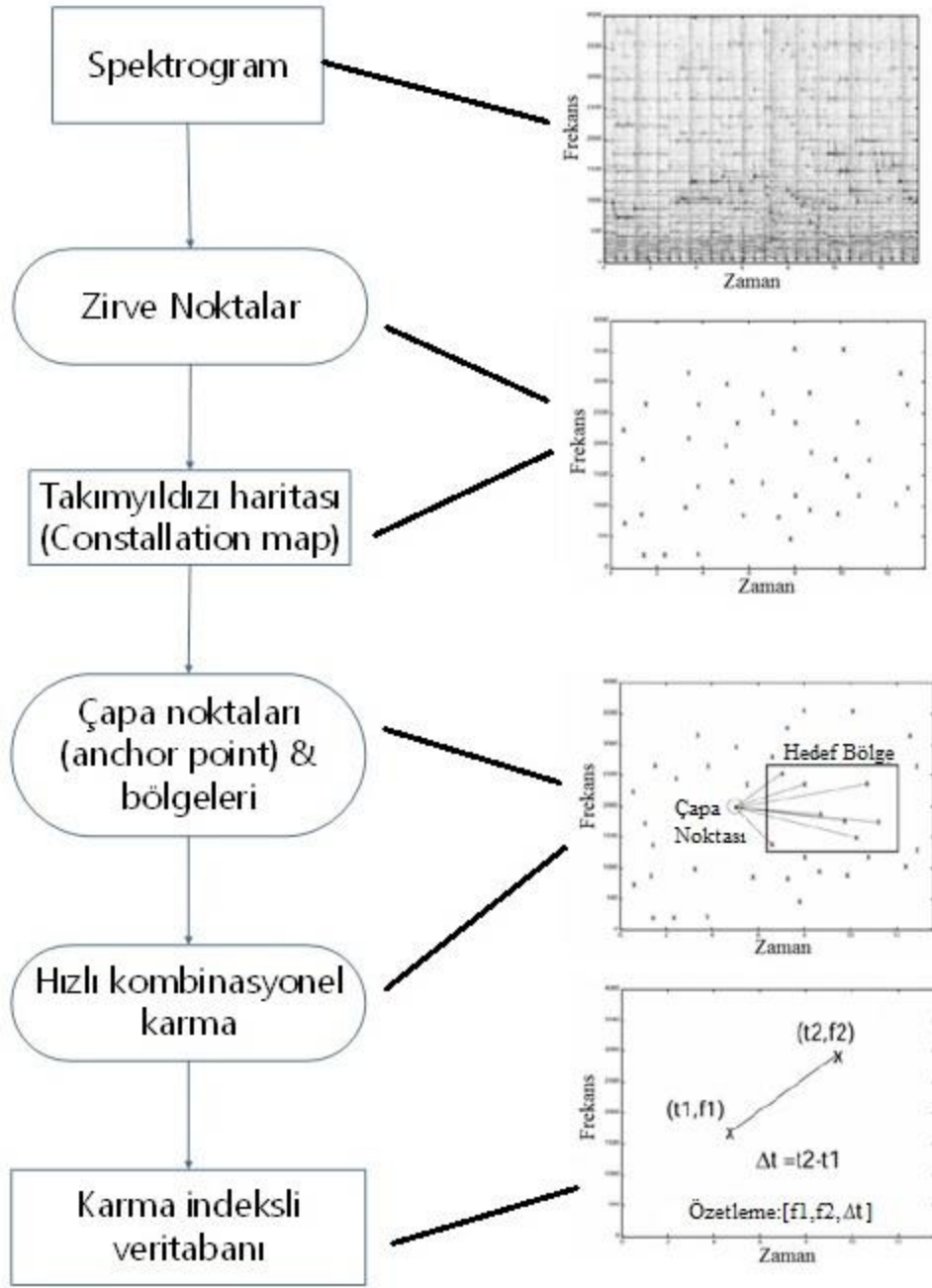
Parmak izleri çıkarıldıktan sonra, veritabanında benzer parmak izlerini aramamız gerekir. Burada benzerlik, iki parmak izinin de ne kadarının benzer olduğunun ölçütüdür ve bir mesafe olarak tanımlanır. Küçük mesafe yüksek benzerlik derecesini gösterir. Mesafenin nasıl hesaplanacağı, büyük ölçüde, parmak izinin tasarımına bağlıdır. Son olarak, veritabanında eşleşen bir ögenin olup olmadığına karar verilir. Veritabanındaki sorgu parmak izi ve diğer referans parmak izleri arasındaki mesafe benzerlik eşik değerinde ise, referans ögesi eşleşen sonuç olarak döndürülür, aksi takdirde sistem eşleşen bir öge olmadığını düşünür. Eşleşen sonuçlara dayanarak, bir ses parmak izi alma sisteminin performansı, test etmek için kullanılan tüm sorguların doğru eşleşme sayısının bir kısmı olarak ölçülür. Çoğu sistem bu tanıma oranını değerlendirme sonuçları olarak bildirmektedir [19, 106, 108, 145]. Eşleşme algoritmasının asıl prosedürü, veritabanını taramak ve benzer

takımyıldızı haritalarını bulmaktır. Bir sorgu ses dosyası, yer işaretleri zirve olarak kabul edilirse, aynı zamanda bir takımyıldızı haritası olan yer işaretleri listesine dönüştürülür. Veritabanı aslında tüm referans parçalarının takımyıldızı haritalarından oluşur. Sorgu parçasının takımyıldızı haritası, referans parça üzerinde soldan sağa kaydırılır ve eşleşen noktalar belirlenen referans aralığına ulaşırsa eşleşir.

Şekil 2.11’de parmak izi oluşturma adımları genel akış şeması ve görünümü gözükmemektedir [19]. Öncelikli olarak spektrogramlar oluşturulur. Daha sonra zirve noktaları ve takımyıldızı haritası oluşturulur. Takımyıldızı haritasından çapa noktaları ve hedef bölgeler seçilir ve kombinasyonel karmalar oluşturulur. En son olarak da özetleme yapılarak karma indeksli veritabanında ses parmak izi tutulur.

2.1.2.2.5. Ses Parmak İzi Oluşturmak İçin Kullanılan Yapılar

Parmak izi metodu için kullanılan yapılardan örnekleme oranı Nyquist teoreme göre 44100 olarak belirlenmiştir. FFT pencere boyutu frekans aralığını daha iyi belirlemek için 4096 olarak belirlenmiştir. Her ardışık pencerenin son ve bir sonraki pencere ile çakıştığı oran 0.5 olarak belirlenmiştir. Daha yüksek çakışma, eşleştirilenin daha yüksek oranda olmasını sağlar ancak daha fazla parmak izi oluşturacağı için daha fazla yer kaplar. Bir parmak izinin komşularıyla eşleştirileceği derece 15 olarak belirlenmiştir. Spektrogram oluşturulduktan sonra bir noktanın zirve nokta olarak kabul edilebilmesi için minimum genlik 15 olarak belirlenmiştir. Parmak izi sayısını azaltmak için minimum genlik yükseltilebilir, ancak doğruluğun azalmasına sebep olabilir. Spektral bir tepe noktası olarak görmesi için spektrogramdaki genlik tepe etrafındaki hücre sayısı 10 olarak belirlenmiştir. Daha yüksek değerler daha az parmak izi ve daha hızlı eşleşme anlamına gelir, ancak doğruluğu etkileyebilir. Parmak izi olarak eşleştirilebilmesi için parmak izlerine ne kadar yakın veya uzak olduğunu dair belirlenen eşik değerler ise 0-200 aralığındadır. Parmak izi hesaplamasında kullanılan SHA-1 özetleme algoritmasının önünden atılacak bit sayısı ise 20 olarak belirlenmiştir. Daha fazla attığımız zaman daha az depolama olacaktır ancak eşleştirme yaparken daha çok çarpışma olacağı için doğruluk azalacaktır. Veritabanında ise ses ve ses parmak izi adlarında iki adet tablo oluşturulmuştur. Öncelikli olarak sesin parmak izi oluşturularak kaydedilir, daha sonra yeni bir ses verisi eşleştirilerek veritabanından ilgili sesin özellikleri çıkartılır.



Şekil 2.11. Ses parmak izi akış şeması [19]

2.1.2.2.6. Kullanılan Kütüphaneler

Bu program için birkaç Python kütüphanesi kullanılmıştır. İki standart Python kütüphanesi dahil edilmiştir. Sys kütüphanesi komut satırı parametrelerini elde etmek için kullanılmakta ve os kütüphanesi de dosya sistemindeki dosyaları bulmak için kullanılmaktadır. Scipy, zirve noktaları bulmak için kullanılmıştır. Numpy, sayısal

hesaplamalar ve ses verisi üzerinde çeşitli matematiksel işlemler gerçekleştirmek için kullanılmaktadır. Matplotlib 2 boyutlu çizim kütüphanesidir. Pyaudio, mikrofondan ses almak için kullanılmıştır.

2.1.3. Diğer Metotlar

Bilgisayarda ses tanıma işleminin gerçekleştirilebilmesi için sesin algılanmasından itibaren tanıma sürecine kadar pek çok öznelik yöntemi uygulanması gerekmektedir. Ses tanıma işlemi de yapılacak tanıma işleminin durumuna göre uygun ve sağlıklı öznelik yönteminin önceden belirlenmesi, başarılı bir ses tanıma işlemi için çok önemli bir aşamadır.

2.1.3.1. Sıfır Geçiş Oranı

Ses sinyalinin sıfırı kesme miktarı, sıfır geçiş oranı (SGO) olarak bilinir. Sıfır noktası; ses işleme, elektronik uygulamalar, matematik ve görüntü işlemede yaygın olarak kullanılan bir terimdir. Matematiksel anlamda, sıfır noktası, fonksiyon işaretini değiştiren bir noktadır.

SGO sinyal boyunca sinyal değişikliklerinin oranını, daha doğrusu, sinyal değişikliklerinin pozitiften negatife ya da tam tersi durumlar için oranlar. SGO bunu ses dalgalarının biçiminin sıfır eksenini kesmesini sayarak hesaplar. SGO (2.21) formüle göre, giriş sinyali $s(t)$ uzunluğu ile normalize edilir [146]:

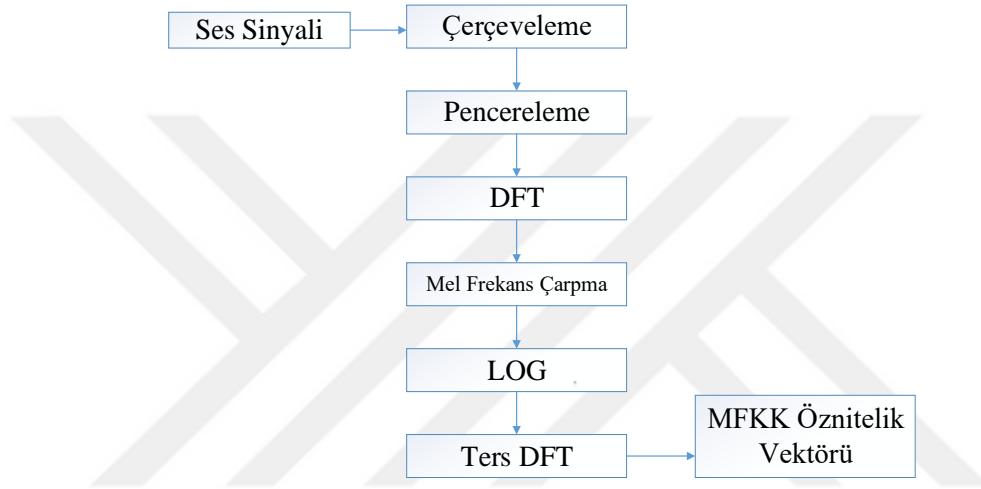
$$SGO = \frac{1}{2} \left(\sum_{t=1}^{T-1} |sign(s(t)) - sign(s(t-1))| \right) \left| \frac{F_s}{T} \right| \quad (2.21)$$

Burada T , $s(t)$ deki f örneklerinin toplam sayısıdır ve F_s örnekleme frekansıdır. Bu özellik, hem konuşma tanıma, müzik türü sınıflandırma hem de multimedya içerik analizlerinde yoğun olarak kullanılmaktadır.

2.1.3.2. Mel Frekansı Kepstrum Katsayıları

Mel frekansı kepsrum katsayıları (MFKK) ses tanıma uygulamalarındaki başarısının yüksek olmasından dolayı en fazla kullanılan öznelik çıkarım yöntemlerinden biridir.

MFKK konuşma araştırması ve konuşmacı tanımlama sistemlerinde yaygın olarak kullanılmaktadır. MFKK, Fourier dönüşümünü temel alır. Öz vektörler, ses sinyallerinin her ses çerçevesinin frekanslarından çıkarılır. MFKK öznelik çıkarma adımları Şekil 2.12’de gösterilmiştir. MFKK yönteminde öncelikli olarak sinyalin enerjisini düşük frekanslarda azaltılır. Bu, ses sinyalinin yüksek frekanslı bileşenlerinin düşük frekans bileşenlerine göre daha az güce sahip olduğu için uygulanır. Ses sinyalleri sürekli değiştiği için bunları zaman alanında kısa periyotlarla analiz edilir.



Şekil 2.12. MFKK adımları

Giriş işaretimizi $x(n)$, pencere fonksiyonunu $w(n)$ ve çıkış işaretimizi ise $y(n)$ ile ifade edecek olursak, çıkış işaretimiz Eşitlik (2.22) ile ifade edilir.

$$y(n) = x(n)w(n) \quad (2.22)$$

Kısa zaman aralıklarında ses dalgası çerçevelere ayrılır. Spektral etkinin azalması için pencereleme uygulanır. Pencereleme işleminde Eşitlik (2.23)'teki Hamming pencere uygulanmıştır.

$$w(n) = 0.54 - 0.46 \cos\left(2\pi \frac{n}{N}\right), 0 \leq n < N \quad (2.23)$$

Daha sonra her çerçevenin zaman bölgesinden, frekans bölgesine çevrilmesi için FFT uygulanır. Zaman etki alanı örneği FFT kullanılarak frekans alanına dönüştürülür. FFT, DFT'nin verimli bir algoritmasıdır. Her bir çerçeve için büyüklük frekans cevabını elde etmek için genellikle FFT uygulanır. N örnekli bir set için Ayrık Fourier Dönüşümünün matematiksel ifadesi Eşitlik (2.24)'de görüldüğü gibidir [147].

$$X_n = \sum_{k=0}^{N-1} x_k e^{-2\pi jkn/N}, \quad n=0,1,2, \dots, N-1 \quad (2.24)$$

Mel filtre bankası üçgen filtrelerden oluşmaktadır. Buradaki filtreler, Eşitlik (2.25) denkleminde mel ölçeği ile frekans ölçeği arasındaki dönüşüm sağlanmaktadır ve ortak doğrusal frekans f ile ilgili olan Mel frekansı boyunca eşit aralıklıdır [147].

$$F(Mel) = [2595 * \log_{10} \left[1 + \frac{f}{700} \right]] \quad (2.25)$$

Mel frekansı, lineer frekansın logaritması ile orantılıdır, insanın öznel işitsel algısındaki benzer etkileri yansıtır. Her çerçeve ters Fourier transformasyona uğrattılır ve frekans uzayından tekrar zaman uzayına döndürülür. Bu işlemin sonucu olarak Eşitlik (2.26)'da görüldüğü gibi MFKK katsayıları elde edilir. Burada M , kepsrum katsayılarının sayısını, X_k ise k . filtrenin log enerji çıktısını göstermektedir.

$$MFKK_i = \sum_{k=1}^{20} X_k \cos \left[i \left(\frac{k-1}{2} \right) \frac{\pi}{20} \right], \quad i = 1, 2, \dots, M \quad (2.26)$$

MFKK, müzik özelliğini çıkarmada kullanılan en popüler tekniktir ve müziği bulmada geniş çapta başarılarla yapılan müzik tanıma işlemi için oldukça güçlüdür. Bunun sebebi MFKK'lerin değerlerinin insanın işitme sistemindeki değerlerine yakın olmasıdır. MFKK'ler, bir log güç spektrumunun lineer olmayan bir frekans mel-ölçeği üzerindeki lineer kosinüs dönüşümüne dayanarak müziğin kısa süreli güç spektrumunu göstermektedir. Bu MFKK algoritması kolektif bir şekilde bir mel-frekans kepsralini oluşturan katsayıları hesaplar. Bu katsayılar, müzik sinyalinin kepsral gösteriminin bir tipinden türetilir. Frekans

bandı, normal kepsralinde kullanılan lineer aralıklı frekans bantlarından daha çok insanın işitme sisteminin tepkisine yakın olan mel ölçeğinde eşit aralıklara sahiptir [96, 148, 149].

2.1.3.3. Spektral Altbant Merkezleri

Doğru akustik özelliklerin seçilmesi doğrudan bir konuşma tanıyıcının performansını etkilediği için, bir konuşma tanıma sisteminin tasarımında en önemli görevdir. Bu özellikler, konuşma tanıma için gerekli olan en üst düzeyde bilgiyi içermesi ve aynı zamanda alakasız bilgileri yani gürültüyü silecek şekilde seçilmelidir. Spektral altbant merkezleri (SAM'lar) konuşma sinyalinin güç spektrumundan kolaylıkla ve tahmin hataları olmadan güvenilir şekilde çıkarılabilir.

Paliwal [150] bir spektrumun daha yüksek genlik kısımlarının gürültüye nispeten daha az etkilendiğini göz önüne alarak SAM'ı önerdi. Altbantlar, konuşma tanımada merkez frekansı temsil eden SAM'lar şeklinde temsil edilebilir. SAM'lar, altbantlarda baskın frekansı kullanan nispeten yeni özelliklerdir. Altbantın üst kenarı h_m , alt kenarı l_m , filtre şekli $w_m(f)$ 'dir. m. spektral altbant merkez C_m , Eşitlik (2.27) ile hesaplanır [150].

$$C_m = \frac{\int_{l_m}^{h_m} f w_m(f) P^y(f) d_f}{\int_{l_m}^{h_m} w_m(f) P^y(f) d_f} \quad (2.27)$$

SAM'lar, altbantların merkez frekansları olarak hesaplanır ve kısa vadeli spektrumun baskın frekanslarını yakalarlar. SAM'lar altbant spektrumlarının merkez frekansları olarak hesaplanır ve güç spektrumunun yerel maksimum değerlerinin yerlerini verirler. SAM'lar konuşma tanıma, kişi tanıma ve ses parmak izi uygulamalarında kullanılmaktadır ve konuşma sinyalinin güç spektrumundan gelen her bir alt bant için frekans merkezleri olarak hesaplanır. Gürültülü ortamlarda neredeyse değişmeyen spektral zirveleri yakaladığı için, gürültülü koşullar altında bile güvenilir şekilde elde edilebilirler.

2.1.3.4. Çizgisel Spektral Frekanslar

Konuşma tanıma uygulamalarında kullanılan ses öznitelik vektörlerinden birisi de Çizgisel Spektral Frekanslar (ÇSF) parametreleridir. ÇSF parametreleri sinyal

spektrumunun yüksek deęerler aldıęı yerlere yoğunlaşırlar. SF parametreleri bozulma baęımsızlıęı özellięini de gstermektedir [151, 152]. Bu özellik, bir SF parametresindeki herhangi bir deęişiklięin küresel olarak etkiler yaratmayacaęı, sadece ona yakın frekans spektrumunu etkileyeceęi anlamına gelir. Bu özellięi kullanarak, insan kulaęının daha yüksek frekanslara karşı çok hassas olmamasından dolayı, daha yüksek frekanslardaki SF parametreleri daha az bitele temsil edilebilir.

SF'ler, konuşma spektrum bilgisini dięer iletim parametrelerinden daha verimli olarak kodlamaktadır [153, 154]. SF'ler, konuşma sinyallerini algılamada önemli olduęu bilinen spektral özellikleri hesaba katarak nicelleştirilebilir.

2.1.3.5. Doğrusal Kestirim Katsayıları

Doęrusal Öngörölü Kepstral Katsayıları, kepstrum alanında temsil edilen Doğrusal Kestirim Katsayılarıdır (DKK). DKK fikri, ses yolunun karakteristięi bir tüm-kutup filtresi ile modellenebilen konuşma üretim modeline dayanır. DKK, basitçe bu çok kutuplu filtrenin katsayılarıdır ve konuşmanın log spektrumunun yumuşıatılmış zarfıyla eşdeęerdir. DKK, doğrudan otokorelasyon veya kovaryans yöntemleri kullanılarak pencerelemiş konuşma bölümünden hesaplanabilir. DKK, gemiş örneklerin doğrusal bir kombinasyonundan ses sinyalinin geçerli örneęini öngörür.

Buna göre doğrusal ayırık bir sistemin řu anki ıkıřı, önceki p adet ıkıřının bir kombinasyonu olarak öngörülebilir: [47]

$$\hat{y}(n) = \sum_{i=1}^p a_i y(n - i) \quad (2.28)$$

Burada a_i katsayıları doğrusal kestirim katsayılarıdır. Katsayıların $y(n)$ ile $\hat{y}(n)$ arasındaki karesel ortalama hatayı minimize edecek řekilde belirlenmesi gereklidir. Bu alıřmada DKK'ler Levinson-Durbin algoritması ile hesaplanmıřtır.

DKK modeli bir kiřinin ses yolunu karakterize eder ve log alanı oranı katsayıları olarak adlandırılan dięer katsayılara dnüşürülebilir. Log alanı oranı analizinde, bir kiřinin ses yolu, farklı enine kesit alanlarına sahip, eşit uzunlukta oluşturulmuş düzgün olmayan bir akustik yapı olarak modellenmiřtir [156, 157].

2.1.3.6. Doğrusal Öngörülü Kepstral Katsayıları

Doğrusal Öngörülü Kepstral Katsayıları (DÖKK), pratik ses tanıma sistemlerinin uygulanmasında yaygın olarak kullanılan spektral özelliktir. Bir konuşma sinyalinin temel parametrelerini tahmin etmek için, DÖKK baskın tekniklerden biri haline gelmiştir. Bu yöntemin arkasındaki temel tema şu anki bir konuşma örneğinin geçmiş konuşma örneklerinin doğrusal bir birleşimi olarak tahmin edilebilmesidir. DÖKK elde etmek için öncelikle DKK elde edilmelidir Ses sinyali $s(n)$ 'nin kısa süreli penceresi, DKK olarak Eşitlik (2.29)'da modellenmiştir [155]. Bu eşitlikte $a(k)$ doğrusal kestirim katsayıları ve $e(n)$ tahmin hatasıdır. Tahmin katsayıları, anlam bilgisi anlamında tahmin hatasını en aza indirerek belirlenir. Bu DKK'ler, tipik olarak, Eşitlik (2.30) özyinelemeli denklemi kullanarak, DKKK olarak bilinen sağlam bir forma dönüştürülür [47].

$$s(n) = \sum_{k=1}^q a(k)s(n-k) + e(n) \quad (2.29)$$

$$c_m = a_m + \sum_{k=1}^{m-1} \binom{k}{m} c_k a_{m-k}, \quad 1 < m \leq q \quad (2.30)$$

DÖKK hesabı q tane DKK üzerinden hesaplanır ve bir işaretin q tane DÖKK'sine işaret eder. q değeri, Nyquist frekansına kadar frekans aralığında gerçek ve karmaşık kutupları etkili bir şekilde yakalayabilecek şekilde seçilir. DÖKK, Konuşma sinyalinin frekans alanı konuşma üretme modeli ya da konuşma algılama modelinin frekans alanıyla eşleştirerek çıkarılırlar ve bir konuşma sinyalinin spektrumunun logaritmik spektrumuyla ilişkilidirler. DÖKK konuşma üretim modelini temel almaktadır. Filtre katsayıları, bitişik konuşma örneklerinde bulunan korelasyon özelliklerini araştırarak, doğrusal tahmin kullanılarak tahmin edilir. Bu katsayılar nadiren özellikler olarak kullanılır, ancak daha güçlü ve daha az korelasyonlu özelliklere dönüşmüşlerdir ve bunlardan biri de DÖKK'dür.

2.1.3.7. Log Alanı Oranı

Log alanı oranı (LAO), bir kanal üzerinden iletim için yansıma katsayılarını (doğrusal tahmin katsayıları için başka bir form) temsil etmek için kullanılabilir. ÇSF'ler kadar etkili olmasa da, log alanı oranlarını hesaplamak daha kolaydır. DKK'lerden elde edilen log alanı oranı katsayıları, konuşma uygulamalarında kullanılan iyi bilinen bir öznitelik çıkarma tekniğidir. LAO katsayıları doğrusal tahmin katsayılarından elde edilir.

LAO, bir kanal üzerinden iletim için yansıma katsayılarından hesaplanır. Eğer k bir yansıma katsayısı ise, LAO'ler, Eşitlik (2.31) kullanılarak hesaplanabilir.

$$LAO = \log \left(\frac{1+k}{1-k} \right) \quad (2.31)$$

LAO, insanlar tarafından öznel olarak değerlendirildiğinde konuşma kalitesi ile daha fazla korelasyona sahip olduğu gösterilen objektif bir konuşma kalitesi ölçüsüdür [158].

2.1.3.8. Sınıflandırma

Sınıflandırma aşamasında K-en yakın komşu (K-EYK) ve destek vektör makineleri (DVM) kullanılmıştır. Bilgisayar ortamında SGO, MFKK, SAM, ÇSF, DÖKK, DKK, LAO öznitelikleri ayrı ayrı hesaplandıktan sonra K-EYK VE DVM sınıflandırıcıları kullanılarak SPİ yöntemi ile karşılaştırılmıştır.

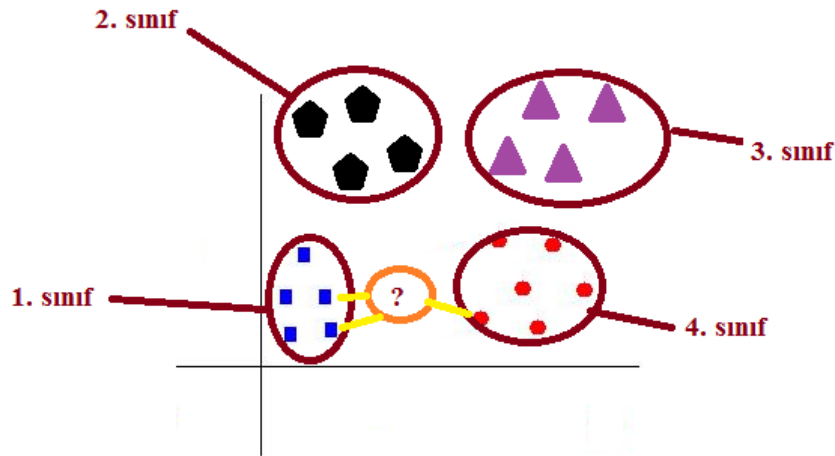
2.1.3.8.1. K-En Yakın Komşu

K-EYK, eğitilmiş öğrenme algoritmasıdır ve amacı, yeni bir örnek geldiğinde var olan öğrenme verisi üzerinde sınıflandırma yapmaktır. Algoritma, yeni bir örnek geldiğinde, onun en yakın K komşusuna bakarak örneğin sınıfına karar verir. K-EYK amacı, yeni bir veriyi özelliklerinden faydalanarak önceden sınıflandırılmış örnekler yardımıyla sınıflandırmaktır. Birçok araştırmacı, K-EYK algoritmasının farklı veri setleri üzerindeki deneylerinde çok iyi bir performans elde ettiğini bulmuştur [159-161]. K-EYK algoritmasında yeni bir olayı sınıflandırmak için, sistem eğitim seti arasında en yakın

komşuları bulur ve kategori adaylarını ağırlıklandırmak için en yakın komşuların kategorilerini kullanır. Bir test belgesini eğitim setindeki tüm örneklerle karşılaştırması gerektiğinden, K-EYK algoritmasının dezavantajlarından biri verimliliğidir. Buna ek olarak, bu algoritmanın performansı büyük ölçüde iki faktöre, yani uygun bir benzerlik fonksiyonuna ve K değerinin seçimine bağlıdır. Sınıflandırma kuralları, ilave örnek veriler olmadan eğitim örnekleri tarafından oluşturulur. K-EYK sınıflandırma algoritması, test örneğinin kategorisini, test örneğine en yakın komşu olan K eğitim numunelerine göre öngörür ve en büyük kategori olasılığına sahip olan kategoriye göre değerlendirir.

Şekil 2.13'te görüldüğü gibi 4 sınıflı bir problemde yeni bir örneğin sınıfına karar verilmektedir. En yakın 3 komşuya bakılarak yeni gelen örnek mavi sınıfa ait olduğu söylenebilir. Yöntem uzaklıkların hesaplanmasında Öklit uzaklığı kullanılmaktadır. Öklit mesafe hesaplaması Eşitlik (2.32)'de görüldüğü gibidir.

$$d(i, j) = \sqrt{\sum_{k=1}^p (x_{ik} - x_{jk})^2} \quad (2.32)$$



Şekil 2.13. Yeni gelen örneğin sınıfının K-EYK ile belirlenmesi

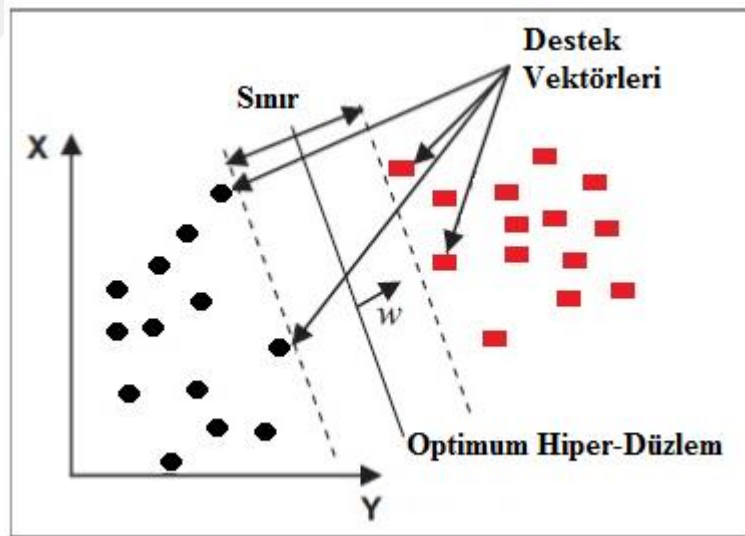
2.1.3.8.2. Destek Vektör Makineleri

DVM, ses tanıma uygulamaları için literatürde uygulanan başarılı bir algoritmadır. DVM, örüntü tanıma ve sınıflandırma problemlerinin çözümü için Cornes ve Vapnik

tarafından geliştirilmiştir [162]. DVM son yıllarda önemli derecede popüler olan güçlü ayırt edici bir sınıflandırıcıdır. Örneğin bir konuşmacı ve farklı konuşmacılardan oluşan bir dizi arasındaki sınırları modelleyebilir. Bu yaklaşım, konuşmacı ve genel popülasyonun olasılık dağılımlarını ayrı ayrı modelleyen geleneksel konuşma yöntemlerinden farklıdır. DVM'ler, girdileri yüksek boyutlu bir alana haritalandırır ve sınıfları bir hiper düzlem ile ayırır. DVM, sınıflandırma görevini, daha yüksek boyutsal bir alanda, verileri en iyi şekilde iki kategoriye ayıran hiper düzlemi oluşturarak gerçekleştirir. DVM'nin en önemli özelliği, doğrusal olmayan karar bölgelerini ayırt edici bir şekilde oluşturulabilme yeteneğidir.

2.1.3.8.2.1. Doğrusal Ayrılabilir Veriler İçin DVM

2 sınıflı bir problemde verileri sınıflandırmak için bu 2 sınıf arasından sonsuz sayıda düzlem geçirebiliriz. DVM'nin amacı Şekil 2.14'de görüldüğü gibi kendisine en yakın noktalar arasındaki uzaklığı maksimuma çıkaran hiper-düzlemi bulabilmektir.



Şekil 2.14. Doğrusal ayrılabilir veriler için hiper-düzlem seçimi [163]

Şekil 2.14'de görüldüğü gibi en uygun hiper düzlemin belirlenmesi için $\|w\|$ ifadesinin minimum hale getirilmesi gerekmektedir. Eşitlik (2.33) kullanılarak minimize edilir [163].

$$\min \left[\frac{1}{2} \|w\|^2 \right] \quad (2.33)$$

Sınırlamaları ise Eşitlik (2.34)'de gösterilmiştir [163].

$$y_i(w \cdot x_i + b) - 1 \geq 0 ; y_i \in \{1, -1\} \quad (2.34)$$

Eşitlik (2.34) denklemini Lagrange denklemi kullanılarak Eşitlik (2.35) elde edilir [163].

$$L(w, b, a) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^k a_i y_i (w \cdot x_i + b) + \sum_{i=1}^k a_i \quad (2.35)$$

Sonuç olarak, karar fonksiyonu Eşitlik (2.36) yazılabilir [163].

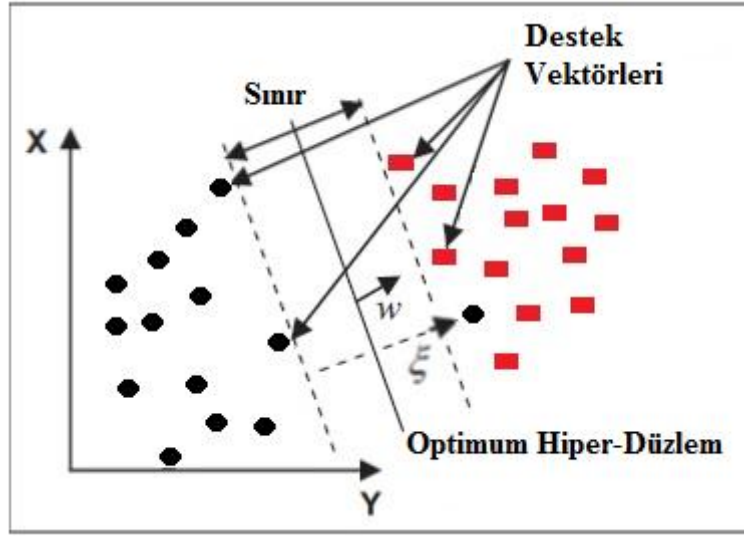
$$f(x) = \text{sign} \left(\sum_{i=1}^k \lambda_i y_i (x \cdot x_i) + b \right) \quad (2.36)$$

2.1.3.8.2.2. Doğrusal Ayrılamayan Veriler İçin DVM

Şekil 2.15'de görüldüğü gibi birçok gerçek dünya uygulamasında veriler doğrusal olarak birbirinden ayrılamazlar. Bu durumda veriler üzerinde sınıflandırma işlemi çekirdek fonksiyonları kullanılarak verinin yüksek boyutlu başka bir uzaya taşınması ve bu uzayda doğrusal ayırıcı hiper-düzlemin aranması ile gerçekleştirilir.

Sınır genişliğinin maksimum hale gelmesi ve yanlış sınıflandırmanın en aza indirilmesi için Eşitlik (2.37)'deki gibi C ($0 < C < \infty$) düzenleme parametresi eklenir [163].

$$\min \frac{1}{2} \|w\|^2 + C \sum_{i=1}^r \xi_i \quad (2.37)$$



Şekil 2.15. Doğrusal ayrılamayan veriler için hiper-düzlem seçimi [163]

DVM matematiksel olarak uygun çekirdek fonksiyonu yardımıyla doğrusal olmayan dönüşümler yaparak verilerin ayrılabilmesini sağlar. Çekirdek fonksiyon kullanılarak doğrusal olarak ayrılamayan verilerde karar kuralı Eşitlik (2.38)'de gösterildiği gibi yazılabilir [163].

$$f(x) = \text{sign} \left(\sum_i \alpha_i y_i \phi(x) \cdot \phi(x_i) + b \right) \quad (2.38)$$

Literatürde kullanılan çekirdek fonksiyonları doğrusal, çokterimli, polinom, normalleştirilmiş polinom, radyal tabanlı fonksiyon, sigmoid olarak söylenebilir.

2.2. Yön Bulma Çalışması

2.2.1. Sistemin Yapısı

Geliştirdiğimiz titreşim tabanlı giyilebilir cihazımız duyma engelliler için önemli bir problem olan hangi yönden ses geldiğine çözüm olmaktadır. Duyma engelli bir kişi ses varsa

bunu algılamalı ve ani bir durum için yönünü de anlamalıdır. Bu çalışma ile sesin yönü tespit edilmiş ve bu şekilde duyma engellilerin sesin nereden geldiğini algılaması sağlanmıştır. Özellikle arkadan gelen veya göremediği sesler duyma engelliye tedirgin etmektedir. Geliştirdiğimiz cihaz sayesinde duyma engelli birey arkasından gelen sesi algılayabilmekte ve daha güvenli bir şekilde seyahat edebilmektedir. Bu çalışma ile aynı zamanda duyma engelli birisinin yanında bağırarak konuşulup, konuşulmadığı da tespit edilmiştir. Bir kişi panik durumu varken duyma engelliye bunu anlatırken ses tonu yükselebilir. Bunun gibi panik durumunda olan bir durumu duyma engellinin daha erken fark etmesi hedeflenmiştir. Duyma engelli geliştirdiğimiz sistem sayesinde kişinin bağırıp bağırmadığını anlayabilecektir ve önemli bir durum varsa bunu algılaması hızlanmaktadır. Örneğin, duyma engelli bir insan zil sesini, arkasından yaklaşan köpek sesini, ona seslenen birisini anlayabilecektir. Bu sayede toplum içinde daha rahat, güvenilir ve daha az stresle bir hayat yaşamaları sağlanacaktır. Sağır kişiler özellikle göremediği bölgede önemli bir ses varsa bunu algılayabilmeleri güvenilirlik hissi vermektedir. Aynı zamanda, ev ortamında özellikle uyurken ona seslenen kişileri gerçek zamanlı olarak algılayabilmesi duyma engelli bireyleri daha huzurlu hissetmelerini sağlamıştır.

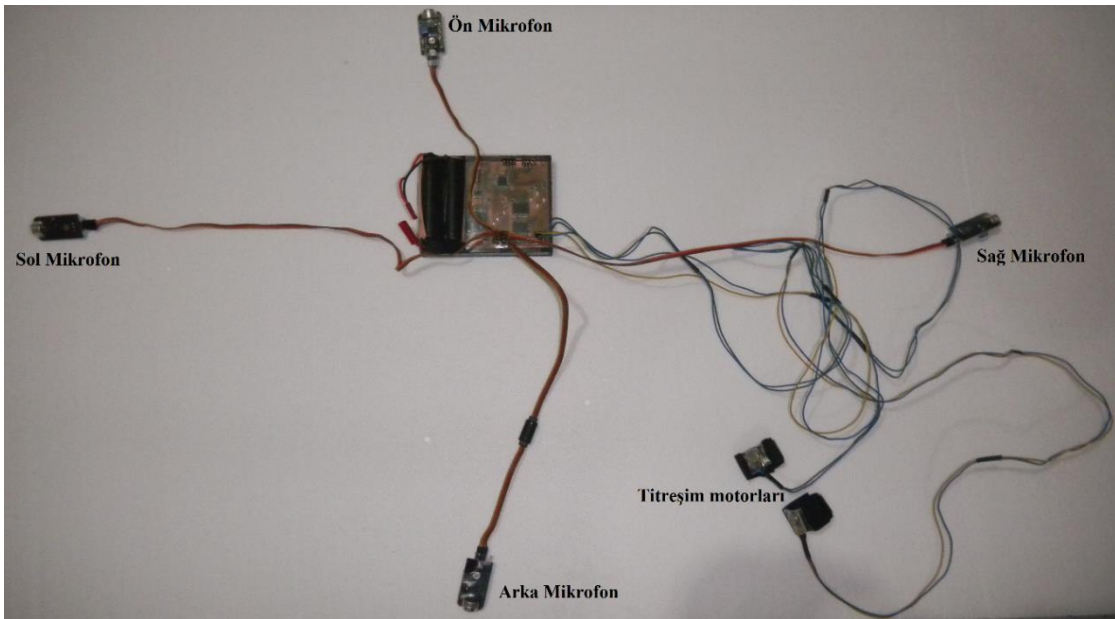
Şekil 2.16’da görüldüğü gibi geliştirilen sistemde dört mikrofon bulunmaktadır. Dört temel yönün ayırt edilebilir ölçüde birbirinden ayrılması için dört mikrofon kullanılmıştır. Her bir yönden gelen veriler dört adet yön tablosuna eklenir. Dört adet veri tablosu kullanılarak yeni gelen bir ses verisi tahmin edilmiştir. Öncelikli olarak üç mikrofon kullanılarak sistem test edilmiştir, ancak başarı oranı düşük çıktığı için dört temel yön olduğundan dört mikrofon kullanılmıştır. Mikrofonlar kişinin sağ, sol, ön ve arkasına geliştirilen insan-bilgisayar ara yüzü sistemi sayesinde yerleştirilmiştir [164].

Geliştirilen sistemde iki adet titreşim motoru kullanılmıştır. Sağ ve sol parmak uçlarına yerleştirilen titreşim motorları sayesinde, sesin hangi yönden geldiği belli titreşim sıklıkları ile gösterilmiştir. Parmak uçlarının tercih edilmesindeki en önemli sebep, parmak uçlarındaki sinir sayılarının fazla olmasıdır. Ayrıca parmak uçlarına yerleştirilen titreşim motorları daha kolay kullanılmakta ve kişiyi rahatsız etmemektedir.

Geliştirilen sistemde dört adet LED çıkışı bulunmaktadır. Ses algılandığı zaman titreşim hangi yönden geldiğini algılasa o çıkışa ait olan LED de yanmaktadır. Bu sayede hem titreşimlerle hem de LED ile yönünü kullanıcının daha belirgin bir şekilde algılamasını sağlamaktadır. LED kullanmamızın sebebi ise görsel uyarı vermesidir. Aynı zamanda dört farklı yön için dört farklı LED yanması düşünülmüştür. Kullanan kişi titreşimleri

anlamaması durumunda LED'lere bakarak anlayabilmektedir. Bu çalışmada titreşimin rolü ise duyma engelli bireylerin dokunma duyusunu harekete geçirmektir. Dokunma yoluyla duyma engelli kişi daha kolay anlayabilecek ve rahat bir yaşam kalitesine sahip olacaktır.

Giyilebilir cihazımız için şarj edilebilir piller kullanılmıştır. Piller yaklaşık 10 saat çalışabilmektedir. Gerçek zamanlı olarak sistemimiz yeni gelen bir veriyi tahmin edip hemen kullanıcıya titreşim ile bilgi vermektedir. Giyilebilir cihazımız sistem yanıt verdikçe yeniden yönü tahmin etmekte ve kullanıcıya iletmektedir. Cihaz ARM-tabanlı 32-bit MCU Flash bellek, 3.6 V, 7 sayıcı, 2 analog/dijital dönüştürücü gibi teknik özelliklere sahiptir.



Şekil 2.16. Yön tanıma çalışması için geliştirdiğimiz giyilebilir cihaz

Titreşimde kullanan kişiler 60 ms farkla gelen sesi algılayabiliyor, titreşim verdikten sonra 60 ms sonra gelen ses için sesin yönünü tespit edebiliriz. Yani 60 ms sonra gelen sesler için kişi ayırt edebilmektedir.

Parmağa 3 farklı kademe titreşim şiddeti uygulanmıştır:

1. Kademe düşük sesli sesleri algılaması için 0,5 V-1 V
2. Kademe orta sesli sesleri algılaması için 1 V-2 V
3. Kademe yüksek sesli sesleri algılaması için 2 V-3V

Bu şekilde titreşimin şiddetini ayarlayarak kişinin yüksek sesle mi, düşük sesle mi konuştuğunu da algılanabilecektir. Geliştirdiğimiz giyilebilir titreşim tabanlı cihaz ile duyma engelli kişi yanında yüksek sesle konuşan biri olursa bunu algılayabilecek ve olaylara tepki

süresi hızlanacaktır. Yanında bağırarak konuşan birinin kendisine bir sorundan bahsetmiş olması ve buna daha dikkatli davranması gerekebilir. Kişinin bağırıp bağırmadığını bu şekilde ayırt edebiliyoruz. Bunu yapmamızın temel amacı duyma engelli birisinin olaylara tepki süresini hızlandırmaktır. Yanında bağırarak konuşan birinin kendisine bir sorundan bahsetmiş olması ve buna daha dikkatli davranması gerekebilir.

Kişiler test edilirken 8 yön baz alınmıştır. Sol, sağ, ön, arka ve bu yönlerin kesiştiği dört yönde ses verilerek başarılı olup olmadığı test edilmiştir.

2.2.2. Eğitim Verisi

Dört sınıflı bir problem bulunmaktadır. Dört adet sınıfa ait eğitim verileri alınarak matrislere aktarılmıştır. Her sınıfa ait eğitim verileri öznitelikler çıkartılarak test için ayrılan veriler için tahmin edilmiştir. Dört farklı mikrofondan alınan veriler dört farklı sınıfı temsil etmektedir.

2.2.3. Ön İşleme

Ön işleme aşamasında normalizasyon kullanılmıştır. Normalizasyon için Eşitlik (2.38) kullanılarak veriler arasındaki uç noktalar azaltılır. Eşitlikte, girdi verisinden girdi verisinin ortalaması çıkartıldıktan sonra, girdi setinin standart sapmasına bölünerek normalize edilmiş veri elde edilir.

$$x' = \frac{x_i - \mu_i}{\sigma_i} \quad (2.39)$$

2.2.4. Kullanılan Öznitelik Yöntemleri

Çeşitli öznitelik çıkarma yöntemleri vardır. Bunlar; bağımsız bileşenler analizi, dalgacık dönüşümü, fourier analiz, ortak uzamsal örüntü, toplam, ortalama, varyans, standart sapma, etkinlik, basıklık (kurtosis), çarpıklık (skewness), polinom uydurma gibi sıralanabilir.

2.2.4.1. Çarpıklık

Çarpıklık, dağılımın asimetrik bir ölçüsüdür. Aynı zamanda, normal dağılımda simetrikliğin bozulma derecesidir. Dağılım sağa uzun kuyruklu ise pozitif çarpık veya sağa çarpık, dağılım sola uzun kuyruklu ise negatif çarpık veya sola çarpık olarak adlandırılır. Çarpıklık, Eşitlik (2.40) ile ifade edilebilir:

$$\zeta = \frac{\frac{1}{L} \sum_{i=1}^L (x_i - \bar{x})^2}{\left(\frac{1}{L} \sum_{i=1}^L (x_i - \bar{x})^2 \right)^3} \quad (2.40)$$

2.2.4.2. Basıklık

Basıklık, aykırı-eğimli bir dağılımın nasıl olduğunun bir ölçüsüdür. Basıklık ölçüsü dağılımı Eşitlik (2.41) ile ifade edilebilir [165]:

$$b = \frac{E(x - \mu)^4}{\sigma^4} \quad (2.41)$$

Burada, μ x değerlerinin ortalamasını; σ , x değerlerinin standart sapmasını; E(t); t miktar beklenen değeri temsil eder.

2.2.4.3. Kare Ortalamalarının Karekökü

Kare Ortalamalarının Karekökü (KOK), sinyalin toplamının ortalamasının kareköküdür. KOK hesaplanmasında kullanılan matematiksel denklem Eşitlik (2.42) formülünde görüldüğü gibidir:

$$X_{KOK} = \sqrt{\frac{x_1^2 + x_2^2 + \dots + x_N^2}{N}} \quad (2.42)$$

X iki nokta arasındaki dikey mesafeyi temsil eder ve N ise kıyaslanan iki yüzey arasındaki referans noktalarının toplam sayısını gösterir.

2.2.4.4. Varyans ve Standart Sapma

Varyans, standart sapmanın karesidir. Varyans, Eşitlik (2.43) ile hesaplanır. Dağılımın ölçüsüdür, elimizdeki veri setindeki değerlerin ortalamaya göre dağılımını gösterir. O an ki değerimiz ile ortalama arasındaki sapmaya göre değişim gösterir. Standart sapma ile de verilerin ne kadarının ortalamaya yakın olduğu bulunur.

$$var(x) = \frac{1}{n} \sum_{i=1}^n (x_i - \mu)^2 \quad (2.43)$$

2.2.4.5. Yerel Maksimum ve Yerel Minimum

Yerel maksimum (Ymak) ve Yerel minimum (Ymin) noktalarına yerel ekstremum noktaları denir. Yerel maksimumların en büyüğüne mutlak maksimum yerel minimumların en küçüğüne mutlak minimum noktaları adı verilir. Bu noktaları şu yöntemlerle bulabiliriz.

- Fonksiyonun türevi alınır varsa kökleri bulunur, işareti incelenir.
- Türev işareti değişirse ekstremum vardır. Aksi halde yoktur.
- Türevin kökleri ekstremum noktalarının apsiseridir. Bunlar fonksiyonda yerine yazılır koordinatlar bulunur.
- Türevin işaretinin (+) dan (-) ye geçtiği nokta yerel minimum noktasıdır. (-) den (+) ya geçtiği nokta yerel maksimum noktasıdır.

2.2.5. Öznitelik Seçme

Öznitelik seçme aşamasında ReliefF kullanılmıştır. ReliefF, Relief istatistiksel modelinin geliştirilmiş versiyonudur. ReliefF, yaygın olarak kullanılan öznitelik seçme algoritmasıdır [166]. ReliefF metodu, veri setinden bir örnek ele alarak ilgili örneğin, kendi sınıflarındaki diğer örneklerle yakınlığını ve farklı sınıflarla olan uzaklığına bağlı bir model oluşturarak öznitelik seçme işlemini gerçekleştirmektedir [167]. Bu çalışmada, her bir öznitelik için diğer özniteliklere göre değerlendirilmesi ve hangi özniteliklerin sınıflandırma aşamasında daha etkili olduğunu belirleyebilmek için ReliefF öznitelik seçim yöntemi uygulanmıştır. ReliefF yöntemi, özniteliklerin değerini aralarında bulunan ya da

bulunmayan bağımlılıkları ortaya çıkarmaya çalışarak bulmayı hedefler. ReliefF, Eşitlik (2.44) ile hesaplanır [166]:

$$W[A] = W[A] - \frac{diff(A,R,H)}{m} + \sum_{C \neq sınıf(R)} [P(C) \times diff(A, R, M(C))]/m \quad (2.44)$$

Burada P(C); C sınıfının gözlem uzayında bulunma olasılığını, sınıf (R); R gözleminin sınıfını, A gözleme ait özellik vektörü, R rasgele seçilen örnek, m gözlem sayısıdır.

2.2.6. Mikrofon Sınıfı Kullanan Lokalizasyon Algoritmaları

Literatürdeki konum kestirim yöntemleri genellikle varış zamanı (TOA), varış zaman farkı (TDOA) ve alınan sinyal gücü (RSS) tabanlı yöntemlerdir [168]. TDOA tabanlı yöntemler yüksek kesinlikle kestirim yapabildiği için oldukça avantajlıdır. En büyük olabilirlik yaklaşımını kullanan TDOA tabanlı yöntemler bir başlangıç değerine ihtiyaç duyarlar ve iteratif bir şekilde optimum sonuca ulaşmaya çalışırlar [169].

Bu yöntem, sinyalin hedeften, sinyalin alındığı zaman ve sinyalin hangi hızda geçtiği zamanını gerektirmez. Sinyal iki referans noktasında alındıktan sonra, varış zamanı farkı, hedef ve iki referans noktası arasındaki mesafeler arasındaki farkı hesaplamak için kullanılabilir. Bu fark Eşitlik (2.45) kullanılarak hesaplanabilir:[170,171]

$$\Delta d = c * (\Delta t) \quad (2.45)$$

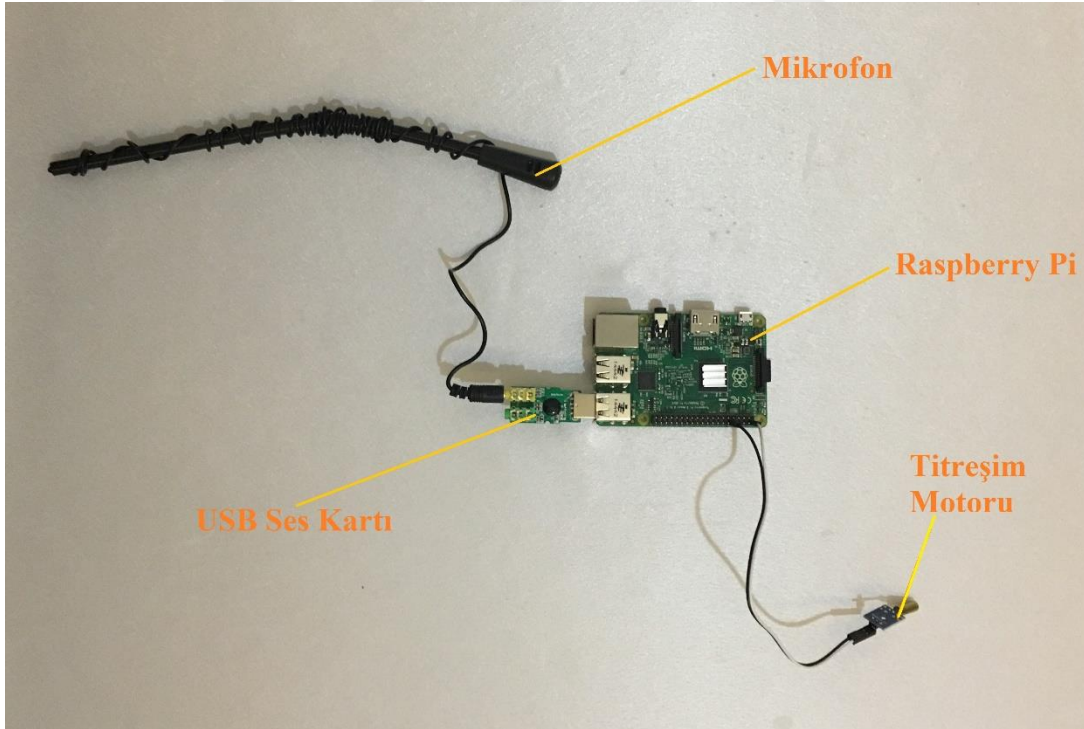
Burada c ışık hızıdır ve Δt her referans noktasında varış zamanlarındaki farktır. İki boyutta, Eşitlik (2.46) ile hesaplanır:[171]

$$\Delta d = \sqrt{(x_2 - x)^2 - (y_2 - y)^2} - \sqrt{(x_1 - x)^2 - (y_1 - y)^2} \quad (2.46)$$

Burada (x1, y1) ve (x2, y2) işaretlerin bilinen konumlarıdır. Doğrusal olmayan regresyon kullanarak, bu denklem hiperbol formuna dönüştürülebilir. Yeterince hiperbol hesaplandıktan sonra, hedefin konumu bulunarak hesaplanabilir.

2.3. Ebeveyn Sesi Tahmini Çalışması

Duyma engelli kişilerin ev ortamında kullanımını kolaylaştırmak için giyilebilir bir cihaz geliştirilmiştir. Cihazımız kullanıcının sırt bölgesine yerleştirilmiştir ve sadece yüzük büyüklüğünde bir titreşim motoru kişinin parmağına bağlanmıştır. Cihazımız Raspberry Pi, USB ses kartı, mikrofon, güç kaynağı ve titreşim motorundan oluşmaktadır. Kişilerin sırt bölgesine monte edilen cihazımız titreşimler ile duyma engelli bireye bilgi vermektedir. Bu sayede duyma engelli birey çok rahatlıkla ebeveyn sesini tahmin edebilmiştir. Uyurken ise yanına konulan cihazımız ebeveyn sesini duyduğu anda kişiyi titreşimlerle uyarmaktadır. Bu çalışmada kullanılan giyilebilir cihazımız Şekil 2.17’de görüldüğü gibidir. Mikrofon gelen sesleri algılayıp Raspberry Pi cihazına aktarır. Raspberry Pi de ise öznitelik çıkarma ve sınıflandırma işlemi yapılarak sesin kime ait olduğu tespit edilir. En son olarak da titreşimlerle ev ortamında kimin konuştuğu bilgisi duyma engelli kişiye aktarılır [172].



Şekil 2.17. Ebeveyn sesi tahmini için geliştirdiğimiz giyilebilir cihaz

Öncelikle ses mikrofon ile alınmaktadır ve örnekleme yapılmaktadır. Öznitelik çıkarma aşamasında MFKK, sınıflandırma aşamasında K-EYK kullanılmıştır. MFKK ses

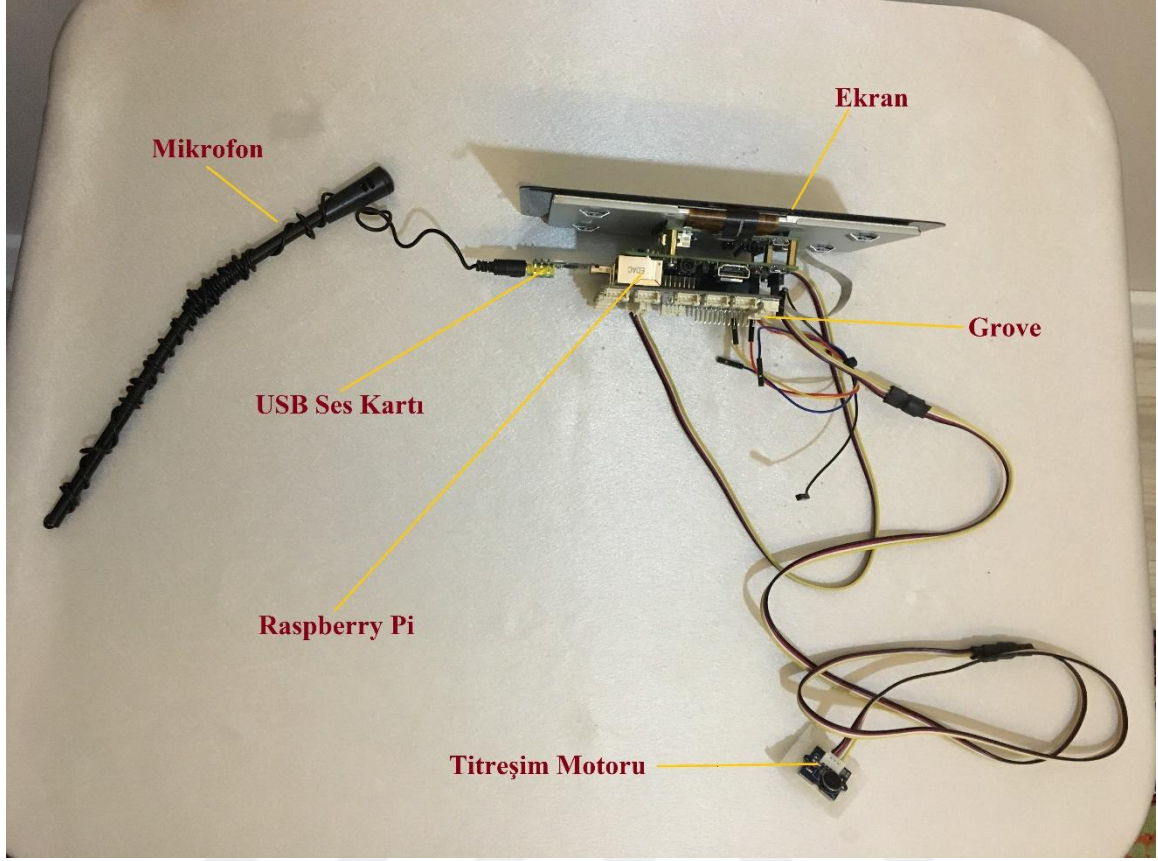
tanıma uygulamalarında sıklıkla kullanılan öznitelik çıkarım yöntemlerinden biridir. MFKK Ses sinyalinin kısa süreli güç spektrumunu temsil eder ve insan kulağının algılama şeklini modellemektedir. K-EYK, eğitilmiş öğrenme algoritmasıdır ve amacı, yeni bir örnek geldiğinde var olan öğrenme verisi üzerinde sınıflandırma yapmaktır. Öznitelik çıkarma ve sınıflandırma aşamaları ile mikrofon ile alınan ses verisi tahmin edilmiş ve sağır kişiye gerçek zamanlı olarak titreşimlerle bu sesin kime ait olduğu bilgisi verilmiştir [172]. Bu çalışmada aynı zamanda SPİ yöntemi kullanılarak da test edilmiş ve sonuçlar karşılaştırılmıştır.

2.4. Diğer Uygulamalar

2.4.1. Horlama Sesi Tahmini Çalışması

Şekil 2.18'de görüldüğü gibi horlama sesi tahmini için kullandığımız cihazımız 6 parçadan oluşmaktadır. Ekran, Raspberry Pi üzerindeki DSI konektörünü kullanır. Yaptığımız uygulama sonuçlarını görsel olarak gösterir. Ekran üzerinde horlama sesleri anlık olarak görüntülenebilir. Raspberry Pi, kredi kartı büyüklüğünde mini bilgisayardır. Grove, Raspberry Pi'yi Grove sensörlerine bağlamamızı sağlayan karttır. Mikrofon, ortamdaki sesi almamız ve işlememiz için kullanılır. Titreşim motoru ise kullanıcıya dokunma duyusunda algılaması için uyarı göndermektedir [173].

Geliştirdiğimiz giyilebilir cihaz, bireylerin uyku sağlığını bozan en sık, en önemli nedenlerinden olan horlamanın önlenmesi için tasarlanmıştır. Cihazımız gerçek zamanlı olarak çalışmaktadır. Horlama sesi algılandığı zaman ekran üzerinde horlama sesi ve şiddeti gösterilmektedir. Horlama sesi SPİ yöntemi ile tespit edilmiştir. Özellikle horlama problemi olan kişiler hem kendilerine hem de yanında bulunan kişilere zarar vermektedir. Horlama problemi olan insanlar uykuda solunum durması da yaşıyorsa kendini çok yorgun hissetmekte aynı zamanda ciddi sağlık problemleri yaşamaktadırlar. Geliştirdiğimiz cihaz ile horlama sesi tespit edildikten sonra titreşimler verilerek kişinin horlamadan çıkması sağlanmıştır.



Şekil 2.18. Horlama sesi tahmini için geliştirdiğimiz giyilebilir cihaz

3. BULGULAR

3.1. Değerlendirme Ölçütleri

Model başarımının ölçülmesinde en çok kullanılan, basit ve belirleyici ölçüt, modelin doğruluk değeridir. Doğruluk değeri, doğru sınıflandırılmış örnek sayısının toplam örnek sayısına oranı şeklinde ifade edilir. Değerlendirme ölçütlerinden diğer bir tanesi hata matrisidir. Hata matrisi gösterimi Tablo 3.1’de görüldüğü gibidir. Bu matriste, sadece sınıflandırma doğruluğunu vermek yerine açık açık Doğru Pozitif (TP), Yanlış Negatif (FN), Yanlış Pozitif (FP), Doğru Negatif (TN) sayıları verilmektedir. Bu sayılara göre doğruluk, duyarlılık ve özgüllük hesaplanır.

Tablo 3.1. Hata matrisi gösterimi

		Belirlenen Sınıf	
		Doğru Ses	Yanlış Ses
Gerçek sınıf	Doğru Ses	Doğru Pozitif (TP)	Yanlış Negatif (FN)
	Yanlış Ses	Yanlış Pozitif (FP)	Doğru Negatif (TN)

Bu tez çalışmasında değerlendirme ölçütü olarak doğruluk, duyarlılık ve özgüllük kullanılmıştır. Bunların hesaplanması (3.1) , (3.2) ve (3.3) eşitlikleri ile ifade edilmiştir.

$$Doğruluk = \frac{TP+TN}{TP+TN+FP+FN} \quad (3.1)$$

$$Duyarlılık = \frac{TP}{TP+FN} \quad (3.2)$$

$$Özgüllük = \frac{TN}{TN+FP} \quad (3.3)$$

Bu tez çalışmasında bir diğer değerlendirme ölçütü olarak Kappa katsayısı kullanılmıştır. Kappa kategorik değişkenlerin uyumluluğunu değerlendirmesinde kullanılır.

Kappa değeri (1, -1) arasındadır. Kappa hesaplanması, Eşitlik (3.4) ile ifade edilmiştir. Bu eşitlikte P_o gözlenen uyumluluk oranı, P_e rasgele beklenen uyumluluk oranıdır. Burada P_o ve P_e değerleri ise Eşitlik (3.5) ve Eşitlik (3.6) ile hesaplanır.

$$k = \frac{P_o - P_e}{1 - P_e} \quad (3.4)$$

$$P_o = \frac{TP + TN}{TP + TN + FP + FN} \quad (3.5)$$

$$P_e = \frac{(TP + FP) \times (TP + FN) + (TN + FP) \times (TN + FN)}{(TP + TN + FP + FN)^2} \quad (3.6)$$

Bu tez çalışmasında bir diğer değerlendirme ölçütü olarak ROC eğrileri kullanılmıştır. ROC analizi, hem model seçimi hem de model değerlendirmesi için tahmini modelleri karşılaştırmak için kullanılır. ROC eğrisi, tanıma testlerinde sürekli işaretlerin doğruluğunun bir ölçütü olarak kullanılır. Algoritma performanslarını değerlendirmek için kullanılabilen, ROC eğrisi ve özellikle ROC eğrisi altında kalan alan (AUC) en çok kullanılan ölçütlerdir [174]. Bu yöntem, model hassasiyeti ve özgüllük arasındaki dengeyi değerlendirmek için tıp, mühendislik ve sosyal bilimlerde sıklıkla uygulanır. ROC eğrisi ne kadar dikleşirse o kadar iyi sınıflandırma elde edilir. Sınıflar içerisinde deneme sayısı farklı olması durumunda faydalı olur. ROC eğrisinde duyarlılık ve özgüllük değerleri de görülür. (ROC eğrisi Şekillerinde Sensitivity: Duyarlılık; Specificity : Özgüllük ; AUC (Area Under Curve - Eğri altındaki alan))

3.2. Deney Ortamı

Gerçek zamanlı olarak gelen seslerin tahmini, geliştirdiğimiz giyilebilir cihaz vasıtasıyla yapılmıştır. Ses tanıma çalışmasında, Şekil 3.1'de gösterildiği gibi kullanıcının sırtına yerleştirilen giyilebilir cihazımız, kullanıcıya anında cevap vermektedir. Bu sayede, işitme engelli kullanıcı çevrelerindeki sesleri kolaylıkla anlayabilir ve böylece daha yüksek bir yaşam standardına sahip olabilir.

50 adet normal duyan bireylerin kulakları kulaklık ile kapatılmıştır ve deniz veya rüzgâr sesi verilmiştir. Bu sesleri kullanan kişileri rahatsız etmeyecek bir seviyede (16 dB-60 dB) ayarlanmıştır. Bu şekilde kullanan kişilerde dikkat dağıtıcı bir işlev göstermemektedir. Beş adet sağır kişi ve iki adet duyma güçlüğü olan kişi de giyilebilir cihazımızı kullanmıştır. Geliştirdiğimiz cihaz kullanımı kolay, hızlı sonuç üreten ve giyilebilir bir cihazdır.



Şekil 3.1. Ses tanıma çalışması için kullanılan giyilebilir cihaz test ortamı

Gerçek zamanlı olarak çalışan giyilebilir cihazımız, sesi öncelikle mikrofon ile algılamakta, daha sonra bu sesin hangi ses olduğunu bulmakta ve kullanıcıya titreşim yoluyla sesin hangi ses olduğunu bilgisini aktarmaktadır. Geliştirdiğimiz giyilebilir cihazımız dış ortamda ve iç ortamda kişinin üzerine monte edilmiştir. Gerçek zamanlı olarak çalışan sistemimiz dış ortamda iken örneğin köpek havlaması sesi duyduğu zaman anında bu sesi anlayabilmektedir. Aynı zamanda ev ortamında da kapı zili çaldığında yüksek başarı oranı ile anlayabilmekte ve kendisi için önemli olan sesleri çok rahat ayırt edebilmektedir. Sistemimiz her ses için farklı bir titreşim seviyesi ve şiddeti vermektedir. Öncelikle kullanıcı 10 dk eğitilmiştir. Eğitim aşamasında hangi ses için nasıl bir titreşim verildiği anlatılmıştır.

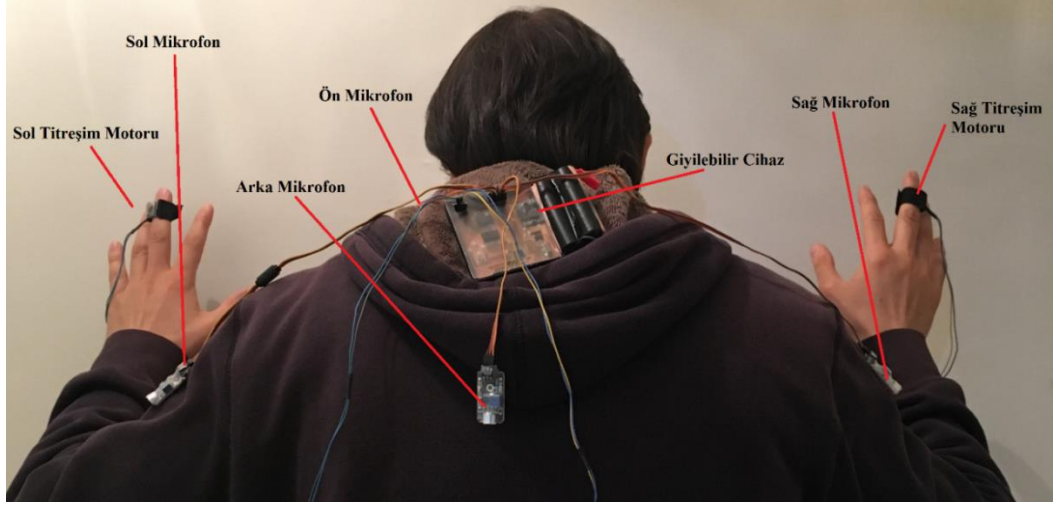
Daha sonra ise kullanıcı çok rahatlıkla ses farklılıklarını ayırt etmektedir. Ses örneklerinin yarısı eğitim için yarısı da test için ayrılmıştır.

Ses tanıma çalışmasında her ses için farklı bir titreşim seviyesi ve büyüklüğü verilmiştir. Bu sayede, sekiz önemli ses tanımlanan farklı titreşimler sayesinde ayırt edilmiştir. Eğitim aşamasında bu titreşimler sağır kişilere adım adım anlatılmış ve hangi sesin hangi titreşime uygun olduğunu hızla anlamışlardır. Örneğin, telefon, alarm ve kapı zili gibi sesler için sırasıyla bir, iki ve üç kez titreşim verilmiştir. Farklılıklar, titreşimler arasındaki büyüklüğe ve periyota dayanılarak yapılmıştır. Bu titreşimler sayesinde sağır birey dış ve iç ortamdaki sesleri kolaylıkla ayırt edebilmektedir. Ses tanıma çalışmasında giyilebilir sistemimiz gerçek zamanlı uygulamasında ses parmak izi yöntemi kullanılmıştır.

Ses tanıma çalışmasında bilgisayar ortamı testlerinde denemelerimiz eğitim ve test aşamaları için yarı yarıya bölünmüştür. 2000 deneme test için ayrılmıştır. Ses örneklerimiz 2-15 sn. aralığındadır. Bazı sesler farklı tonda olabilir. Bu bakımdan 14 farklı köpek havlaması sesi, 5 farklı fren sesi, 10 farklı korna sesi, 10 farklı kapı zili sesi, 5 farklı alarm sesi, 10 farklı telefon sesi, 10 farklı insan sesi, 10 farklı diğer ses eğitim aşamasında aynı ses olarak tanımlanmıştır.

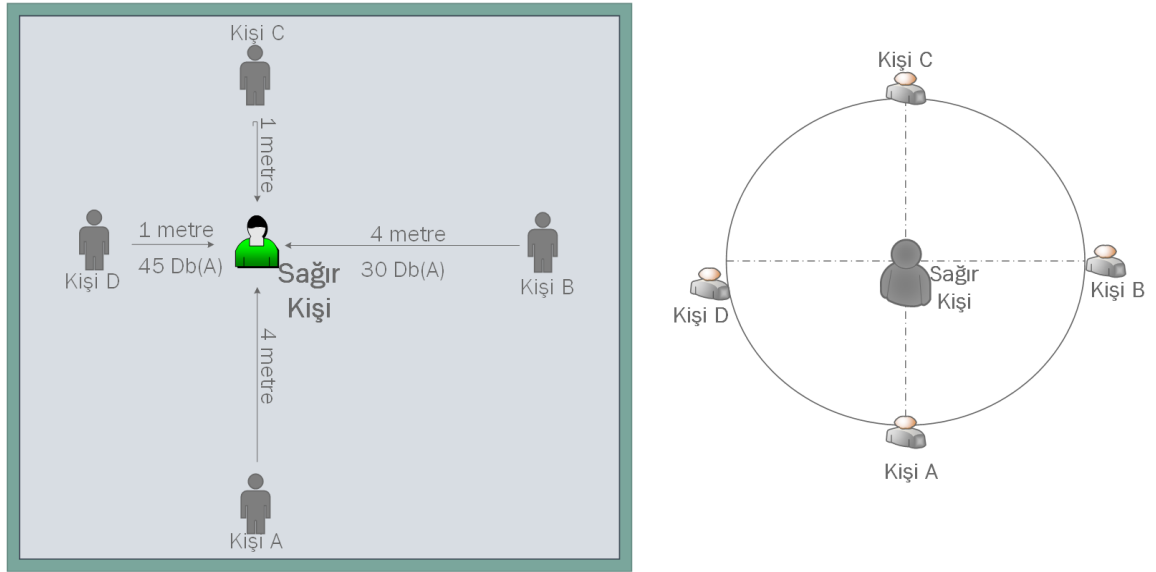
Yön bulma çalışmasında ise ilk olarak, deneysel çalışmalarda en iyi öznelik ve sınıflandırma yöntemi seçilmiştir. Ardından hem gerçek zamanlı hem de bilgisayar testleri gerçekleştirilmiştir. Testlerimiz gürültülü ortamda (30-40 dB(A)) 1-4 metre mesafede yapılmıştır. Deneyler hem kapalı ortamlarda hem de dışarı ortamda yapılarak test edilmiştir. Oda, koridor ve dış ortamlarda 1-4 metre mesafe ile testler yapılmıştır. Normal sesli veya bağırarak konuşan biri de çalışmalarımızda test edilmiştir. Prototip olarak geliştirdiğimiz giyilebilir cihaz ile sağır insanların daha rahat yaşam sürmesi sağlanmıştır.

Yön bulma çalışmasında bilgisayar ortamı testlerinde denemelerimiz eğitim ve test aşamaları için yarı yarıya bölünmüştür. 2400 deneme test için ayrılmıştır. Her bir yön için 600 deneme test aşamasında sınıflandırılmıştır. Yön bulma çalışmasında Şekil 3.2'de görüldüğü gibi dört mikrofon, dört LED ve iki adet titreşim motoru kullanılmıştır. Dört mikrofon dört yönden gelen verileri almamız için kullanılmıştır. Dört LED her yön için farklı görsel uyarı vermek için kullanılmıştır. Sağ ve sol ellerde bulunan titreşim motorlarında, sağ yönden ses alınmışsa sadece sağ titreşim motoru, sol yönden ses alınmışsa sadece sol titreşim motoru çalışmıştır. Arka ve önden gelen sesler için ise her iki titreşim motoru çalıştırılmış sadece titreşim şiddeti ve uzunluğu artırılarak arka ve ön yönden gelen sesler ayırt edilmiştir.



Şekil 3.2. Yön bulma çalışması için kullanılan giyilebilir cihaz test ortamı

Şekil 3.3’de kullanılan giyilebilir cihazın yön bulma sistemi test durumu gözükmemektedir. Duyma engelli kişinin arkasından biri seslendiğinde çok rahatlıkla kendisine seslenen kişinin yönünü algılayabilmektedir. Şekil 3.3’de görüldüğü gibi duyma engelli bireyin arkasında bulunan Kişi C’nin kendisine seslendiğinde duyma engelli birey bunu algılayıp algılayamadığı test edilmiştir.



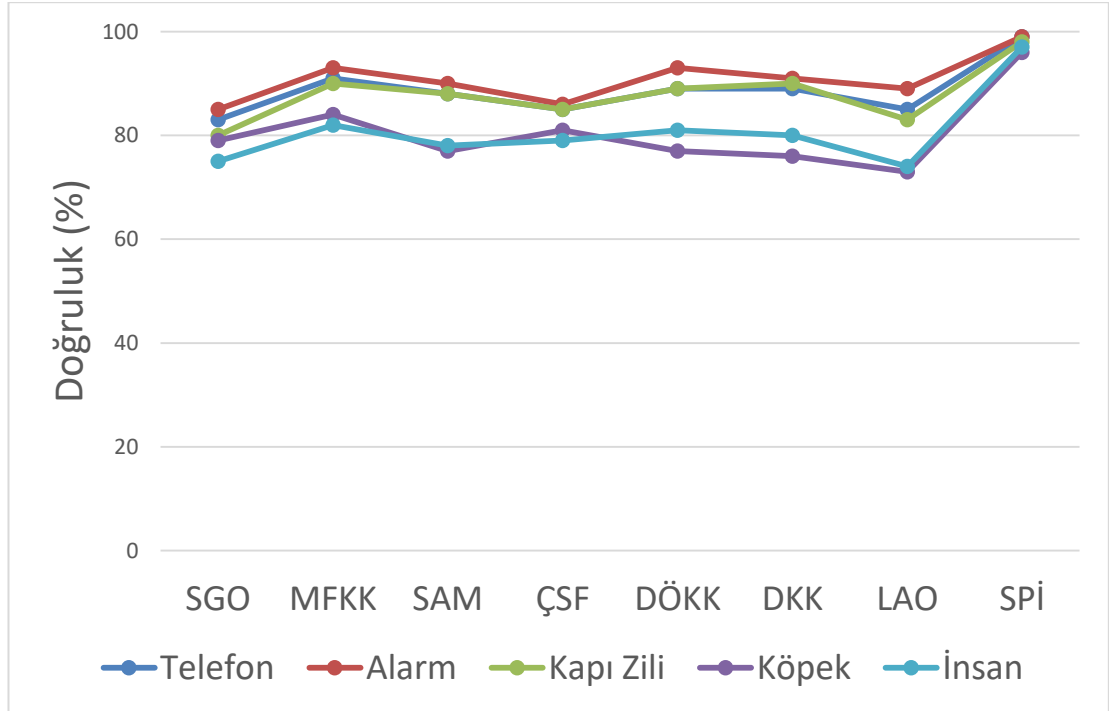
Şekil 3.3. Giyilebilir test sistemi

Normal kullanıcıların kulakları dışarıdan gelen sesi duyamayacak şekilde kulaklık ile kapatılmıştır. Geliştirdiğimiz giyilebilir cihaz kişinin sırtına yerleştirilmiş ve test edilmiştir. Kullanan kişi kulağındaki sese rağmen sesin nerden geldiğini anlamış ve o yöne yönelmiştir.

Gerçek zamanlı yapılan yön bulma çalışmasında ses kaynağı olarak insanlar kullanılmıştır. Dışarda yürürken arkasından biri gelip ona seslenerek, cihazı kullanan kişinin onu tespit edip etmemesi değerlendirilmiştir.

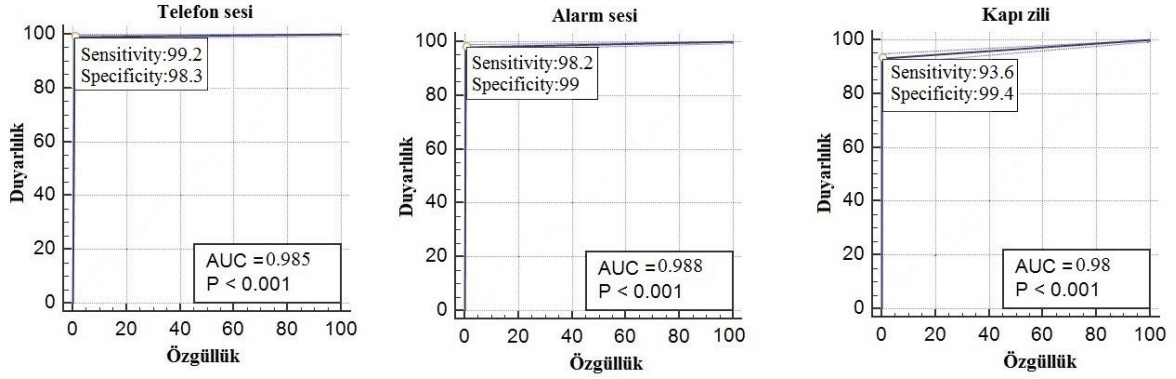
3.3. Ses Tanıma Çalışması

Duyuma engelliler için önemli sesleri algılamak önemli bir problemdir. Örneğin dışarıda yürürken aniden fren olabilir ya da arkasından köpek gelebilir. Bu gibi durumlarda duyma engelli kişi korkmakta ve yaşam kalitesi düşmektedir. Şekil 3.4 ve Şekil 3.6’da görüldüğü gibi gerçek-zamanlı çalışan giyilebilir sistemimizde en iyi sonuçlar SPİ yöntemi ile elde edilmiştir. SPİ yöntemi kullanılarak iç ve dış ortamda yüksek bir başarı elde edilmiştir. Şekil 3.4’de iç ortam seslerinin karşılaştırılması görülmektedir. İç ortamda telefon, alarm, kapı zili, köpek havlama sesi ve insan sesleri karşılaştırılmıştır ve en iyi sonuçlar SPİ yöntemi ile elde edilmiştir.



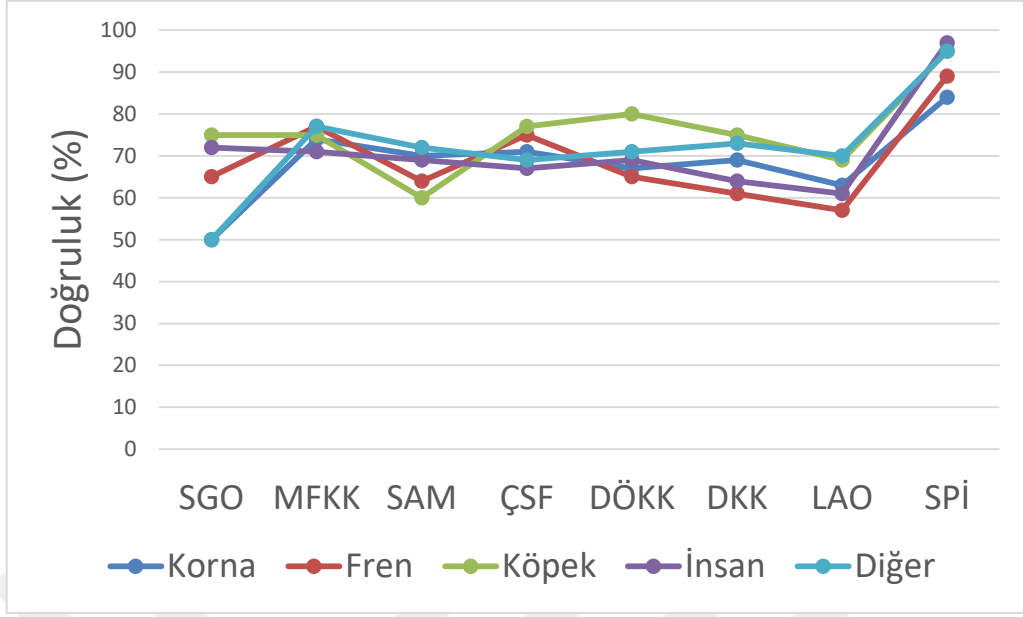
Şekil 3.4. İç ortam seslerinin karşılaştırılması

Gerçek zamanlı elde edilen iç ortam seslerinden telefon, alarm ve kapı zili seslerinin elde edilen sonuçlarının ROC eğrileri Şekil 3.5’de görüldüğü gibidir. Şekilden görüldüğü gibi iç ortam sesleri yüksek oranda tespit edilmiştir.



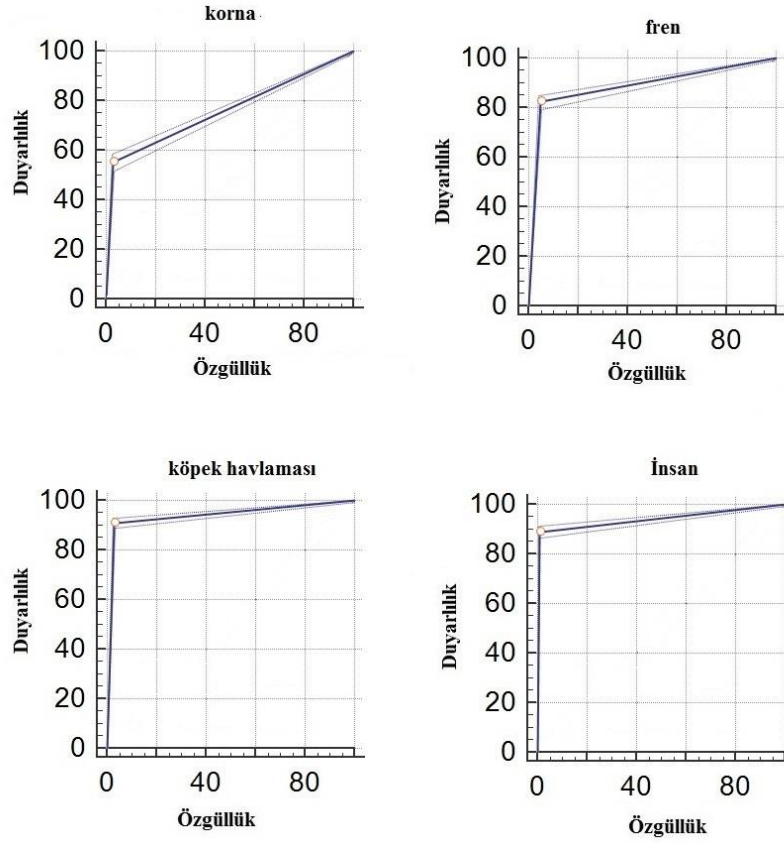
Şekil 3.5. İç ortam sesleri ROC eğrileri

Şekil 3.6’da dış ortam seslerinin karşılaştırılması görülmektedir. Örneğin korna, fren ve köpek havlaması sesi gibi dış ortam seslerinde şehrin genel gürültüsü ve diğer etkenlerle birlikte (ortalama 55 Db(A)) kişinin algılanması sağlanmıştır. Her ses için ses parmak izi oluşturulmuştur. Yeni bir ses geldiğinde bunun hangi ses olduğu tespit edilmiştir ve titreşimlerle kullanıcıya aktarılmıştır. Şekil 3.6’da görüldüğü gibi dış ortam seslerinde de korna sesinde %84, fren sesinde %87, köpek sesinde %93’lük bir başarı elde edilmiştir. SPİ metodundan sonra en iyi sonuç MFKK metodu ile elde edilmiştir.



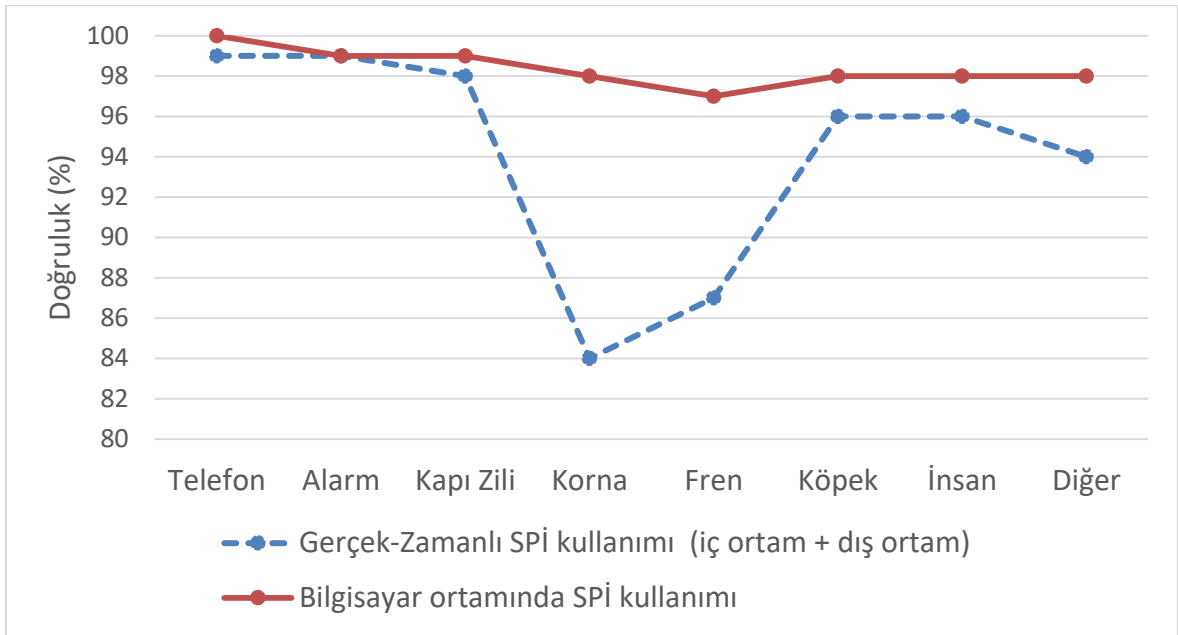
Şekil 3.6. Dış ortam seslerinin karşılaştırılması

Gerçek zamanlı elde edilen dış ortam seslerinden korna, fren, köpek havlaması ve insan seslerinin elde edilen sonuçlarının ROC eğrileri Şekil 3.7’de görüldüğü gibidir.



Şekil 3.7. Dış ortam sesleri ROC eğrileri

Çalışmamız aynı zamanda bilgisayar ortamında hoparlör yardımıyla sesler verilerek test edilmiştir. Gürültünün olmadığı bu testlerde ise geliştirdiğimiz yöntem tüm seslerde yüksek bir başarı elde etmiştir. Şekil 3.8’de gerçek-zamanlı test ile bilgisayar ortamında yapılan sonuçlar karşılaştırılmıştır. Bilgisayar ortamında yapılan çalışmalarda hoparlör ile ses verilmiştir ve bu sesin tahmini yapılmıştır. İç ortam seslerinde (telefon, alarm, kapı zili) %1 lik bir fark oluşmuştur. Bunun sebebi evin içinde oluşan farklı ses örnekleri ve sesin uzaklığı olarak söylenebilir. Şekil 3.8’de görüldüğü gibi özellikle dış ortamda oluşan başarı oranı düşüklüğü artmıştır. Bunun en temel sebebi gürültü ve mesafe olarak söylenebilir.



Şekil 3.8. Farklı test ortamında SPI yöntemi karşılaştırılması

Sinyal/Gürültü Oranı (SNR), arka plandaki gürültü seviyesiyle istenilen sinyalin düzeyini karşılaştıran bir ölçüdür. Genelde çalışmalarda dış ortam gürültüsü ve arka plan gürültüsü eklenerek sonuçlar değerlendirilmiştir. Tablo 3.2’de görüldüğü gibi çevresel sesler kullanarak yapılan çalışmalarda en iyi sonuçlar elde edilmiştir. Tabloda yapılan çalışmaların gürültülü ortamda ya da gerçek kişiler ile yapılan çalışmalar sonucu görülmektedir. Aynı zamanda tabloda kullanılan metotlar, yapılan deney ve deneme sayısı da gösterilmiştir. Sınıf sayımız diğer çalışmalardan az olmasına rağmen, deneme sayımız diğer çalışmalardan çoktur. Tablo 3.2’de görüldüğü gibi SNR 15 dB ve SNR 20 dB altında çalışmamız daha iyi sonuç üretmiştir. Aynı zamanda gürültüsüz ortamda da diğer çalışmalardan daha iyi sonuç

üretmiştir. Çevresel seslerin tanımlanması çalışmalarında başarı oranlarını Eronen ve diğerleri [175] %69, Ma ve diğerleri [141] %92, Su ve diğerleri [176] %81, Wang ve diğerleri [177] %85 olarak göstermişlerdir. Bu başarı oranları gerçek zamanlı olarak yapıldığında ise ciddi düşüşler gözükmemektedir. Örneğin Su ve diğerleri [176] çalışmasında gürültülü sesler olduğunda %28.6 ya, Eronen ve diğerleri [175] çalışmasında test işlemini gerçek kişiler üzerinde gerçekleştirdiğinde başarı oranı %58'e düşmektedir. Ma ve diğerleri [141] çalışmasında, insan dinleme testleri % 35'lik genel bir doğruluk hesaplanmıştır. Wang ve diğerleri [120] çalışmalarında 15 dB, 10 dB ve 5 dB olmak üzere 3 tane SNR seviyesi kullanmışlardır. Wang ve diğerlerinin [120] başarı oranı 15 dB gürültülü vakalarda % 84.25'e düşmüştür. Wang ve diğerlerinin. [120] başarı oranı, 5 dB gürültülü vakalarda % 77.44'e düştü. Çalışmamızda, kapalı ve açık alanda gerçek insanlara uygulanan gerçek zamanlı testlerde başarı oranı ortalama % 94'a düşmüştür. Çalışmamızda SNR ortalama 20 dB olarak hesaplanmıştır. Kapalı ve açık ortamlarda gerçek zamanlı ortalamaların ortalaması alınarak hesaplanmıştır. SNR 15 dB'deki denemelerde başarı % 90 oranında elde edilmiştir.

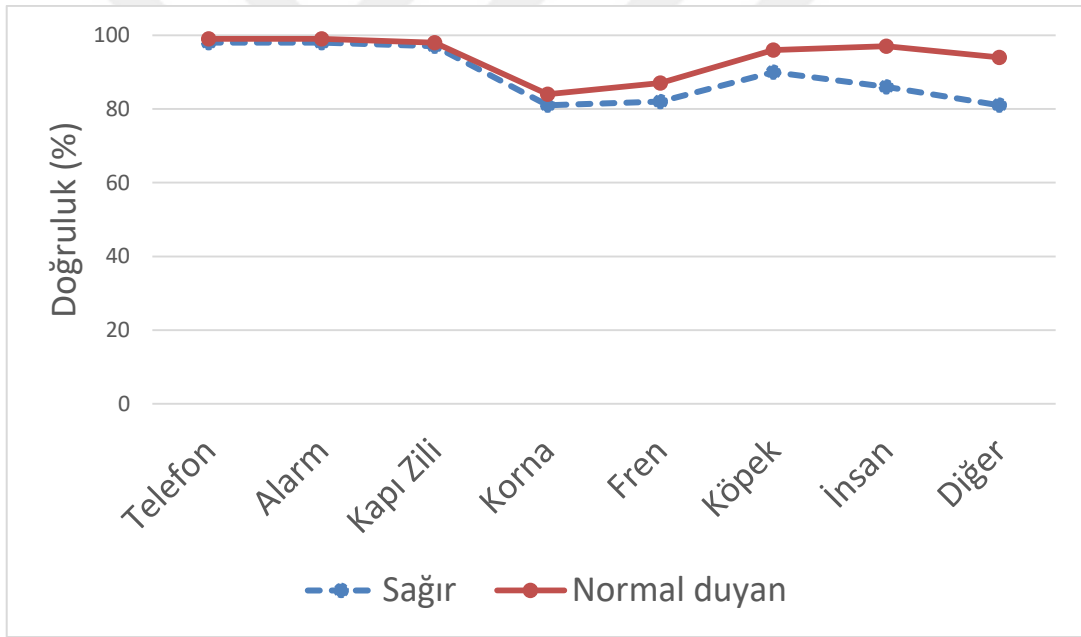
Tablo 3.2. Çevresel ses tanıma çalışmaları karşılaştırması

	Genel Doğruluk	Doğruluk (Gerçek-zamanlı, gerçek kişiler veya gürültülü)	Metot	Yapılan deney	Deneme sayısı
Ma vd. [141]	92%	35%	GMM + MFKK	Günlük rutinlere dayalı 12 çeşit akustik çevre sınıflandırması (10 farklı akustik ortam)	880 deneme (Eğitim 660 – Test 220)
Eronen vd. [175]	69%	58%	GMM , K-EYK	225 gerçek dünya kaydı, 24 farklı durum için ses tabanlı bağlam tanıma sistemi	Eğitim: 24 durumdan 155 deneme, 16 durumdan 70 deneme. Test: 18 farklı durumdan 45 deneme

Tablo 3.2'nin devamı

	Genel Doğruluk	Doğruluk (Gerçek-zamanlı, gerçek kişiler veya gürültülü)	Metot	Yapılan deney	Deneme sayısı
Su vd. [176]	81%	%28.6	YDB	21 çevresel ses tanıma çalışması(10 sesli sahneler)	Veri kümesi: internet televizyon ve filmlerden alınmıştır.
Wang vd. [177]	85.1%		DVM +K-EYK	12 yaygın kullanılan çevresel ses tanıma çalışması	527 deneme(eğitim:%50, test: %50)
Wang vd. [120]	89.78% (gürültüsüz)	84.25% (SNR:15 dB)	MFKK	10 sınıflı çevresel ses tanıma sistemi	469 deneme (eğitim:%50, test: %50)
Wang vd. [120]	90.63% (gürültüsüz)	88.51% (SNR:15 dB)	Dalgacık Altuzay tabanlı	10 sınıflı çevresel ses tanıma sistemi	469 deneme (eğitim:%50, test: %50)
Chu vd. [178]	83.9%	82% (insan dinleyiciler)	MFKK, DVM, MP	14 sınıflı çevresel ses sınıflandırması	140 ses klipi(Ses rastgele seçildi)
Wang vd. [7]	86.21%	60.24% (SNR:15 dB)	TBA, SVM	17 sınıflı çevresel ses sınıflandırması	749 çevresel ses (eğitim:%50, test: %50)
Malkin ve Waibel [179]	77%	70% - 80% (insan dinleyiciler)	GMM	11 çevresel ses	Eğitim: 7299, Test:1814
Tez çalışması	98% (gürültüsüz)	90% (SNR:15 dB)	SPİ	8 sınıflı çevresel ses tanıma sistemi (5 iç ortam, 5 dış ortam sesi)	10000 deneme
Tez çalışması	98% (gürültüsüz)	94% (SNR: ortalama 20 dB)	SPİ	8 sınıflı çevresel ses tanıma sistemi (5 iç ortam, 5 dış ortam sesi)	10000 deneme

Şekil 3.9'da duyma engelliler ile normal duyan kişilerin karşılaştırılması gösterilmiştir. Duyma engelliler için önemli sesleri algılamak önemli bir problemdir. Örneğin dışarıda yürürken aniden fren olabilir ya da arkasından köpek gelebilir. Bu gibi durumlarda duyma engelli kişi korkmakta ve yaşam kalitesi düşmektedir. Şekil 3.9'da görüldüğü gibi sağır kişilerle yapılan çalışmada iç ortam seslerinde çok yüksek bir başarı elde edilmiştir. Geliştirdiğimiz giyilebilir cihaz sayesinde sağır kişiler %98 telefon sesi, %98 alarm sesi, %97 kapı zili sesini algılayabilmektedir. Sağır kişi uyurken kapı zili ya da telefon çalabilir. Bu sayede sağır birey normal insanlar gibi ev ortamında daha kaliteli bir yaşam sürmekte ve önemli sesleri rahatlıkla algılayabilmektedir. Dış ortam seslerinde de sağır kişilerde korna sesinde %81, fren sesinde %82, köpek sesinde %90'lık bir başarı elde edilmiştir.



Şekil 3.9. Duyma engelliler ile normal duyan kişilerin karşılaştırılması

Bilgisayar ortamında yapılan ses tanıma işleminin hata matrisi Tablo 3.3'de görüldüğü gibidir. Bilgisayar ortamında her sınıf için 250 deneme yapılarak, toplamda 2000 deneme yapılmıştır.

Tablo 3.3. Ses tanıma çalışması bilgisayar ortamı sonuçları hata matrisi

	Telefon	Alarm	Kapı Zili	Korna	Fren	Köpek	İnsan	Diğer
Telefon	248	2	0	0	0	0	0	0
Alarm	1	247	2	0	0	0	0	0
Kapı Zili	2	1	245	0	0	0	1	1
Korna	0	0	0	221	19	8	0	2
Fren	0	0	0	30	214	5	0	1
Köpek	0	0	0	2	1	217	25	5
İnsan	2	1	1	2	2	4	224	14
Diğer	0	0	1	0	5	8	19	217

Gerçek zamanlı testlerde yapılan iç ortam seslerinin ses tanıma işleminin hata matrisi Tablo 3.4’de görüldüğü gibidir. Gerçek zamanlı testlerde yapılan dış ortam seslerinin ses tanıma işleminin hata matrisi Tablo 3.5’de görüldüğü gibidir. Gerçek zamanlı testlerde her sınıf için 1250 deneme yapılarak, toplamda 10000 deneme yapılmıştır.

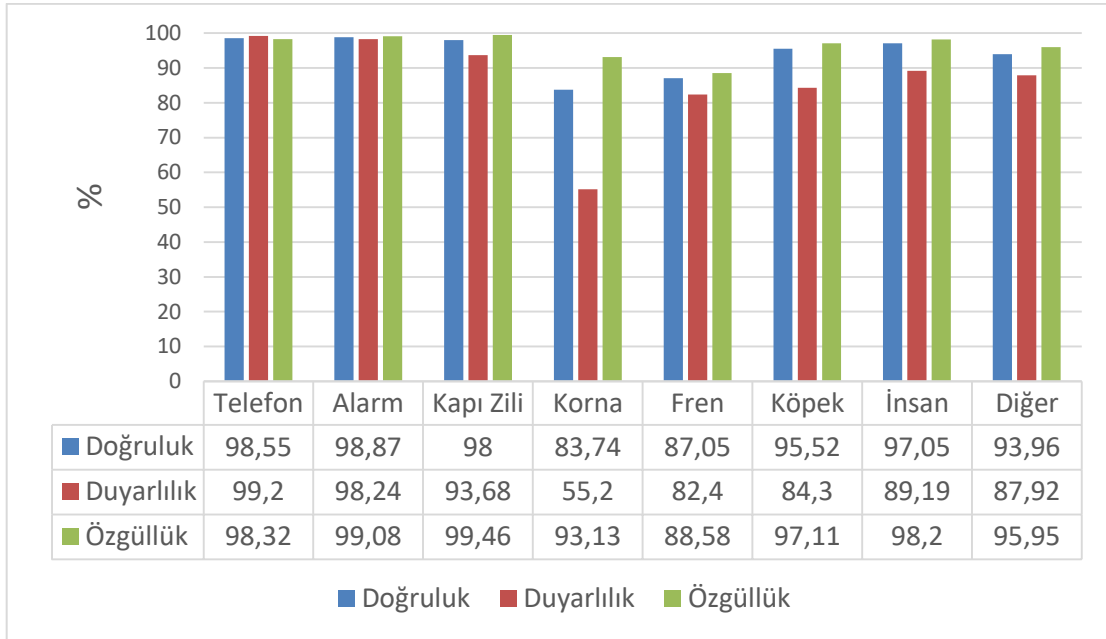
Tablo 3.4. İç ortam sesleri gerçek zamanlı test sonuçları hata matrisi

	Telefon	Alarm	Kapı Zili	Köpek	İnsan
Telefon	1240	3	7	0	0
Alarm	12	1228	10	0	0
Kapı Zili	48	30	1171	0	1
Köpek	0	0	2	514	84
İnsan	2	1	1	26	570

Tablo 3.5. Dış ortam sesleri gerçek zamanlı test sonuçları hata matrisi

	Korna	Fren	Köpek	İnsan	Diğer
Korna	690	380	123	7	50
Fren	153	1030	29	8	30
Köpek	75	14	539	5	17
İnsan	18	22	11	542	57
Diğer	15	18	65	53	1099

Gerçek zamanlı testlerden elde edilen sonuçların doğruluk, duyarlılık ve özgüllük sonuçları Şekil 3.10’da görüldüğü gibidir. Şekilden görüldüğü gibi en düşük doğruluk ve duyarlılık korna sesinden elde edilmiştir. Bunun en temel sebebi korna sesinin kısa bir ses örneği olduğu için özellikle köpek havlaması ve fren sesleri ile karışmaktadır. Aynı zamanda gürültülü ortamdan dolayı korna sesi diğer sesler ile karışabilmektedir. Telefon ve alarm seslerinde ise yüksek başarı, duyarlılık ve özgüllük elde edilmiştir.



Şekil 3.10. Ses tanıma çalışması gerçek zamanlı test sonuçları

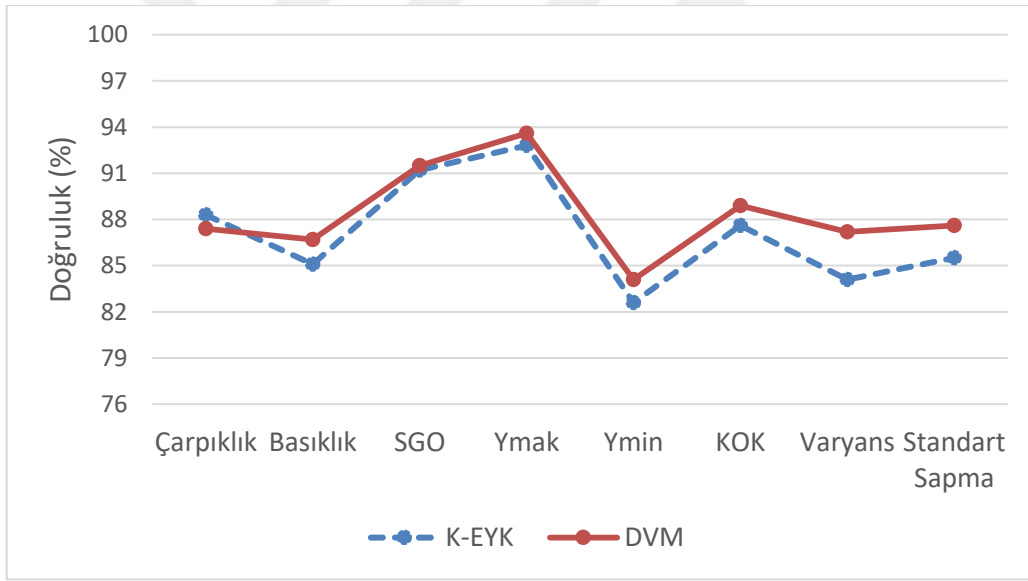
Gerçek zamanlı testlerden elde edilen sonuçların kappa katsayısı Tablo 3.6’da görüldüğü gibidir. Bu çalışmada tablodan görüldüğü gibi; telefon, alarm, kapı zili, köpek, insan ve diğer sesler çok iyi düzeyde uyumlu çıkmıştır. Fren sesi iyi düzeyde uyumlu, korna sesi ise orta düzeyde uyumlu çıkmıştır.

Tablo 3.6. Ses tanıma çalışması gerçek zamanlı test kappa değerleri

Telefon	Alarm	Kapı Zili	Korna	Fren	Köpek	İnsan	Diğer
0.97	0.97	0.95	0.58	0.72	0.84	0.90	0.86

3.4. Yön Bulma Çalışması

Bu tez çalışmasında sesin yönünü algılamak için sınıflandırma tekniği kullanılmıştır. Giyilebilir cihazımız 0.68 Saniyede cevap vermektedir. Kişi hareket halindeyken de sistemimiz çalışmaktadır ve ortalama %94 başarı elde etmektedir. Gerçek zamanlı yapılan deneylerde ses kaynağı olarak insanlar kullanılmıştır. Dışarda yürürken arkasından biri gelip ona seslenerek, cihazı kullanan kişinin onu tespit edip etmemesi değerlendirilmiştir. Şekil 3.11'de yapılan yön bulma çalışması için öznelik yöntemlerinin ve sınıflandırma yöntemlerinin başarısını göstermektedir. 1-4 metre mesafelerin ortalama sonuçları alınarak yapılan testlerde en iyi öznelik yöntemi yerel maksimumdur. Bu çalışmada ReliefF yöntemi kullanılarak en iyi iki öznelik seçilmiştir. En iyi iki öznelik yöntemi Ymak ve SGO çıkmıştır.



Şekil 3.11. Öznelik yöntemlerinin karşılaştırılması

Tablo 3.7'de gerçek-zamanlı çalışma sonuçları ile bilgisayar ortamında sonuçların karşılaştırılması görülmektedir. Sonuçlar 1-4 metre mesafelerde yapılan sonuçların ortalaması alınarak hesaplanmıştır. Tablo 3.7'de görüldüğü gibi gerçek zamanlı olarak elde edilen sonuçlar, verilerin bilgisayar ortamına aktarılıp elde edilen sonuçlar ile karşılaştırılmıştır.

Tablo 3.7. Yön bulma çalışması gerçek-zamanlı ve bilgisayar ortamı başarı sonuçları

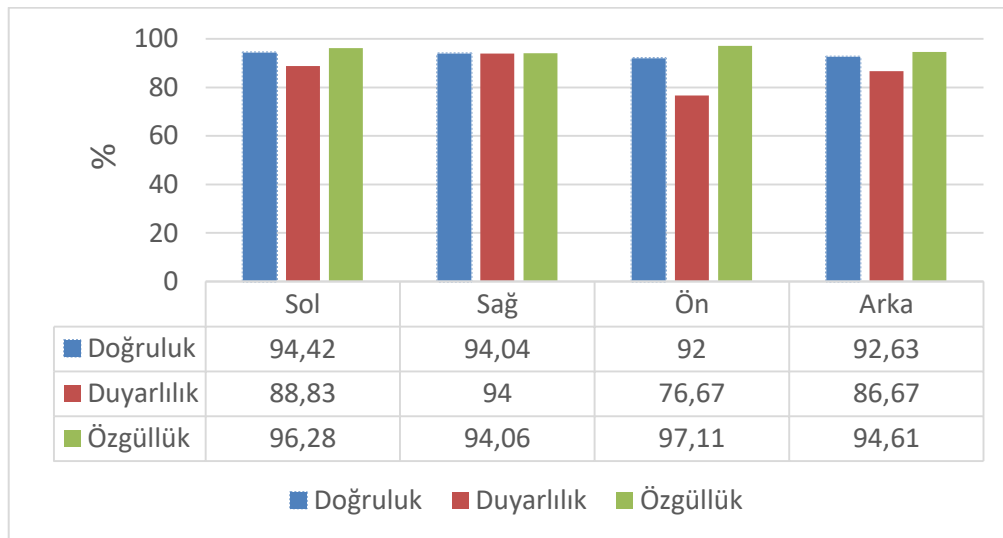
Test Ortamı	Sol	Sağ	Ön	Arka
Bilgisayar ortamı	94.42%	94.04%	92%	92.63%
Gerçek-zamanlı	92.88%	92.96%	90.88%	90.96%

Tablo 3.8'de bilgisayar ortamında elde edilen hata matrisi görülmektedir. Bu çalışmada her yön için 600 deneme yapılmıştır. Toplamda 2400 deneme yapılmıştır.

Tablo 3.8. Yön bulma çalışması bilgisayar ortamı sonuçları hata matrisi

	Sol	Sağ	Ön	Arka
Sol	533	62	3	2
Sağ	27	564	4	5
Ön	25	25	460	90
Arka	15	20	45	520

Şekil 3.12'de bilgisayar ortamında elde edilen sonuçların doğruluk, duyarlılık ve özgüllük değerleri gösterilmektedir. Şekilden görüldüğü gibi sol ve sağ yönler %94 lük başarı ile tahmin edilmiştir. Ön yönden gelen seslerde ise duyarlılık ve doğruluk değerleri düşmektedir.



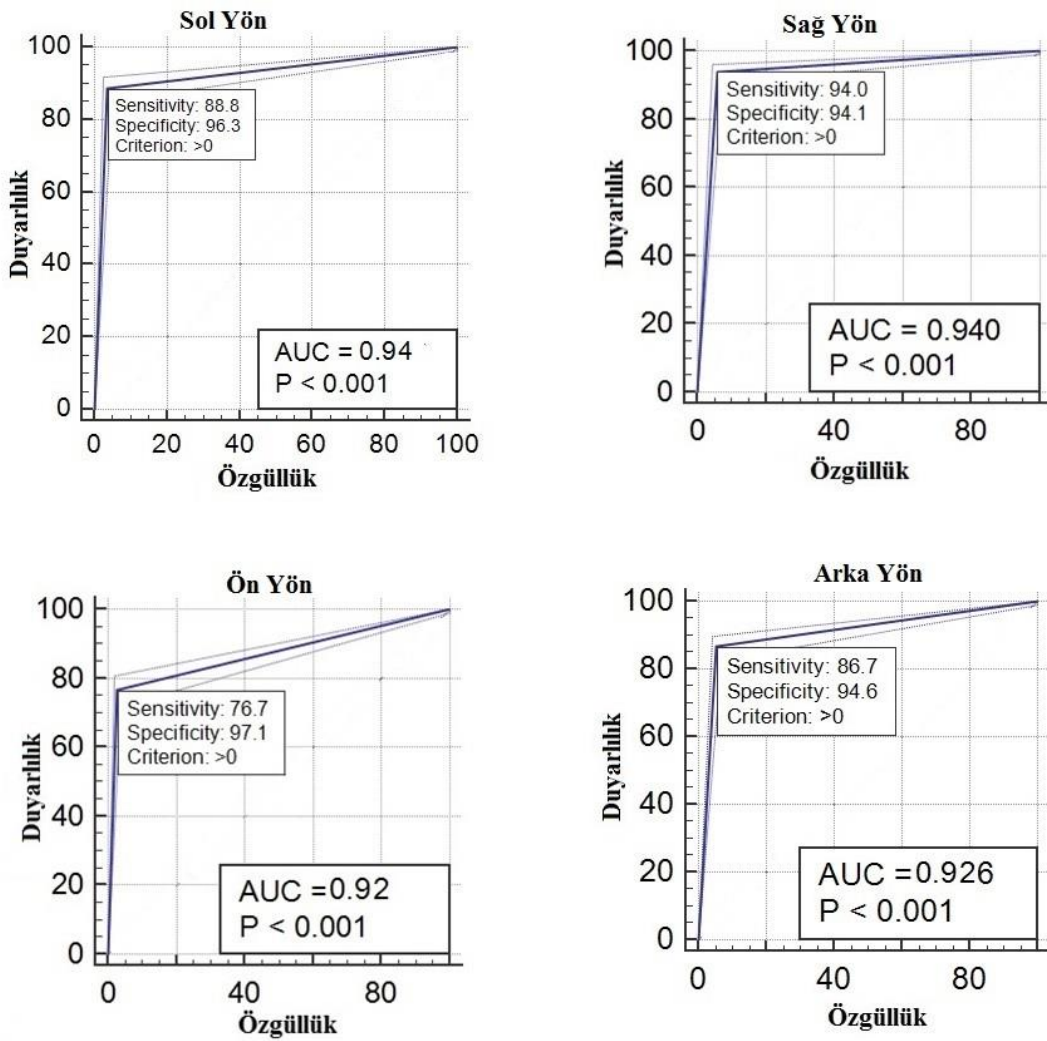
Şekil 3.12. Yön bulma çalışması bilgisayar ortamında elde edilen sonuçlar

Bilgisayar ortamında elde edilen sonuçların kappa katsayısı Tablo 3.9’da görüldüğü gibidir. Bu çalışmada tablodan görüldüğü gibi; sol, sağ ve arka yönler çok iyi düzeyde uyumlu, ön yön tespiti ise iyi düzeyde uyumlu çıkmıştır.

Tablo 3.9. Bilgisayar ortamında elde edilen kappa değerleri

Sol	Sağ	Ön	Arka
0.85	0.85	0.78	0.81

Bilgisayar ortamında elde edilen sonuçların ROC eğrileri ise Şekil 3.13’de görüldüğü gibidir. Eğrinin altındaki alan doğruluk değerini vermektedir.



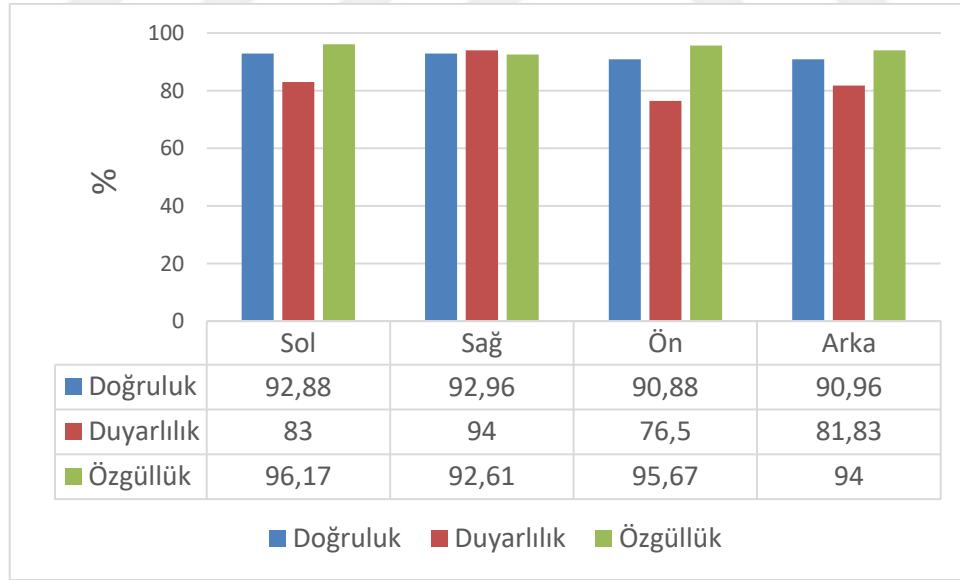
Şekil 3.13. Sol, sağ, ön ve arka yönler için ROC eğrisi

Tablo 3.10’da gerçek-zamanlı olarak elde edilen hata matrisi görülmektedir. Bu çalışmada her yön için 600 deneme yapılmıştır. Toplamda 2400 deneme yapılmıştır.

Tablo 3.10. Yön bulma çalışması gerçek zamanlı elde edilen hata matrisi

	Sol	Sağ	Ön	Arka
Sol	498	94	5	3
Sağ	27	564	4	5
Ön	25	16	459	100
Arka	17	23	69	491

Şekil 3.14’de gerçek-zamanlı olarak elde edilen sonuçların doğruluk, duyarlılık ve özgüllük değerleri gösterilmektedir. Sol ve sağ yönde başarı oranları yüksek çıkmıştır. Ön ve arka yönde ise sol ve sağ yöne göre düşüş görülmektedir. En düşük duyarlılık ön yönde elde edilmiştir.



Şekil 3.14. Yön bulma çalışması gerçek zamanlı elde edilen sonuçlar

Gerçek-zamanlı olarak elde edilen sonuçların kappa katsayısı Tablo 3.11’de görüldüğü gibidir. Bu çalışmada tablodan görüldüğü gibi; sol ve sağ yönler çok iyi düzeyde uyumlu, ön ve arka yönler ise iyi düzeyde uyumlu çıkmıştır.

Tablo 3.11. Yön bulma çalışması gerçek zamanlı elde edilen kappa değerleri

Sol	Sağ	Ön	Arka
0.81	0.82	0.75	0.76

Tablo 3.12’de görüldüğü gibi yapılan ölçümler oda ortamında, koridorda ve dış ortamda 1-4 metrelik mesafeler alınarak test edilmiştir. Tabloda görüldüğü gibi gürültülü ortama geçildikçe başarı oranı düşmektedir.

Tablo 3.12. Farklı ortamlarda elde edilen sonuçlar

Uygulanan ortam	Ölçüm	Yön Tahmini (1 m)	Yön Tahmini (4 m)	Ses Algılama
Oda	30 Db (A)	98%	97%	100%
Koridor	34 Db (A)	96%	93%	100%
Dış Mekan	40 Db (A)	93%	89%	99%

Tablo 3.13’de görüldüğü gibi 1 metre mesafede sesin yönünü algılaması %90 üzerinde çıkmıştır. En iyi başarıyı sağ ve sol yönünden gelen seslerde elde edilmiştir. Mesafe arttıkça başarı azalmıştır. Mesafe arttıkça sesin algılanmasında iki yön birden algılanmıştır, o bakımdan başarı azalmıştır. Başarının en az alındığı yön ise önden gelen sesler için alınmıştır. Bunun en temel sebebi, geliştirilen giyilebilir sistemin kişinin sırt kısmına yerleştirilmesindedir. Önden gelen sesler için soldan ya da sağdan gelen ses ile karışmaktadır.

Tablo 3.13. Mesafeye göre yön bulma başarı oranları

Mesafe	Sol (%)	Sağ (%)	Ön (%)	Arka (%)
1 m	96.21	97.13	92.04	91.38
2 m	95.10	96	90.20	90
3 m	93.30	93.50	90	89.50
4 m	90.38	91.04	84.13	85.54

1 metre mesafede elde edilen doğruluk, duyarlılık ve özgüllük değerleri Tablo 3.14'de görüldüğü gibidir. Tablodan görüldüğü gibi sağ yön %97, sağ yön ise %96, ön yön %92, arka yön %91 doğruluk ile tahmin edilmiştir.

Tablo 3.14. 1 metre mesafe test sonuçları

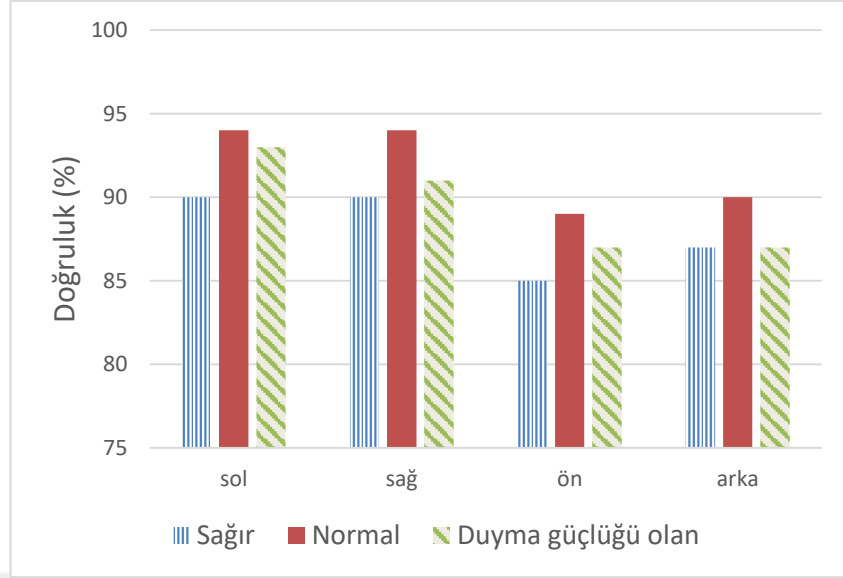
	Sol (%)	Sağ (%)	Ön (%)	Arka (%)
Doğruluk	96.21	97.13	92.04	91.38
Duyarlılık	92.83	99.17	82	79.5
Özgüllük	97.33	96.44	95.39	95.33

4 metre mesafede elde edilen doğruluk, duyarlılık ve özgüllük değerleri Tablo 3.15'de görüldüğü gibidir. Tablodan görüldüğü gibi sağ yön %91, sol yön %90, ön yön %84, arka yön %86 doğruluk ile tahmin edilmiştir.

Tablo 3.15. 4 metre mesafe test sonuçları

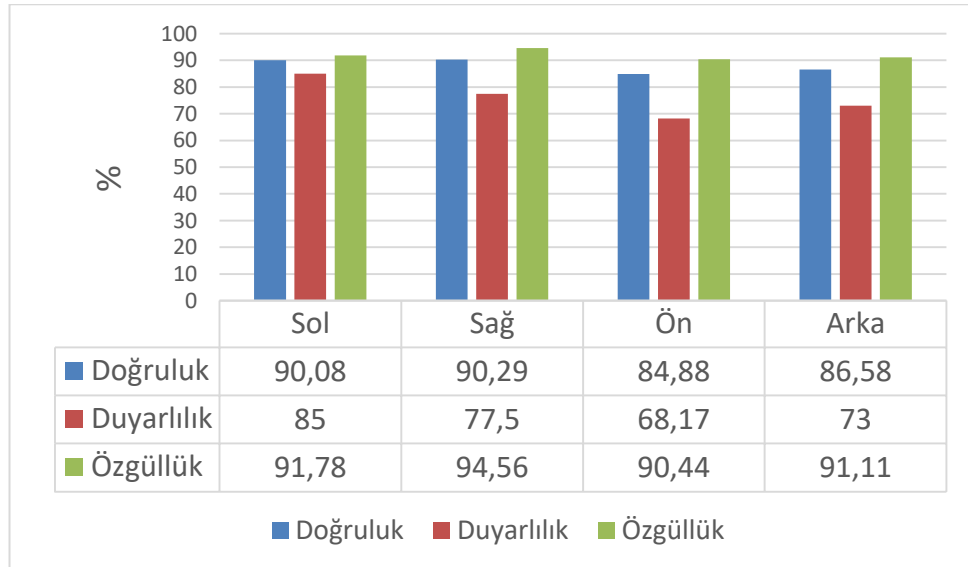
	Sol (%)	Sağ (%)	Ön (%)	Arka (%)
Doğruluk	90.38	91.04	84.13	85.54
Duyarlılık	85.83	80	68.17	68.17
Özgüllük	91.89	94.72	89.44	91.33

Şekil 3.15'de görüldüğü gibi sağır bireyler, normal bireyler ve duyma güçlüğü olan bireylerin karşılaştırılması gözükmemektedir. Gerçek zamanlı olarak gerçekleştirilen uygulama sonucunda sağır bireyin sağ ve sol yönü algılamasında %90 başarı elde edilmiştir.



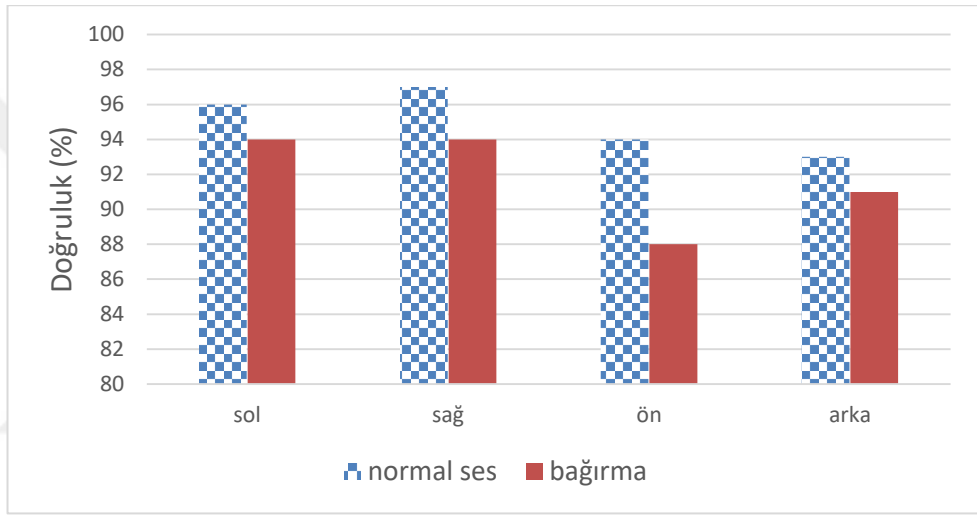
Şekil 3.15. Farklı bireylerin ses algılama başarı oranları

Sağır kişilere ait doğruluk, duyarlılık ve özgüllük değerleri Şekil 3.16'da görüldüğü gibidir. Şekilden görüldüğü gibi sağır bireylerde sol ve sağ yönlerde %90 başarı elde edilmiştir. Arka yönden gelen sesler %86.59, önden gelen sesler ise %84.88 başarı elde edilmiştir. Ön düşük duyarlılık önden gelen seslerde elde edilmiştir.



Şekil 3.16. Sağır birey yön bulma test sonuçları

Şekil 3.17’de görüldüğü gibi sistemimizde gerçek zamanlı olarak konuşan bireyin bağırarak mı yoksa normal sesle konuştuğu tespit edilmiştir. Bağıran kişinin ya da normal konuşan kişinin yönüne de tespit edilerek sınıflandırma başarısı gösterilmiştir. Tabloda görüldüğü gibi kişi yüksek sesle mi konuşuyor düşük sesle mi konuşuyor bunun algılanması sağlanmıştır. Kişi bağırarak konuşuyorsa bu tespit edilmiştir. Duyma engelli kullanıcı kişinin sesinin şiddeti algılayarak, bir sorun olduğunda daha hızlı cevap verebilmektedir. Sonuç olarak, sağ yönden gelen sesler için normal sesleri algılaması %97 iken bağırarak konuşan bireyi algılaması %94 olarak hesaplanmıştır.



Şekil 3.17. Normal ses ve bağıрма sesi algılama başarı oranları

Tablo 3.16’da görüldüğü gibi dört yön ile sekiz yön baz alınarak yapılan sonuçlar karşılaştırılmıştır. Sekiz yönde daha az başarı oranı çıkmasının sebebi, ara yönlerde nereden sesin geldiğinin belirlenememesidir. Sekiz yöne bakılarak yapıldığı için, iki yönün ortalarında konuşulduğunda iyi bir ayırt etme yapılamamaktadır. Sadece dört yöne bakıldığında ise daha iyi bir başarı oranı gözükmektedir.

Tablo 3.16. Dört ve sekiz yön başarı oranları

	Sol	Sağ	Ön	Arka
4 Yön	94%	94%	92%	92%
8 Yön	91%	90%	88%	85%

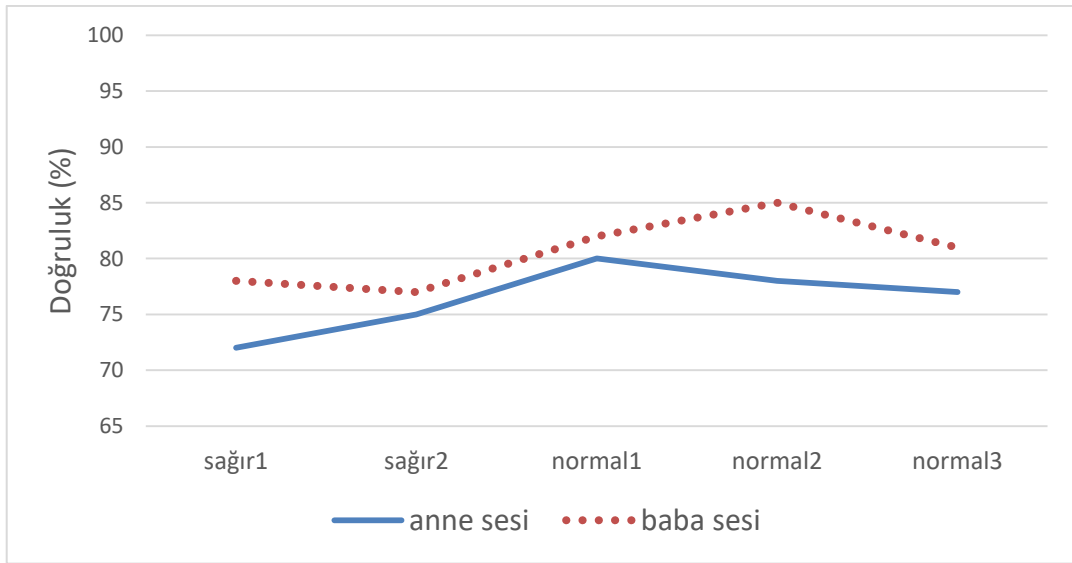
Tablo 3.17’de görüldüğü gibi geliştirdiğimiz cihaz TDOA algoritmasına göre daha hızlıdır. Aynı zamanda daha az mikrofon kullandığı için daha az maliyetlidir.

Tablo 3.17. Sistemimizin mikrofon sınıfı kullanan algoritma ile karşılaştırılması

	Hız	Maliyet	Doğruluk
TDOA	2.33 s	16 mikrofon	97%
Tez çalışması	0.68 s	4 mikrofon	98%

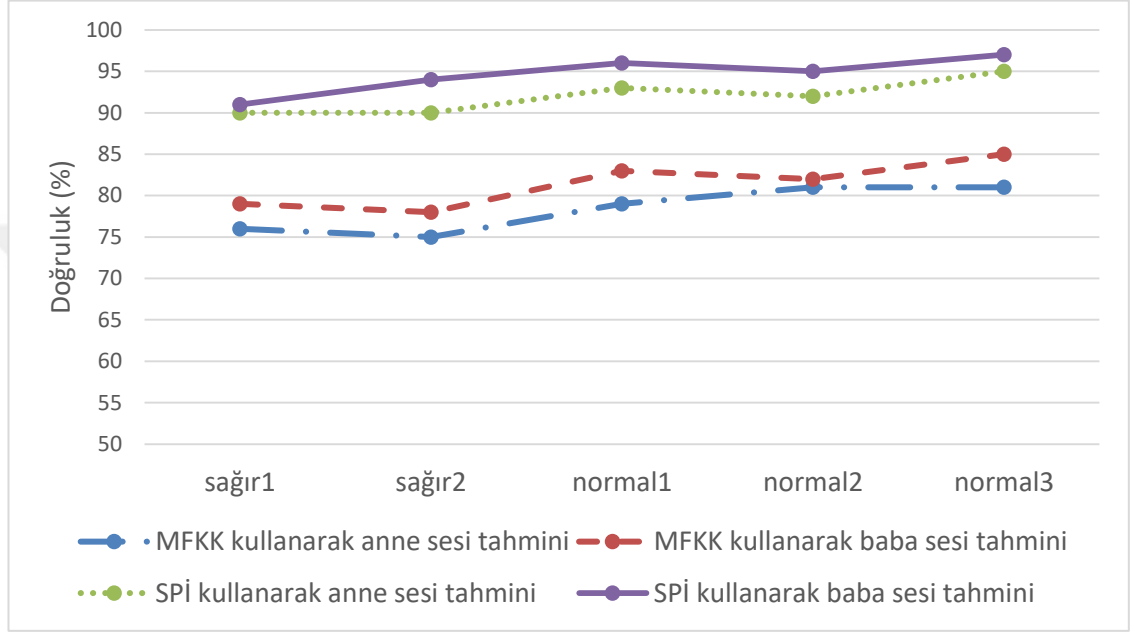
3.5. Ebeveyn Sesi Tahmini Çalışması

Bu çalışmada iki sağır, üç normal kişi giyilebilir sistemimizi test etmiştir ve öznitelik yöntemi olarak MFKK kullanılmıştır. Şekil 3.18’de ev ortamında sağır bireyler ve normal duyan kişilerin başarı oranları görülmektedir. Sağır bireylerin anne sesini tahmin etme oranı ortalama %74 iken normal bireylerin anne sesini tahmin etme oranı ortalama %79 çıkmaktadır. Sağır bireylerin baba sesini tahmin etme oranı ortalama %78 iken, normal bireylerin baba sesini tahmin etme oranı ortalama %83 çıkmaktadır. Genel ortalamaya bakıldığında ise ev ortamında anne sesi %76 oranında baba sesi ise %81 oranında tahmin edilmiştir.



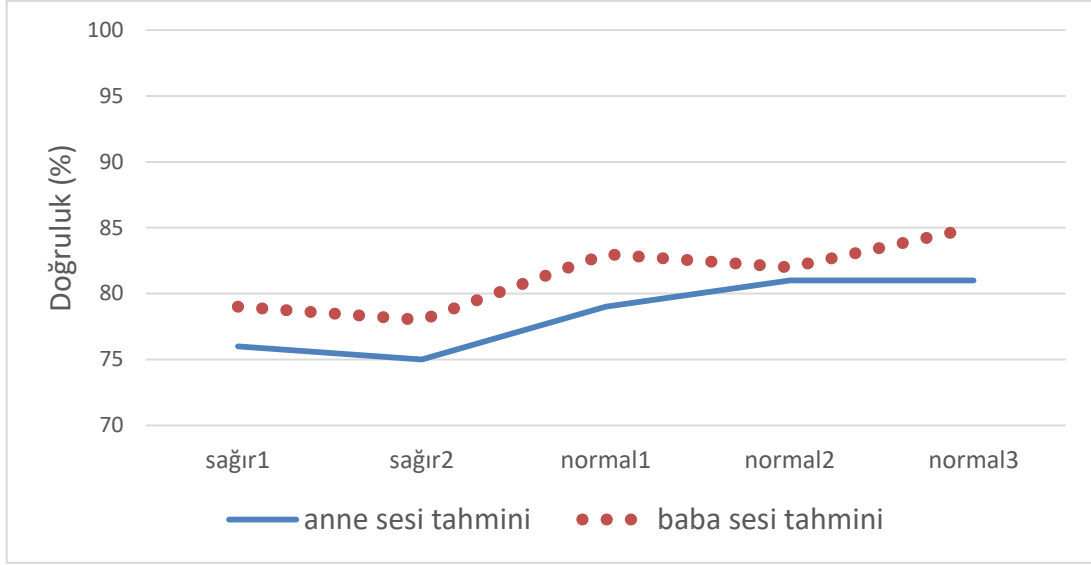
Şekil 3.18. Ev ortamında sağır ve normal kişilerin ebeveyn sesi tahmin oranları

Bu çalışmada aynı zamanda ev ortamında SPİ yöntemi kullanılarak da ebeveyn sesi tahmini yapılmıştır. Şekil 3.19'da görüldüğü gibi SPİ yöntemi ile daha yüksek doğruluk elde edilmiştir. SPİ yöntemi kullanarak sağır kişiler ortalama %88 anne sesini, ortalama %92 baba sesini tahmin etmişlerdir. Normal kişiler ise SPİ yöntemi ile ortalama %91 anne sesini, ortalama %94 baba sesini tahmin etmişlerdir.



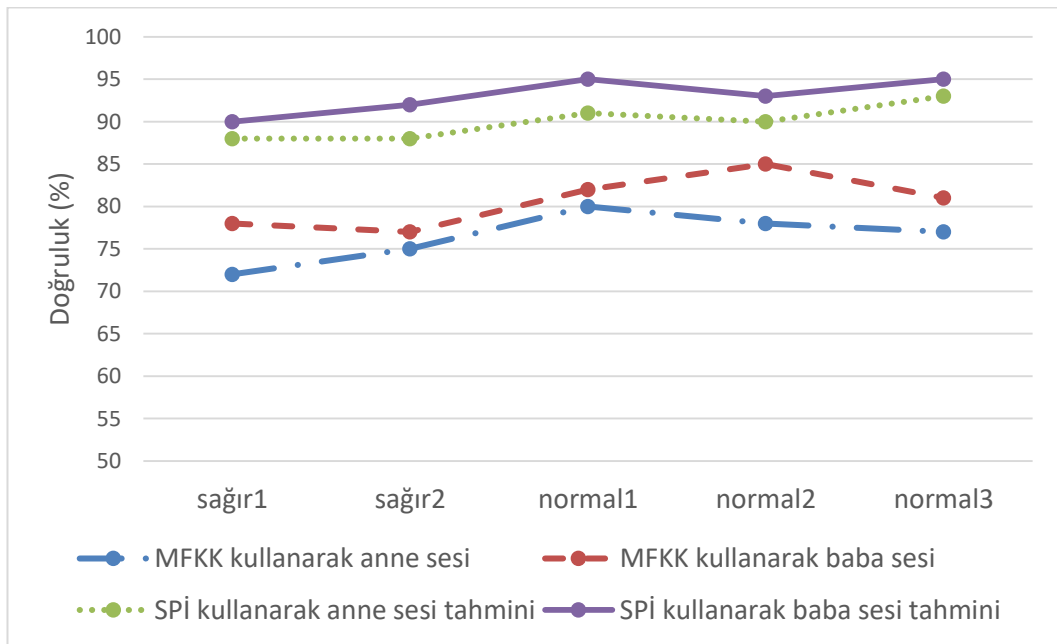
Şekil 3.19. Ev ortamında MFKK ve SPİ yöntemleri karşılaştırması

Şekil 3.20'de uyurken sağır bireyler ve normal duyan kişilerin başarı oranları gözükmemektedir. Sağır bireylerin anne sesini tahmin etme oranı ortalama %76 iken, normal bireylerin anne sesini tahmin etme oranı ortalama %80 çıkmaktadır. Sağır bireylerin baba sesini tahmin etme oranı ortalama %79 iken, normal bireylerin baba sesini tahmin etme oranı ortalama %83 çıkmaktadır Genel ortalamaya bakıldığında ise uyurken anne sesi %78 oranında baba sesi ise %83 oranında tahmin edilmiştir.



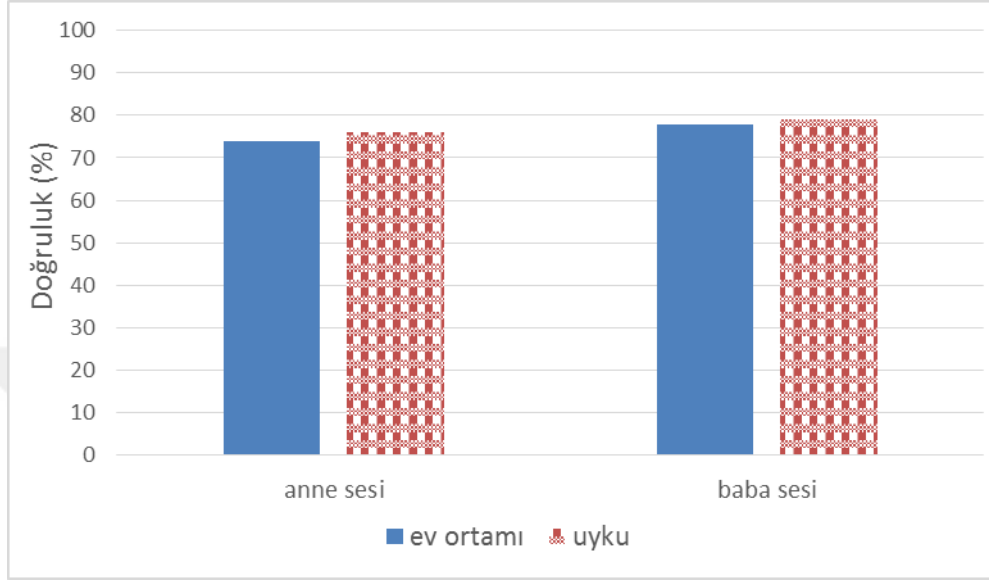
Şekil 3.20. Uyurken sağır ve normal kişilerin ebeveyn sesi tahmin oranları

Bu çalışmada aynı zamanda uyurken SPİ yöntemi kullanılarak da ebeveyn sesi tahmini yapılmıştır. Şekil 3.21’de görüldüğü gibi SPİ yöntemi ile daha yüksek doğruluk elde edilmiştir. SPİ yöntemi kullanarak sağır kişiler ortalama %90 anne sesini, ortalama %93 baba sesini tahmin etmişlerdir. Normal kişiler ise SPİ yöntemi ile ortalama %93 anne sesini, ortalama %96 baba sesini tahmin etmişlerdir.



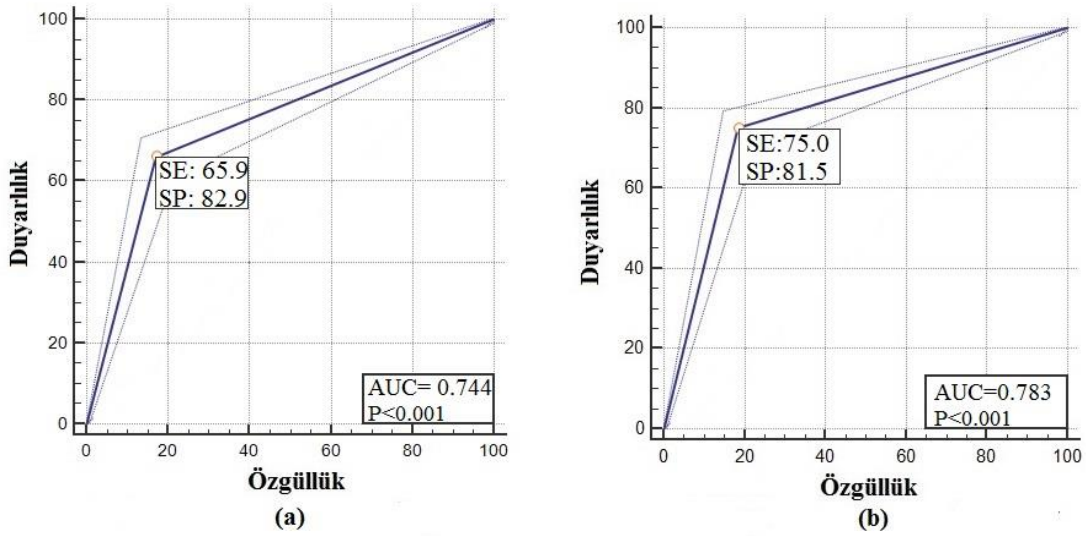
Şekil 3.21. Uyurken MFKK ve SPİ yöntemleri karşılaştırması

Sadece sağır kişilerin başarı oranları ise Şekil 3.22’de görüldüğü gibidir. Ev ortamında sağır kişiler sırasıyla %74 ve %76 ile anne sesini tahmin etmiştir. Uyurken ise sağır kişiler sırasıyla %78 ve %79 oranında baba sesini tahmin etmiştir.



Şekil 3.22. Sağır kişilerin ebeveyn sesi tahmin oranları

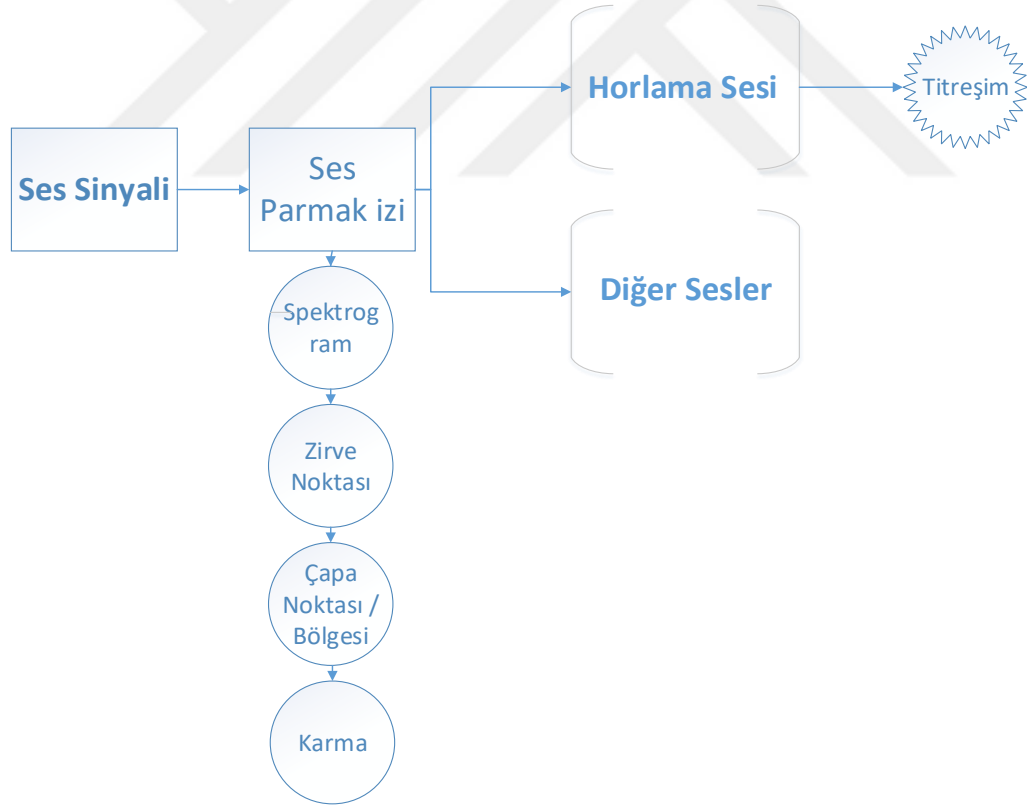
Sağır kişilerin ev ortamında elde edilen anne ve baba sesi tahmini sonuçlarının ROC eğrileri ise Şekil 3.23 (a) ve Şekil 3.23 (b) ’de görüldüğü gibidir.



Şekil 3.23. Sağır bireylerin ev ortamında (a) anne sesi tahmini ROC eğrisi, (b) baba sesi tahmini ROC eğrisi

3.6. Horlama Sesi Tahmini Çalışması

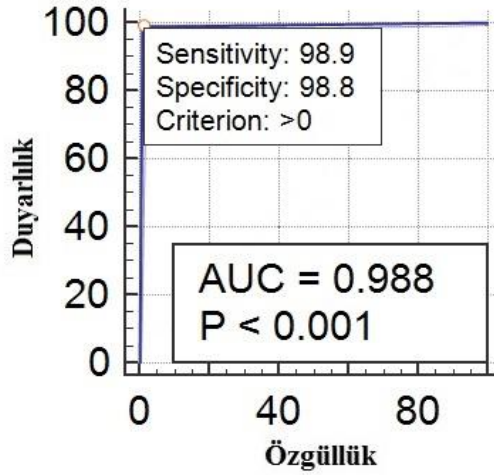
Bu çalışmada horlama sesi tespit için Ses Parmak izi yöntemi kullanılmıştır. SPI, ses içeriğinin bir parçasının algısal özetini çıkarır. Ses veri geldiğinde onun parmak izi hesaplanır ve veritabanında saklanan verilerle eşleştirilir. Parmak izi ve eşleşme algoritmalarını kullanarak, gelen ses içeriğinin horlama sesi olup olmadığı tespit edilebilir. Bu çalışmada kullanılan sistem Şekil 3.24’de görüldüğü gibidir. Gelen ses verisi eğer horlama sesi olarak tespit edilmişse kullanıcıya titreşim vererek horlamadan çıkması sağlanmıştır. Bu çalışmanın eğitim aşamasında 10 farklı horlama sesi sisteme tanıtılmıştır. Bu seslerden herhangi bir benzer ses gelirse sistem horlama sesi olarak algılamaktadır. Diğer sesler için herhangi bir titreşim verilmemektedir. Sistem her seferinde ses gelene kadar uykuda horlama sesi araması yapmakta ve horlayan kişiye titreşimler göndermektedir [173].



Şekil 3.24. Horlama sesi tahmin etme sistemi

Horlama sesi doğruluk değeri %98.9 bulunmuştur. Duyarlılık %98.9, özgüllük ise %98.8 olarak hesaplanmıştır. Horlama sesi kappa değeri 0.976 olarak hesaplanmıştır. Horlama sesinin bu değer sonucuna göre çok iyi derece uyumlu çıkmıştır.

Gerçek zamanlı elde edilen horlama seslerinden elde edilen sonuçlarının ROC eğrisi Şekil 3.25’de görüldüğü gibidir. Şekilden de görüldüğü gibi horlama sesleri yüksek oranda tespit edilmiştir.



Şekil 3.25. Horlama sesleri ROC eğrisi

Uyku esnasında horlama sesini tespit etmek için birçok yöntem geliştirilmiştir. Tablo 3.18’de görüldüğü gibi çalışmamız diğer çalışmalara göre daha iyi başarı elde etmiştir.

Horlama, orta yaş erkek popülasyonunun %44 ve kadın popülasyonunun’ da %28’ inde görülen önemli bir yakındır. Toplumda utanmaya sebep olan sosyal bir sorun olup, evlilik hayatını bile etkilemektedir. Çalışmamızda tasarladığımız horlama sesine sensitivitesi oldukça yüksek olan giyilebilir cihazının uygulaması sayesinde horlama atakları, horlamanın şiddeti ve horlama sayısında azalmalar meydana gelmiştir.

Tablo 3.18. Horlama sesi çalışmaları karşılaştırması

	Doğruluk	Metot	Yapılan deney	Deneme sayısı
Wang vd. [180]	94%	Örnek Entropi+ DVM	Horlama seslerini orijinal uyku seslerinden otomatik olarak sınıflandırmak	Eğitim: 2287 Test:1923
Dafna vd. [181]	98.2%	AdaBoost-tabanlı metod	Horlama sesi otomatik analizi	76600 akustik ses bölümü
Yadollahi ve Moussavi [182]	93.2%	Fisher Lineer Ayırıcı+ Bayes	Otomatik horlama sınıflandırma	3040 horlama
Karunajeewa vd. [183]	90.7%	SGO + Sinyalin enerjisi + gürültü giderme teknikleri	Horlama ile ilişkili seslerden sessizlik- nefes alma-horlama sınıflandırması.	4410 deneme
Cavusoglu vd. [184]	86.8%	Enerji , SGO	Uyku seslerinden horlama sınıflandırması	Eğitim:300 horlama bölümü, Test:150 horlama bölümü
Duckitt vd. [185]	89%	GMM + spectral tabanlı özellikler	Ortam akustik verilerinden horlamanın otomatik algılanması, bölütlenmesi ve değerlendirilmesi	Eğitim:4006 Test:1554
Tez çalışması	98.9%	SPİ	Horlama sesi tahmini	5000 deneme

4. SONUÇLAR

Bu çalışmada duyma engelliler ve yaşlılar için büyük problem olan duyma güçlükleri geliştirdiğimiz giyilebilir titreşim tabanlı cihaz ile giderilmeye çalışılmıştır. Telefon sesi, alarm sesi, kapı zil sesi, korna sesi, fren sesi ve köpek havlaması sesi gibi önemli seslerin duyma engellinin algılamasını sağlamak duyma engelliler için önem arz etmektedir. Bu tarz sesleri algıladığı zaman kişi daha hızlı tepki verebilir ve panik durumu önlenmiş olabilir. Bunun için çıkışa yerleştirilen titreşim motorları sayesinde, her ses için farklı bir titreşim şiddeti verilerek duyma engelli kişinin iç ve dış ortamdaki önemli sesleri anlaması sağlanmıştır.

Bu çalışmada konuşan kişinin yönü tespit edilmektedir. Örneğin duyma engelli bireyin arkasından konuşan bir kişiyi fark etmesi sağlanmıştır. Geliştirdiğimiz sistemde dört adet mikrofon bulunmaktadır. Konuşan kişinin sesi bu dört mikrofon tarafından alındıktan sonra, hangi mikrofondan daha belirgin olduğu hesaplandıktan sonra ona ait LED çıkışı yanmakta ve titreşim yoluyla duyma engellinin algılaması sağlanmıştır. Sağırarda en önemli sorunlardan biri, sesin nereden geldiğini anlayamamalarıdır. Bu çalışma, işitme engelli kişilerin sesin gerçek zamanlı olarak nereden geldiğini anlamalarına yardımcı olmuştur. Sağırarın uyarıcı seslerin yönünü bulabilmeleri onlar için çok yararlı olacaktır. Bazen gelen bir arabanın ve köpeğin havlayarak gelmesini duymamaları çok tehlikeli olabilir. Bu çalışmada, gerçek zamanlı testlerde sağır bireyler, giyilebilir cihazımız sayesinde sesin yönünü belirleyebilmektedir. Özellikle duyma engelliler için arkasından gelen seslerden tedirgin olmaktadır. Giyilebilir sistemimiz sayesinde duyma engelli birey daha refah bir yaşam sürmesi hedeflenmiştir. Bu çalışmada öncelikle en iyi öznelik yöntemleri ve sınıflandırma yöntemi seçilmiştir. Daha sonra ise prototip cihaz üzerinde bu seçilen yöntemler kullanılarak test edilmiştir. Aynı zamanda bilgisayar ortamında da testler yapılmış ve karşılaştırmalı olarak gösterilmiştir. Geliştirdiğimiz prototip giyilebilir cihaz ile elde edilen veriler kaydedilmiştir. Bu veriler bilgisayar ortamında da test edilerek gerçek zamanlı ile bilgisayar ortamı için sonuçlar karşılaştırılmıştır.

Bir kişi panik durumu varken duyma engelliye bunu anlatırken ses tonu yükselebilir. Duyma engelli geliştirmeyi hedeflediğimiz sistem sayesinde kişinin bağırp bağırmadığını anlayabilecektir ve önemli bir durum varsa bunu algılaması hızlanmaktadır. Örneğin, duyma

engelli bir insanın zil sesini, arkasından yaklaşan ayak seslerini, ona seslenen birisini duymasını, konuşmaları anlamasını ve daha az stresle televizyon izlemesini sağlayacaktır.

Duyma engelli kişiler ev ortamında uyurken ya da evde tek kaldığı zaman önemli sesleri duymadığı için rahat bir yaşam sürememektedir. Bu çalışmada duyma engelliler için anne ve baba sesi tahmin sistemi önerilmiştir. Ebeveynlerden hangisinin konuştuğunu duyma engelli kişilere titreşim ile bilgi verilmektedir. Bu sayede duyma engelli kişi kendisine kimin seslendiğini ya da kimin konuştuğunu gerçek zamanlı olarak algılamaktadır. Geliştirdiğimiz giyilebilir cihaz gerçek zamanlı olarak uyurken ve ev ortamındayken çok rahatlıkla ebeveyn sesini algılamakta ve duyma engelli kişiye aktarmaktadır.

İşitme güçlüğü çeken kişilerin evde, işyerinde veya dışarda seslerin nereden geldiğini ve hangi ses olduğunu anlayarak uygun bir yaşam sürmelerini hedefliyoruz. Bu tez çalışmasında, işitme engelli bireylerin sesin yönü ve sesin hangi ses olduğunu algılamaları için bir titreşim tabanlı giyilebilir sistem önerilmiştir.

Bu şekilde geliştirilen komple sistem duyma engellilerin daha refah bir şekilde yaşamalarını sağlayıp, uygulamalarda özgünlük sağlamaktadır. Bu çalışmanın yaygın etkisi, duyma engelliler için toplumsal bir sorun olan durumlara kolaylık getirmesidir.

Bu çalışma, işitme engelli insanlar için, giyilebilir titreşen tabanlı yeni bir fikir sunmaktadır. Burada bulunan bilgilerin, giyilebilir hesaplama, giyilebilir bilgisayarlar ve insan-bilgisayar etkileşimi alanlarında sistem geliştirme için çalışan ve işitme sorunu olan insanlar için faydalı olacağına inanıyoruz. Sonuç olarak, bu tez çalışmasında sesin yönü büyük ölçüde algılanmıştır. Aynı zamanda, konuşmacının bağırp bağırmadığı da belirlenmiştir. Bu tez çalışmasında önemli seslerin tespiti ve ebeveyn sesi tespiti de yapılarak geliştirilen gerçek zamanlı giyilebilir cihaz sayesinde duyma engelliliklerin yaşam standardını artırması düşünülmektedir.

5. ÖNERİLER VE TARTIŞMA

Sağır insanlar günlük olarak sesleri algılayamamaları veya tanımlayamamaları, bu kişiler için büyük problemdir. Bu problemin otomatik olarak çözümünü sağlayan teknikleri ve algoritmaları incelemek sağır insanlara büyük yararlar sağlayacak önemli teknolojik ilerleme olarak görülmektedir. Sağır kişilerin yaşam kalitesini iyileştirmek için bir evde ortaya çıkabilir çeşitli sesleri algılayabilen ve sınıflandırabilen bir uygulama geliştirmek yardıma yönelik temel bir gereklilik olarak kabul edilmektedir. Bu çalışma, yüksek doğruluk ve güvenilirlik ile önemli seslerin sınıflandırılmasına ilişkin bir dizi parametre elde etmek için yürütülen çalışmanın sonuçlarını sunmaktadır.

Yapılan yön bulma çalışmasında, eğitim ve test verilerinin elde edilmesi ve en iyi öznitelik çıkarma yönteminin seçimi zaman almaktadır. Özellikle öznitelik bulunurken diğer yöntemler kullanılarak en ayırt edici yöntem bulunması gerekmektedir. Birden çok öznitelik beraber kullanılarak sınıflandırma başarısı artırılabilir.

Bu çalışma, bir kişinin sağır bir bireyin yanında yüksek sesle konuştuğunu belirlemektedir. Kişiler panik halinde sağır bireyle konuşurken ses tonunu artırabilir. Sağır kişilerin panik halini daha erken fark edebilmesi için bu durum hedeflenmiştir. Sağır kişilerin birilerinin bağırp bağırmadığını algılayacak ve önemli bir durum varsa algılama gecikmesi, geliştirdiğimiz sistem sayesinde azalacaktır. Örneğin, sağır birey uyurken annesinin acilen bağırarak onu çağırmasını bu sayede algılayabilmektedir. Bu nedenle, bu daha az stresli, daha rahat ve güvenli bir şekilde yaşam sürmesine yardımcı olabilir. Cihazımızı kullanan sağır bireylerden gelen geribildirimde, cihazın çok yararlı olduğunu gördüklerini vurguladılar. Görmedikleri birinden gelen önemli bir sesi olup olmadığını anlamaları daha güvenilir hissettirmektedir. Aynı zamanda, ebeveynlerinin seslerini özellikle evde uyurken gerçek zamanlı olarak seslendirdiklerini anladıkları için kendilerini daha rahat hissetmişlerdir.

Bu çalışmada, iç ve dış ortamdaki seslerin algılanışını analiz edilmiştir. Sonuçlar, geliştirdiğimiz yöntemin, gerçek zamanlı sonuçlar üretip yüksek başarı oranına ulaşması nedeniyle geleneksel yöntemlere göre avantajlara sahip olduğunu göstermektedir. İşitme engelli kişinin arkasına, Raspberry Pi, mikrofon sensörü ve titreşim motorundan oluşan giyilebilir cihazımızı yerleştirilmiştir. Bu cihaz düşük bir üretim maliyetinden ve yüksek başarı oranından dolayı günlük yaşamda kolaylıkla kullanılmasına olanak tanır.

Çalışmamızın en önemli özelliklerinden biri, geliştirdiğimiz cihazın gerçek zamanlı çalışması ve uzun süre giyilebilmesidir. Cihazımız, sesi aldıktan sonra kısa süre içinde yanıt verir. Sistemimizin dezavantajı, bir ses gelmesi ve aynı anda başka bir ses çıkmasıdır. Bu durumda, önce gelen ses algılanır. Gelecekteki çalışmalarında, bu tür vakaları önlemek için gelen sesleri kaydetmek suretiyle kullanıcıya daha önemli ses iletmesi düşünülmektedir. Özellikle korna, fren ve köpek sesi, engelliler için büyük önem taşır. Bu açıdan, bu seslerin hızlı bir şekilde yüksek başarı ile belirlenmesi düşünülmektedir.

Deneysel çalışmalarımızda, gürültülü ortamlarda başarı oranları düşmüştür. Örneğin, korna sesi SPİ yöntemi kullanarak bilgisayar ortamında doğruluk oranı %97 olarak belirlenirken, bu oran gerçek zamanlı bir ortamda %84'e düşmektedir. Bu düşüşün temel nedeni, sesin uzaklığı ve gürültülü ortamda olunmasıdır.

Peltonen ve diğerleri [186] özellik olarak MFKK ve sınıflandırıcı olarak GMM ve sinir ağını kullanmışlardır. 17 doğal sesi sınıflandırırken, MFKK'yi özellikler ve GMM'yi de sınıflandırıcı olarak kullanırken %68.4'lük bir ortalama tanımlama oranı belirtmişlerdir. Chu ve diğerleri [178] çalışmalarında 14 doğal seslik bir diziyi sınıflandırabilmek için eşleştirme takip algoritmasını kullanırken çıkarılan bir dizi özellik ile MFKK'nin birlikte kullanılmasını ortaya koymuşlardır. GMM tercihen bir sınıflandırıcı olarak kullanıldığında da yaklaşık %83.9'luk bir doğruluk belirtmişlerdir. Adiloğlu ve diğerleri [187] sesler arasındaki mesafeyi hesaplayabilmek için bir benzeşme fonksiyonunu geliştirmiş ve sınıflandırma amacı için DVM kullanmışlardır. Bu çalışmada ise SPİ yöntemi kullanılarak ortalama %94 lük başarı elde edilmiştir.

Ke ve diğerleri [101] çalışmasında ses tanınması için bilgisayarla görü teknikleri kullanılmıştır. Ses sinyali, zaman-frekans olarak dönüştürülmüş ve Bark frekansı kepsral katsayıları hesaplanmıştır. Bu çalışmada ise zaman-frekans olarak dönüştürüldükten sonra doruk noktaları bulunmuş ve parmak izi oluşturulmuştur.

Baluja ve Covell [74] çalışmalarında spektrogram hesaplanıp Haar dalgacık uygulanmıştır. Ellis ve diğerleri [78] çalışmasında spektrogram 8 banda bölünmüş, her bant için ayrı ayrı nota başlangıçları ve parmak izleri hesaplanmıştır. Bu çalışmada ise spektrogramlar oluşturulduktan sonra, doruk noktaları, parmak izi karmaları, takımyıldızı haritasındaki zirve noktalarının ikili kombinasyonlar şeklinde eşleştirilmesiyle elde edilmektedirler.

Horlama sesi bulma çalışmalarında birçok yöntem kullanılmıştır [180-185]. Bu çalışmalarda genelde ses için öznel bulurken sesin ayırt edici özelliği bulunmaya çalışılmış ve sınıflandırıcı ile sesin horlama sesi olduğu tespit edilmiştir. Bu çalışmada ise horlama sesleri parmak izi alınarak diğer seslerden ayırt edici özelliği tespit edilmiş, aynı zamanda gürültülü durumlarda bile iyi bir sınıflandırma başarısı göstermesi sağlanmıştır.

Bu tez çalışmasında her ses için farklı bir titreşim verilmiştir. Buradaki farklı titreşimler özellikle sağır bireylerde bazen karışmaktadır. Bu bakımdan bu sistemdeki titreşim olayları ilerleyen çalışmalarda biraz daha iyileştirecek ve daha kullanıcı dostu bir sisteme dönüştürülecektir. Sağır kişilerde dokunma duyusu aktif olduğu için titreşimler verilmiştir. Sağır bireyin sadece parmağına sarılı madeni para büyüklüğünde bir titreşim motoru bulunmaktadır. Giyilebilir cihazımız kişilerin sırt bölgesinde ve kişiyi rahatsız etmeyecek seviyededir. Cihazımız uzun süre kullanıma da uygundur.

Sistemimiz öncelikle duyduğu sese göre işlem yapmaktadır. Örneğin, köpek ve insan sesleri için, köpek sesi ilk önce gelirse bunu hissedebilir. Bir ses karışımı bir araya geldiğinde, sesin spektrogramını dikkate alarak önemli bir ses karışımı olup olmadığına karar verilir. Örneğin, insan sesi ve köpek havlaması aynı anda geldiğinde, spektrogram (daha yüksek veya daha düşük) frekans büyüklük seviyeleri ile belirlenebilir. Bu durumda, sağır kişinin, onları hissetmesine yardımcı olabilecek titreşimler vererek bu sesler algılanabilir.

Geliştirdiğimiz giyilebilir cihaz, hızlı çıktı sağlayan kullanımı kolay ve giyilebilir bir cihazdır. Sağır insanlar için tasarlanmış cihazımız anlaşılır ve sağır insanlar için önemli sesler için hızlı bir karar verme mekanizmasına sahiptir. Mesela, sağır bir kimse uyurken ve evde yalnız kalınca kapı zili ve telefon sesini duymaz. Geliştirdiğimiz cihaz sayesinde, sağır bir kişi bu sesleri çok rahat hissedecek ve normal insanlar gibi bir yaşam standardına sahip olacaktır. Ayrıca bu sistemin geliştirilmesi durumunda, sağır olan kişiler için bu çalışmanın faydalı olacağı düşünülmektedir. Bu çalışma, sağır bireyler için yaşam kalitesi artırmasından dolayı ana hedeflerimizden birini karşılamaktadır.

Duyuma engelliler için sesi anlamak önemli bir sorundur. Geliştireceğimiz ürün sayesinde duyma engelliler toplum içinde daha rahat ve güvenilir bir hayat yaşamaları sağlanacaktır. Örneğin duyma engelli kişi yolda yürürken korna çalarsa bunu duyması sağlanacak ve bu sayede daha güvenli bir yaşam standardına sahip olacaktır. Bu çalışmanın en temel amacı ise titreşimler sayesinde ses tanıma sistemi geliştirmektir.

Bu çalışmanın bir diğer amacı duyma engelli bireylerin sokakta yürürken arkasından gelen sesi, fren sesi, korna sesi, köpek havlaması vb. gibi sesleri fark etmesini sağlamaktır. Özellikle arkasından gelen sesler duyma engelli bireyleri tedirgin etmektedir. Bunun yanında korna ve fren sesi gibi önem arz eden seslerin nereden geldiğini de anlayarak duyma engelli bireyin daha sağlıklı bir seyahat etmesi hedeflenmektedir. Duyma engelli bireylerin gerçek zamanlı olarak olayları hemen algılamaları ve ani karar vermeleri sağlanacaktır. Duyma engellilerin günlük hayatta kullanabilecekleri bir ürün yapılması hedeflenmektedir.

Konuşma tanıma uygulamalarında en önemli fayda duyma engellilerin anlayabileceği bir sistem tasarlayabilmektir. Bu tür uygulamalarda, duyma engelli bir kişinin yaşantısını normalden daha iyi hale getirmek amaçlanmıştır. Bu çalışmada ise duyma engelli kişilerin anne ve baba sesini algılayabileceği bir sistem tasarlanmıştır. Duyma engelli bir kişi kendisine seslenen anne babasını duymamakta ve yaşam standardı düşmektedir. Özellikle acil durumlarda örneğin uyurken anne veya baba kendisine seslendiğinde anlayabilecektir. Kişilerin sırt bölgesine monte edilen cihazımız titreşimler ile duyma engelli bireye bilgi vermektedir. Bu sayede duyma engelli birey çok rahatlıkla ebeveyn sesini tahmin edebilmiştir. Uyurken ise yanına konulan cihazımız ebeveyn sesini duyduğu anda kişiyi uyarmaktadır.

Cihazın duyma engelli kişilere anlatılması ve eğitim aşaması zaman almaktadır. Kullanacağı cihazı anlatırken ve kişinin cihazı ilk kullanımlarında anlaşılabilirlik düşmüştür. Ancak cihazı kullandıktan sonra daha iyi algılayabilmiştir. Duyma engelli birey sayısı az olduğu için ve eğitilmesi zaman alması ve zor olmasından dolayı normal bireylerden daha az çıkmıştır. Kişi sayısı ve daha iyi bir eğitim verilerek bu oran artırılabilir. Ev ortamında gürültüden dolayı başarı oranları daha düşük çıkmaktadır. Özellikle televizyon sesi, telefon sesi gibi seslerden dolayı başarı oranları daha düşük çıkmaktadır. Ancak uyurken fazla gürültü olmadığı için başarı oranları artmıştır. Bu çalışmada duyma engelli bireylerin yaşam standartlarını artırmak için giyilebilir cihaz tasarlanmıştır. Duyma engelli bireyler ebeveyn sesinden kimin konuştuğunu algılayabilmiştir. Özellikle acil durumlarda sağır bireye seslendiğinde veya uyurken kendisine seslendiğini algılayabilecektir.

Bu tez çalışmasında uzun süre giyilebilirliğe uygun, kullanıcının rahatlıkla anlayabileceği bir cihaz tasarlanmıştır. Giyilebilir cihazımız duyma engellinin sırt bölgesine dikilmiştir ve sadece vücudunda titreşim motoru takılıdır. Titreşim motoru, DC motorlu olup bir madeni para büyüklüğündedir. Önceki çalışmalardan en önemli farkımız giyilebilir,

ucuz, hızlı ve yüksek başarılı bir cihaz geliştirmiş olmamızdır. Sağır insanlar için tasarlanmış cihazımız anlaşılır ve sağır insanlar için önemli olan sesler için hızlı bir karar mekanizmasına sahiptir.

Gelecekteki çalışmalarda derin öğrenme kullanılacak ve sistemin başarısı artırılmaya çalışılacaktır. Ayrıca, sınıflandırma aşamasındaki diğer yöntemler en iyi sonucu elde etmek için kullanılacaktır. Daha ileri çalışmalar için, mikrofonlar için en uygun mesafeler hesaplanacak ve ses tanıma en iyi sınıflandırıcı ile yapılacaktır. Aynı zamanda, ilerleyen çalışmalarda gerçek zamanlı görselleştirme yapılması düşünülmektedir. Seslerin çıktığı yönü ileten bir giyilebilir cihaz, sağır bir kişinin giyebileceği gözlük ile görselleştirilmesi düşünülmektedir.

6. KAYNAKLAR

1. Chen, J., Kam, A.H., Zhang, J., Liu, N., ve Shue, L., Bathroom activity monitoring based on sound, in International Conference on Pervasive Computing, Mayıs 2005, Münih, Springer,47-61.
2. Wang, J.C., Lee, H.P., Wang, J.F., ve Lin, C.B., Robust environmental sound recognition for home automation, IEEE Transactions on Automation Science and Engineering, 5,1 (2008) 25-31.
3. Fleury, A., Noury, N., Vacher, M., Glasson, H., ve Serignat, J.F., Sound and speech detection and classification in a health smart home, in 30th Annual International Conference of the IEEE, Engineering in Medicine and Biology Society, Ağustos 2008, Vancouver, IEEE, 4644-4647.
4. Shaikh, M.A.M., Rebordao, A.R.F., Nakasone, A., Helmut, P., ve Hirose, K., An automatic approach to virtual living based on environmental sound cues, in 3rd International Conference on Affective Computing and Intelligent Interaction and Workshops, Eylül 2009, Amsterdam, IEEE, 1-6.
5. Vacher, M., Portet, F., Fleury, A., ve Noury, N., Development of audio sensing technology for ambient assisted living: Applications and challenges, Digital Advances in Medicine, E-Health, and Communication Technologies, (2013) 148.
6. Vacher, M., Istrate, D., Portet, F., Joubert, T., Chevalier, T., Smidtas, S., Meillon, B., Lecouteux, B., Sehili, M., ve Chahuara, P., The sweet-home project: Audio technology in smart homes to improve well-being and reliance, in Annual International Conference of the IEEE Engineering in Medicine and Biology Society, Ağustos 2011, Boston, IEEE, 5291-5294.
7. Wang, J.-C., Lin, C.-H., Chen, B.-W., ve Tsai, M.-K., Gabor-based nonuniform scale-frequency map for environmental sound classification in home automation, IEEE Transactions on Automation Science and Engineering, 11,2 (2014) 607-613.
8. Bukhari, J., Rehman, M., Malik, S.I., Kamboh, A.M., ve Salman, A., American sign language translation through sensory glove; signspeak, Int. J. u-and e-Service, Science and Technology, 8,1 (2015) 131-142.
9. Adams, J.W. ve Rohring, P.S., Handbook to service the deaf and hard of hearing: a bridge to accessibility, Emerald Group Pub Ltd., 2004.
10. Scheetz, N.A., Psychosocial aspects of deafness, Allyn & Bacon, 2004.
11. Zhang, M., Luo, M., Nie, R., ve Zhang, Y., Technical attributes, health attribute, consumer attributes and their roles in adoption intention of healthcare wearable technology, International journal of medical informatics, 108, (2017) 97-109.

12. Burges, C.J., Platt, J.C., ve Jana, S., Distortion discriminant analysis for audio fingerprinting, IEEE Transactions on Speech and Audio Processing, 11,3 (2003) 165-174.
13. Haitsma, J. ve Kalker, T., A highly robust audio fingerprinting system, in 3rd International Conference on Music Information Retrieval, Ekim 2002, Paris, Bildiriler Kitabı, 107-115.
14. Herre, J., Hellmuth, O., ve Cremer, M., Scalable robust audio fingerprinting using MPEG-7 content description, in IEEE Workshop on Multimedia Signal Processing, Aralık 2002, USA, IEEE, 165-168.
15. Mapelli, F., Pezzano, R., ve Lancini, R., Robust audio fingerprinting for song identification, in 12th European Signal Processing Conference, Eylül 2004, Vienna, IEEE, 2095-2098.
16. Neuschmied, H., Mayer, H., ve Batlle, E., Content-based identification of audio titles on the internet, in First International Conference on Web Delivering of Music, Kasım 2001, Florence, IEEE, 96-100.
17. Sukittanon, S., Atlas, L.E., ve Pitton, J.W., Modulation-scale analysis for content identification, IEEE Transactions on Signal Processing, 52,10 (2004) 3023-3035.
18. Venkatachalam, V., Cazzanti, L., Dhillon, N., ve Wells, M., Automatic identification of sound recordings, IEEE Signal Processing Magazine, 21,2 (2004) 92-99.
19. Wang, A., An Industrial Strength Audio Search Algorithm, in 4th International Conference on Music Information Retrieval (ISMIR), Ekim 2003, Washington, Bildiriler Kitabı, 7-13.
20. Cano, P., Batle, E., Kalker, T., ve Haitsma, J., A review of algorithms for audio fingerprinting, in Multimedia Signal Processing, 2002 IEEE Workshop on, Aralık 2002, St. Thomas, IEEE, 169-173.
21. Wang, A., The Shazam music recognition service, Communications of the ACM, 49,8 (2006) 44-48.
22. Mont-Reynaud, B., Master, A., Stonehocker, T.P., ve Mohajer, K., System and methods for continuous audio matching, Google Patents, 2016.
23. Son, W., Cho, H.-T., Yoon, K., ve Lee, S.-P., Sub-fingerprint masking for a robust audio fingerprinting system in a real-noise environment for portable consumer devices, IEEE Transactions on Consumer Electronics, 56,1 (2010).
24. Casey, M.A., Veltkamp, R., Goto, M., Leman, M., Rhodes, C., ve Slaney, M., Content-based music information retrieval: Current directions and future challenges, Proceedings of the IEEE, 96,4 (2008) 668-696.

25. Unal, E., Chew, E., Georgiou, P.G., ve Narayanan, S.S., Challenging uncertainty in query by humming systems: a fingerprinting approach, IEEE Transactions on Audio, Speech, and Language Processing, 16,2 (2008) 359-371.
26. Shirali-Shahreza, S., Abolhassani, H., ve Shirali-Shahreza, M.H., Fast and scalable system for automatic artist identification, IEEE Transactions on Consumer Electronics, 55,3 (2009).
27. Lim, S.C., Lee, J.S., Jang, S.J., Lee, S.P., ve Kim, M.Y., Music-genre classification system based on spectro-temporal features and feature selection, IEEE Transactions on Consumer Electronics, 58,4 (2012).
28. Blume, H., Bischl, B., Botteck, M., Igel, C., Martin, R., Roetter, G., Rudolph, G., Theimer, W., Vatolkin, I., ve Weihs, C., Huge music archives on mobile devices, IEEE Signal Processing Magazine, 28,4 (2011) 24-39.
29. Lu, L., Jiang, H., ve Zhang, H., A robust audio classification and segmentation method, in Proceedings of the ninth ACM international conference on Multimedia, Eylül 2001, Ottawa, ACM, 203-211.
30. Foote, J.T., Content-based retrieval of music and audio, in Multimedia Storage and Archiving Systems II, Ekim 1997, Dallas, Bildiriler Kitabı, 138-148.
31. Haitisma, J., Kalker, T., ve Oostveen, J., Robust audio hashing for content identification, in International Workshop on Content-Based Multimedia Indexing, Eylül 2001, Brescia, Bildiriler Kitabı, 117-124.
32. Sukittanon, S. ve Atlas, L.E., Modulation frequency features for audio fingerprinting, in IEEE International Conference on Acoustics, Speech, and Signal Processing, Mayıs 2002, Orlando, IEEE, II-1773-II-1776.
33. Herre, J., Allamanche, E., ve Hellmuth, O., Robust matching of audio signals using spectral flatness features, in IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics, Ekim 2001, New Platz, IEEE, 127-130.
34. Diamantaras, K.I. ve Kung, S.Y., Principal component neural networks: theory and applications, John Wiley & Sons, Inc., 1996.
35. Burges, C.J., Platt, J.C., ve Jana, S., Extracting noise-robust features from audio data, in IEEE International Conference on Acoustics, Speech, and Signal Processing, Mayıs 2002, Orlando, IEEE, I-1021-I-1024.
36. Duda, R.O., Hart, P.E., ve Stork, D.G., Pattern classification and scene analysis 2nd ed, Wiley Interscience, 1995.
37. Malvar, H., A modulated complex lapped transform and its applications to audio processing, in IEEE International Conference on Acoustics, Speech, and Signal Processing, Mart 1999, Phoenix, IEEE, 1421-1424.

38. Cano, P., Batlle, E., Kalker, T., ve Haitzma, J., A review of audio fingerprinting, Journal of VLSI signal processing systems for signal, image and video technology, 41,3 (2005) 271-284.
39. Billinghamurst, M. ve Starner, T., Wearable devices: new ways to manage information, Computer, 32,1 (1999) 57-64.
40. Agaskar, A.P.V.A., Mithagari, A., Mhatre, A., ve Shetty, N., Internet of things: Home automation and surveillance system, International Educational Scientific Research Journal, 2,4 (2016).
41. Swan, M., The quantified self: Fundamental disruption in big data science and biological discovery, Big Data, 1,2 (2013) 85-99.
42. Oh, J. ve Lee, U., Exploring UX issues in Quantified Self technologies, in Eighth International Conference on Mobile Computing and Ubiquitous Networking (ICMU), Ocak 2015, Hakodate, IEEE, 53-59.
43. Mann, S., Wearable computing as means for personal empowerment, in Int. Conf. on Wearable Computing (ICWC), Mayıs 1998, Fairfax VA, Bildiriler Kitabı, 51-59.
44. Scheirer, E. ve Slaney, M., Construction and evaluation of a robust multifeature speech/music discriminator, in IEEE International Conference on Acoustics, Speech, and Signal Processing, Nisan 1997, Washington, IEEE, 1331-1334.
45. Brown, J., Musical instrument identification using autocorrelation coefficients, in International Symposium on Musical Acoustics, Haziran 1998, Washington, Bildiriler Kitabı, 291-295.
46. Martin, K.D. ve Kim, Y.E., Musical instrument identification: A pattern-recognition approach, The Journal of the Acoustical Society of America, 104,3 (1998) 1768-1768.
47. Rabiner, L.R. ve Juang, B.H., Fundamentals of speech recognition, PTR Prentice Hall Englewood Cliffs, 1993
48. Vassis, D., Belsis, P., Skourlas, C., Marinagi, C., ve Tsoukalas, V., Secure mobile assessment of deaf and hard-of-hearing and dyslexic students in higher education, in 17th Panhellenic Conference on Informatics, Eylül 2013, Thessaloniki, ACM, 311-318.
49. Shivakumar, B. ve Rajasenathipathi, M., A New Approach for Hardware Control Procedure Used in Braille Glove Vibration System for Disabled Persons, Research Journal of Applied Sciences, Engineering and Technology, 7,9 (2014) 1863-1871.
50. Arato, A., Markus, N., ve Juhasz, Z., Teaching morse language to a deaf-blind person for reading and writing SMS on an ordinary vibrating smartphone, in International Conference on Computers for Handicapped Persons, Temmuz 2014, Paris, Springer, 393-396.

51. Nanayakkara, S.C., Wyse, L., Ong, S.H., ve Taylor, E.A., Enhancing musical experience for the hearing-impaired using visual and haptic displays, Human-Computer Interaction, 28,2 (2013) 115-160.
52. Gollner, U., Bieling, T., ve Joost, G., Mobile Lorm Glove: introducing a communication device for deaf-blind people, in Proceedings of the sixth international conference on tangible, embedded and embodied interaction, Şubat 2012, Kingston, ACM, 127-130.
53. Schmitz, B. ve Ertl, T., Making digital maps accessible using vibrations, in International Conference on Computers for Handicapped Persons, Temmuz 2010, Vienna, Springer, 100-107.
54. Ketabdar, H. ve Polzehl, T., Tactile and visual alerts for deaf people by mobile phones, in 11th international ACM SIGACCESS conference on Computers and accessibility, Ekim 2009, Pittsburgh, ACM, 253-254.
55. Nanayakkara, S., Taylor, E., Wyse, L., ve Ong, S.H., An enhanced musical experience for the deaf: design and evaluation of a music display and a haptic chair, in SIGCHI Conference on Human Factors in Computing Systems, Nisan 2009, Boston, ACM, 337-346.
56. Caetano, G. ve Jousmäki, V., Evidence of vibrotactile input to human auditory cortex, Neuroimage, 29,1 (2006) 15-28.
57. Damper, R.I. ve Evans, M.D., A multifunction domestic alert system for the deaf-blind, IEEE Transactions on Rehabilitation Engineering, 3,4 (1995) 354-359.
58. Heracleous, P., Beautemps, D., ve Aboutabit, N., Cued Speech automatic recognition in normal-hearing and deaf subjects, Speech Communication, 52,6 (2010) 504-512.
59. Jeyalakshmi, C., Krishnamurthi, V., ve Revathi, A., Development Of Speech Recognition System For Hearing Impaired In Native language, Journal of Engineering Research, 2,2 (2014) 81-89.
60. Lozano, H., Hernáez, I., Picón, A., Camarena, J., ve Navas, E., Audio classification techniques in home environments for elderly/dependant people, in International Conference on Computers for Handicapped Persons, Temmuz 2010, Vienna, Springer, 320-323.
61. Beskow, J., Granström, B., Nordqvist, P., Al Moubayed, S., Salvi, G., Herzke, T., ve Schulz, A., Hearing at home-communication support in home environments for hearing impaired persons, in INTERSPEECH, 2008, 2203-2206.
62. Carrera, A., Alonso, A., de la Rosa, R., ve Abril, E.J., Sensing Performance of a Vibrotactile Glove for Deaf-Blind People, Applied Sciences, 7,4 (2017) 317.
63. Sakajiri, M., Miyoshi, S., Nakamura, K., Fukushima, S., ve Ifukube, T., Voice pitch control using tactile feedback for the deafblind or the hearing impaired persons to

- assist their singing, in IEEE International Conference on Systems, Man and Cybernetics, Ekim 2010, İstanbul, IEEE, 1483-1487.
64. Shull, P.B. ve Damian, D.D., Haptic wearables as sensory replacement, sensory augmentation and trainer—a review, Journal of neuroengineering and rehabilitation, 12,1 (2015) 59.
 65. Mirri, S., Prandi, C., Salomoni, P., ve Monti, L., Fitting like a GlovePi: A wearable device for deaf-blind people, in 14th IEEE Annual Consumer Communications & Networking Conference (CCNC), Ocak 2017, Las Vegas, IEEE, 1057-1062.
 66. Jadhav, A.J. ve Joshi, M.P., AVR based embedded system for speech impaired people, in International Conference on Automatic Control and Dynamic Optimization Techniques (ICACDOT), Eylül 2016, Pune, IEEE, 844-848.
 67. Zhong, X., Wang, S., Dorman, M., ve Yost, W., Sound source localization from tactile aids for unilateral cochlear implant users, The Journal of the Acoustical Society of America, 134,5 (2013) 4062-4062.
 68. Wang, S., Zhong, X., Dorman, M.F., Yost, W.A., ve Liss, J.M., Using tactile aids to provide low frequency information for cochlear implant users, The Journal of the Acoustical Society of America, 134,5 (2013) 4235-4235.
 69. Ho-Ching, F.W.-l., Mankoff, J., ve Landay, J.A., From data to display: the design and evaluation of a peripheral sound display for the deaf, in Proceedings of CHI, Nisan 2003, Florida.
 70. Kumari, P., Goel, P., ve Reddy, S., PiCam: IoT based Wireless Alert System for Deaf and Hard of Hearing, in International Conference on Advanced Computing and Communications (ADCOM), Eylül 2015, Chennai, IEEE, 39-44.
 71. Matthews, T., Fong, J., Ho-Ching, F.W.-L., ve Mankoff, J., Evaluating non-speech sound visualizations for the deaf, Behaviour & Information Technology, 25,4 (2006) 333-351.
 72. Gao, W., Ma, J., Wu, J., ve Wang, C., Sign language recognition based on HMM/ANN/DP, International journal of pattern recognition and artificial intelligence, 14,05 (2000) 587-602.
 73. Liu, Y., Yun, H.S., ve Kim, N.S., Audio fingerprinting based on multiple hashing in DCT domain, IEEE signal processing letters, 16,6 (2009) 525-528.
 74. Baluja, S. ve Covell, M., Waveprint: Efficient wavelet-based audio fingerprinting, Pattern recognition, 41,11 (2008) 3467-3480.
 75. Zhu, B., Li, W., Wang, Z., ve Xue, X., A novel audio fingerprinting method robust to time scale modification and pitch shifting, in 18th ACM international conference on Multimedia, Ekim 2010, Firenze, ACM, 987-990.

76. Chung, M. ve Ko, I., An algorithm that minimizes audio fingerprints using the difference of Gaussians, Journal of Zhejiang University SCIENCE C, 12,10 (2011) 836.
77. Rein, S. ve Reisslein, M., Identifying the classical music composition of an unknown performance with wavelet dispersion vector and neural nets, Information Sciences, 176,12 (2006) 1629-1655.
78. Ellis, D.P., Whitman, B., ve Porter, A., Echoprint: An open music identification service, in 12th International Society for Music Information Retrieval Conference (ISMIR 2011), Ekim 2011, Miami, Bildiriler Kitabı, 121-126.
79. Serra, J., Gómez, E., Herrera, P., ve Serra, X., Chroma binary similarity and local alignment applied to cover song identification, IEEE Transactions on Audio, Speech, and Language Processing, 16,6 (2008) 1138-1151.
80. Chung, M.-b. ve Ko, I.-j., Identical-video retrieval using the low-peak feature of a video's audio information, Journal of Zhejiang University SCIENCE C, 11,3 (2010) 151-159.
81. Pan, X., Yu, X., Deng, J., Yang, W., ve Wang, H., Audio fingerprinting based on local energy centroid, in IET International Communication Conference on Wireless Mobile and Computing (CCWMC 2011), Kasım 2011, Shanghai, IEEE, 351-354.
82. Jiang, T., Wu, R., Li, J., Xiang, K., ve Dai, F., A real-time peak discovering method for audio fingerprinting, in Fifth International Conference on Internet Multimedia Computing and Service, Ağustos 2013, Huangshan, ACM, 368-371.
83. Ouali, C., Dumouchel, P., ve Gupta, V., Fast audio fingerprinting system using GPU and a clustering-based technique, IEEE/ACM Transactions on Audio, Speech and Language Processing (TASLP), 24,6 (2016) 1106-1118.
84. Badii, A. ve Thiemert, D., A framework towards a multi-modal fingerprinting scheme for multimedia assets, International Journal of Business Information Systems, 6,1 (2010) 133-149.
85. Saracoglu, A., Esen, E., Ates, T.K., Acar, B.O., Zubari, U., Ozan, E.C., Ozalp, E., Alatan, A.A., ve Ciloglu, T., Content based copy detection with coarse audio-visual fingerprints, in Seventh International Workshop on Content-Based Multimedia Indexing, Haziran 2009, Chania, IEEE, 213-218.
86. Cheung, S.-S. ve Zakhori, A., Efficient video similarity measurement with video signature, IEEE Transactions on Circuits and Systems for video Technology, 13,1 (2003) 59-74.
87. Oostveen, J., Kalker, T., ve Haitsma, J., Feature extraction and a database strategy for video fingerprinting, in 5th International Conference on Recent Advances in Visual Information Systems, Mart 2002, Springer, 117-128.

88. Hampapur, A. ve Bolle, R., VideoGREP: Video copy detection using inverted file indices, nine, 3, (2002) 6.
89. Seo, J.S., Jin, M., Lee, S., Jang, D., Lee, S., ve Yoo, C.D., Audio fingerprinting based on normalized spectral subband moments, IEEE Signal Processing Letters, 13,4 (2006) 209-212.
90. Yuan, J., Duan, L.-Y., ve Tian, Q., Fast Video Segment Identification from Large Video Collection, in Pacific-Rim Conference on Multimedia (PCM'04), Kasım 2004, Tokyo.
91. Park, M., Kim, H.R., ve Yang, S.H., Frequency-Temporal Filtering for a Robust Audio Fingerprinting Scheme in Real-Noise Environments, ETRI journal, 28,4 (2006) 509-512.
92. Kim, H.-G., Cho, H.-S., ve Kim, J.Y., Robust audio fingerprinting using peak-pair-based hash of non-repeating foreground audio in a real environment, Cluster Computing, 19,1 (2016) 315-323.
93. Fenet, S., Richard, G., ve Grenier, Y., A Scalable Audio Fingerprint Method with Robustness to Pitch-Shifting, in 12th International Society for Music Information Retrieval Conference (ISMIR 2011), Ekim 2011, Miami, Bildiriler Kitabı, 121-126.
94. Anguera, X., Garzon, A., ve Adamek, T., MASK: Robust local features for audio fingerprinting, in IEEE International Conference on Multimedia and Expo, Temmuz 2012, Melbourne, IEEE, 455-460.
95. Kim, H.G. ve Kim, J.Y., Robust audio fingerprinting method using prominent peak pair based on modulated complex lapped transform, ETRI Journal, 36,6 (2014) 999-1007.
96. Panagiotou, V. ve Mitianoudis, N., PCA summarization for audio song identification using Gaussian mixture models, in 18th International Conference on Digital Signal Processing (DSP), Temmuz 2013, Fira, IEEE,1-6.
97. Lebossé, J., Brun, L., ve Pailles, J.C., A robust audio fingerprint extraction algorithm, in Fourth conference on IASTED International Conference: Signal Processing, Pattern Recognition, and Applications, Şubat 2007, Innsbruck, ACM, 269-274.
98. Younessian, E., Anguera, X., Adamek, T., Oliver, N., ve Marimon, D., Telefonica Research at TRECVID 2010 Content-Based Copy Detection, in TRECVID 2010.
99. Mukai, R., Kurozumi, T., Hiramatsu, K., Kawanishi, T., Nagano, H., ve Kashino, K., NTT Communication Science Laboratories at TRECVID 2010 Content Based Copy Detection, in TRECVID, 2010.
100. Nagano, H., Mukai, R., Kurozumi, T., ve Kashino, K., A fast audio search method based on skipping irrelevant signals by similarity upper-bound calculation, in IEEE

- International Conference on Acoustics, Speech and Signal Processing (ICASSP), Nisan 2015, Brisbane, IEEE, 2324-2328.
101. Ke, Y., Hoiem, D., ve Sukthankar, R., Computer vision for music identification, in IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), Haziran 2005, San Diego, IEEE, 597-604.
 102. Ramalingam, A. ve Krishnan, S., Gaussian mixture modeling using short time fourier transform features for audio fingerprinting, in IEEE International Conference on Multimedia and Expo, Temmuz 2005, Amsterdam, IEEE, 1146-1149.
 103. Allamanche, E., Herre, J., Hellmuth, O., Fröba, B., Kastner, T., ve Cremer, M., Content-based Identification of Audio Material Using MPEG-7 Low Level Description, in 2nd Annual International Symposium on Music Information Retrieval (ISMIR), Ekim 2001, Indiana, Bildiriler Kitabı, 197-204.
 104. Cotton, C.V. ve Ellis, D.P., Audio fingerprinting to identify multiple videos of an event, in IEEE International Conference on Acoustics, Speech and Signal Processing, Mart 2010, Dallas, IEEE, 2386-2389.
 105. Ibarrola, A.C. ve Chávez, E., A robust entropy-based audio-fingerprint, in IEEE International Conference on Multimedia and Expo, Temmuz 2006, Toronto, IEEE, 1729-1732.
 106. Baluja, S. ve Covell, M., Audio fingerprinting: Combining computer vision & data stream processing, in IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07, Nisan 2007, Honolulu, IEEE, II-213-II-216.
 107. Kim, S. ve Yoo, C.D., Boosted binary audio fingerprint based on spectral subband moments, in IEEE International Conference on Acoustics, Speech and Signal Processing - ICASSP '07, Nisan 2007, Honolulu, IEEE, 241,244.
 108. Jang, D., Yoo, C.D., Lee, S., Kim, S., ve Kalker, T., Pairwise boosted audio fingerprint, IEEE transactions on information forensics and security, 4,4 (2009) 995-1004.
 109. Uchida, Y., Sakazawa, S., Agrawal, M., ve Akbacak, M., KDDI LABS and SRI International at TRECVID 2010: Content-Based Copy Detection, in TRECVID, 2010.
 110. Jégou, H., Delhumeau, J., Yuan, J., Gravier, G., ve Gros, P., Babaz: a large scale audio search system for video copy detection, in IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), Mart 2012, Kyoto, IEEE, 2369-2372.
 111. Khemiri, H., Petrovska-Delacretaz, D., ve Chollet, G., Detection of repeating items in audio streams using data-driven ALISP sequencing, in International Conference on Advanced Technologies for Signal and Image Processing (ATSIP), Mart 2014, Sousse, IEEE, 446-451.

112. Ngo, C.-W., Jiang, Y.-G., Wei, X.-Y., Zhao, W., Liu, Y., Wang, J., Zhu, S., ve Chang, S.-F., Vireo/dvmm at trecvid 2009: High-level feature extraction, automatic video search, and content-based copy detection, in TRECVID2009, Kasım 2009, Gaithersburg, Bildiriler Kitabı, 415-432.
113. Seo, J.S., Jin, M., Lee, S., Jang, D., Lee, S., ve Yoo, C.D., Audio fingerprinting based on normalized spectral subband centroids, in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05), Mart 2005, Philadelphia, IEEE, 213-216.
114. Six, J. ve Leman, M., Panako: a scalable acoustic fingerprinting system handling time-scale and pitch modification, in 15th International Society for Music Information Retrieval Conference (ISMIR-2014), Ekim 2014, Taipei, Bildiriler Kitabı, 259-264.
115. Sonnleitner, R. ve Widmer, G., Quad-Based Audio Fingerprinting Robust to Time and Frequency Scaling, in 17th Int. Conference on Digital Audio Effects (DAFx-14), Eylül 2014, Erlangen, Bildiriler Kitabı, 173-180.
116. George, J. ve Jhunjhunwala, A., Scalable and robust audio fingerprinting method tolerable to time-stretching, in IEEE International Conference on Digital Signal Processing (DSP), Temmuz 2015, Singapore, IEEE, 436-440.
117. Shi, Y.Z., Zhang, W.Q., ve Liu, J., Robust audio fingerprinting based on local spectral luminance maxima scheme, in 12th Annual Conference of the International Speech Communication Association, Ağustos 2011, Florence, Bildiriler Kitabı, 2485-2488.
118. Dupraz, E. ve Richard, G., Robust frequency-based audio fingerprinting, in IEEE International Conference on Acoustics, Speech and Signal Processing, Mart 2010, Dallas, IEEE, 281-284.
119. Ouali, C., Dumouchel, P., ve Gupta, V., A robust audio fingerprinting method for content-based copy detection, in 12th International Workshop on Content-Based Multimedia Indexing (CBMI), Haziran 2014, Klagenfurt, IEEE, 1-6.
120. Wang, J.-C., Lee, Y.-S., Lin, C.-H., Siahaan, E., ve Yang, C.-H., Robust environmental sound recognition with fast noise suppression for home automation, IEEE Transactions on Automation Science and Engineering, 12,4 (2015) 1235-1242.
121. Salamon, J. ve Bello, J.P., Deep convolutional neural networks and data augmentation for environmental sound classification, IEEE Signal Processing Letters, 24,3 (2017) 279-283.
122. Kugler, M., Tossavainen, T., Nakatsu, M., Kuroyanagi, S., ve Iwata, A., Real-time hardware implementation of a sound recognition system with in-field learning, IEICE TRANSACTIONS on Information and Systems, 99,7 (2016) 1885-1894.

123. Kingsbury, B.E., Morgan, N., ve Greenberg, S., Robust speech recognition using the modulation spectrogram, Speech communication, 25,1-3 (1998) 117-132.
124. Reynolds, D.A., Quatieri, T.F., ve Dunn, R.B., Speaker verification using adapted Gaussian mixture models, Digital signal processing, 10,1-3 (2000) 19-41.
125. Oberle, S. ve Kaelin, A., Recognition of acoustical alarm signals for the profoundly deaf using hidden Markov models, in IEEE International Symposium on Circuits and Systems (ISCAS '95), Nisan 1995, Seattle, IEEE, 2285-2288.
126. Seoane, F., Mohino-Herranz, I., Ferreira, J., Alvarez, L., Buendia, R., Ayllón, D., Llerena, C., ve Gil-Pita, R., Wearable biomedical measurement systems for assessment of mental stress of combatants in real time, Sensors, 14,4 (2014) 7120-7141.
127. Reed, C.M. ve Delhorne, L.A., The reception of environmental sounds through wearable tactual aids, Ear and hearing, 24,6 (2003) 528-538.
128. Alías, F., Socoró, J.C., ve Sevillano, X., A review of physical and perceptual feature extraction techniques for speech, music and environmental sounds, Applied Sciences, 6,5 (2016) 143.
129. Eronen, A., Comparison of features for musical instrument recognition, in IEEE Workshop on the Applications of Signal Processing to Audio and Acoustics, Ekim 2001, New Platz, IEEE, 19-22.
130. Barbedo, J.G.A. ve Tzanetakis, G., Musical instrument classification using individual partials, IEEE Transactions on Audio, Speech, and Language Processing, 19,1 (2011) 111-122.
131. Lin, R.S. ve Chen, L.H., A new approach for audio classification and segmentation using Gabor wavelets and Fisher linear discriminator, International Journal of Pattern Recognition and Artificial Intelligence, 19,06 (2005) 807-822.
132. Chen, J.K. ve Soong, F.K., An N-best candidates-based discriminative training for speech recognition applications, IEEE Transactions on Speech and Audio Processing, 2,1 (1994) 206-216.
133. El-Ramly, S.H., Abdel-Kader, N., ve El-Adawi, R., Neural networks used for speech recognition, in Nineteenth National Radio Science Conference (NRSC 2002), Mart 2002, Alexandria, IEEE, 200-207.
134. Yaniv, R. ve Burshtein, D., An enhanced dynamic time warping model for improved estimation of DTW parameters, IEEE Transactions on Speech and Audio Processing, 11,3 (2003) 216-228.
135. Lin, Y.L. ve Wei, G., Speech emotion recognition based on HMM and SVM, in International Conference on Machine Learning and Cybernetics, Ağustos 2005, Guangzhou, IEEE, 4898-4901.

136. Debyeche, M., Haton, J., ve Houacine, A., A New Vector Quantization front-end Process for Discrete HMM Speech Recognition System, International Journal of Signal Processing, 3,1 (2006) 46-51.
137. Kruger, S.E., Schaffoner, M., Katz, M., Andelic, E., ve Wendemuth, A., Mixture of support vector machines for hmm based speech recognition, in 18th International Conference on Pattern Recognition (ICPR'06), Ağustos 2006, Hong Kong, IEEE, 326-329.
138. Lam, H.-K. ve Leung, F.H., Design and training for combinational neural-logic systems, IEEE transactions on industrial electronics, 54,1 (2007) 612-619.
139. De Wachter, M., Matton, M., Demuyne, K., Wambacq, P., Cools, R., ve Van Compernelle, D., Template-based continuous speech recognition, IEEE Transactions on Audio, Speech, and Language Processing, 15,4 (2007) 1377-1390.
140. Al-Haddad, S.A.R., Samad, S.A., Hussain, A., ve Ishak, K.A., Isolated Malay digit recognition using pattern recognition fusion of dynamic time warping and hidden Markov models, American Journal of Applied Sciences, 5,6 (2008) 714-720.
141. Ma, L., Milner, B., ve Smith, D., Acoustic environment classification, ACM Transactions on Speech and Language Processing (TSLP), 3,2 (2006) 1-22.
142. Guzman-Zavaleta, Z.J., Feregrino-Urbe, C., Menendez-Ortiz, A., ve Garcia-Hernandez, J.J., A robust audio fingerprinting method using spectrograms saliency maps, in 9th International Conference for Internet Technology and Secured Transactions (ICITST-2014), Aralık 2014, London, IEEE, 47-52.
143. Sankaran, S.K.S., Analysis of Audio Fingerprinting Techniques, Northern Illinois University, 2017.
144. Liu, F., Audio fingerprinting for speech reconstruction and recognition in noisy environments, 2017.
145. Kastner, T., Allamanche, E., Herre, J., Hellmuth, O., Cremer, M., ve Grossmann, H., MPEG-7 scalable robust audio fingerprinting, in Audio Engineering Society Convention, Mayıs 2002, Munich.
146. Wang, Y., Liu, Z., ve Huang, J.-C., Multimedia content analysis-using both audio and visual clues, IEEE signal processing magazine, 17,6 (2000) 12-36.
147. Goyal, N. ve Purwar, R., Analyzing Mel Frequency Cepstral Coefficient for Recognition of Isolated English Word using DTW Matching, IJRCCT, 3,4 (2014) 436-443.
148. Krishna, A. ve Sreenivas, T.V., Music instrument recognition: from isolated notes to solo phrases, in IEEE International Conference on Acoustics, Speech, and Signal Processing, Mayıs 2004, Montreal, IEEE, 265-268.

149. Marques, J. ve Moreno, P.J., A study of musical instrument classification using gaussian mixture models and support vector machines, Cambridge Research Laboratory Technical Report Series CRL, 4, (1999).
150. Paliwal, K.K., Spectral subband centroid features for speech recognition, in IEEE International Conference on Acoustics, Speech and Signal Processing, Mayıs 1998, Seattle, IEEE, 617-620.
151. Soong, F. ve Juang, B., Line spectrum pair (LSP) and speech data compression, in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'84), Mart 1984, San Diego, IEEE, 37-40.
152. Sugamura, N. ve Farvardin, N., Quantizer design in LSP speech analysis-synthesis, IEEE journal on selected areas in communications, 6,2 (1988) 432-440.
153. Bäckström, T. ve Magi, C., Properties of line spectrum pair polynomials—A review, Signal processing, 86,11 (2006) 3286-3298.
154. McLoughlin, I.V., Line spectral pairs, Signal processing, 88,3 (2008) 448-467.
155. Yucesoy, E. ve Nabiyev, V.V., Comparison of MFCC, LPCC and PLP features for the determination of a speaker's gender, in 22nd Signal Processing and Communications Applications Conference (SIU), Nisan 2014, Trabzon, IEEE, 321-324.
156. Gulzar, T., Singh, A., ve Sharma, S., Comparative analysis of LPCC, MFCC and BFCC for the recognition of Hindi words using artificial neural networks, International Journal of Computer Applications, 101,12 (2014) 22-27.
157. Adeniyi, K.I., Comparative Study on the Performance of Mel-Frequency Cepstral Coefficients and Linear Prediction Cepstral Coefficients under different Speaker's Conditions, International Journal of Computer Applications, 90,11 (2014).
158. Glowacz, A., Diagnostics of DC and induction motors based on the analysis of acoustic signals, Measurement Science Review, 14,5 (2014) 257-262.
159. Yang, Y. ve Liu, X., A re-examination of text categorization methods, in 22nd annual international ACM SIGIR conference on Research and development in information retrieval, Ağustos 1999, California, ACM, 42-49.
160. Joachims, T., Text categorization with support vector machines: Learning with many relevant features, in 10th European Conference on Machine Learning, Nisan 1998, London, Springer, 137-142.
161. Baoli, L., Yuzhong, C., ve Shiwen, Y., A comparative study on automatic categorization methods for Chinese search engine, in Eighth Joint International Computer Conference, 2002, Hangzhou: Zhejiang University Press, 117-120.

162. Cortes, C. ve Vapnik, V., Support-vector networks, Machine learning, 20,3 (1995) 273-297.
163. Kavzođlu, T. ve ölkesen, İ., Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi, Harita Dergisi, 144,7 (2010) 73-82.
164. Yađanođlu, M. ve Köse, C., Wearable Vibration Based Computer Interaction and Communication System for Deaf, Applied Sciences, 7,12 (2017) 1296.
165. Yađanođlu, M., Bozkurt, F., ve Günay, F.B., Eeg Tabanlı Beyin-Bilgisayar Arayüzü Sistemlerinde Öznitelik ıkarma Yöntemleri, Mühendislik Bilimleri ve Tasarım Dergisi, 2,3 (2014) 313-318.
166. Zhang, J., Chen, M., Zhao, S., Hu, S., Shi, Z., ve Cao, Y., Relieff-based EEG sensor selection methods for emotion recognition, Sensors, 16,10 (2016) 1558.
167. Bolón-Canedo, V., Sánchez-Marono, N., Alonso-Betanzos, A., Benítez, J.M., ve Herrera, F., A review of microarray datasets and applied feature selection methods, Information Sciences, 282, (2014) 111-135.
168. So, H.C., Source localization: Algorithms and analysis, Handbook of Position Location: Theory, Practice, and Advances, (2011) 25-66.
169. Dogañay, K., Emitter localization using clustering-based bearing association, IEEE Transactions on Aerospace and Electronic Systems, 41,2 (2005) 525-536.
170. Roberts, R., Tdoa localization techniques, Harris Corporation, 4, (2004).
171. Shi, G. ve Ming, Y., Survey of indoor positioning systems based on ultra-wideband (UWB) technology, in Wireless Communications, Networking and Applications, 2016, Springer, 1269-1278.
172. Yađanođlu, M. ve Köse, C., Real-time Parental Voice Recognition System For Persons Having Impaired Hearing, Bilge International Journal of Science and Technology Research, 2,1 (2018) 40-46.
173. Yađanođlu, M., Kayabekir, M., ve Köse, C., SNORAP: A Device for the Correction of Impaired Sleep Health by Using Tactile Stimulation for Individuals with Mild and Moderate Sleep Disordered Breathing, Sensors, 17,9 (2017) 2006.
174. Garrido, F., Verbeke, W., ve Bravo, C., A Robust profit measure for binary classification model evaluation, Expert Systems with Applications, 92, (2018) 154-160.
175. Eronen, A.J., Peltonen, V.T., Tuomi, J.T., Klapuri, A.P., Fagerlund, S., Sorsa, T., Lorho, G., ve Huopaniemi, J., Audio-based context recognition, IEEE Transactions on Audio, Speech, and Language Processing, 14,1 (2006) 321-329.

176. Su, F., Yang, L., Lu, T., ve Wang, G., Environmental sound classification for scene recognition using local discriminant bases and HMM, in 19th ACM international conference on Multimedia, Kasım 2011, Arizona, ACM, 1389-1392.
177. Wang, J.-C., Wang, J.-F., He, K.W., ve Hsu, C.-S., Environmental sound classification using hybrid SVM/KNN classifier and MPEG-7 audio low-level descriptor, in IEEE International Joint Conference on Neural Network Proceedings, Temmuz 2006, Vancouver, IEEE, 1731-1735.
178. Chu, S., Narayanan, S., ve Kuo, C.-C.J., Environmental sound recognition with time–frequency audio features, IEEE Transactions on Audio, Speech, and Language Processing, 17,6 (2009) 1142-1158.
179. Malkin, R.G. ve Waibel, A., Classifying user environment for mobile applications using linear autoencoding of ambient audio, in IEEE International Conference on Acoustics, Speech, and Signal Processing (ICASSP'05), Mart 2005, Philadelphia, IEEE, 509-512.
180. Wang, C., Peng, J., Song, L., ve Zhang, X., Automatic snoring sounds detection from sleep sounds via multi-features analysis, Australasian physical & engineering sciences in medicine, 40,1 (2017) 127-135.
181. Dafna, E., Tarasiuk, A., ve Zigel, Y., Automatic detection of whole night snoring events using non-contact microphone, PloS one, 8,12 (2013) e84139.
182. Yadollahi, A. ve Moussavi, Z., Automatic breath and snore sounds classification from tracheal and ambient sounds recordings, Medical Engineering and Physics, 32,9 (2010) 985-990.
183. Karunajeewa, A.S., Abeyratne, U.R., ve Hukins, C., Silence–breathing–snore classification from snore-related sounds, Physiological Measurement, 29,2 (2008) 227.
184. Cavusoglu, M., Kamasak, M., Erogul, O., Ciloglu, T., Serinagaoglu, Y., ve Akcam, T., An efficient method for snore/nonsnore classification of sleep sounds, Physiological measurement, 28,8 (2007) 841.
185. Duckitt, W., Tuomi, S., ve Niesler, T., Automatic detection, segmentation and assessment of snoring from ambient acoustic data, Physiological measurement, 27,10 (2006) 1047.
186. Peltonen, V., Tuomi, J., Klapuri, A., Huopaniemi, J., ve Sorsa, T., Computational auditory scene recognition, in IEEE International Conference on Acoustics, Speech, and Signal Processing, Mayıs 2002, Orlando, IEEE, II-1941-II-1944.
187. Adiloglu, K., Annies, A.s., Wahlen, E., Purwins, H., ve Obermayer, K., A graphical representation and dissimilarity measure for basic everyday sound events, IEEE Transactions on Audio, Speech, and Language Processing, 20,5 (2012) 1542-1552.

ÖZGEÇMİŞ

28.09.1985 yılında Erzurum'da doğdu. İlk, orta ve lise öğrenimini Erzurum'da tamamladı. 2004 yılında Erciyes Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği Bölümü'nde lisans programına başladı ve 2009 yılında bu bölümden mezun oldu. Aynı yıl Atatürk Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda yüksek lisans öğrenimine başladı. 2010 yılında Atatürk Üniversitesi Mühendislik Fakültesi Bilgisayar Mühendisliği bölümünde Araştırma Görevlisi olarak göreve başladı. 2012 yılında Atatürk Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda yüksek lisans öğrenimini tamamladı. Yabancı dil olarak iyi seviyede İngilizce bilmektedir.

SCI/SCI-E indekslerine giren dergilerde yayınlanan makaleler

1. Yağanoğlu, M. ve Köse, C., Wearable Vibration Based Computer Interaction and Communication System for Deaf, Applied Sciences, 7,12 (2017) 1296.
2. Yağanoğlu, M. ve Köse, C. Real-Time Detection of Important Sounds with a Wearable Vibration Based Device for Hearing-Impaired People. Electronics 2018, 7, 50.
3. Yağanoğlu, M., Kayabekir, M., ve Köse, C., SNORAP: A Device for the Correction of Impaired Sleep Health by Using Tactile Stimulation for Individuals with Mild and Moderate Sleep Disordered Breathing, Sensors, 17,9 (2017) 2006.

Diğer dergilerde yayınlanan makaleler

1. Yağanoğlu, M. ve Köse, C., Real-time Parental Voice Recognition System For Persons Having Impaired Hearing, Bilge International Journal of Science and Technology Research, 2,1 (2018) 40-46.