

**SES ANALİZİNDE AKUSTİK PARAMETRELERİN
TESPİTİ VE ANKSİYETE BOZUKLUĞUNUN
AKUSTİK PARAMETRELERLE
İLİŞKİSİNİN ARAŞTIRILMASI**

**2017
DOKTORA TEZİ
BİLGİSAYAR MÜHENDİSLİĞİ**

Turgut ÖZSEVEN

**SES ANALİZİNDE AKUSTİK PARAMETRELERİN TESPİTİ VE
ANKSİYETE BOZUKLUĞUNUN AKUSTİK PARAMETRELERLE
İLİŞKİSİNİN ARAŞTIRILMASI**

Turgut ÖZSEVEN

**Karabük Üniversitesi
Fen Bilimleri Enstitüsü
Bilgisayar Mühendisliği Anabilim Dalında
Doktora Tezi
Olarak Hazırlanmıştır**

**KARABÜK
Şubat 2017**

Turgut ÖZSEVEN tarafından hazırlanan “SES ANALİZİNDE AKUSTİK PARAMETRELERİN TESPİTİ VE ANKSİYETE BOZUKLUĞUNUN AKUSTİK PARAMETRELERLE İLİŞKİSİNİN ARAŞTIRILMASI” başlıklı bu tezin Doktora Tezi olarak uygun olduğunu onaylarım.

Yrd. Doç. Dr. Muharrem DÜĞENÇİ
Tez Danışmanı, Endüstri Mühendisliği Anabilim Dalı



Prof. Dr. Ali DORUK
Tez Danışmanı, Sağlık Bilimleri Üniversitesi, Psikiyatri Anabilim Dalı

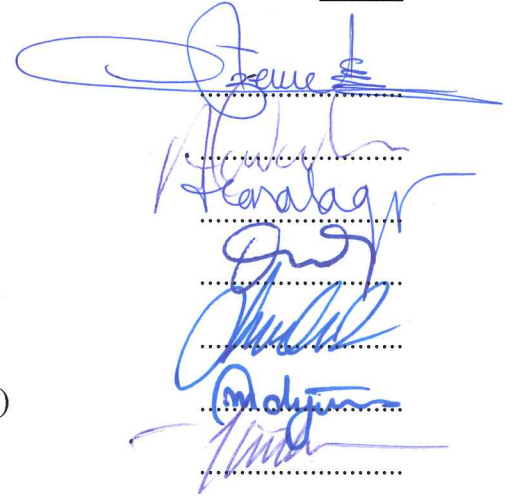


Bu çalışma, jürimiz tarafından oy birliği ile Bilgisayar Mühendisliği Anabilim Dalında Doktora tezi olarak kabul edilmiştir. 03/02/2017

Ünvanı, Adı SOYADI (Kurumu)

İmzası

Başkan : Prof. Dr. Ercan ÖZTEMEL (MÜ)
Üye : Prof. Dr. Ali DORUK (SBÜ)
Üye : Prof. Dr. Derviş KARABOĞA (ERÜ)
Üye : Doç. Dr. Mustafa AKTAŞ (KBÜ)
Üye : Doç. Dr. Oğuz FINDIK (KBÜ)
Üye : Yrd. Doç. Dr. Muharrem DÜĞENÇİ (KBÜ)
Üye : Yrd. Doç. Dr. İlker TÜRKER (KBÜ)



...../...../2017

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

Prof. Dr. Nevin AYTEMİZ
Fen Bilimleri Enstitüsü Müdürü



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

Turgut ÖZSEVEN

ÖZET

Doktora Tezi

SES ANALİZİNDE AKUSTİK PARAMETRELERİN TESPİTİ VE ANKSİYETE BOZUKLUĞUNUN AKUSTİK PARAMETRELERLE İLİŞKİSİNİN ARAŞTIRILMASI

Turgut ÖZSEVEN

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

Tez Danışmanı:

Yrd. Doç. Dr. Muharrem DÜĞENCİ

Prof. Dr. Ali DORUK

Şubat 2017, 142 sayfa

Ses, bir iletişim aracı olmasının dışında kişinin kimliği, ruhsal durumu ve fiziksel sağlığı gibi konularda da gösterge oluşturmaktadır. İnsan seslerinden psikolojik tanıda, duygu durumu tespitinde ve kişi tanımada nesnel ve öznel yöntemler kullanılmaktadır. Algısal değerlendirme öznel bir değerlendirme yöntemidir ve konuşmaların uzman kişiler tarafından dinlenip yorumlanması ile gerçekleştirilir. Bu nedenle sonuçlar uzmanın tecrübesine göre değişiklik göstermektedir. Bu sorunun üstesinden gelmek için nesnel değerlendirme yöntemleri kullanılmaktadır. Akustik analiz, sesteki çeşitli parametreler elde edilerek ses yapısının nesnel olarak değerlendirmesi için kullanılır.

Duygusal durum üzerine yapılan alıřmalardaki en buyk sınırlılık duygusal durumun tam olarak tetiklendiđi verilerin edinilmesidir. Ayrıca, anksiyete temel bir duygu olmayıp birden fazla duygunun kombinasyonu olduđu iin sadece anksiyete ieren hazır veri seti bulunmamaktadır. ođunlukla korku duygusu ile birlikte incelenmiřtir.

Duygusal durumların akustik parametreler zerindeki etkileri eřitli alıřmalarda arařtırılmıřtır. Ancak, yapılan alıřmalar ya sınırlı sayıda akustik parametre zerinden gerekleřtirilmiř ya da hazır veri setleri kullanılmıřtır. Ayrıca, anksiyetenin akustik parametreler zerindeki etkisi ayrıntılı olarak arařtırılmamıř ve sınırlı sayıda parametre kullanılmıřtır.

Bu tez alıřmasının konusu, seslerden akustik parametrelerin tespit edilmesi ve anksiyetenin akustik parametreler zerindeki etkilerinin ayrıntılı incelenmesidir. Bu amala konuřma iřleme ve akustik analiz iin MATLAB ortamında SPAC isimli bir yazılım geliřtirilmiřtir. SPAC'ın literatrdeki yazılımlara gre stnlđ kullanıcı dostu grafiksel arayze ve modler yapıya sahip olması, son iřleme ve sınıflandırma modlerini iermesi ve seslerden elde ettiđi akustik parametre sayısı ve eřitliliđidir. SPAC ile znelik kmelerinin katsayılarına bađlı olarak 723 ve zeri akustik parametre elde edilebilmektedir.

Geerlilik testi sonularına gre, SPAC ile elde edilen akustik parametreler diđer aralarla benzerlik gstermektedir. Sınıflandırma testi sonularına gre SPAC ile elde edilen akustik parametreler her sınıfı tanımlayabilmektedir. Sonu olarak hazırlanan yazılım ile elde edilen parametreler ses verisini dođru řekilde tanımlamaktadır. Ayrıca, SPAC ile edilen dalgacık katsayıları diđer aralarla elde edilememekte ve parametrelerin Δ ve Δ^2 deđerleri her ara tarafından desteklenmemektedir.

Anksiyetenin akustik parametreler zerindeki etkisini arařtırmak iin Glhane Eđitim ve Arařtırma Hastanesi Psikiyatri Anabilim Dalı Bařkanlıđında yatan ve ayaktan hastalardan ses kayıtları toplanmıřtır. Olgu-kontrol niteliđindeki bu deneysel alıřmada, anksiyete bozukluđu olan hastalar ile sađlıklı kontrol grubunun ntr ve

anksiyöz olmak üzere 2 duyguya ilişkin 3 ayrı ses kaydı alınmıştır. Olguların anksiyete düzeyleri uluslararası standartlarda olan psikometrik ölçeklerle değerlendirilmiştir.

Toplanan ses kayıtlarının akustik analizi SPAC ile yapılmıştır. SPAC ile elde edilebilen 9 kategorideki akustik parametrelerin istatistiksel olarak ortalama ve standart sapma değerleri analizlere dahil edilmiştir. Bu nedenle 723 parametre 122'ye düşürülmüştür. Bu parametrelerin anksiyete ile ilişkisini tespit etmek için istatistiksel analiz yöntemleri kullanılmıştır. Analiz sonuçlarına göre 122 parametrenin 42'si anksiyete durumunda değişim göstermiştir. Duygunun tetiklenmesi için metin okutmanın da anksiyeteyi ortaya çıkarmada etkili olduğu görülmüştür.

Yapılan çalışma sonucunda, konuşma işleme ve akustik analiz için geniş öznitelik kütüphanesine sahip ve modüler bir yazılım geliştirilmiştir. Analizler sonrasında, literatüre ek olarak formant frekanslarının bant genişliği, Mel-Frekans Kepstral Katsayıları, Doğrusal Öngörü Kepstral Katsayıları ve dalgacık katsayılarının da anksiyeteden etkilendiği tespit edilmiştir. Anksiyete üzerindeki etkili parametreler önem sırasına göre F0, F1, F2, F3, jitter, LPCC, MFCC ve dalgacık katsayıları şeklinde sıralanabilir. Ayrıca, olguların hasta ve sağlam ayrımında temel frekans, formant frekansı, konuşma oranı ve dalgacık katsayıları ön plana çıkmaktadır.

Anahtar Sözcükler : Akustik analiz, konuşma işleme, anksiyete, duygu akustik parametre ilişkisi.

Bilim Kodu : 902.1.067

ABSTRACT

Ph. D. Thesis

**THE DETECTION OF ACOUSTIC PARAMETERS IN THE VOICE
ANALYSIS AND THE INVESTIGATION OF RELATIONSHIP WITH
ACOUSTIC PARAMETERS OF ANXIETY DISORDERS**

Turgut ÖZSEVEN

Karabük University

Graduate School of Natural and Applied Sciences

Department of Computer Engineering

Thesis Advisor:

Assist. Prof. Dr. Muharrem DÜĞENÇİ

Prof. Dr. Ali DORUK

February 2017, 142 pages

Voice is an indicator of the identity, mental state, and physical health of an individual other than being a communication tool. Psychological diagnosis, emotional state detection and person recognition from human voices are realized by objective and subjective methods. Perceptual evaluation is a subjective evaluation method and is carried out by listening and interpreting the speeches by the experts. Therefore, the results vary according to the experience of the expert. Objective evaluation methods are used to overcome this problem. Acoustic analysis is used in order to obtain several parameters form voice and objectively evaluate voice disorders.

The greatest limitation in studies on the emotional state is the acquisition of data in which the emotional state is precisely triggered. In addition, since anxiety is not a basic emotion but rather a combination of multiple emotions, there are no ready data sets containing only anxiety. It has often been studied with fear.

The effects of emotional states on acoustic parameters have been investigated in various studies. However, studies have been performed on a limited number of acoustic parameters, or have used preliminary data sets. In addition, the effect of anxiety on acoustic parameters has not been investigated in detail and a limited number of parameters have been used.

The purpose of this thesis is to determine acoustic parameters from sounds and to investigate the effects of anxiety on acoustic parameters. For this purpose, a software named SPAC was developed in the MATLAB environment for speech processing and acoustic analysis. The superiority of SPAC over software in the literature is that it has a user-friendly graphical interface and modular structure, including the finishing and classification modalities and the number of acoustic parameters obtained from the sounds. With SPAC, acoustic parameters can be obtained over 723 depending on the coefficients of the feature sets.

According to the results of validation test, acoustic parameters obtained with SPAC similar to other tools. According to the results of classification test, acoustic parameters obtained with SPAC can define every class. As a result, parameters obtained by the prepared software correctly describe the voice data. Moreover, the wavelet coefficients with SPAC cannot be obtained with other software, and the values of Δ and Δ^2 of the parameters are not supported by each software.

To investigate the effect of anxiety on acoustic parameters, voice records were collected from outpatient and inpatient in the Gülhane Education and Research Hospital Psychiatry Department. In this case-control experimental study, three separate voice recordings were taken for two emotions, being neutral and anxious from patients with anxiety disorder and healthy control group. The anxiety levels of the cases were assessed with psychometric scales that are at international standards.

Acoustic analysis of collected sound recordings was performed with SPAC. Statistical mean and standard deviation values of the acoustic parameters in the 9 categories obtained by SPAC were included in the analyzes. For this reason, 723 parameters have been reduced to 122. The statistical analysis methods were used to determine the anxiety associations of these parameters. According to results of the analysis, 42 parameters showed change in anxiety case. It has also been shown that the text reader is also effective in generating anxiety for the triggering of feelings.

As a result of this work, a modular software has been developed with extensive feature library for speech processing and acoustic analysis. After the analyzes, in addition to literature, the bandwidth of the formant frequencies, Mel-Frequency Cepstral Coefficients, Linear Prediction Cepstral Coefficients and wavelet coefficients has also been found to be affected by anxiety. The effective parameters on the anxiety can be listed as F0, F1, F2, F3, jitter, LPCC, MFCC and wavelet coefficients according to importance order. In addition, the fundamental frequency, formant frequency, speech rate and wavelet coefficients are the most important parameters for separation of patient and health of cases.

Key Word : Acoustic analysis, speech processing, anxiety, relationship between emotions and acoustic parameters.

Science Code : 902.1.067

TEŐEKKÜR

Bu tez alıőmasının planlanmasında, araőtırılmasında, yürütülmesinde ve oluşumunda ilgi ve desteęini esirgemeyen, engin bilgi ve tecrübelerinden yararlandığım, yönlendirme ve bilgilendirmeleriyle alıőmamı bilimsel temeller ışığında őekillendiren sayın hocalarım Yrd. Do. Dr. Muharrem DÜĐENCİ ve Prof. Dr. Ali DORUK'a sonsuz teőekkürlerimi sunarım.

Ses verisinin toplanmasında yardımlarını esirgemeyen, baőta Psikolog Hilal KAHRAMAN olmak üzere Gülhane Eęitim ve Araőtırma Hastanesi Ruh ve Sinir Hastalıkları Anabilim Dalı personeline, istatistiksel analiz konusunda bana yol gösteren Do. Dr. Seyit Ali KAYIŐ ve Do. Dr. Filiz ERSÖZ hocalarıma teőekkür ederim.

Sevgili aileme manevi hiçbir yardımı esirgmeden yanımda oldukları için tüm kalbimle teőekkür ederim.

İÇİNDEKİLER

	<u>Sayfa</u>
KABUL	ii
ÖZET.....	iv
ABSTRACT.....	vii
TEŞEKKÜR.....	x
İÇİNDEKİLER	xi
ŞEKİLLER DİZİNİ.....	xv
ÇİZELGELER DİZİNİ	xviii
SİMGELER VE KISALTMALAR DİZİNİ	xix
BÖLÜM 1	1
GİRİŞ	1
BÖLÜM 2	7
GENEL BİLGİLER	7
2.1. SES VE SESİN OLUŞUMU	7
2.2. SİNYAL İŞLEME.....	9
2.2.1. Örnekleme Teoremi	12
2.2.2. Sinyallerin Frekans Analizi	14
2.2.3. Pencereleme.....	17
2.2.4. Dalgacık Dönüşümü	19
2.2.5. DFT Uygulaması	21
2.3. DUYGU VE KONUŞMA	22
2.4. ANKSİYETE.....	25
2.5. YAPILAN ÇALIŞMALAR	26
2.5.1. Akustik Parametrelerin Tespiti İçin Kullanılan Araçlar	26
2.5.2. Anksiyetenin Akustik Parametrelerle İlişkisi	27
BÖLÜM 3	30

	<u>Sayfa</u>
AKUSTİK ANALİZ	30
3.1. VERİ TOPLAMA	32
3.2. KONUŞMA İŞLEME	33
3.2.1. Stereo-Mono Dönüşümü.....	34
3.2.2. Örnekleme Değerini Değiştirme.....	36
3.2.3. DC-Offset Kaldırma	37
3.2.4. Pre-Emphasis/De-Emphasis Filtre.....	38
3.2.5. Gürültü Giderme	39
3.2.6. Çerçeveleme	39
3.2.7. Pencereleme.....	40
3.2.8. Öznitelik Çıkartımı	41
3.2.9. Son İşleme	41
3.2.10. Sınıflandırma	42
3.3. AKUSTİK PARAMETRELER.....	43
3.3.1. Pitch	44
3.3.2. Formant Frekansları.....	45
3.3.3. Jitter	46
3.3.4. Shimmer.....	47
3.3.5. Konuşma Oranı.....	47
3.3.6. Ses Kalitesi (HNR Ve NHR).....	48
3.3.7. Ses Şiddeti	48
 BÖLÜM 4	 50
KONUŞMA İŞLEME YAZILIMI.....	50
4.1. SPAC TASARIMI.....	50
4.1.1. Modül 1: Konfigürasyon.....	52
4.1.2. Modül 2: Ön İşleme.....	54
4.1.3. Modül 3: Öznitelik Çıkartımı	55
4.1.4. Modül 4: Son İşleme.....	56
4.1.5. Modül 5: Sınıflandırma.....	57
4.1.6. Modül 6: Toplu Dosya İşleme	58

	<u>Sayfa</u>
BÖLÜM 5	60
MATERYAL VE METOD	60
5.1. SİSTEMİN GENEL YAPISI.....	60
5.2. ANKSİYÖZ KONUŞMA VERİSİNİN EDİNİMİ.....	61
5.3. KODLAMA VE YAZILIM GELİŞTİRME.....	63
5.3.1. Pitch Tespiti.....	64
5.3.2. Formant Frekanslarının Hesaplanması Ve LPCC.....	66
5.3.3. Sinyal Gürültü Oranının Hesaplanması	68
5.3.4. Jitter Ve Shimmer Değerlerinin Hesaplanması	68
5.3.5. MFCC	70
5.3.6. Zero-Crossing Rate.....	71
5.3.8. Enerji Ve Konuşma İle İlgili Öznitellikler	72
5.3.9. Dalgacık Katsayıları	72
5.3.10. Temel Bileşen Analizi (PCA).....	73
5.3.11. Lineer Diskriminant Analizi.....	74
5.3.12. Fisher Oranı	74
5.3.13. Çapraz Doğrulama.....	75
5.4. İSTATİSTİKSEL ANALİZ	76
5.5. SINIFLANDIRMA	80
5.5.1. Destek Vektör Makineleri.....	80
5.5.2. k-En Yakın Komşu	84
5.5.3. Çok Katmanlı Algılayıcı.....	85
BÖLÜM 6	87
BULGULAR.....	87
6.1. SPAC ÖZELLİK KÜMESİNİN BAŞARI ANALİZİ.....	87
6.1.1. Test Verisi.....	87
6.1.2. Test Yöntemleri	88
6.1.2.1. MSE Ve MPE Sonuçları	88
6.1.2.2. Sınıflandırma Sonuçları	89
6.2. OLGULARIN SOSYODEMOGRAFİK ÖZELLİKLERİ	91
6.3. AKUSTİK PARAMETRELERİN ANKSİYETE İLE İLİŞKİSİ.....	91

	<u>Sayfa</u>
6.3.1. Benzerlik Gösteren Parametreler	95
6.3.2. Duygular Üzerinde Etkin Olan Akustik Parametreler	96
6.3.3. Duygu Durumları Üzerinde Etkin Olan Parametrelerin Değişim Yönü Ve Şiddeti	100
6.3.4. Hasta Ve Sağlam Olguları Ayırt Etmede Etkin Olan Parametreler.....	103
6.3.5. Hasta Ve Kontrol Gruplarının Duygularını Sese Yansıtmadaki Farklılıkları	106
BÖLÜM 7	109
SONUÇLAR VE ÖNERİLER	109
KAYNAKLAR	113
EK AÇIKLAMALAR A. DFT UYGULAMASI MATLAB KODLARI.....	127
EK AÇIKLAMALAR B. ETİK KURUL İZİNİ	129
EK AÇIKLAMALAR C. GÖNÜLLÜ OLUR FORMU	131
EK AÇIKLAMALAR D. SOSYODEMOGRAFİK VERİ FORMU.....	134
EK AÇIKLAMALAR E. KULLANILAN METİNLER	136
EK AÇIKLAMALAR F. ANKSİYETE ŞİDDETİNİ BELİRLEME ÖLÇEKLERİ	139
ÖZGEÇMİŞ	142

ŞEKİLLER DİZİNİ

	<u>Sayfa</u>
Şekil 1.1. Tez içeriği.....	4
Şekil 2.1. Ses yolunun yapısı.....	7
Şekil 2.2. Konuşma oluşum süreci	8
Şekil 2.3. Sürekli zamanlı ve ayrık zamanlı sinyal örneği	9
Şekil 2.4. DSP uygulama alanları.....	10
Şekil 2.5. DSP sisteminin blok şeması	11
Şekil 2.6. ADC'nin temel bölümleri.....	11
Şekil 2.7. Sıfır dereceli tutmalı basamak yaklaşımlı DAC.....	12
Şekil 2.8. 100 Hz ve 40 Hz örnekleme frekansı ile bir sinüzoidal dalganın sürekli ve ayrık örnekleri.....	14
Şekil 2.9. Bir sinyalin fourier dönüşümü frekans domenine taşınması ile sinüzoidal bileşenleri.....	14
Şekil 2.10. 4 kategoride fourier dönüşümü sonuçları.....	16
Şekil 2.11. Tek bir frekansa sahip sinüzoidal bir sinyal.....	17
Şekil 2.12. Çeşitli pencereleme fonksiyonlarının zaman ve frekans boyutundaki tepkileri.....	18
Şekil 2.13. Ayrık dalgacık dönüşümünde alt bantlara ayırışım; $g[n]$ yüksek geçiren filtre, $h[n]$ alçak geçiren filtre	20
Şekil 2.14. Örnek bir konuşma sinyali üzerinde gerçekleştirilen tekrar örnekleme, pencereleme ve alçak geçiren FIR filtreleme sonucu sinyal üzerinde meydana değişiklikler	22
Şekil 2.15. Duyguların boyutsal gösterimi	24
Şekil 3.1. Örnek bir konuşma sinyali	30
Şekil 3.2. Akustik analiz tabanlı duygu tanıma akış diyagramı	31
Şekil 3.3. Ön işleme sürecine ait akış diyagramı.....	34
Şekil 3.4. Stereo ses örneği ve stereo sesin monoya dönüştürülmüş hali.....	35
Şekil 3.5. Örnekleme frekansı değiştirildiğinde sinyaldeki değişim.....	37
Şekil 3.6. DC offset içeren ve içermeyen sinyal örneği	37
Şekil 3.7. Konuşma sinyali için çerçeveleme örneği.....	40
Şekil 3.8. Örnek bir ses sinyalinin orijinal hali ve pencerelenmiş hali	40
Şekil 3.9. Örnek bir konuşma için F0 grafiği	45

Şekil 3.10. Örnek bir konuşma için formant frekansları	46
Şekil 3.11. Örnek bir konuşma için sesli ve sessiz bölümlerin gösterimi	48
Şekil 3.12. Örnek bir konuşma için enerji dağılımı.....	49
Şekil 4.1. SPAC modüler yapısı	51
Şekil 4.2. SPAC ile örnek bir ses dosyasının işlenmesi	52
Şekil 4.3. Modül 1 grafiksel arayüzü.....	53
Şekil 4.4. Örnekleme değerini düşürme sonrası sinyaldeki değişim (mavi: orijinal sinyal, siyah: ön işleme sonrası).....	54
Şekil 4.5. Pre-emphasis sonrası sinyaldeki değişim (mavi: orijinal sinyal, kırmızı: ön işleme sonrası)	55
Şekil 4.6. Gürültü giderme sonrası sinyaldeki değişim (mavi: orijinal sinyal, cyan: ön işleme sonrası)	55
Şekil 4.7. DC-offset kaldırma sonrası sinyaldeki değişim (mavi: orijinal sinyal, yeşil: ön işleme sonrası)	55
Şekil 4.8. Modül 5'e ait ekran görüntüsü	57
Şekil 4.9. Modül 6'ya ait ekran görüntüsü	58
Şekil 5.1. Sistemin genel yapısı.....	60
Şekil 5.2. Otokorelasyon tabanlı pitch tespiti akış diyagramı	65
Şekil 5.3. Kepstrum tabanlı pitch tespiti akış diyagramı.....	66
Şekil 5.4. LPC blok diyagramı	67
Şekil 5.5. Formant frekanslarının tespiti	68
Şekil 5.6. Sinyal gürültü oranı ve enerjisinin hesaplanması için akış diyagramı ...	68
Şekil 5.7. Jitter değerlerinin hesaplanması için akış diyagramı	69
Şekil 5.8. Shimmer değerlerinin hesaplanması için akış diyagramı.....	69
Şekil 5.9. MFCC akış diyagramı	71
Şekil 5.10. 5 seviye dalgacık dönüşümü	73
Şekil 5.11. 5-fold çapraz doğrulama için verinin dağılımı.....	76
Şekil 5.12. İki sınıflı bir problem için oluşturulabilecek hiper-düzlemler, optimum hiper-düzlem ve destek vektörleri.....	81
Şekil 5.13. Doğrusal olarak ayırlamayan sınıflar ve doğrusal ayırlamayan sınıflar için hiper düzlemin belirlenmesi	82
Şekil 5.14. Çekirdek fonksiyonu ile verinin daha yüksek bir boyuta dönüştürülmesi	83
Şekil 5.15. EMO-DB verisetinde sınıflandırma için kullanılan MLP modeli.....	86
Şekil 6.1. SPAC, Praat ve OpenSMILE öznetelik kümelerinde SVM sınıf bazlı başarı	90

	<u>Sayfa</u>
Şekil 6.2. Deneysel çalışma akış diyagramı.....	93
Şekil 6.3. Öznitelik çıkartımı aşamasında kullanılan ön tanımlı ayarlar.	94
Şekil 6.4. Anksiyetenin etkilediği en önemli 10 parametrenin içerdiği değerlere göre grup bazlı dağılım grafiği.	102
Şekil Ek A.1. DFT uygulaması MATLAB kodları	128
Şekil Ek B.1. Etik kurul izni.	130
Şekil Ek E.1. Endişe duygu durumu için kullanılan metin	137
Şekil Ek E.2. Nötr duygu durumu için kullanılan metin.....	138

ÇİZELGELER DİZİNİ

	<u>Sayfa</u>
Çizelge 2.1. Çeşitli uygulamalar için örnekleme oranları	13
Çizelge 2.2. Pencereleme fonksiyonları	18
Çizelge 2.3. Öznitelik çıkartımı araçları	27
Çizelge 2.4. Anksiyetenin akustik parametreler üzerindeki etkisini inceleyen çalışmaların özeti	28
Çizelge 2.5. Akustik parametrelerin anksiyete ve korku üzerindeki etkileri	29
Çizelge 3.1. Duygusal konuşma veritabanları	32
Çizelge 4.1. SPAC ile elde edilebilecek öznitelikler.....	56
Çizelge 5.1. Korelasyon aralıklarına göre ilişki düzeyleri	79
Çizelge 5.2. Uzaklık ölçütleri.....	85
Çizelge 6.1. MSE ve MPE sonuçları.....	88
Çizelge 6.2. SPAC öznitelik kümelerinin sınıflandırma başarısı.....	89
Çizelge 6.3. SPAC, Praat, OpenSMILE öznitelik kümeleri üzerinde genel başarı.	90
Çizelge 6.4. Olguların sosyodemografik verileri	91
Çizelge 6.5. Olguların sosyodemografik verilerinin karşılaştırması.....	92
Çizelge 6.6. Kullanılan akustik parametreler	94
Çizelge 6.7. Korelasyon analizi sonrası benzer olduğu kabul edilen parametre.	96
Çizelge 6.8. Friedman testi sonuçları	97
Çizelge 6.9. Wilcoxon testi sonuçları.....	98
Çizelge 6.10. Parametrelerin duygulara göre değişim yönü ve şiddeti.....	100
Çizelge 6.11. Anksiyetenin etkilediği parametrelerin etkilenme derecesine göre sıralaması.....	101
Çizelge 6.12. Hasta ve sağlam olguları ayırt etmede etkin parametreler.....	103
Çizelge 6.13. Çalışma gruplarının duygularını sese yansıtmadaki farklılıkları	107
Çizelge Ek C.1. Asgari bilgilendirilmiş gönüllü olur formu örneği	132
Çizelge Ek D.1. Sosyodemografik veri formu	135
Çizelge Ek F.1. Hamilton anksiyete değerlendirme ölçeği.....	140
Çizelge Ek F.2. Beck anksiyete envanteri.....	141

SİMGELER VE KISALTMALAR DİZİNİ

SİMGELER

Amp	: genlik
$A(z)$: LPC analiz süzgeci
b_i	: LPC katsayıları
E	: enerji
F	: frekans
F_{max}	: maksimum frekans
f_{MEL}	: mel frekans ölçeği
F_s	: örnekleme frekansı
F0	: fundamental frequency (temel frekans)
F1	: birinci formant frekansı
F2	: ikinci formant frekansı
F3	: üçüncü formant frekansı
FO	: fisher oranı
$g[.]$: yüksek geçiren filtreye ait çıkış
$h[.]$: alçak geçiren filtreye ait çıkış
J_{mutlak}	: mutlak jitter
$J\%$: yüzde jitter
L_p	: p. norm seviyesi
N	: örnek sayısı
p	: LPC analizinin derecesi
$r(\tau)$: otokorelasyon fonksiyonu
S_{db}	: mutlak shimmer
$S\%$: yüzde shimmer
t	: örnekleme aralığı
T	: periyot
T0	: temel periyot

- $x(n)$: ayrık zamanlı sinyal
 $X(F)$: sinyalin fourier dönüşümü
 $x[n]$: sayısal sinyal
 $X(w)$: sinyalin ayrık zamanlı fourier dönüşümü
 $x_0(t)$: analog sinyal
 X_{mono} : mono ses örneği
 $x_q(n)$: nicemlenmiş sinyal
 X_{stereo} : stereo ses örneği
 $w(n)$: pencereleme fonksiyonu
 $y[n]$: filtreleme sonrası çıkış sinyali
 μ : ortalama değer
 σ : standart sapma

KISALTMALAR

ADC	: Analog to Digital Converter (Analog Sayısal Çevirici)
ANN	: Artificial Neural Network (Yapay Sinir Ağları)
AY	: Anksiyöz Yaşantı
DAC	: Digital to Analog Converter (Sayısal Analog Çevirici)
DCT	: Discrete Cosine Transform (Ayrık Kosinüs Dönüşümü)
DFT	: Discrete Fourier Transform (Ayrık Fourier Dönüşümü)
DSM-5	: Diagnostic and Statistical Manual of Mental Disorders, Fifth Edition (Ruhsal Bozuklukların Tanısal ve İstatistiksel El Kitabı, Beşinci Versiyon)
DSP	: Digital Signal Processing (Sayısal Sinyal İşleme)
DTFT	: Discrete Time Fourier Transform (Ayrık Zamanlı Fourier Dönüşümü)
EMO	: Endişeli Metin Okuma
EMO-DB	: Berlin Database of Emotional Speech
FFT	: Fast Fourier Transform (Hızlı Fourier Dönüşümü)
FS	: Forward Selection (İleri Seçim)
GCI	: Glottal Closure Instant (Gırtlığın Kapanma Anı)
GMM	: Gaussian Mixture Model (Gauss Karışım Modeli)
GUI	: Graphical User Interface (Kullanıcı Grafik Arayüzü)
HMM	: Hidden Markov Model (Saklı Markov Modeli)
HNR	: Harmonics to Noise Ratio (Harmonik Gürültü Oranı)
HTK	: Hidden Markov Toolkit
IDFT	: Inverse Discrete Fourier Transform (Ters Ayrık Fourier Dönüşümü)
IDTFT	: Inverse Discrete Time Fourier Transform (Ters Ayrık Zamanlı Fourier Dönüşümü)
k-NN	: k-Nearest Neighbor (k-En Yakın Komşu)
LDA	: Linear Discriminant Analysis (Lineer Diskriminant Analizi)
LFPC	: Log Frequency Power Coefficient (Log Frekans Güç Katsayıları)
LPC	: Linear Predictive Coding (Lineer Kestirimci Kodlama)
LPCC	: Linear Predictive Cepstral Coefficient (Doğrusal Öngörü Kepstral Katsayıları)

MFCC	: Mel Frequency Cepstral Coefficients (Mel-Frekans Kepstral Katsayıları)
MLP	: Multilayer Perceptron (Çok Katmanlı Algılayıcı)
MPE	: Mean Percentage Error (Yüzde Ortalama Hata)
MSE	: Mean Square Error (Karesel Ortalama Hata)
NHR	: Noise to Harmonics Ratio (Gürültü Harmonik Oranı)
NMO	: Nötr Metin Okuma
PCA	: Principal Component Analysis (Temel Bileşen Analizi)
PCM	: Pulse Code Modulation (Darbe Kod Modülasyonu)
RASTA-PLP	: Relative Spectral Transform - Perceptual Linear Prediction (Bağıl Spektral Dönüşüm-Algısal Doğrusal Tahmin)
SAVEE	: Surrey Audio-Visual Expressed Emotion
SNR	: Signal to Noise Ratio (Sinyal Gürültü Oranı)
SPAC	: Speech Acoustic
SPL	: Sound Pressure Level (Ses Basınç Seviyesi)
STFT	: Short Time Fourier Transform (Kısa Zamanlı Fourier Dönüşümü)
SVM	: Support Vector Machine (Destek Vektör Makineleri)
TEO	: Teager Energy Operator (Teager Enerji Operatörü)
ZCR	: Zero-Crossing Rate

BÖLÜM 1

GİRİŞ

Ses oluşumu fizyolojik bir yapı olduğu için insanların içerisinde buldukları duygusal durum bu yapıda değişikliğe neden olmakta doğal olarak da duygular sese yansımaktadır. İnsanoğlunu diğer canlılardan ayıran en önemli özelliklerden birisi sesin konuşmaya dönüştürülerek iletişim kurabilmesidir [1]. Ses, bir iletişim aracı olmasının dışında kişinin kimliği, ruhsal durumu ve fiziksel sağlığı gibi konularda da gösterge oluşturmaktadır. İnsan sesleri ve konuşmadan kişi tanıma, duygu tespiti, cümle ve kelime tanıma üzerine yapılmış çeşitli çalışmalar mevcuttur. İnsan seslerinden psikolojik tanı, duygu durumu tespiti ve kişi tanımda nesnel ve öznel yöntemler kullanılmaktadır. Algısal değerlendirme öznel bir değerlendirme yöntemidir ve konuşmaların uzman kişiler tarafından dinlenip yorumlanması ile gerçekleştirilir. Bu nedenle sonuçlar uzmanın tecrübesine göre değişiklik göstermektedir. Bu sorunun üstesinden gelmek için nesnel değerlendirme yöntemleri kullanılmaktadır [2].

Akustik analiz, sesteki çeşitli parametreler elde edilerek seslerin nesnel olarak değerlendirmesi için kullanılır. Bu amaçla kullanılacak hazır yazılımlar mevcuttur [3]. Akustik analiz ile elde edilen akustik parametreler konuşmadan duygu tanıma çalışmalarında yoğun olarak kullanılmaktadır. İnsanların konuşmalarından duygusal durumların tespiti insan makine etkileşimi çalışmalarında önemli bir yer almaktadır. Duygusal durumların tespit edilebilmesi için konuşma, sinyal işleme yöntemleri ile sayısallaştırılmakta ve akustik analiz ile konuşmaya ait çeşitli akustik parametreler tespit edilmektedir. Ses sinyali ve ses yolunun özellikleri kişiye, yaşa, cinsiyete, ses yolu uzunluğuna, boya, ağırlığa ve içinde bulunan duygusal duruma göre farklılık göstermektedir. Bundan dolayı bir insanın içinde bulunduğu duygusal durum ses analizi yöntemleri ile elde edilecek öznitelik vektörlerinden hareketle belirlenebilir. Konuşmadan duygu tespiti yeni bir çalışma alanı olmayıp ilk olarak

1980'lerin ortalarında bazı akustik parametrelerin istatistiksel özellikleri kullanılarak gerçekleştirilmiştir. Sonraki yıllarda bilgisayar mimarisinin gelişimi ile birlikte daha karmaşık duygu tanıma algoritmaları kullanılmaya başlamıştır [4]. Akustik özniteliklerle birlikte dil ve söylem bilgileri de duygu tespiti için kullanılmıştır [5].

Akustik analiz veya başka bir ifadeyle seslerden öznitelik çıkartımı için hazır yazılımlar veya araştırmacılar tarafından geliştirilen kodlar kullanılabilir. Mevcut yazılımlar belirli çalışma alanı için geliştirilmiş olup birçoğu kullanıcı dostu bir arayüze sahip değildir. Ayrıca, mevcut yazılımların büyük çoğunluğu sadece öznitelik çıkartımı gerçekleştirmekte ve sınıflandırma yapamamaktadır. Mevcut yazılımlar arasındaki bir diğer farklılık ise her öznitelik her yazılım ile elde edilememektedir. Örneğin, delta ve delta-delta katsayıları her araç ile elde edilememektedir. Ayrıca, mevcut yazılımlar dalgacık katsayılarını da elde edilememektedir.

Anksiyete sıkıntı, endişe ve/veya korkuyu içeren nahoş bir duygudur. Tehdit veya tehlike durumlarında ortaya çıkmakta ve kişiyi tehlike anında eyleme geçmeye hazırlamaktadır. Normal anksiyete mantıklı, gerekli, hatta yaşamsaldır. Olası tehlike uyarılarına tüm canlılar kalıtsal ve doğumsal bir düzenek olan anksiyete ile yanıt vermektedirler. Ancak, beklenenden fazla ya da şiddetli olduğunda, gereğinden daha uzun sürdüğünde ya da tekrarladığında, zihni sürekli meşgul ettiğinde, kısaca kişinin günlük işlevlerini bozduğunda patolojik anksiyeteden söz edilmektedir.

Gerek duygu tanıma gerekse de duyguların sese yansımaları üzerine yapılan çalışmalarda en büyük sınırlılık duyguların doğru tetiklendiği başka ifadeyle sese yansıtıldığı verilerin edinilmesidir. Bu sorunun üstesinden gelmek için literatürde hazır veri setleri bulunmaktadır. Ancak, bu veri setlerinin önemli bir çoğunluğu duygunun simule edilmesi yani aktörler tarafından ifade edilmesi ile elde edilmiş olup gerçek duygular değildir.

Sesin duygusal durumu yansıtma derecesi ve şekli konusunda yapılan araştırmalar oldukça sınırlıdır. Bu sınırlılık anksiyete ve ses ilişkisinde daha belirgin hale gelmektedir. Ayrıca, yapılan çalışmaların sonuçları arasında çelişkiler mevcuttur.

Duygulanımın bir parçası olan sesin nitelik ve niceliklerinin psikiyatrik hastalıklar açısından incelenmesi tanı ve tedavi süreci açısından katkılar sağlayacaktır.

Bu çalışmanın amacı, seslerden akustik parametrelerin tespit edilmesi ve anksiyete bozukluğu hastalarında sesin bu parametreler üzerindeki etkisinin incelenmesidir. Seslerden akustik parametrelerin tespit edilmesi için SPeech ACoustic (SPAC) isimli bir yazılım geliştirilmiştir. Anksiyetenin akustik parametreler üzerindeki etkisini incelemek için Gülhane Eğitim ve Araştırma Hastanesi Psikiyatri Anabilim Dalı Başkanlığında yatan ve ayaktan hastalardan 01.09.2015 ile 31.06.2016 tarihleri arasında toplanan gerçek ses kayıtları kullanılmıştır.

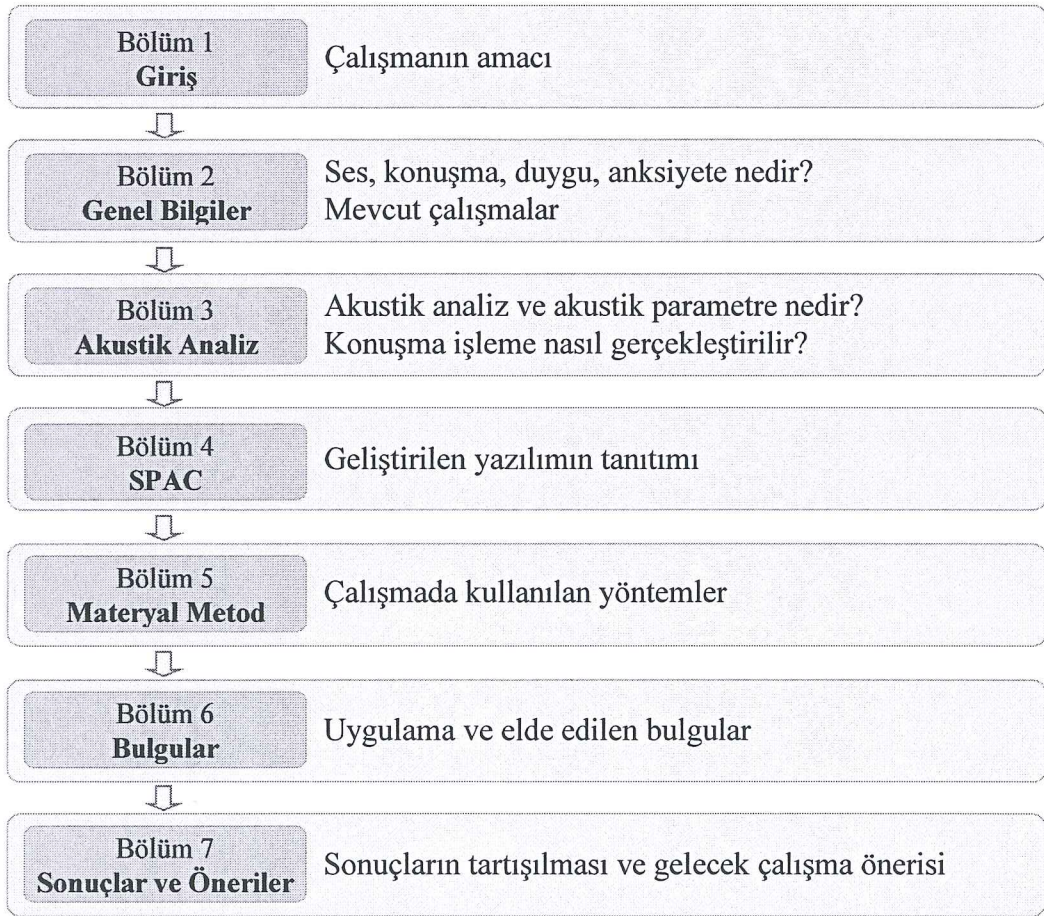
SPAC, modüler yapısı, geniş öznitelik kümesi, kullanıcı dostu arayüzü, ön işleme, son işleme ve sınıflandırma modülleri ile mevcut yazılımlardaki eksiklikleri gidermektedir. Ayrıca, dalgacık katsayıları, delta ve delta-delta öznitelik kümelerini de içermektedir. SPAC ile hangi akustik parametrelerin tespit edileceğini belirlerken hazır yazılımlar ve literatürde yer alan çalışmalar incelenmiştir. Bu inceleme sonucunda temel frekans, formant frekansları ve ses kalitesi parametrelerinin çok sayıda çalışmada kullanıldığı ve mevcut yazılımların tamamı ile ilde edilebildiği görülmüştür. Literatürde dikkat çeken diğer önemli bir nokta ise son yıllarda yapılan çalışmalarda fourier dönüşümüne göre üstünlükleri olan dalgacık dönüşümünün yoğun olarak kullanılmasıdır. Ayrıca, insan işitme sistemini modelleyen Mel-Frekans Kepstral Katsayıları (MFCC) ve Doğrusal Öngörü Kepstral Katsayıları (LPCC) öznitelik kümeleri de duygu tanımada yoğun olarak kullanılmaktadır. Bu durumlar göz önünde bulundurularak temel frekans, formant frekansları, ses kalitesi parametreleri, Zero-Crossing Rate (ZCR), enerji, konuşma yapısı ve dalgacık katsayıları SPAC öznitelik kümesine dahil edilmiştir.

Veri toplama çalışması kesitsel nitelikte bir olgu-kontrol çalışmasıdır. Çalışmaya 23'ü anksiyete bozukluğu (Hasta grup) ve 20'si sağlıklı (kontrol grubu) toplam 43 olgu alınmıştır. Anksiyete bozukluğu olgularının 2'si panik bozukluk ve 21'i yaygın anksiyete bozukluğundan oluşmaktadır. Olguların 26'sı bayan, 17 'si erkektir. Olguların her birinden ilki nötr diğer ikisi anksiyöz durumda olmak üzere üç adet ses kaydı alınmıştır. Ses kayıtları toplanırken nötr durum ve anksiyöz durumun birisi için

hazır metin okutulmuştur. Diğer anksiyöz durum için ise olgular son yaşadığı anksiyete atağını 3 dakika anlatmıştır.

Toplanan ses kayıtlarının akustik analizi SPAC ile gerçekleştirilmiştir. Akustik analiz sürecinde SPAC üzerinde yer alan öznelik kümelerinin ortalama değeri ve bazı parametrelerin standart sapma değerleri çalışmaya dahil edilmiştir. Böylece, her bir olgunun ses kaydına ait 122 akustik parametre SPAC ile elde edilmiştir. Anksiyöz durumun elde edilen akustik parametreler üzerindeki etkileri istatistiksel olarak araştırılmıştır.

Hazırlanan bu çalışma, yazılım geliştirme ve deneysel çalışma olmak üzere iki kısımdan oluşmuştur ve çalışma planı aşağıda verilmiştir.



Şekil 1.1. Tez içeriği.

Bölüm 2’de, ses ve sesin oluşumu, sinyal işleme, duygu/konuşma ve anksiyete kavramlarına ait kuramsal bilgiler verilmiştir.

Bölüm 3’de, akustik analiz detaylı bir şekilde açıklanmış, konuşma işleme ve akustik parametreler bu bölümde detaylandırılmıştır.

Bölüm 4’de, çalışma kapsamında hazırlanan SPAC yazılımı tanıtılmış ve yazılımın yetenekleri açıklanmıştır.

Bölüm 5’de, çalışma tamamlanana kadar geçen süreçte kullanılan tüm materyaller ve yöntemler açıklanmıştır. Ayrıca, çalışmada kullanılan akustik parametrelerin tespit edilmesinde kullanılan formüller ve akış diyagramları verilmiştir.

Bölüm 6’da, çalışmadan elde edilen bulgular verilmiştir. SPAC yazılımı ile elde edilen öznitelik kümelerinin geçerliliği test edilmiştir. Gerçek veriler ile gerçekleştirilen istatistiksel analiz ve akustik analiz sonuçları verilmiştir.

Bölüm 7’de, çalışmanın bulguları deneysel çalışmanın amacına uygun bir biçimde yorumlanarak tartışılmıştır.

Bu tez kapsamında literatüre kazandırılan katkılar aşağıda belirtilmiştir.

1. Anksiyöz durum ile ilgili gerçek ses kayıtları içeren yeni bir veri kümesi oluşturulmuştur.
2. Akustik analiz için kolay kullanıma ve kullanıcı dostu bir arayüze sahip, modüler yapıda, sonuçların grafiksel gösterimine olanak sunan, sınıflandırma modülünü içeren ve gerek kendi ürettiği gerekse de farklı araçlar üzerinden elde edilen veriler ile çalışmaya olanak sağlayan yeni bir yazılım geliştirilmiştir.
3. Duygunun tetiklenmesi için metin okutmanın da anksiyeteyi ortaya çıkartmada etkili olduğu tespit edilmiştir.

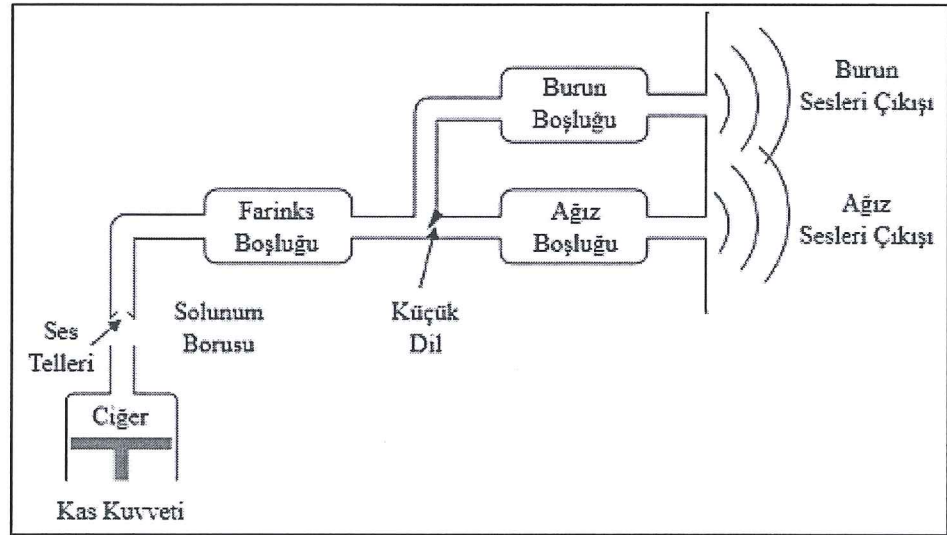
4. Kullanılan 122 parametreden 42'si anksiyöz durum ile ilişkilidir. Anksiyetenin F0, F1, jitter, shimmer gibi genel parametrelere ek olarak MFCC, LPCC ve dalgacık katsayılarını da önemli oranda etkilediği tespit edilmiştir.
5. Anksiyete F0 değerini artırmakta doğal olarak da sesin üretiminde subglottal basıncı artırmaktadır.
6. Anksiyöz durum formant frekanslarına bağlı olarak konuşmada sesli harflerin dil vuruşu azalmaktadır.
7. Anksiyetenin sese yansıdığı ve insanların konuşmayı dinlerken anksiyeteyi algılayabilecekleri tespit edilmiştir.
8. Anksiyöz durum jitter ve shimmer parametrelerinde artış oluşturmaktadır.
9. Anksiyete durumunda ses tellerinin ve ses şiddetinin düzensiz olduğu tespit edilmiştir.
10. Anksiyöz durumda konuşma enerjisinde değişim gözlemlenmemiştir.
11. Konuşmadaki duraksamalar anksiyöz durum hakkında bilgi vermemektedir.

BÖLÜM 2

GENEL BİLGİLER

2.1. SES VE SESİN OLUŞUMU

Ses, anatomik yapıların istemli hareketleri sonucunda oluşan, akustik bir basınç dalgasıdır. Bu sistemin ana bölümleri ciğerler, nefes borusu, gırtlak, boğaz, ağız boşluğu ve burun boşluğudur. Teknik terim olarak boğaz ve ağız boşluğu ses yolu olarak tanımlanır. Dolayısıyla ses yolu, gırtlak çıkışından başlayıp, dudaklarda sona erer. Burun yolu ise damaktan başlar burun deliklerinde sona erer. Ses üretimi için kritik olan anatomik yapılar, ses telleri, damak, dil, dişler ve dudaklardır. Ses yolunun yapısı Şekil 2.1’de görülmektedir [6].

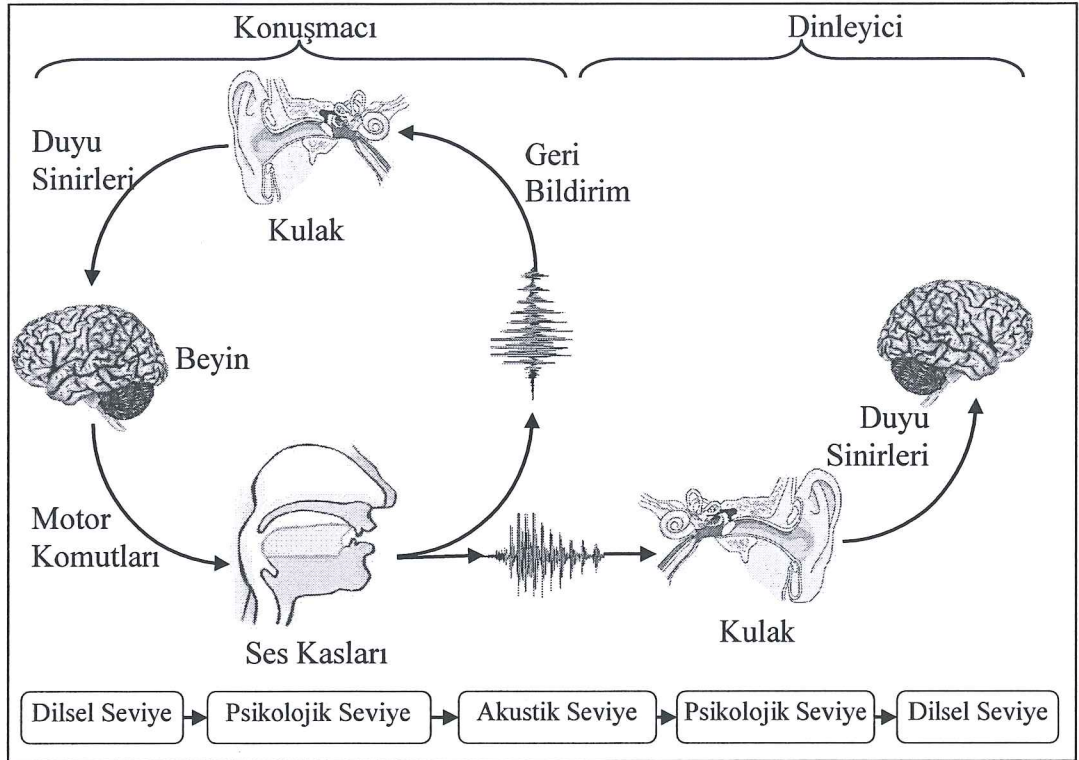


Şekil 2.1. Ses yolunun yapısı [6].

İnsan sesi, ses kıvrımlarının titreşimi ile ortaya çıkar. Konuşma fonasyon, rezonans ve artikülasyon olmak üzere üç ögeden oluşur. Fonasyon, ses kıvrımlarının titreşimine bağlı olarak sesin oluşumudur. Rezonans, oluşan sesin uzatılması,

artırılması ve filtrelenmesidir. Artikülasyon ise ses yolunun değişken hareketleri sonucu sesin konuşmaya dönüşmedir [7].

Konuşma ses organlarında motor komutlarının oluşturulması, artikülasyon ve akciğerler tarafından hava ile ses yolunun uyarılması olmak üzere Şekil 2.2’de gösterildiği gibi üç aşamada gerçekleşir [8].



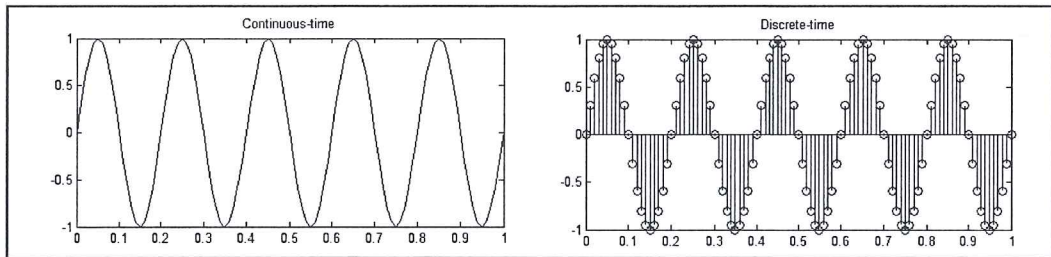
Şekil 2.2. Konuşma oluşum süreci [9].

Ses üzerine kayıtlı ilk çalışmalar milattan önce 5.yüzyıla kadar uzanmaktadır. Hipokrat akciğer, trakea ve dilin fonasyon için önemini belirtmiştir. Aristo sesin duygu ile olan ilişkisini tanımlamıştır. Cladius Galen konuşma ile sesi birbirinden ayırmıştır. Teknolojinin gelişmesiyle birlikte ses oluşumunda farklı anatomik yapılar ve fizyolojik sistemlerin yer aldığı kompleks bir sistem olduğu anlaşılmıştır [7]. Ses bir iletişim aracı olmasının dışında kişinin kimliği, ruhsal durumu ve fiziksel sağlığı gibi konularda bilgi sağlamaktadır [10].

2.2. SİNYAL İŞLEME

Sinyal, zaman gibi belirli bir veya birkaç bağımsız değişkenin değerine göre değişiklik gösteren fiziksel bir niceliktir [11]. Örneğin, bir sinüzoidal sinyal zaman değişkenine göre fonksiyonel olarak ifade edilebilirken bir konuşma sinyali fonksiyonel bir ifade ile tanımlanamaz ama farklı genlik ve farklı frekanslarda birden çok sinüzoidal frekansın toplamı olarak ifade edilebilir. Sinyaller üç kategoriye ayrılır: sürekli zamanlı sinyaller (analog sinyal), ayrık zamanlı sinyaller ve sayısal sinyaller.

Günlük hayatta sürekli karşılaştığımız sinyaller çoğunlukla analog sinyallerdir. Bu sinyaller zamanda sürekli olup sonsuz sayıda genlik değerine sahiptir. Ayrık zamanlı sinyaller zamanın belirli örneklerinde tanımlıdır. Sayısal sinyaller zaman ve genliğin her ikisinde de ayrık değerlere sahiptir, yani bilgisayarlar veya mikroişlemciler ile işlenebilir [12]. Sürekli zamanlı sinyaller zaman eksenini boyunca her zaman değer aldığından, sinyal üzerinde belirli zaman aralıklarında örnekleme yapılırsa ayrık zamanlı sinyal elde edilir. Şekil 2.3'de sürekli ve ayrık zamanlı sinyal örnekleri verilmiştir.

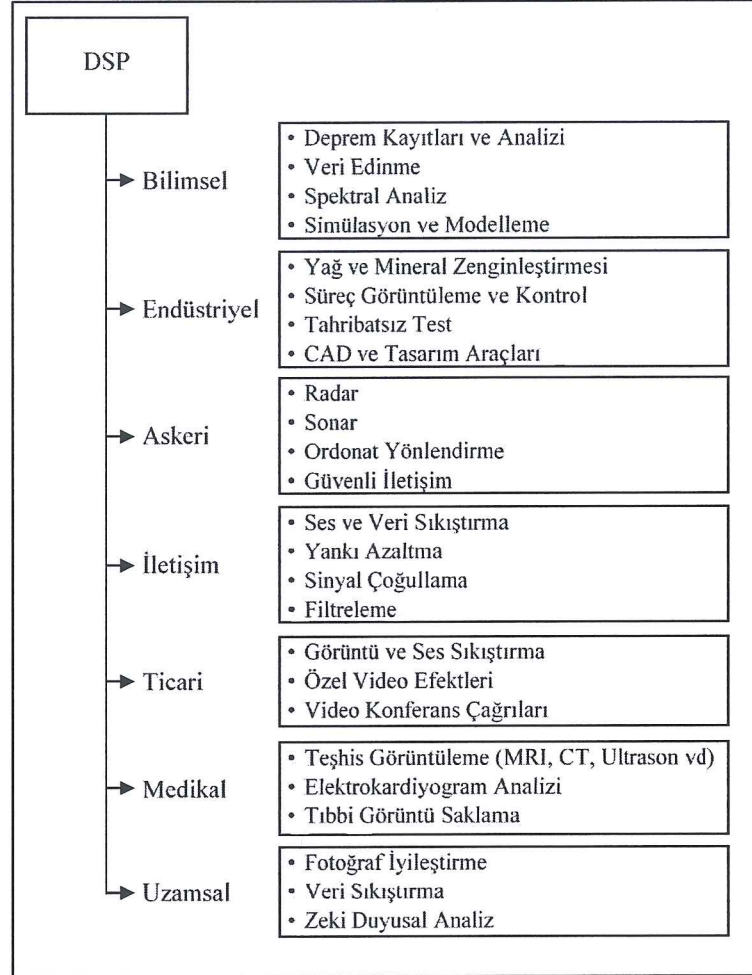


Şekil 2.3. Sürekli zamanlı ve ayrık zamanlı sinyal örneği.

Analog sinyaller, özelliklerini değiştirmek veya bazı istenen bilgileri ayıklamak amacıyla frekans çoğaltıcı veya uygun analog sistemler (filtreler veya frekans analizi gibi) ile doğrudan işlenebilir [11].

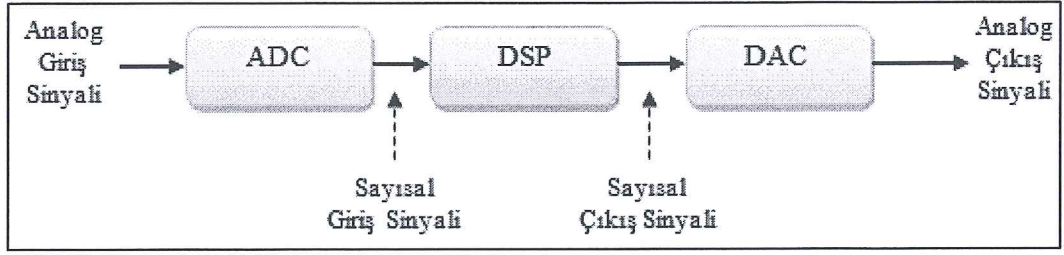
Sayısal Sinyal İşleme (DSP) bilim ve mühendislik açısından günümüzde güçlü bir teknolojidir ve sinyalin sayısal gösterimi ile ilgilidir. DSP görüntü işleme, konuşma

tanıma ve oluşturma, veri sıkıştırma ve taşıma gibi çeşitli alanlarda kullanılmaktadır. Şekil 2.4’de DSP kullanım alanlarının bir kısmı gösterilmiştir [13].



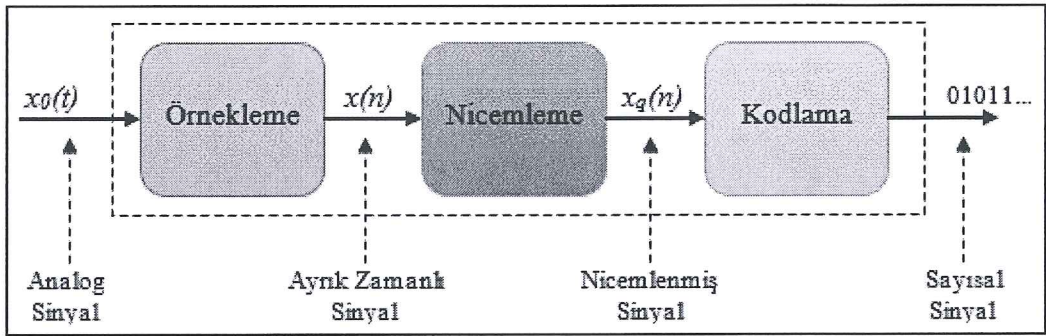
Şekil 2.4. DSP uygulama alanları [13].

DSP analog sinyal işleme sürecine Şekil 2.5’de gösterildiği gibi bir alternatif sunar. DSP, analog sinyal ve sayısal işleme arasında bir ara yüze ihtiyaç duyar ve bu arayüz Analog Sayısal Çevirici (ADC) olarak adlandırılır. Sayısal sinyal ve analog sinyal arasındaki dönüşüm için de Sayısal Analog Çevirici (DAC) kullanılır [11].



Şekil 2.5. DSP sisteminin blok şeması [11].

Konuşma, biyolojik sinyaller, sismik sinyaller, radar sinyalleri, ses ve video sinyalleri gibi birçok sinyal analog sinyaldir. Analog sinyal işlemenin ilk adımı sinyalin sayısala dönüştürülmesidir ve bu amaçla ADC kullanılır. Kavramsal olarak ADC işlemi Şekil 2.6’da gösterildiği gibi üç adımda gerçekleştirilir [11].



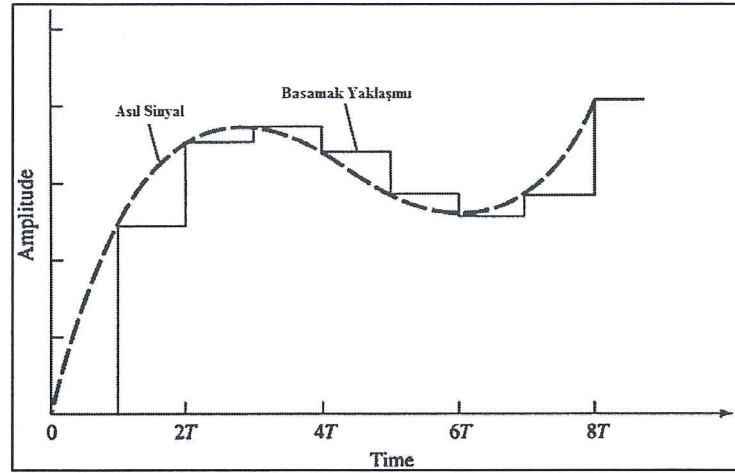
Şekil 2.6. ADC’nin temel bölümleri [11].

Örnekleme, sürekli zamanlı sinyalin ana sinyalden ayrık zamanlarda örnekler alınarak ayrık zamanlı sinyale dönüştürülmesidir. Çıkış $x_0(t) \Rightarrow x[n]$ olacaktır. t örnekleme aralığıdır [11].

Nicemleme, ayrık zamanlı sürekli değerli sinyalin ayrık zamanlı ayrık değerli (sayısal) sinyale dönüştürülmesidir. Giriş sinyalinin sürekliliğini, sınırlı sayıdaki tepki değerine dönüştürmektedir. Nicemlemeden dolayı çıkış, hatalı olacağından sinyalin üzerinde bir gürültü taşıdığı düşünülür ve bu gürültü nicemleme gürültüsü olarak adlandırılır [11,14].

Kodlama, analog sinyalin bit dizisine dönüşümüdür. Her ayrık $x_q(n)$ değeri, ikilik tabanda değerlerle temsil edilir [11,14].

Birçok uygulamada işlenmiş sinyalin tekrar analog sinyale dönüştürülmesi gerekir. Bu amaçla DAC sayısal sinyalin noktalarını birleştirerek bir çeşit ara değerlendirme (enterpolasyon) yapar. Şekil 2.7’de sıfır dereceli basamak yaklaşımı adı verilen basit bir DAC verilmiştir [11].



Şekil 2.7. Sıfır dereceli tutmalı basamak yaklaşımli DAC [11].

2.2.1. Örnekleme Teoremi

Analog bir sinyal verildiği zaman örnekleme periyodu (T) veya örnekleme hızının (F_s) neye göre belirleneceği önemlidir. Bu değerler belirlenirken sinyal karakteristiği hakkında bilgi sahibi olmak gereklidir. Örneğin, konuşma sinyalinin temel frekans bileşenlerinin 3000 Hz’in altında olduğu bilinir. Diğer taraftan televizyon sinyalleri 5 MHz’e kadar önemli frekans bileşenlerine sahiptir [11]. Örnekleme teoreminde Nyquist frekansı ve Nyquist oranı iki önemli terimdir [13].

Nyquist örnekleme teoremi, analog bir sinyal sayısallaştırıldıktan sonra tekrar kayıpsız olarak analog sinyale dönüştürülebilmesi için sinyal içerisindeki en yüksek frekanslı (F_{max}) bileşenden en az iki kat frekansta örneklenmesi gerektiğini ifade eder (Eşitlik 2.1).

$$F_s \geq 2F_{max} \quad (2.1)$$

Minimum örnekleme oranı örnekleme teoremi ile belirlenir, yani $F_s = F_{max}$ Nyquist oranı olarak adlandırılır. F_s 'nin gelişigüzel değerleri için $F_s / 2$ Nyquist frekansı veya katlama frekansı olarak adlandırılır [15] (Eşitlik 2.2).

$$\text{Nyquist Aralığı} = \left[-\frac{F_s}{2}, \frac{F_s}{2} \right] \quad (2.2)$$

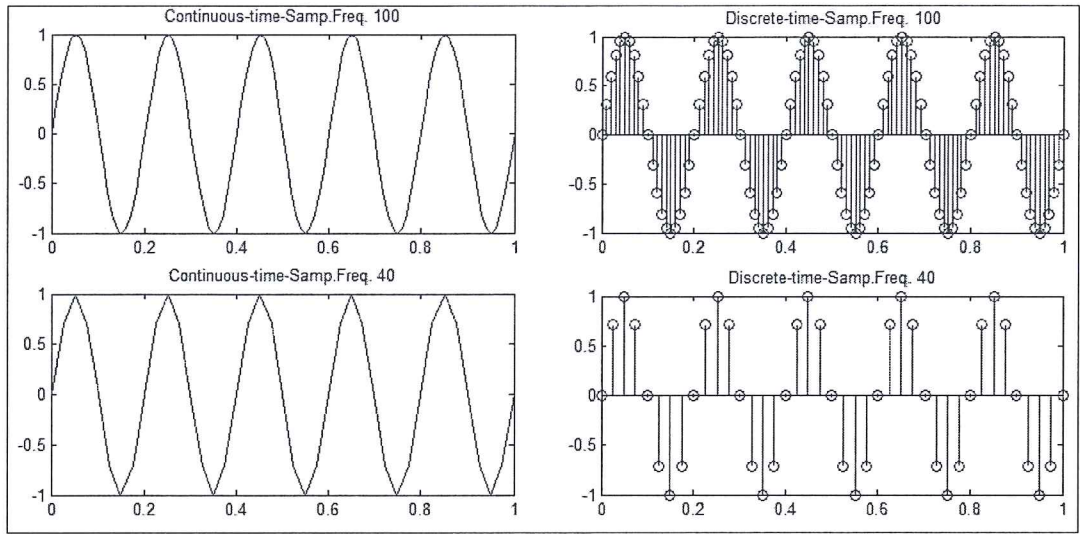
Örneğin maksimum 1000 Hz frekans değerine sahip bir analog sinyal için örnekleme oranı 2000 Hz ve örnekleme hızı 16000 bps (2000*8 bit) olmalıdır. Bunun anlamı en az bozulma ile tekrar analog sinyale dönüşüm için saniyede 16000 örnekleme yapılması demektir. Bazı DSP uygulamaları için örnekleme oranları Çizelge 2.1'de verilmiştir [15].

Çizelge 2.1. Çeşitli uygulamalar için örnekleme oranları [15].

Uygulama	F_{max}	F_s
Jeofizik	500 Hz	1 kHz
Biyomedikal	1 kHz	2 kHz
Mekanik	2 kHz	4 kHz
Konuşma	4 kHz	8 kHz
Ses	20 kHz	40 kHz
Video	4 MHz	8 MHz

Örnekleme işlemi ile ilgili diğer önemli bir konu ise aliasing işlemidir. Aliasing, analog sinyal sayısala dönüştürülürken meydana gelen bir etki olup sayısallaştırılan sinyal tekrar analoğa dönüştürülürken sinyalde kayıplara neden olmaktadır. Eğer örnekleme frekansı örneklenen sinyalin en yüksek frekans bileşeninin iki katından fazla değilse bu olay meydana gelir.

Şekil 2.8'de bir sinüzoidal sinyal üzerinde 100 Hz ve 40 Hz örnekleme frekansı sonucu elde edilen görüntüler verilmiştir.

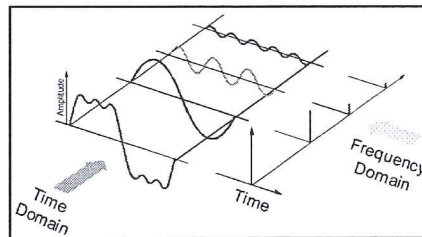


Şekil 2.8. 100 Hz ve 40 Hz örnekleme frekansı ile bir sinüzoidal dalganın sürekli ve ayrık örnekleri.

2.2.2. Sinyallerin Frekans Analizi

Sinyallerin Fourier gösterimi, sürekli zamanlı ve ayrık zamanlı sinyal işlemede son derece önemlidir. Fourier dönüşümü, sinyalin sinüzoidal bileşenlerini elde etmek ve sinyalin zaman boyutundan frekans boyutuna eşleştirilmesi için kullanılır [11,16,17]. Ayrıca, fourier dönüşümü sinyaller ve sistemleri yorumlamak için farklı bir yol sağlar [18]. Tıp, optik, fizik ve elektrik mühendisliği dahil olmak üzere birçok alanda Fourier dönüşümü ortak araç olarak kullanılmaktadır.

19.yüzyılda Fransız matematikçi Joseph Fourier herhangi bir periyodik fonksiyonun sınırsız sayıda karmaşık üstel periyodik fonksiyonun toplamıyla ifade edebileceğini göstermiştir [19]. Şekil 2.9'da periyodik bir sinyalin frekans dönüşümü ile elde edilen sinüs bileşenleri gösterilmiştir.



Şekil 2.9. Bir sinyalin fourier dönüşümü frekans domenine taşınması ile sinüzoidal bileşenleri [20].

Sinyaller sürekli, ayrık, periyodik ve periyodik olmayan şekilde 4 grupta düşünülürse fourier dönüşümü de 4 kategoride ele alınabilir [13,16].

Periyodik olmayan sürekli sinyaller, periyodik desenler içermeyen hem negatif hem de pozitif sonsuza uzayan sinyaller olup bu tip sinyallerin fourier dönüşümü temel olarak fourier dönüşümü olarak adlandırılır. Periyodik olmayan $x(t)$ sinyalinin fourier dönüşümü Eşitlik 2.3'de verilen ve ters fourier dönüşümü Eşitlik 2.4'de verilen eşitliklerle hesaplanmaktadır [11,12].

$$X(F) = \int_{-\infty}^{\infty} x(t)e^{-j2\pi Ft} dt \quad (2.3)$$

$$x(t) = \int_{-\infty}^{\infty} X(F)e^{j2\pi Ft} dF \quad (2.4)$$

Periyodik sürekli sinyaller, periyodik desenler içeren hem negatif hem de pozitif sonsuza uzayan sinyaller olup bu tip sinyallerin fourier dönüşümü temel olarak fourier serisi olarak adlandırılır. $x(t)$ periyodik sinyalinin Fourier serisi aşağıda verilmiştir. Eşitlik 2.5 sentez ve Eşitlik 2.6'da çözümlene eşitliklerini içermektedir [11,12].

$$x(t) = \sum_{k=-\infty}^{\infty} c_k e^{j2\pi kF_0 t} \quad (2.5)$$

$$c_k = \frac{1}{T_p} \int_{T_p} x(t)e^{-j2\pi kF_0 t} dt \quad (2.6)$$

Periyodik olmayan ayrık sinyaller, pozitif ve negatif sonsuzluk arasında ayrık noktalar olarak tanımlı olup periyodik değildir. Bu tip sinyaller için fourier dönüşümü Ayrık Zamanlı Fourier Dönüşümü (DTFT) olarak adlandırılır. Ayrık bir sinyal olan $x(nT)$ 'nin DTFT'si Eşitlik 2.7'de ve Ayrık Zamanlı Ters Fourier Dönüşümü (IDTFT) Eşitlik 2.8'de verilmiştir [12].

$$X(\omega) = \sum_{n=-\infty}^{\infty} x(n)e^{-j\omega n} \quad (2.7)$$

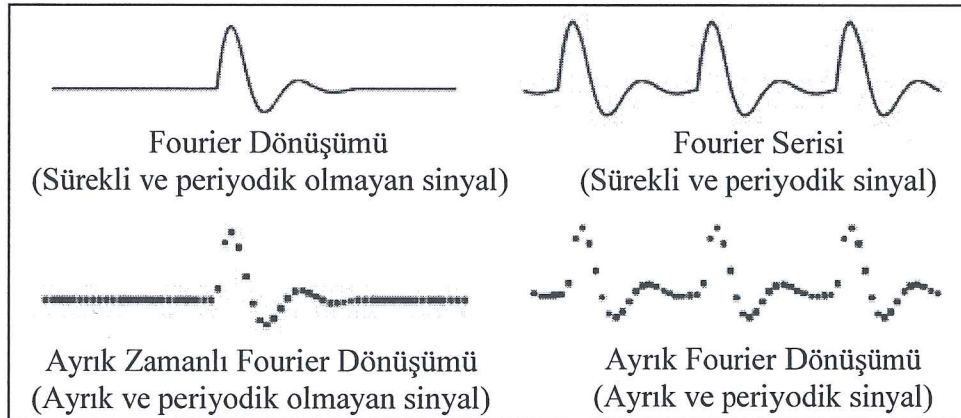
$$x(n) = \frac{1}{2\pi} \int_{2\pi} X(\omega)e^{j\omega n} d\omega \quad (2.8)$$

Periyodik ayrık sinyaller, pozitif ve negatif sonsuzluk arasında ayrık noktalar olarak tanımlı olup periyodiktir. Bu tip sinyaller için fourier dönüşümü Ayrık Fourier Dönüşümü (DFT) olarak adlandırılır. Hızlı Fourier Dönüşümü (FFT), DFT uygulamak için kullanılan zeki bir algoritmadır. FFT, DFT ile aynı sonuçları verir ama algoritma verimliliğinden dolayı çok daha hızlıdır [17]. DSP 'de fourier dönüşümü tiplerinden yalnızca DFT kullanılır [13]. N uzunluğunda sonlu bir x[n] sinyali için DFT Eşitlik 2.9'da ve Ters Ayrık Fourier Dönüşümü (IDFT) Eşitlik 2.10'da verilmiştir [19].

$$X(F) = \sum_{n=0}^{N-1} x[n]e^{-j2\pi F_n/N} \quad (2.9)$$

$$x(n) = \frac{1}{N} \sum_{F=0}^{N-1} X(F)e^{j2\pi F_n/N} \quad (2.10)$$

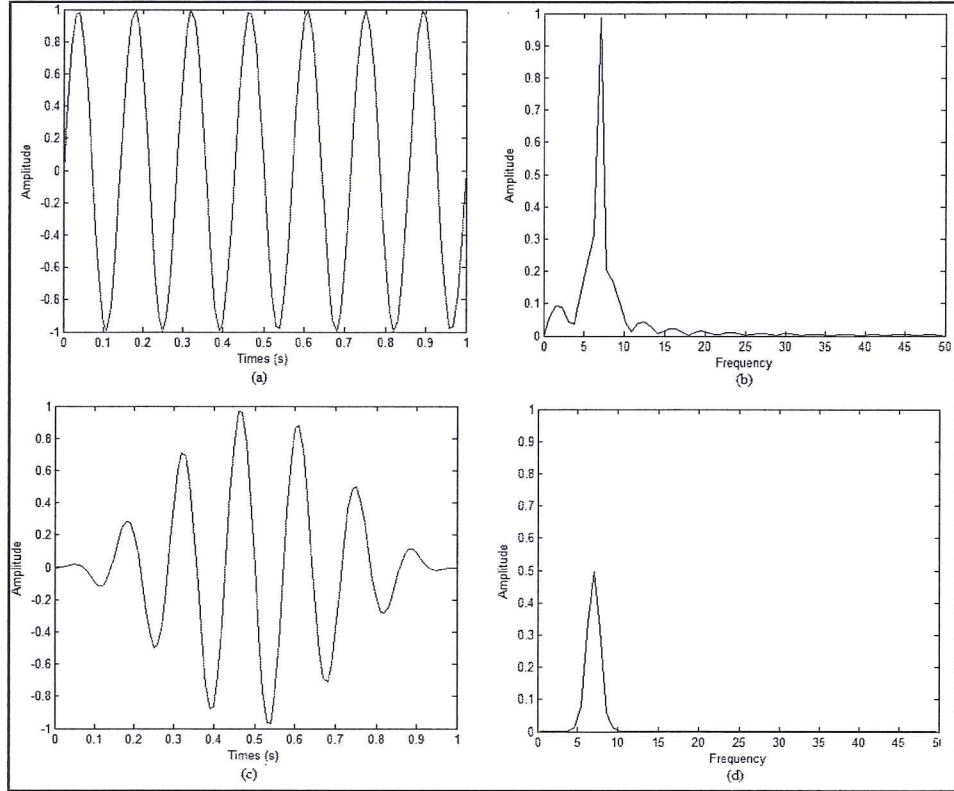
4 kategoride fourier dönüşümü sonrası sinyal durumları Şekil 2.10'da verilmiştir.



Şekil 2.10. 4 kategoride fourier dönüşümü sonuçları [13].

2.2.3. Pencereleme

Tek bir frekansa sahip bir sinyale DFT uygulanması sonucunda sadece tek bir frekansa sahip düz bir çizgi olması beklenir ancak spektrum incelendiğinde sinyalin sahip olduğu frekanslara yakın bölgelerde eğriler oluşmaktadır. Bu durum spektral sızma olarak adlandırılmaktadır. DFT hesabı sırasında spektral sızmayı azaltmanın bir yolu, sinyalin son değerleri ile ilk değerleri arasındaki devamsızlığın giderilmesidir. Bu amaçla sinyalde, DFT öncesi pencereleme kullanılmaktadır [19,21]. Örnek bir sinyalin fourier dönüşümü ve pencereleme sonrası durumu Şekil 2.11’de verilmiştir.



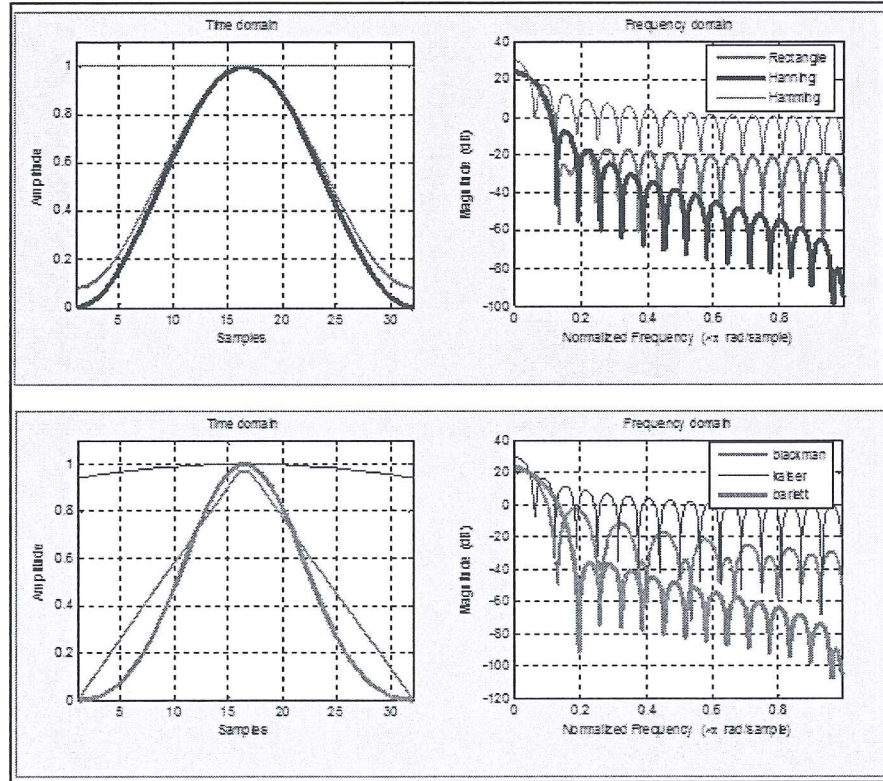
Şekil 2.11. a) Tek bir frekansa sahip sinüzoidal bir sinyal b) Sinyalin FFT dönüşümü c) Pencereleme sonrası sinyalin durumu d) Sinyalin pencerelemesi sonrası FFT dönüşümü.

Çizelge 2.2’de çoğunlukla kullanılan pencereleme fonksiyonları (Eşitlik 2.11) verilmiştir.

Çizelge 2.2. Pencereleme fonksiyonları.

Pencereleme Türü	Pencereleme Fonksiyonu $N-1 \geq n \geq 0$
Hamming	$w(n) = 0.54 - 0.46 \cos \frac{2\pi n}{N-1}$ (2.11a)
Hanning	$w(n) = \frac{1}{2} \left(1 - \cos \frac{2\pi n}{N-1} \right)$ (2.11b)
Bartlett	$w(n) = 1 - \frac{2 \left n - \frac{N-1}{2} \right }{N-1}$ (2.11c)
Blackman	$w(n) = 0.42 - 0.5 \cos \frac{2\pi n}{N-1} + 0.8 \cos \frac{4\pi n}{N-1}$ (2.11d)
Rectangular	$w(n) = \begin{cases} 1, & n \leq \frac{N-1}{2} \\ 0, & \text{diğer_durumlarda} \end{cases}$ (2.11e)

Pencereleme fonksiyonlarının zaman boyutu ve frekans boyutundaki tepkileri Şekil 2.12’de verilmiştir.



Şekil 2.12. Çeşitli pencereleme fonksiyonlarının zaman ve frekans boyutundaki tepkileri.

2.2.4. Dalgacık Dönüşümü

FFT, sinyalin zaman boyutundan frekans boyutuna eşleştirilmesi ve frekans boyutundaki özniteliklerin tespiti için kullanılmaktadır. Ancak, FFT ile sinyal frekans boyutuna aktarıldığında sinyale ait zaman çözünürlüğü kaybolmaktadır. Yani, FFT 'si alınan bir sinyalde önemli olan bir noktanın hangi zamanda gerçekleştiği anlaşılamaz. Durağan sinyallerde (karakteristik özelliği zaman boyunca aynı olan sinyal) zamana göre frekans değişmediği için bu dezavantaj önemli değildir. Eğer sinyal durağan değilse (zaman göre frekansın değişmesi) FFT önemli sayılabilecek özniteliklerin başlangıç ve bitiş zamanlarının kaybolmasına neden olacaktır.

FFT'nin zaman bilgisini kaybetmesi dezavantajının üstesinden gelmek için pencereleme özelliği içeren Kısa Zamanlı Fourier Dönüşümü (STFT) geliştirilmiştir.

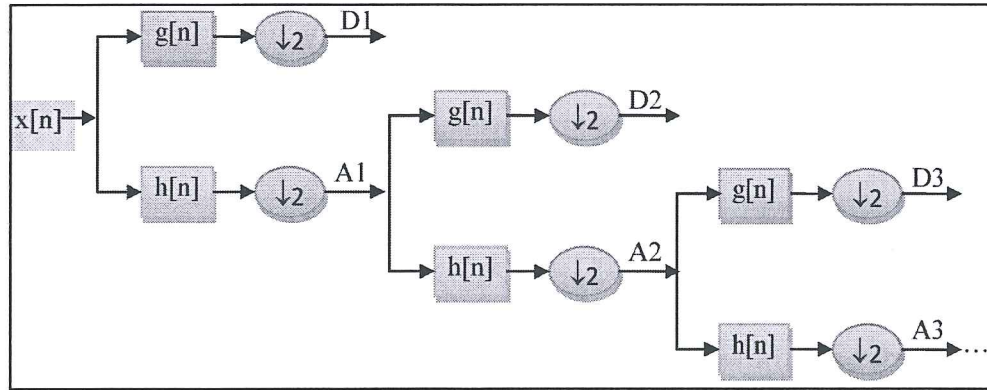
STFT, sinyalin ne zaman ve hangi frekansla oluştuğu hakkında bilgi verir. Ancak, pencere boyutu sabit tutulduğu için elde edilen bilgilerin doğruluğu sınırlıdır. Çünkü pencere boyutu frekans çözünürlüğünü etkilemektedir. Pencere boyutu yüksek tutulursa iyi bir frekans çözünürlüğü, düşük tutulursa frekans çözünürlüğü düşük olacaktır [22]. Fourier dönüşümünün bu dezavantajlarının üstesinden gelmek için dalgacık dönüşümü geliştirilmiştir.

Dalgacık ifadesi ilk olarak 1909 yılında Alfred Haar tarafından ortaya atılmış ve Jean Marlet, Y. Meyer ve arkadaşları ve Stephane Mallat bu yöntemi geliştirmişlerdir. Daha sonra Ingrid Daubechies, Ronald Coifman gibi uluslararası araştırmacılar yöntemi geliştirerek bugünkü şekline getirmişlerdir [22,23].

Dalgacık dönüşümü, pencere boyutunu değişken tutarak yüksek pencere genişliğinde alçak frekans bilgisini, düşük pencere genişliğinde yüksek frekans bilgisini belirlemeye yardımcı olur. Böylece, bütün frekans aralıklarında optimum zaman-frekans çözünürlüğü sağlanabilmektedir. Bunun yanı sıra, fourier dönüşümünün aksine sadece zaman-frekans boyutunu değil, aynı zamanda zaman-ölçek boyutunu kullanır [24].

Dalgacık dönüşümü sürekli ve ayrık olmak üzere iki şekilde kullanılabilir. Sürekli dalgacık dönüşümü, bütün ölçekleri kullanarak sinyale ait tüm detayları verdiği için hesaplama zamanı yüksektir. Ayrık dalgacık dönüşümünde ise ayrık zaman aralıkları ve ölçekleri kullanarak hesaplama zamanı düşürülmektedir. Buna karşın detay kaybı fazla değildir. Bu nedenle, hesaplama süresinin önemli olduğu ve ayrıntılı detaya ihtiyaç olmayan sinyaller için ayrık dalgacık dönüşümü kullanılabilir. Ayrık dalgacık dönüşümünde öteleme ve ölçeklerin ikili kuvveti alınmasına ikili dalgacık dönüşümü adı verilir. İkili dalgacık dönüşümü ile sinyaller kademeli olarak yüksek ve düşük frekanslara ayrılmaktadır. Yüksek frekans kısmında D (Ayrıntılar) ve düşük frekans kısmında A (Temel İşaret) yer almaktadır [25].

Şekil 2.13'de örnek bir $x[n]$ işareti için ayrık dalgacık dönüşümünün bantları verilmiştir.



Şekil 2.13. Ayrık dalgacık dönüşümünde alt bantlara ayrışım; $g[n]$ yüksek geçiren filtre, $h[n]$ alçak geçiren filtre.

Şekil 2.13'de A katsayıları sinyalin alçak frekanslı bileşenlerini, D katsayıları ise sinyalin yüksek frekanslı bileşenlerini göstermektedir. İlk yüksek geçiren filtreye ($g[.]$) ve alçak geçiren filtreye ($h[.]$) ait olan örneklenmiş çıkışlar sırası ile ayrıntılı D1 ve yaklaşık A1 alt bantlarını oluşturur. A1 yaklaşım bandı tekrar ayrıştır ve bu işlem Şekil 2.13'de görüldüğü gibi devam eder.

Dalgacık dönüşümünün bir diğer avantajı ise sinyalin orijinalliğini bozmadan sıkıştırma gürültü giderme işlemlerinin yapılabilmesidir [26]. Bu yöntemle frekansı

zamanla deęişen sistemlerin analizi ve geçici durum analizleri oldukça hassas bir şekilde yapılmaktadır [27].

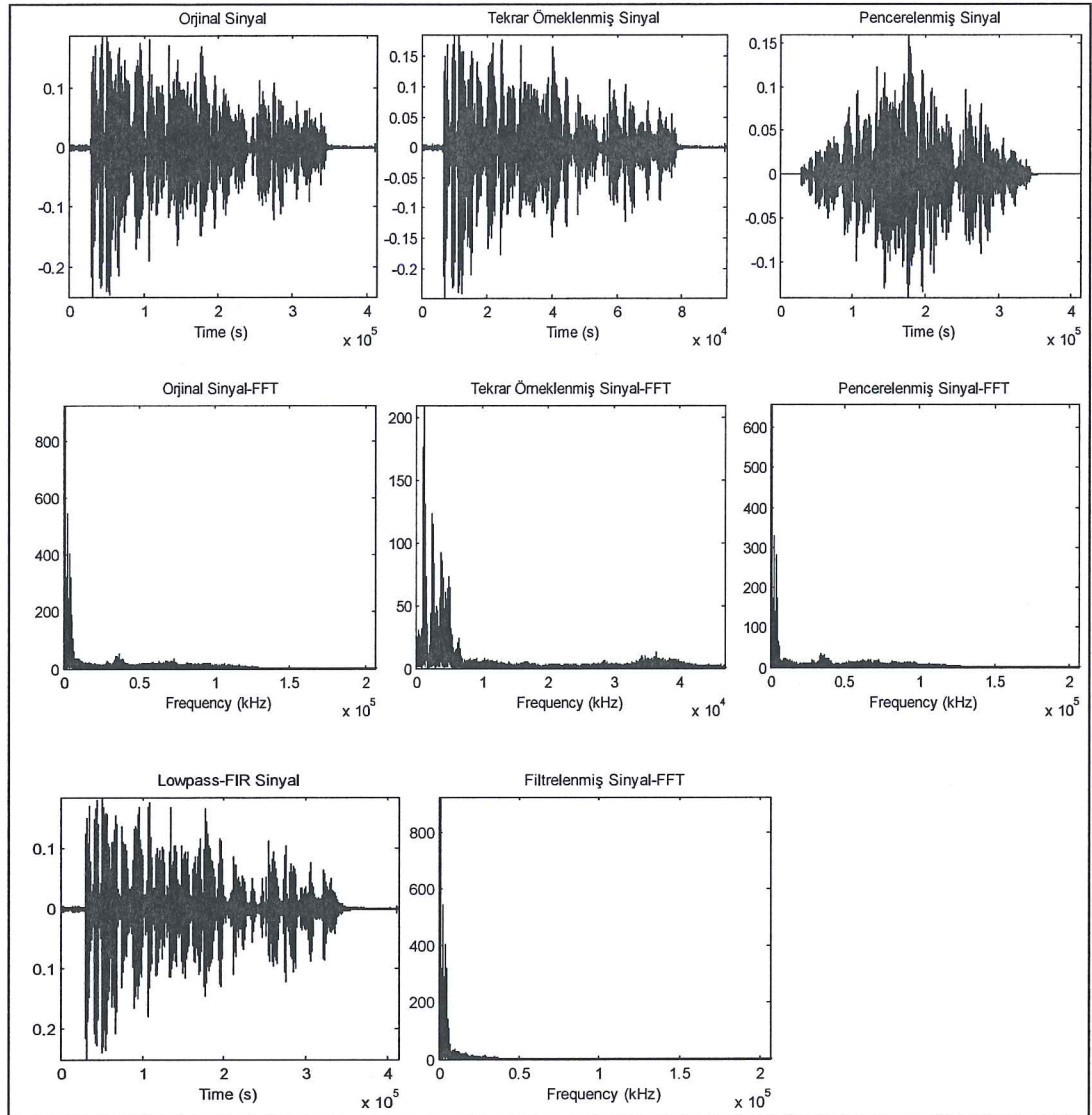
Dalgacık yöntemi kullanarak bir sinyalin ayrıştırılması ve tekrar oluşturulması genel olarak üç aşamadan oluşur [28]:

1. Ayrık dalgacık dönüşümü kullanarak sinyalin bileşenlerine ayrılması,
2. Ortaya çıkan katsayıların yumuşatma metotlarıyla yumuşatılması,
3. Yumuşatılmış ayrık dalgacık katsayılarından tekrardan orijinal sinyalin oluşturulması.

Dalgacık dönüşümü metodu ile EEG işaretlerinden duygu tanıma çalışmalarında Daubechies 4, Daubechies 8, Symlets 8 ve Coiflet 5 en çok kullanılan dalgacık fonksiyonlarıdır [28–30].

2.2.5. DFT Uygulaması

DFT sayısal sinyal işleme için en önemli araçlardan birisidir. DFT ile yapılacak uygulamalarda DFT ile sinyalin frekans spektrumu hesaplanabilir. Böylece sinyalin sinüzoidal bileşenlerinin frekans, faz ve genliğinde kodlanmış bilgiler incelenebilir. Ayrıca, DFT ile impuls yanıtında frekans yanıtı veya tersi tespit edilebilir [13]. Bir sinyalin zaman boyutundaki dalga şekli herhangi bir bilgi sağlamaz. Frekans, faz ve genliğinde kodlanmış önemli bilgiler frekans boyutunda elde edilebilir. Şekil 2.14’de sinyal üzerinde gerçekleştirilen bazı işlemlere ait sonuçlar verilmiştir.



Şekil 2.14. Örnek bir konuşma sinyali üzerinde gerçekleştirilen tekrar örnekleme, pencereleme ve alçak geçiren FIR filtreleme sonucu sinyal üzerinde meydana gelen değişiklikler.

Şekil 2.14'de verilen sonuçlara ait MATLAB kodları Ek Açıklamalar A'da verilmiştir.

2.3. DUYGU VE KONUŞMA

Duygu tanıma çalışmalarının temelini duygunun ne anlama geldiği ve konuşma sinyalini nasıl etkilediği oluşturmaktadır. Duygu, bireyin ruh halinde biyokimyasal (içsel) ve çevresel tesirlerle etkileşiminden doğan kompleks psikofizyolojik bir değişimdir. Duygu kişinin bir acı hissetmesi, sevindirici bir haber alması veya bir

olay karşısında korkması gibi olaylar karşısında tetiklenen ve kişinin sesine ve yüzüne yansıyan fizyolojik tepkidir. Geniş anlamda, duygular insanların hayatını korumak ve hayatta kalmak için evrimleşmiş biyodüzenleyici mekanizmanın bir parçasıdır. Duygular uyaran ve davranışsal tepki arasında yer alan bir ara katman yardımıyla olaylara karşı verilecek tepki yumuşatılmaktadır [31,32]. Aynı zamanda uyaranların sürekli değerlendirmesi ve bu uyaranlara davranışsal tepkilerin hazırlanması duygunun temel fonksiyonlarından birisi olarak ileri sürülmüştür [33,34]. Duygusal reaksiyonlar, organizmalar arasında bir sinyalizasyon sistemi olarak görev yapar ve yeni davranış kalıplarının edinimini sağlar. Bu davranış kalıpları öğrenme için bir ön koşuldur [36,37]. Duygular ve duygusal reaksiyonlar sonucu öğrenmeler bireye özgü olsa da temel özellikler tüm insanlar arasında tutarlıdır [38]. Tipik olarak bir duygu kan dolaşımı veya nöronlar aracılığıyla vücuda kişinin durumunda genel değişikliğe neden olan komutlar gönderir. Damasio tarafından 2000 yılında sunulan bu görüş Ekman gibi diğer yazarların çalışmalarında da ortaya konulmuştur [39,40].

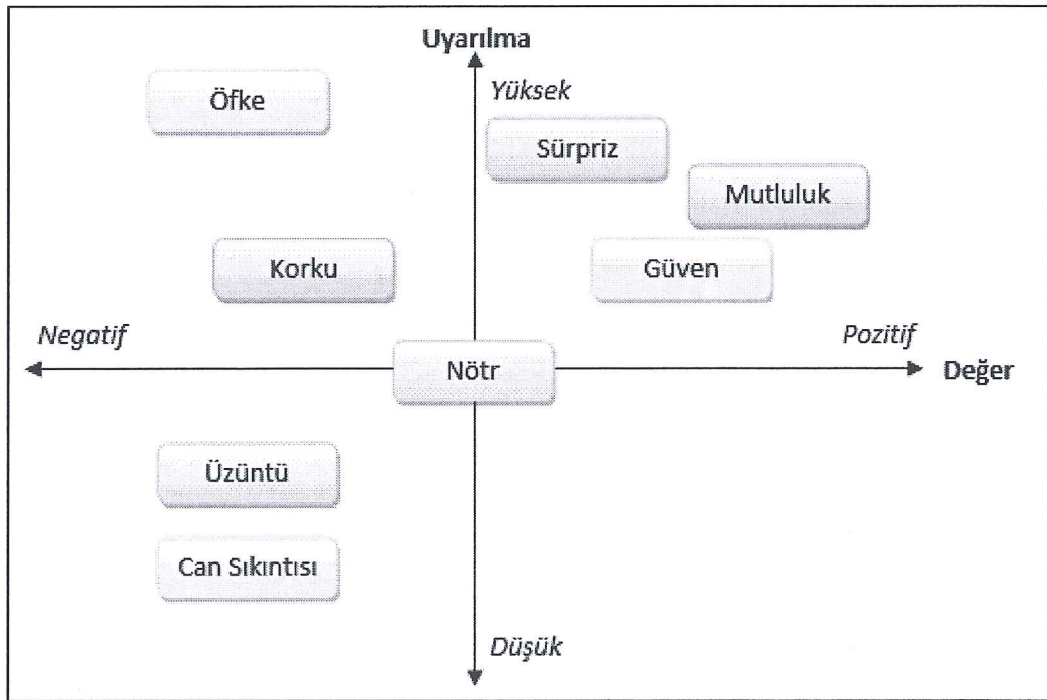
Duygusal durum kişinin bulunduğu ortam ve şartlara karşı içsel ve istem dışı verdiği tepki olarak tanımlanabilir. Kişilerin duygusal durumlarına göre nabız ve nefes almada, kan akış hızında, hormonal dengelerde, ses özelliklerinde, yüz ifadesinde ve el kol hareketlerinde çeşitli fiziksel ve fizyolojik değişiklikler gözlemlenebilir. Bu değişikliklerin çeşitliliği, boyutu ve derecesi farklı duygusal durumlara göre şekillenir [41].

İnsanların öğrenmişlikleri kişinin konuşmasından duygu durumunu anlamaya yardımcı olur. Örneğin, özellikle çok mutlu olduğunda yüksek sesle konuşulur veya korku/üzüntü durumunda seste titreme oluşur. Duygu tanımada sesle birlikte yüz ifadeleri içeren çalışmalar olduğu gibi sadece ses veya sadece yüz ifadelerinden tanıma gerçekleştirilen çalışmalar bulunmaktadır. Bu durum kişinin duygusunun tespit edilmesi için konuşmasının yeterli olacağını göstermektedir.

Duyguların fizyolojik tanımlarına göre duygular sesi hem istemli hem de istemsiz olarak etkileyebilir ve vokal ifadeler bu etkilerin sonucu oluşur. Ayrıca, bedensel hareketleri etkileyen faaliyetler ses üretim mekanizmasını dolayısıyla sesi de

etkileyecektir. Fonagy 1981 yılında yaptığı çalışmada ses ve fiziksel hareketler arasındaki paralelliği inceleyerek bu görüşü doğrulamıştır. Örneğin, boğazdaki gerginlik bütün vücuda yansımakta ve buna paralel olarak da sesin şiddetini artırmaktadır [42].

Konuşmadan duygu tanıma çalışmalarının büyük bir kısmı korku, üzüntü ve mutluluk gibi ana duygulara odaklanırken, bazı çalışmalar negatif-pozitif [5,6] ve valence-arousal [17,23] gibi duyguları boyutsal yapıda ele almıştır. Duyguların boyutsal incelenmesi duygu durumlarının şiddetine göre boyutsal yapının oluşturulmasını kapsamaktadır. Duyguların arousal ve valence boyutu Şekil 2.15’de verilmiştir.



Şekil 2.15. Duyguların boyutsal gösterimi.

Şekil 2.15’e göre sürpriz duygusunun uyarılma değeri mutluluktan yüksek ve değerlik değeri mutluluktan düşüktür. Değerlik ve uyarılma ekseninde duygular yoğunluklarına göre yer almaktadır.

2.4. ANKSİYETE

Anksiyete sıkıntı, endişe ve/veya korkuyu içeren nahoş bir duygudur. Tehdit veya tehlike durumlarında ortaya çıkmakta ve kişiyi tehlike anında eyleme geçmeye hazırlamaktadır. Tehlike anında kaç ya da dövüş yanıtının oluşmasını sağlamaktadır. Normal anksiyete mantıklı, gerekli, hatta yaşamsaldır. Olası tehlike uyarılarına tüm canlılar kalıtsal ve doğumsal bir düzenek olan anksiyete ile yanıt vermektedirler. Bu tehlikelerin mutlaka yaşamı tehdit eder nitelikte olması şart değildir. Diğer insanlarla ilişkilerimizi olumsuz etkileyecek, zarar verecek olaylar veya durumlarda anksiyeteyi uyarmaktadır. Günlük yaşamda karşılaştığımız birçok durumda (sınavlar, yeni bir insanla tanışma, yabancı olduğumuz ortamlar...) anksiyetenin meydana gelmesi olağandır. Ancak, beklenenden fazla ya da şiddetli olduğunda, gereğinden daha uzun sürdüğünde ya da tekrarladığında, zihni sürekli meşgul ettiğinde, kısaca kişinin günlük işlevlerini bozduğunda patolojik anksiyeteden söz edilmektedir.

Kaç ya da dövüş yanıtı otonom sinir sisteminin uyarılması sonrasında ortaya çıkar. Sempatik sinir sisteminin uyarılmasına bağlı olarak ortaya çıkan terleme, çarpıntı, hızlı ve yüzeysel nefes alma, ellerde ve yüzde yanma gibi belirtiler, tehlikenin bitmesi durumunda parasempatik sistemin uyarılması ile normale dönmektedir. Anksiyete bozukluklarında otonom sinir sisteminde ki dengesizlikten söz edilmektedir. Anksiyete durumunda meydana gelen bedensel belirtiler, özellikle solunum sistemi ile ilgili belirtiler kişinin ses sistemini de etkilemektedir.

Sesin duygusal durumu yansıtma derecesi ve şekli konusunda yapılan araştırmalar oldukça sınırlıdır. Bu sınırlılık anksiyete ve ses ilişkisinde daha belirgin hale gelmektedir. Ayrıca, yapılan çalışmaların sonuçları arasında çelişkiler mevcuttur.

2.5. YAPILAN ÇALIŞMALAR

2.5.1. Akustik Parametrelerin Tespiti İçin Kullanılan Araçlar

Konuşmadan öznitelik çıkartımı ve akustik parametre tespiti hazır araçlar veya araştırmacılar tarafından geliştirilen kodlar yardımıyla gerçekleştirilmektedir. Hazır araçlar iş gücünü azaltma ve çıkartılan özniteliklerin güvenilirliğini artırmaya yardımcı olmaktadır. Mevcut araçlar belirli çalışma alanı için geliştirilmiş olup birçoğu kullanıcı dostu bir ara yüze sahip değildir. Ayrıca, çıkartılan öznitelikler kullanılan araca göre farklılık oluşturmaktadır.

Bu araçlardan, Hidden Markov Toolkit (HTK) [43] birden fazla konuşma üzerinden çok sayıda öznitelik çıkartma özelliğine sahip olmasına rağmen grafiksel ara yüze sahip olmadığı için kullanım zorluğuna sahiptir. Ayrıca, öznitelik çıkartımı için birden fazla dosya üzerinde çalışmak gerekebilir. Birçok araç bunu desteklerken COLEA [44] kullanım kolaylığı sunan grafiksel ara yüze sahip olmasına rağmen toplu dosya işleme yeteneğine sahip değildir. Grafiksel ara yüze ve toplu dosya işleme yeteneğine sahip olan Praat [45] birçok çalışmada kullanılmaktadır. Praat çok sayıda öznitelik çıkartma yeteneğine ve grafiksel ara yüze sahip olmasına karşın toplu dosya işleme için script hazırlanması gerekir. Bir diğer araç olan OpenSMILE [46] ve OpenEAR [47] çok sayıda öznitelik çıkartma ve toplu dosya işleme sahip olmasına rağmen grafiksel arayüz ayrı bir yazılım yardımıyla kullanılmaktadır. Ancak, öznitelik çıkartımı için gerekli ayarlamalar ve çıkartılacak öznitelik konfigürasyon dosyası üzerinden gerçekleştirilmek zorundadır. Bu araçlar dışında küçük çapta geliştirilmiş çeşitli araçlarda bulunmaktadır [48–51]. Çizelge 2.3’de konuşma işleme ve öznitelik çıkartımı amacıyla geliştirilmiş araçların genel bir özeti verilmiştir.

Çizelge 2.3. Öznitelik çıkartımı araçları.

Araç	Platform	Erişim	Ara yüz	Çıkartılan Öznitelikler
Praat [45]	C++	Ücretsiz	Grafiksel	Enerji, FFT spektrum, kepstral, pitch, ses kalitesi, LPC, PLP, formant, spektral, H1-H2, H1-A1, H1-A2, H1-A3, konuşma oranı
OpenSMILE [46]	C++	Ücretsiz	Komut satırı ve grafiksel	Dalga biçimi, enerji, ses şiddeti, FFT spektrum, ACF, kepstrum, Mel/Bark spektr., pitch, ses kalitesi, LPC, PLP, formant, tonal
OpenEAR [47]				
HTK [43]	C	Ücretsiz	Komut satırı	Enerji, kepstral, PLP, Mel/Bark spektr., LPC
Voicebox [53]	Matlab	Ücretsiz	Matlab fonksiyonları	Dalga biçimi, enerji, pitch, LPC, kepstral, Mel/Bark spektr.
COLEA [54]	Matlab	Ücretsiz	Grafiksel	pitch, formant, spektral, enerji
SPEFT [55]	Matlab	Belirsiz	Grafiksel	Dalga biçimi, enerji, ses kalitesi, pitch, formant, kepstral, PLP, LPC, Mel/Bark spektr.

Çizelge 2.3 'ün içerdiği kepstral öznitelik kümesi MFCC, pitch öznitelik kümesi Temel Frekans (F0), ses kalitesi öznitelik kümesi Harmonik Gürültü Oranı (HNR), jitter ve shimmer, dalga biçimi öznitelik kümesi ZCR, formant öznitelik kümesi formant frekansları ve bant genişliği, spektral öznitelik kümesi enerji ve entropi, tonal öznitelik kümesi CHROMA ve CENS özelliklerini içermektedir.

2.5.2. Anksiyetenin Akustik Parametrelerle İlişkisi

Bir kişi fiziksel ya da psikolojik tehdit içeren bir durum yaşadığında anksiyete oluşur ve fizyolojik yanıt olarak başa çıkma stratejileri tetiklenir. Anksiyete temel bir duygu olarak kabul edilmez ve genellikle korku, belirsizlik, sıkıntı, endişe ve kaygı gibi bir veya daha fazla negatif duygunun kombinasyonudur. Anksiyete çoğunlukla korku gerektirdiğinden korku ailesine ait olduğu kabul edilir [56]. Ancak, korku ile günlük hayatta anksiyeteden daha fazla karşılaşılmaktadır. Ayrıca, anksiyete tehdit kaçınılmaz olduğunda görünürken korku kaçmayı uyarır [57]. Anksiyetenin çeşitli akustik parametreler üzerindeki etkilerini inceleyen çalışmaların elde ettiği sonuçlar Çizelge 2.4'de verilmiştir.

Çizelge 2.4. Anksiyetenin akustik parametreler üzerindeki etkisini inceleyen çalışmaların özeti.

Akustik Parametre	Değişim	Referans
F0 ortalama	Artma	[5–11]
F0 aralık	Düzensiz	[7,8,12]
F0 standart sapma	Düzensiz	[10,13]
Konuşma süresi	Azalma	[4]
Konuşma oranı	Artma	[58]
% duraksama	Azalma	[3,10]
Enerji	Düzensiz	[3,5,7]
HNR	Düzensiz	[58]
Jitter/Shimmer	Artma	[66]

Çizelge 2.4'e göre F0'ın ortalama değerinin genel olarak arttığı tespit edilmiştir [4,58–63]. Konuşmanın sözel olmayan kısımlarını inceleyen çalışmaya göre ise F0'ın ortalama değerinde azalma tespit edilmiştir [56]. F0 aralık değerindeki değişimleri inceleyen çalışmalara göre bu parametrede değişim olmadığı tespit edilirken [64], bazı çalışmalarda bu parametrenin değerinin azaldığı [4,60] ve bazı çalışmalarda arttığı [61] görülmüştür. Bu durum, anksiyete durumunda F0 aralık değerinin düzensiz olduğunu göstermektedir. Aynı durum F0 standart sapma için de geçerli olup hem azalma [65] hem de artma [62] tespit edilen çalışmalar mevcuttur. Konuşmanın fiziksel yapısı ile ilgili parametreleri inceleyen çalışmalara göre konuşma süresi değerinin azaldığı [4], konuşma oranı değerinin çok hızlı arttığı [58] ve bekleme oranının azaldığı [56,62] tespit edilmiştir. Enerji değerini inceleyen çalışmalara göre anksiyete durumunda bu değeri normal [58], yüksek [60] ve sözel olmayan kısımlarda düşük olarak [56] tespit eden çalışmalar mevcuttur. Anksiyete durumunda ses kalitesinde düzensizlikler [58], jitter ve shimmer değerlerinde artış [66] olduğu tespit edilmiştir.

Anksiyete ve korku arasında kesin bir ayrım yapılamadığı için Laukka vd. 2008 yılında bazı akustik parametrelerin anksiyete ve korku durumundaki değişimini incelerken anksiyete için Siegman tarafından 1987 yılında yapılan çalışmayı [67], korku için Juslin ve Laukka'nın 2003 yılında yaptığı çalışmayı [68] referans alarak Çizelge 2.5'de verildiği gibi ifade etmiştir [56].

Çizelge 2.5. Akustik parametrelerin anksiyete ve korku üzerindeki etkileri [56].

Akustik Parametre	Anksiyete'ye Etkisi	Korku'ya Etkisi
F0 ortalama	NA	+
F0 standart sapma	NA	-
F0 maksimum	NA	+
Ortalama enerji	NA	±
Ses kalitesi	NA	±
Konuşma oranı	NA	+
% sessizlik	+	NA

Çizelge 2.5'de kullanılan (+) akustik parametrenin değerinin arttığını, (-) azaldığını, (±) değişim olmadığını ve NA herhangi bir tahmin içermediğini gösterir.

Laukka vd.'nin 2008 yılında yaptıkları kendi çalışmalarında anksiyete için ortalama ve maksimum pitch, enerji spektrumunun yüksek frekans bileşenleri ve sessiz bekleme oranlarında (% slience) azalma gözlemlenmiştir. Bu durum anksiyetenin konuşmanın sözel olmayan kısımlarında etkiye sahip olduğunu göstermektedir [56].

F0'ın ortalama ve aralık değeriyle anksiyete düzeyinin doğru orantılı olduğu ifade edilen çalışmada duyguların ilişkisel çerçevesi ve terapi olmak üzere 2 farklı yöntemle duygular arasındaki ilişki araştırılmıştır [61].

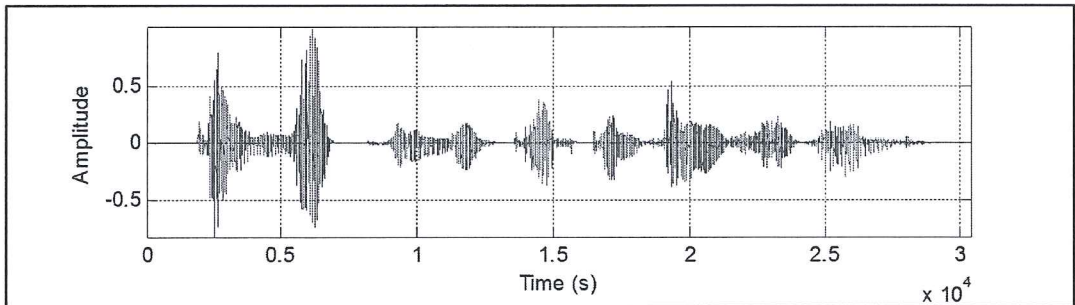
Topluluk önünde konuşma esnasında anksiyetenin etkisinin araştırıldığı çalışmada anksiyetenin konuşmanın başlangıcında daha etkin olduğu ve duraksamaların azaldığı, F0'ın ortalama ve standart sapma değerinin arttığı tespit edilmiştir [62].

BÖLÜM 3

AKUSTİK ANALİZ

Ses ve konuşma analizi bir sayısal sinyal işleme süreci olup günümüzde insanlar farkında olarak veya olmadan sürekli olarak kullanılmaktadır. Ses analizi ses kalitesinin artırılması, gürültü temizleme, kişi tanıma, öznitelik çıkartımı, iletişim gibi birçok alanda karşımıza çıkmaktadır. Bir önceki bölümde de bahsedildiği gibi ses analog bir sinyal olup sayısal olarak dönüştürülmektedir. Bu süreçte dijital kayıt cihazları, DSP, mikrofon ve ses kartı kullanılmaktadır. Bir ses örneği bir sayı dizisi olarak vektörel veya matris şeklinde depolanmaktadır [69]. Ses analizi konuşmacı tanıma, kelime tanıma, konuşma-yazı dönüşümü ve tıp alanında tanı/tedavi sürecinin takibi gibi alanlarda kullanılmaktadır.

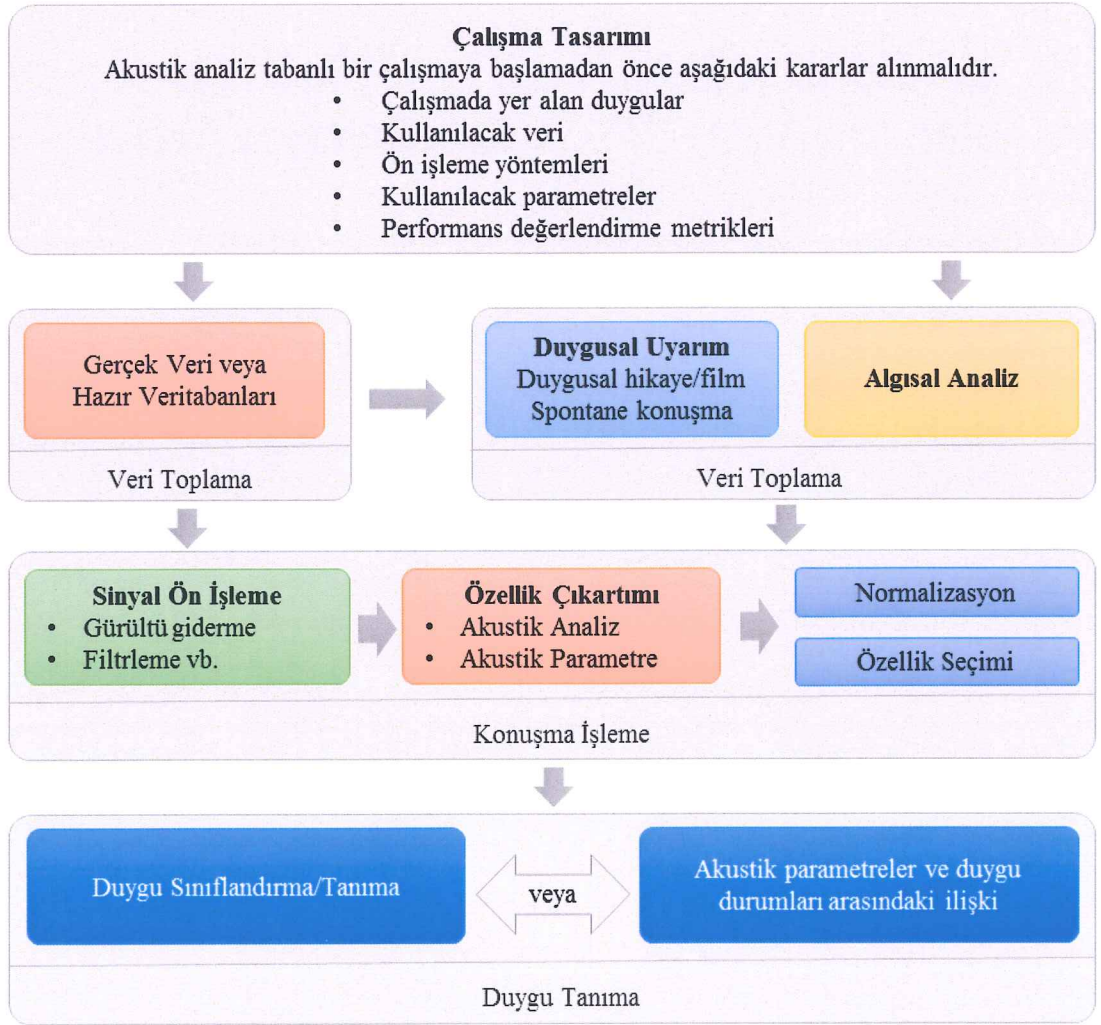
Ses veya konuşmaları sayısal olarak saklamak için kullanılacak çeşitli formatlar mevcuttur. Ancak, Darbe Kod Modülasyonu (PCM) formatı olan wav formatı yüksek örnekleme oranına sahip olması ve veri kaybı oluşturmaması nedeni ile en çok kullanılan formattır [69]. Sesleri 44,1 kHz örnekleme oranı ve 16-bit PCM formatında depolar. Yani, stereo sesler için 1411,2 kbps bit oranına sahiptir [12]. Örnek bir konuşma sinyali Şekil 3.1’de verilmiştir.



Şekil 3.1. Örnek bir konuşma sinyali.

Sayısal bir ses veya konuşma yatayda zaman dikeyde genlik eksenlerine sahiptir. Sesin işlenmesi için mevcut durumdaki zaman boyutu kullanılabilceği gibi frekans boyutu da kullanılabilir. Ses sinyalinin zaman boyutunda frekans boyutuna geçişi için bir önceki bölümde de anlatıldığı gibi fourier dönüşümü veya dalgacık dönüşümü kullanılmaktadır.

Akustik analiz tabanlı duygu tanıma sistemleri genel olarak 5 fazdan oluşmaktadır ve Şekil 3.2’de akış diyagramı verilmiştir.



Şekil 3.2. Akustik analiz tabanlı duygu tanıma akış diyagramı.

Şekil 3.2’de verilen akış diyagramında veri toplama adımı çalışmada kullanılacak duygusal bilgi içeren verilerin edinimini, duyguları tetikleyici durumların belirlenmesini ve algısal analiz ile verilerin duygusal duruma göre sınıflandırılmasını

içermektedir. Eğer elimizdeki veri sınıflandırılmış ise algısal analize gerek yoktur. Konuşma işleme adımında sinyal işleme teknikleri ile konuşma sinyalinin filtreye tabi tutulması, akustik parametrelerin elde edilmesi, elde edilen öznitelikler üzerinde normalizasyon teknikleri ve öznitelik seçim tekniklerinin kullanımını içermektedir. Son adım olan duygu tanıma adımında akustik parametreler kullanılarak duyguların sınıflandırılması veya istatistiksel analiz ile parametreler ve duygular arasındaki ilişki araştırılmaktadır.

3.1. VERİ TOPLAMA

İnsan konuşmalarından duygu tespiti veya duygu sınıflandırması çalışmalarında öncelikle hangi duygular üzerine çalışılacağı belirlenmektedir. Çalışılacak duyguların belirlenmesi için belirli bir kriter bulunmayıp araştırmacının tercihine bağlıdır. Çalışılacak duygu belirlendikten sonra kullanılacak veri tespit edilir.

Duygu tanımada araştırmacıların veri elde etme sıkıntısını ortadan kaldırmak için ücretli ve ücretsiz çeşitli hazır veritabanları mevcuttur. Hazır veritabanları bir yayın/proje kapsamında veya araştırma enstitüleri tarafından oluşturulmuştur. Birçok araştırmacı hazır veritabanlarını ya yetersiz buldukları için ya da özgün çalışma oluşturmak için hazır veritabanı kullanmayıp çalışma kapsamında veri toplamaktadır. Literatürde yer alan hazır veritabanlarına ait özet Çizelge 3.1’de verilmiştir.

Çizelge 3.1. Duygusal konuşma veritabanları.

Veritabanı	Dil	İçerdiği Duygular/Boyutlar
VAM [70]	Almanca	Değerlik, Baskınlık, Aktivasyon
HUMAINE [71]	İngilizce	Öfke, Mutluluk, Üzüntü, Rahat
GEMEP	Fransızca	Eğlence, Anksiyete, Öfke, Umutsuzluk, Korku, İlgi, Sevinç, Zevk, Gurur, Rahatlama, Üzüntü, Hayranlık, Küçümseme, İğrenme, Sürpriz, Şefkat
EMO-DB [72]	Almanca	Öfke, Sevinç, Hüzün, Korkma, İğrenme, Sıkıntı, Nötr
FAU Aibo Em. [73]	Almanca	Öfke, Duygusal, Nötr, Pozitif, Dinlenme

Çizelge 3.1. (devam ediyor).

IEMOCAP [74]	İngilizce	Mutluluk, Öfke, Üzüntü, Hayal Kırıklığı, Nötr
SUSAS [54]	İngilizce	Dört stres türü
SEMAINE [75]	İngilizce	Öfke, İğrenme, Eğlence, Mutluluk, Hüzün, Küçümseme
Danish Em.Data. [76]	Danimarka	Öfke, Sevinç, Hüzün, Sürpriz, Nötr
eINTERFACE'05 [77]	İngilizce	Öfke, İğrenme, Korku, Mutluluk, Hüzün, Sürpriz
BHUDES [78]	Mandarin	Öfke, Sevinç, Hüzün, Tiksinti, Sürpriz

Çizelge 3.1’de verilen sonuçlar incelendiğinde literatürde duygusal durum ile ilgili çok sayıda hazır veritabanı bulunduğu görülebilir. Ancak, anksiyete durumuna odaklanıldığında sadece GEMEP veritabanında yer aldığı görülmüştür. GEMEP veritabanındaki ses kayıtları da aktörler tarafından simüle edilmiş duygulardır.

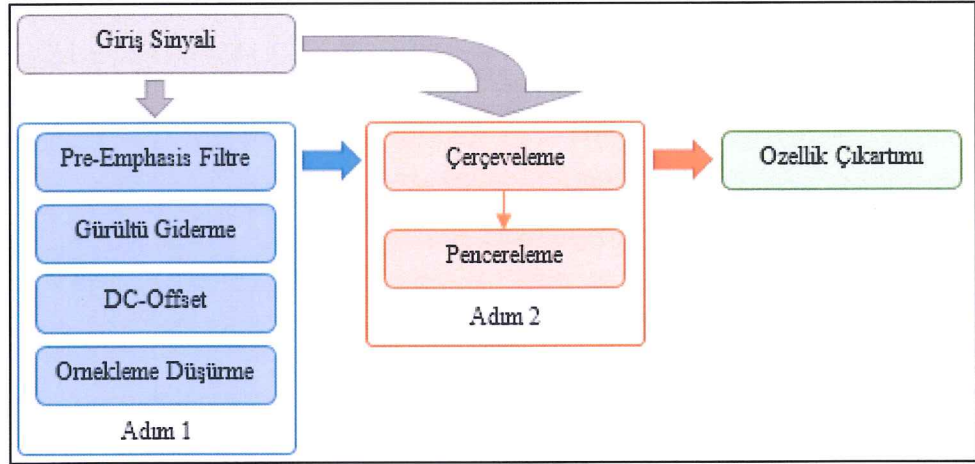
Hazır veritabanları sınıflandırma doğruluğu açısından farklılık göstermektedir. Bu durum duyguların doğru simüle edilmemesinden kaynaklanmaktadır. Ayrıca veritabanını oluşturmak için kullanılan ortam, ses kayıt cihazı, ses kalitesi ve kayıt için kullanılan aktörler de sınıflandırma performansını doğrudan etkilemektedir [79].

Çalışmalarda kullanılacak veriler eğer duygusal duruma göre etiketlenmemiş ise algısal analiz veya dinleyici yorumları ile her bir ses kaydı içerdiği duyguya göre etiketlenmektedir. Bu işleme algısal analiz denilmektedir. Bazı çalışmalarda veri etiketlenmiş olsa dahi algısal analiz kullanılmaktadır. Literatürde duyguların ortaya çıkartılması ve tespiti için duygusal hikaye okuma ve ses kayıtlarının katılımcılara dinletilmesi ile algısal analiz ve dinleme testleri kullanılmıştır [80–85].

3.2. KONUŞMA İŞLEME

Konuşma işleme sayısallaştırılan ses ve konuşmaların DSP ile işlenmesidir. Uzun yıllardır ses kalitesinin artırılması, kişi tanıma, konuşma tanıma, duygu tanıma ve iletişim gibi desen tanıma çalışmalarında kullanılmaktadır.

Şekil 3.2’de verilen akış diyagramında yer alan konuşma işleme adımında sinyal üzerinde filtreleme yöntemleri, örnekleme değerini düşürme, pre-emphasis, gürültü giderme ve DC-offset kaldırma işlemleri gerçekleştirilmektedir. Bu yöntemler dışında stereo-mono dönüşümü de ses kayıtlarının durumuna göre kullanılabilir. Ön işleme adımı Şekil 3.3’de gösterildiği gibi kendi içerisinde de 2 adıma ayrılmaktadır.

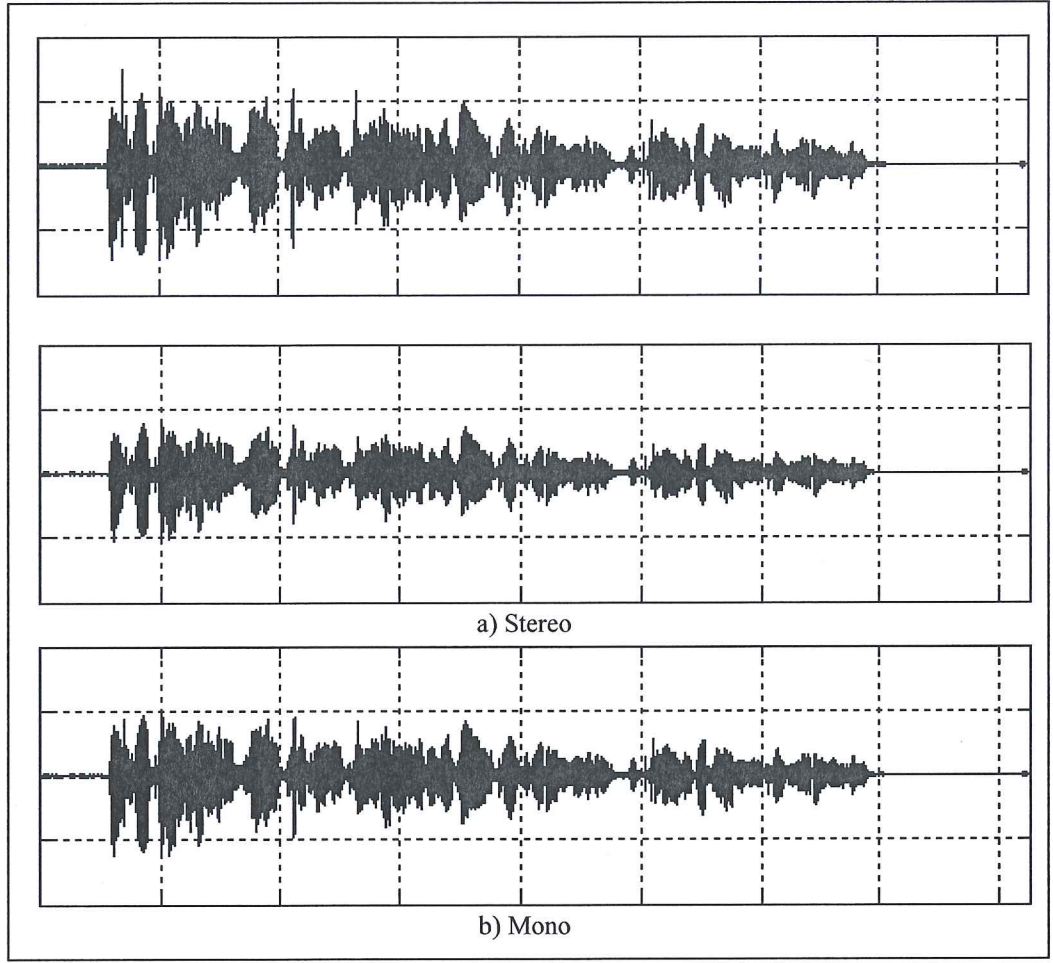


Şekil 3.3. Ön işleme sürecine ait akış diyagramı.

Adım 1’in kullanım zorunluluğu yoktur. Ancak konuşmanın içerdiği ses sinyalleri küçük zaman aralıklarında kararlı kaldığı için sinyal bir bütün olarak işlenmeyip belirli aralıklarla kesitler alınarak işlenmesi gerektiği için Adım 2’de yer alan çerçeveleme ve pencereleme kullanılmak zorundadır.

3.1.1. Stereo-Mono Dönüşümü

Stereo-Mono dönüşümü, ses kayıtlarının stereo olarak kaydedildiği durumlarda kullanılmaktadır. Sayısal ortamlarda kaydedilen konuşmalar kayıt cihazına ve şekline bağlı olarak stereo olarak kaydedilebilmektedir. Ses analiz uygulamalarında mono kayıt tercih edildiği için stereo sesin monoya dönüştürülmesi gerekebilir. Şekil 3.4’de bir konuşmanın stereo ve mono örnekleri verilmiştir.



Şekil 3.4. a) Stereo ses örneği b) Stereo sesin monoya dönüştürülmüş hali.

Bu dönüşüm için hazır araçlar kullanılabileceği gibi kodlama yardımıyla da gerçekleştirilebilir.

Stereo-mono dönüşümü için en basit yöntem stereo ses iki kanal olduğu için kanallardan birisinin seçilmesidir. Ancak bu yöntem ses sinyalinde ciddi kayıplara neden olabileceği için önerilmez. Bu yöntem için kullanılacak formül Eşitlik 3.1’de verilmiştir.

$$X_{mono} = X_{stereo}(1) \vee X_{stereo}(2) \quad (3.1)$$

Kullanılabilecek diğer bir yöntem ise iki kanalın içerdiği değerlerin ortalamasının alınmasıdır. Bu yöntem için kullanılacak formül Eşitlik 3.2’de verilmiştir.

$$X_{mono} = \frac{X_{stereo}(1) + X_{stereo}(2)}{2} \quad (3.2)$$

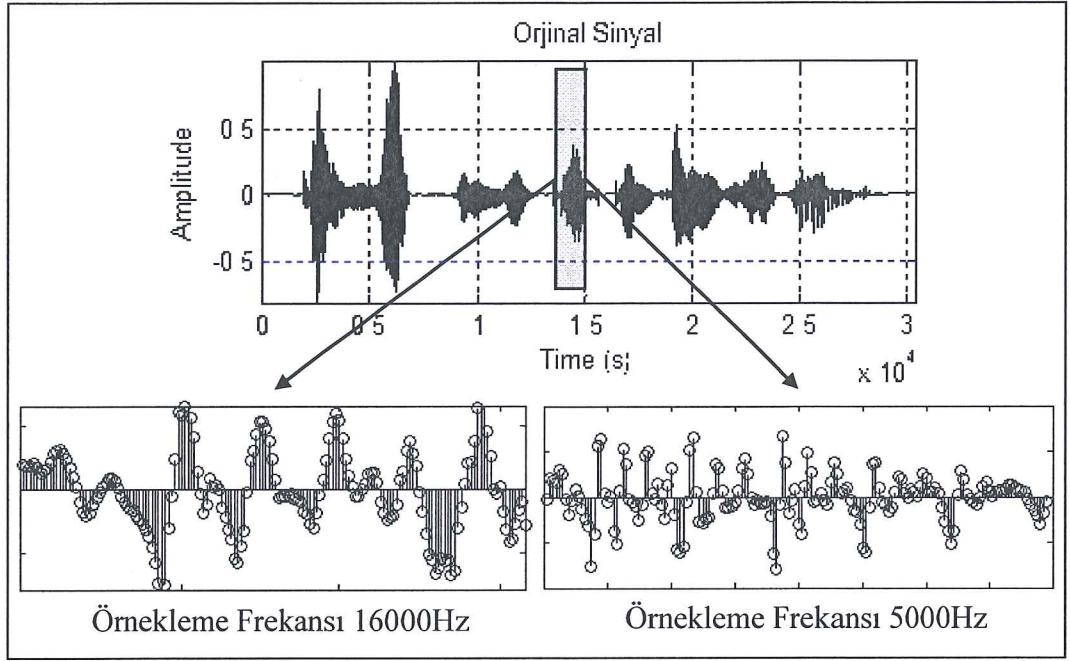
Stereo-mono dönüşümü için kullanılabilir üçüncü yöntem ise iki kanalın içerdiği değerlerin toplamının, toplamın maksimum değerine oranının alınmasıdır. Bu yöntem için kullanılacak formül Eşitlik 3.3'de verilmiştir.

$$X_{mono} = \frac{X_{stereo}(1) + X_{stereo}(2)}{\max(|X_{stereo}(1) + X_{stereo}(2)|)} \quad (3.3)$$

3.1.2. Örnekleme Değerini Değiştirme

Örnekleme, sürekli zamanlı sinyalden ayrık zamanlarda örnekler alınarak ayrık zamanlı sinyale dönüştürülmesidir. Nyquist örnekleme teoremine göre analog bir sinyal sayısallaştırıldıktan sonra tekrar kayıpsız olarak analog sinyale dönüştürülebilmesi için sinyal içerisindeki en yüksek frekanslı bileşenden en az iki kat frekansta örneklenmesi gerekmektedir. Bu amaçla mevcut ses kayıtlarının örnekleme değerleri yüksek ise sistemin iş yükünü azaltmak için örnekleme değeri düşürülebilir. Bu işlem uzmanın deneyimine ve sinyalin kullanım amacına göre belirlenmektedir. Yüksek örnekleme değerleri sorun oluşturmaz ama iş yükünü artırabilir.

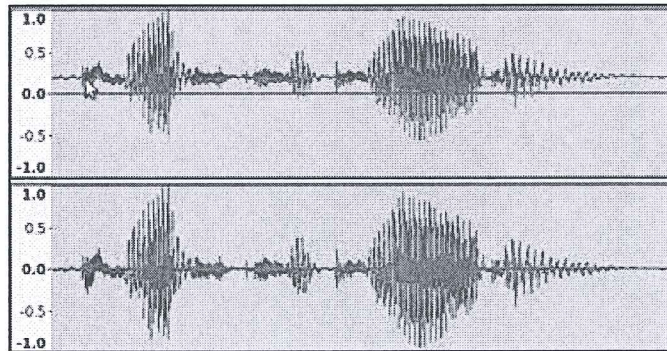
16 kHz örnekleme frekansına sahip bir konuşmanın örnekleme frekansı 5 kHz'e düşürüldüğünde değişim konuşma üzerinden alınan bir kesitle Şekil 3.5'de verilmiştir.



Şekil 3.5. Örnekleme frekansı değiştirildiğinde sinyaldeki değişim.

3.1.3. DC-Offset Kaldırma

Zaman boyutundaki periyodik bir fonksiyon için DC offset dalganın ortalama değeridir. Eğer ortalama genlik değeri 0 ise DC offset yoktur. Yani bir sinyalin pozitif ve negatif bölgedeki sinyal miktarının eşit olmadığı durumdur. Ses kayıtlarında DC offset istenmeyen bir özelliktir [86]. Şekil 3.6'da üstteki sinyal DC offset içerirken alttaki içermemektedir.



Şekil 3.6. DC offset içeren ve içermeyen sinyal örneği [87].

DC Offset Eşitlik 3.4’de gösterildiği gibi sinyalin zaman serisi örneklerinden ortalamanın çıkartılması ile kaldırılabilir [86].

$$x[n] = x[n] - \frac{1}{N} \sum_{n=1}^N x[n] \quad (3.4)$$

3.1.4. Pre-Emphasis/De-Emphasis Filtre

Gürültü sinyal işlemedeki en önemli sorunlardan birisi olmakla birlikte kişinin vurgularının önemli olması gibi bazı durumlarda sorun olmaktan çıkmaktadır. Sinyalin kalitesi başka bir deyişle sinyal ve gürültü arasındaki orana Sinyal Gürültü Oranı (SNR) denilir. SNR değerinin mümkün olduğunca yüksek olması istenen bir durumdur.

Bir konuşma veya ses sinyali temel frekans ve harmoniklerden oluşur. Harmonikler temel frekansa göre daha zayıf ve düşük genlik değerlerine sahip olduğu için gürültüden çok fazla etkilenmektedir. Bu durum SNR değerinin düşmesine neden olur. Bu sorunun üstesinden gelmek için pre-emphasis yöntemi ile harmonikler güçlendirilerek SNR değeri yükseltilir. Harmoniklerin yükseltilmesi SNR değerini de yükseltir ancak sesin karakteristiğinde değişiklikler meydana gelir. Bu nedenle, pre-emphasis ile SNR yükseltildikten sonra de-emphasis uygulanarak harmonikler zayıflatılır ve ses eski haline dönüştürülür. Harmoniklerdeki bu zayıflatma doğal olarak sinyaldeki gürültüyü de zayıflatacaktır. Pre-emphasis işleminin transfer fonksiyonu Eşitlik 3.5’de verilmiştir [69].

$$H(z) = (1 - 0,9375^{-1}) \quad (3.5)$$

Konuşma sinyalinin sesli bölümleri negatif spektral eğime sahiptir. Pre-emphasis filtre spektral analiz öncesi bu eğimi dengeleyerek analiz verimliliği artırması [88], işitme spektrumu 1kHz bölgesinin üstünde daha hassas olduğu için pre-emphasis filtre spektrumun bu alanını güçlendirmesi ve konuşma spektrumunun algısal önemli özniteliklerini modellemede spektral analiz algoritmasına yardımcı olması [89] pre-emphasis filtrenin avantajlarıdır [90].

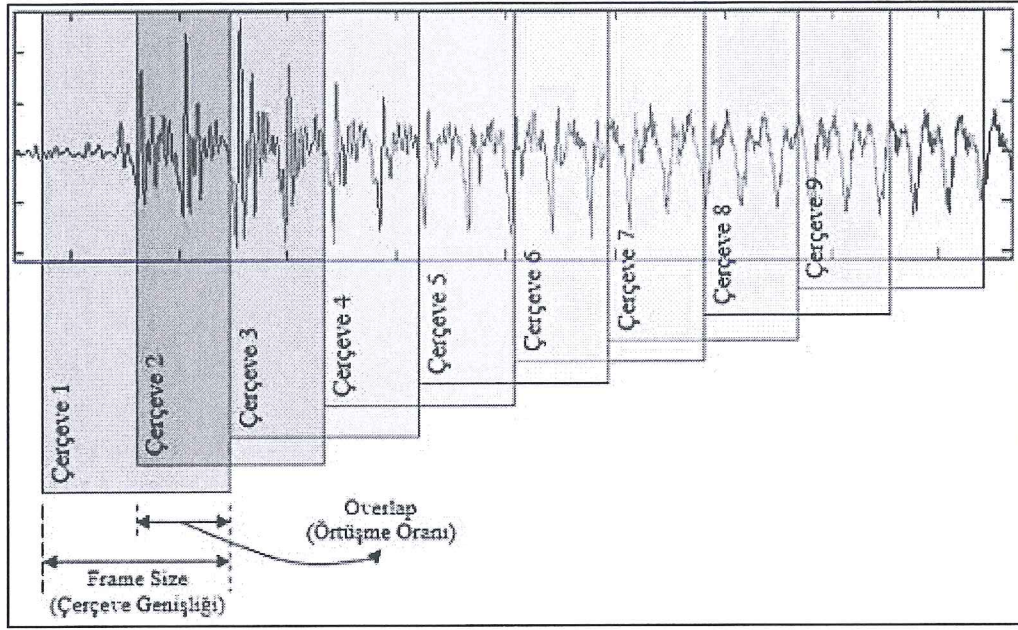
3.1.5. Gürültü Giderme

Gürültü iletişim, ölçme ve sinyal işleme gibi uygulamalarda sinyalde bozulmalara neden olan istenmeyen sinyaller olarak tanımlanmaktadır [91]. Bir gürültünün sinyali etkileme oranı SNR ile tanımlanmaktadır. Bu değerin yüksek olması gürültünün düşük olduğunu gösterir. Gürültü konuşma analizi içeren sistemlerde sistemin performansını düşürmektedir. Gürültü frekans, spektrum veya zaman karakteristiğine bağlı olarak çeşitli kategorilere ayrılabilir [92].

Son 20 yılda gürültü azaltmak için çeşitli filtreleme yöntemleri önerilmiştir. Gürültü giderme yöntemi gürültüyü tamamen kaldırmak için ana sinyalin yapısında değişikliğe neden olmamalıdır. Bu nedenle, birçok yöntem 10-20 dB aralığında gürültünün bastırılmasını sağlar [93].

3.1.6. Çerçeveleme

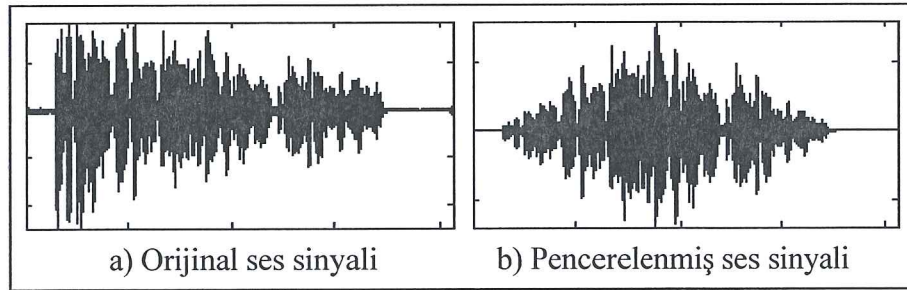
Konuşmanın içerdiği ses sinyalleri küçük zaman aralıklarında kararlı kaldığı için sinyal bir bütün olarak işlenmeyip belirli aralıklarla kesitler alınarak işlenmektedir. Çerçeveleme yapılırken pencere boyutunun yanı sıra örtüşme oranı da belirlenmektedir. Örtüşme oranı bir çerçeveden diğerine geçişi yumuşatmak ve bilgi kaybını önlemek amacıyla kullanılmaktadır. Rabiner ve Schafer'a göre çerçeve boyutu 10-20ms seçilmesi uygundur [88]. Aşağıdaki şekilde konuşma sinyalinin örnek bir çerçevelemesi gösterilmiştir. Şekil 3.7'de verilen örtüşme oranı pencerelerin belirli bir oranda bir öncekini kapsamaları için kullanılır.



Şekil 3.7. Konuşma sinyali için çerçeveleme örneği.

3.1.7. Pencereleme

Pencereleme sinyaldeki spektral sızmayı ve örtüşmeden dolayı oluşan kesişimi düzenlemek amacıyla kullanılmaktadır [21]. Sinyalin çerçevelemesi sonrası her çerçeve pencerelenerek çerçevenin başındaki ve sonundaki süreksizlikler kaldırılmış olur [94]. Çeşitli pencereleme fonksiyonları mevcut olup bunlar bir önceki bölümde verilmiştir. Şekil 3.8'de örnek bir ses için orijinal ve pencereleme sonrası durum verilmiştir.



Şekil 3.8. Örnek bir ses sinyalinin orijinal hali a)'da, pencerelenmiş hali b)'de verilmiştir.

3.1.8. Öznitelik Çıkartımı

Öznitelik çıkartım aşamasında çerçevelere bölünen ses sinyalinin her çerçevesinden öznitelikler çıkartılmaktadır. Bu özniteliklerin tespiti için kullanılacak hazır araçlar olduğu gibi araştırmacılar tarafından geliştirilen kodlar da bulunmaktadır. Konuşma öznitelikleri akustik öznitelikler, dilsel öznitelikler, bağlamsal bilgi ve akustik öznitelikler ile diğer bilgi kaynaklarının birleşimini içeren hibrit öznitelikler olmak üzere dört kategoride gruplandırılabilir [95]. Çoğunlukla kullanılan akustik öznitelikler pitch, formant frekansları, jitter, shimmer, ses şiddeti, Zero-Crossing Rate, konuşma oranı, ses kalitesi, MFCC, LPCC, dalgacık katsayısı ve Teager Enerji Operatörü (TEO) tabanlı özniteliklerdir. Özniteliklerle ilgili ayrıntılı bilgi ilerleyen kısımlarda verilmiştir.

3.1.9. Son İşleme

Öznitelik çıkartımı sonrası son işleme adımında elde edilen öznitelik verisine normalizasyon ve öznitelik seçimi yöntemleri uygulanmaktadır. Sınıflandırıcıların eğitimi ve test aşamasında kullanılacak olan verilerdeki birim farklılıkları ve verinin sayısal büyüklüğü sınıflandırıcı performansını doğrudan etkilemektedir. Bu sorunun üstesinden gelmek için öznitelik normalizasyon teknikleri kullanılmaktadır.

Öznitelik seçim yöntemleri, öznitelik kümesinden en iyi sınıflandırmayı gerçekleştirecek özniteliklerin belirlenmesi amacıyla kullanılır. Öznitelik seçimi ile öznitelik veri kümesinin boyutu azaltılarak sınıflandırma performansı ve sınıflandırma doğruluğu artırılmaktadır. Çalışmalarda çoğunlukla Temel Bileşen Analizi (PCA), Lineer Diskriminant Analizi (LDA) ve İleri Seçim (FS) tabanlı yöntemler kullanılmıştır.

Öznitelik seçim yöntemlerinin kullanımı her zaman sınıflandırıcı başarısını artırmamaktadır. Grimm vd. tüm öznitelik kümeleri ve PCA ile seçilen öznitelik kümeleri üzerinde yaptıkları çalışmada tüm özniteliklerin kullanımında daha yüksek başarı elde etmiştir [96]. Hoque vd. PCA ve LDA kullanarak öznitelikleri düşük boyutlu uzayda yansıtmıştır ve PCA ve LDA 'nın birlikte kullanımının ayrı ayrı

kullanıma göre daha iyi sonuçlar verdiğini tespit etmişlerdir. Çünkü PCA ilintisizleştirilmiş veri üzerinde ve LDA düşük boyutlu veri üzerinde etkindir [97]. İki öznitelik indirgeme yöntemi (Fisher ve PCA) ve iki sınıflandırıcı kullanarak karşılaştırmalı dört deney gerçekleştirilen çalışmanın sonuçlarına göre boyut azaltmada Fisher PCA'dan daha iyidir [95].

3.1.10. Sınıflandırma

Sınıflandırma, veriden elde edilen öznitelikler yardımıyla verinin sınıflandırılması işlemidir. Sınıflandırma aşamasında konuşmadan çıkartılan öznitelikler kullanılarak konuşma metin, duygu, içerik gibi durumlara göre sınıflandırılmaktadır. Sınıflandırma işlemi için Gauss Karışım Modeli (GMM), Destek Vektör Makineleri (SVM), Saklı Markov Modeli (HMM), Yapay Sinir Ağları (ANN), k-En Yakın Komşu (k-NN) ve hibrit sınıflandırıcılar gibi kullanılacak çeşitli sınıflandırıcılar mevcuttur. Geleneksel sınıflandırma teknikleri duygu tanıma sistemlerinin hemen hemen hepsinde kullanılmıştır. Güncel çalışmalar hibrit sınıflandırıcılar ve sınıflandırıcıların akustik öznitelikler üzerindeki etkilerine odaklanmaktadır. Literatürde yer alan çalışmaların büyük bir kısmında SVM sınıflandırıcı kullanılmıştır [94,95,98–111]. SVM sınıflandırıcıyı GMM [81,83,105,111–123] ve HMM [101,109,117,120,124–126] sınıflandırıcılar takip etmiştir.

SVM, istatistiksel öğrenme teorisi temelli sınıflandırma algoritmasıdır. SVM temel olarak iki sınıfı birbirinden ayırabilen en uygun hiper-düzlemin tanımlanması esasına dayanır. SVM temel olarak eğitim ve sınıflandırma olmak üzere iki aşamada gerçekleşmektedir. Eğitim aşamasında sınıfları birbirinden ayıracak özellikteki veriler öznitelik vektörü ile ifade edilir ve kernel fonksiyonları ile sınıfları ayıran hiper-düzlem tespit edilir. Eğitim sonrasında hiper-düzlem bulunduktan sonra sınıflandırma işlemi gerçekleştirilir.

GMM, ayrık olasılık dağılımlarını modellemek için çeşitli Gauss dağılımlarını kullanır, bu yüzden Expectation Maximization algoritması ile optimize edilebilir.

HMM, temel dinamik Bayesian network olarak kabul edilebilir. HMM bir örüntüye ait çerçevelerin spektral öznitelikleri, farklı durumlar arasındaki geçiş olasılıkları ve bu durumlara ilişkin öznitelik vektörleri arasındaki koşullu olasılıkları kullanır. Konuşma sinyalinin üretim mekanizması fiziksel olduğu için HMM sınıflandırıcı yaygın olarak kelime ve konuşma tanımadan izole edilmiş konuşma uygulamalarında kullanılır.

3.2. AKUSTİK PARAMETRELER

Duygu tanımadaki en önemli konu farklı duyguları ifade eden özniteliklerin tespit edilmesi ve bu özniteliklerin duygu durumuna göre değişimlerinin tespit edilmesidir. Bu özniteliklerden akustik özniteliklere ait parametreler duygu tanımada kullanılmaktadır. Belirli duygular için akustik özniteliklerin belirlenmesindeki zorluk yaş ve cinsiyet dahil olmak üzere sesteki bireysel farklılıklardır [127].

Konuşma öznitelikleri birkaç kategoriye ayrılabilir. Steidl vd. akustik parametreleri bürünsel (prosodic), spektral ve ses kalitesi olmak üzere 3 kategoriye ayırmıştır [128]. Ayadi vd. konuşma özniteliklerini sürekli (continuous), nicel (qualitative), spektral ve TEO tabanlı olmak üzere 4 kategoriye ayırmıştır [79].

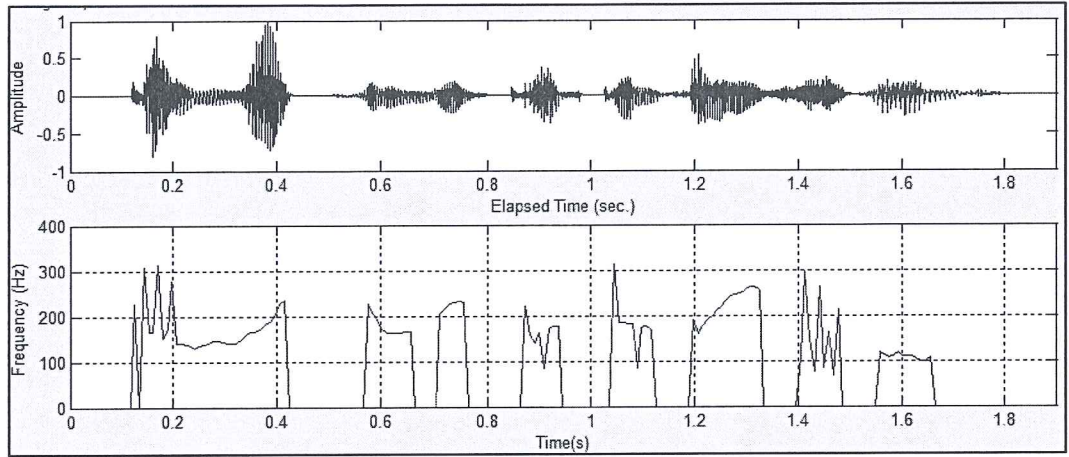
Bürünsel öznitelikler ritim, stres, konuşma tonlamaları ve konuşmanın büyük kesimleri üzerinden elde edilen özniteliklerdir. Pitch, ses şiddeti, konuşma oranı, konuşma süresi, bekleme ve ritim bürünsel öznitelikler kümesinde yer alır. Konuşma sinyalinin bürünsel öznitelikler dışında frekansla ilgili diğer karakteristikleri spektral öznitelik kümesinde yer alır. Formant frekansları, MFCC, Log Frekans Güç Katsayıları (LFPC), LPCC ve Bağlı Spektral Dönüşüm-Algısal Doğrusal Tahmin (RASTA-PLP) spektral öznitelik kümesinde yer alır. Yapılan çalışmalar ses kalitesi öznitelikleri ve kişinin duygu durumu arasında güçlü bir ilişki olduğunu göstermektedir. Sesin üretimi esnasında ses yolunun etkilerini yok etmek ve konuşma sinyalini elde etmek için ters filtreleme kullanılmaktadır. Jitter ve shimmer ters filtreleme için bir alternatiftir. Shimmer tepe veya ortalama genlik değişimini ölçerken, jitter periyot genişliğinin değişimini ölçer. Ses kalitesi öznitelik kümesinde

yer alan diğer öznitelikler Harmonik Gürültü Oranı (HNR) ve Gürültü Harmonik Oranı (NHR) öznitelikleridir [129].

Akustik parametreler ses sinyallerinden DSP yöntemleri ile çıkartılan özniteliklerdir ve en çok kullanılan akustik parametreler aşağıda verilmiştir.

3.2.1. Pitch

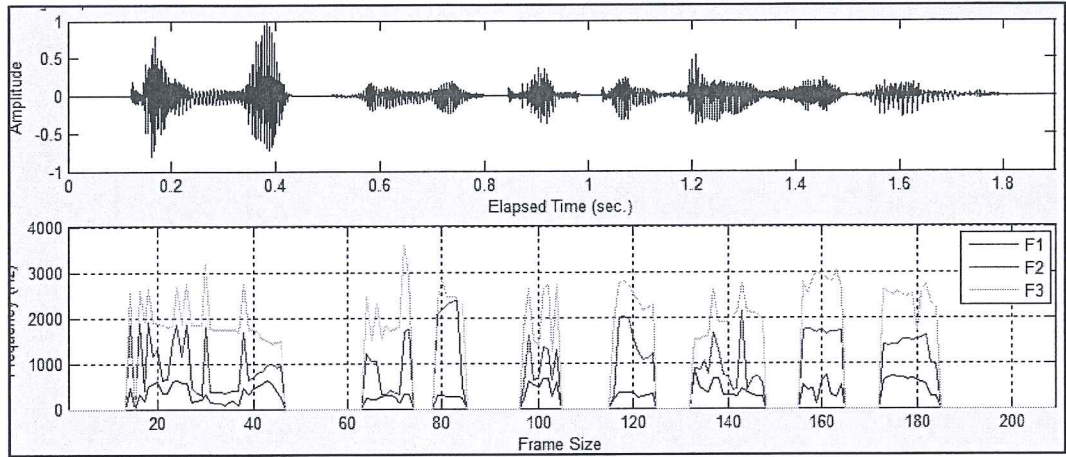
Pitch gırtlaksı uyarılmanın temel frekansıdır ve ses kıvrımları ve alt gırtlak hava basıncına bağlıdır. Sesin kalınlık ve inceliğini bildirir. Periyot, pitch algısında birincil faktördür. Döngü/saniye olarak ölçülür ve Hertz (Hz) ile ifade edilir. Bu değer ergenlik öncesi kız ve erkeklerde 220-240 Hz civarında iken erişkin erkekler ve kadınlarda sırası ile ortalama 100-150 Hz ve 150-250 Hz arasındadır [7]. Pitch duygu tanıma çalışmalarında çoğunlukla kullanılan özniteliklerden birisidir. Pitch, F0 olarak da bilinmektedir [4]. Ses sinyalinin konuşma içeren kısımlarında pitch özelliğinin minimum, maksimum, standart sapma, aralık ve regresyon katsayısı gibi varyasyonları da kullanılmaktadır [130]. F0 ses programları ile tespit edilebildiği gibi spektrografi veya elektroglotografi ile de tespit edilebilir. Literatürde pitch tespiti için geliştirilmiş çeşitli algoritmalar mevcuttur. Pitch tespit algoritmaları zaman boyutu, frekans boyutu veya her ikisinde de kullanılabilir [131]. Zaman boyutunda kullanılan yöntemler; ZCR, otokorelasyon tabanlı yöntemler. Frekans boyutunda kullanılan yöntemler; harmonic product spectrum, kepstrum analizi. Pitch öznitelik kümesinde F0 dışında gırtlak hız düzeyi de duygu tespitinde kullanılmaktadır. Yüksek değerler mutluluk, neşe veya sürpriz duygularını, düşük değerler öfke ve nefret duygularını göstermektedir. Gırtlak hız düzeyini hesaplamak için F0 hesaplama yöntemlerinden otokorelasyonun maksimum değeri kullanılır [132]. Örnek bir konuşma için elde edilen F0 değerleri Şekil 3.9'da verilmiştir.



Şekil 3.9. Örnek bir konuşma için F0 grafiği.

3.2.2. Formant Frekansları

Formant ses yolundaki rezonanstır ve ses yolunun nicel özellikleri ile ilgili spektral bilgi sağlamaktadır. Teorikte sonsuz sayıda format vardır ama pratikte yalnızca ilk 3 veya 4 formant önemli bilgiler içerir. Formantlar F1 (Birinci Formant Frekansı), F2 (İkinci Formant Frekansı) ve F3 (Üçüncü Formant Frekansı) gibi formant sayısı ile tanımlanır. F1, "i" ve "u" gibi yüksek seslilere ait düşük frekansa sahiptir ve "a" ve "ae" gibi düşük sesliler bu formantın yüksek frekanslarına sahiptir. Yani F1 frekansı sesli harflerin dil vuruşu yüksekliği ile ters orantılı olarak değişir. Diğer yandan "a" ve "u" kalın seslileri F2'nin düşük frekanslarını paylaşır ve "i" ve "ae" ince seslileri F2'nin yüksek frekanslarını paylaşır. Yani, F2 sesli artikülasyonunun posterioranterior (akciğer ile ilgili bir ifade) boyutuyla değişiklik gösterir [133]. Her formant kendi merkez frekansı ve bant genişliği ile karakterize edilir ve duygu hakkında önemli bilgiler içerir. Örneğin, insanlar stres ve depresyon altında sesli sesleri üretmez ve aynı durum nötr duygu durumunda da gerçekleşir. Sesteki bu değişim formant bant genişliklerinde farklılıklara neden olur. Formant frekansları ve bant genişliklerinin tespiti için Lineer Kestirimci Kodlama (LPC) en yaygın kullanılan yöntemdir [132]. Örnek bir konuşma için elde edilen formant frekansları Şekil 3.10'da verilmiştir.



Şekil 3.10. Örnek bir konuşma için formant frekansları.

3.2.3. Jitter

Periyotlar arası değişikliği gösteren parametredir. Temel frekanstaki istem dışı ortaya çıkan düzensizlikleri içerir [2]. Bu parametre ardışık titreşimli döngüleri arasındaki temel frekansın değişimleri olarak tanımlanır [134]. Jitterdeki değişimler ses kıvrımlarındaki patolojileri gösterir. Mutlak jitter analiz edilen sesin her periyodunun, kendinden sonraki periyot ile farkının mutlak değerinin ortalaması olup F_0 'a göre değişiklik gösterir (Eşitlik 3.5). Birimi μs 'dir. Yüzde jitter mutlak jitter'in F_0 'a bağlı olarak değişiklik göstermesi sakıncasını ortadan kaldırmak için ortalama periyoda bölünmesi ile elde edilir (Eşitlik 3.6). Normal değeri % 1'in altındadır [7]. Ses hastalığı olanlarda bu değer genellikle artmış olarak bulunur.

$$J_{mutlak} = \frac{1}{T-1} \sum_{t=1}^{T-1} |T_0(i) - T_0(i+1)| \quad (3.5)$$

$$J_{\%} = \frac{J_{mutlak}}{\frac{1}{T} \sum_{i=1}^T T_0(i)} \quad (3.6)$$

Eşitlik 3.6'da T periyot sayısını, T_0 ise temel periyodun süresini göstermektedir.

3.2.4. Shimmer

Temel frekans pertürbasyonunda olduğu gibi ses sinyallerindeki çok kısa süreli genlik değişiklikleri ölçülür [2]. Genlik pikleri arasındaki periyodik varyasyona ise shimmer adı verilir. Bu parametre, ardışık titreşimli döngüler arasındaki gırtlak akış genlik değişiklikleri olarak tanımlanır [135]. Mutlak shimmer (Desibel) ya da yüzde shimmer (%) olarak ifade edilir. Mutlak shimmer 'de her periyodun tepe genliği bir sonraki periyodun tepe genliği ile karşılaştırılır (Eşitlik 3.7) ve dB cinsinden shimmer hesaplanır [2]. Kısa aralıklarla ses dalgasının genlikleri arasındaki rölatif değişikliği göstermektedir. Yüzde shimmer her periyodun kendinden sonraki periyotla arasındaki şiddet farkının mutlak değerinin ortalamasını ortalama periyot şiddetine bölerek elde edilir (Eşitlik 3.8), normal değeri 3'ün altındadır. Ses hastalığı olanlarda bu değer genellikle artmış olarak bulunur.

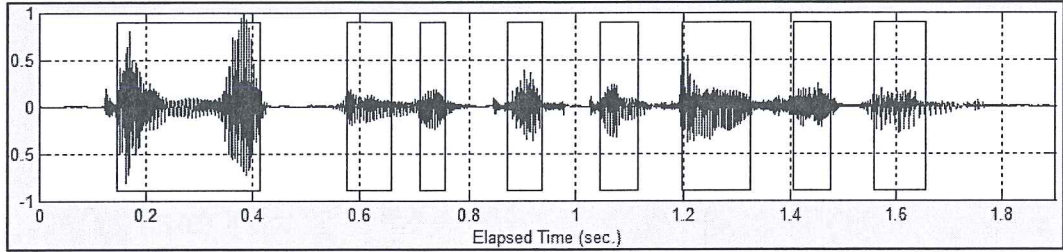
$$S_{dB} = \frac{1}{T-1} \sum_{i=1}^{T-1} \left| 20 \log \left(\frac{Amp(i)}{Amp(i+1)} \right) \right| \quad (3.7)$$

$$S_{\%} = \frac{\frac{1}{T-1} \sum_{i=1}^{T-1} |Amp(i) - Amp(i+1)|}{\frac{1}{T} \sum_{i=1}^T Amp(i)} * 100 \quad (3.8)$$

3.2.5. Konuşma Oranı

Konuşmanın hem sesli hem de sessiz bölümleri ile ilgili zamansal öznitelikler sağlar [132]. Bu öznitelikler konuşma oranı, süresi ve duraksamalar gibi konuşma stili ile ilişkilidir. En çok kullanılan konuşma süresi özelliği konuşma oranıdır ve ses sinyalinin konuşma içeren bölümlerini gösterir [129]. Konuşma gerçekleşirken konuşma süresince duraksamalar olabilir. Yüzde duraksama süresi, 250 ms üzerindeki tüm duraksamaların toplam süreye bölünüp, 1'den çıkartılıp 100 ile çarpılmasıyla bulunur. 250 ms üzerindeki duraksamalar konuşma kalıplarının zihinsel ve duygusal yönlerini yansıttığı düşünülmektedir [136]. Duraksama sayısı, konuşma gerçekleşirken duraksamaların sayısını gösterir. Duraksamanın ortalama

süresi ise toplam duraksama süresinin duraksama sayısına oranıdır [137]. Örnek bir konuşma için elde edilen konuşma oranları Şekil 3.11’de verilmiştir.



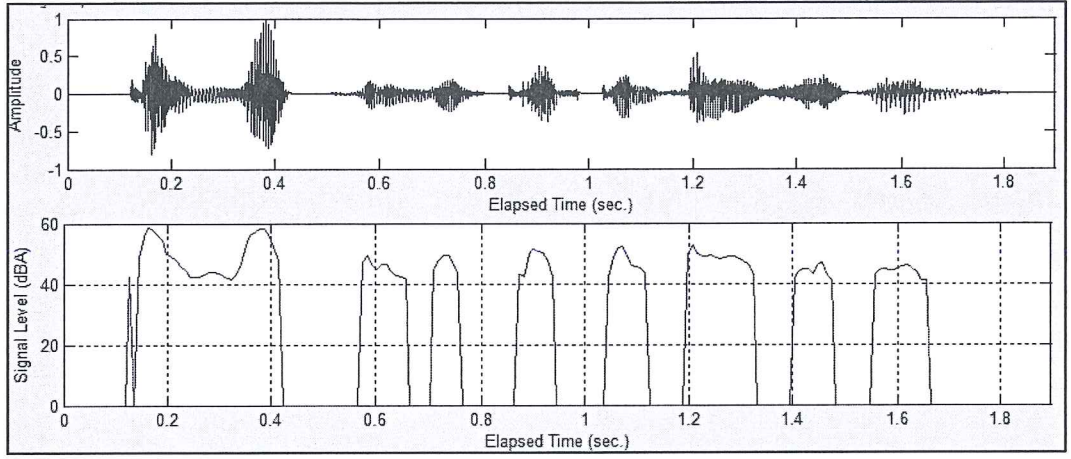
Şekil 3.11. Örnek bir konuşma için sesli ve sessiz bölümlerin gösterimi.

3.2.6. Ses Kalitesi (HNR Ve NHR)

HNR ve NHR, sesin spektrumunda yer alır ve F_0 'ın katlarından oluşan harmonikler dışındaki gürültü sesleri ile ilgilidir. Bir konuşma sinyalinde temel frekansın tam katları harmonikleri diğerleri ise gürültüyü oluşturur. HNR, F_0 ve onun katları olan harmoniklerin toplam enerjisinin gürültü enerjisine oranıdır. Birimi dB'dir. Yüksek değerler sesteki gürültü oranının düşük olduğunu gösterir. NHR, HNR parametresinin modifiye edilmesi ile elde edilir. Değeri HNR'nin aksine, sesteki gürültü miktarı ile doğru orantılı olarak değişir.

3.2.7. Ses Şiddeti

Ses şiddeti konuşmanın içerdiği enerjinin ölçümüdür. Pitch gibi uyarıma bağlı değişir ve ses yolu filtreleme modelinden bağımsızdır. Diğer taraftan bir dinleyici tarafından algılanan konuşmanın ses seviyesi konuşmacı ve dinleyici arasındaki mesafe ve konuşmanın şiddetine ve kulak zarındaki ses dalgalarının Ses Basınç Seviyesi (SPL) değerine bağlıdır. Kaydedilmiş bir ses kullanıldığında ses şiddeti mikrofonun SPL 'sine bağlıdır ve bu yüzden ses şiddeti sinyalin enerjisi ile karakterize edilir [38]. Sonuç olarak, enerji öznelikleri duyuların uyarılma düzeyi hakkında bilgiler içerir ve ses şiddeti uygulanarak elde edilir [132]. Örnek bir konuşma için elde edilen enerji dağılımı Şekil 3.12’de verilmiştir.



Şekil 3.12. Örnek bir konuşma için enerji dağılımı.

BÖLÜM 4

KONUŞMA İŞLEME YAZILIMI

Konuşma işleme bir sinyal işleme süreci olup temel olarak ön işleme ve öznelik çıkartımı süreçlerini içermektedir. Literatürde bu amaçla geliştirilmiş çeşitli yazılımlar mevcut olup Bölüm 2’de açıklanmıştır. Bu yazılımlar incelendiğinde elde edilebilen öznelik kümeleri arasındaki farklılıklar, kullanım zorlukları, toplu dosya işleme yeteneği gibi öznelikler ve bu araçların içermediği öznelik kümelerini de elde etme ihtiyacı göz önünde bulundurularak yeni bir yazılım geliştirme ihtiyacı oluşmuştur. Bu bağlamda, seslerden akustik parametrelerin tespiti için MATLAB ortamında SPAC isminde bir yazılım geliştirilmiştir. SPAC, 6 modülden oluşmaktadır. Modüler yapıya sahip olduğu için her bir modül bağımsız olarak kullanılabilir. SPAC pitch, formant, ses kalitesi, LPCC, MFCC, ZCR, enerji, konuşma ve dalgacık katsayıları olmak üzere 9 farklı ana kategoride öznelik çıkartılabilmektedir.

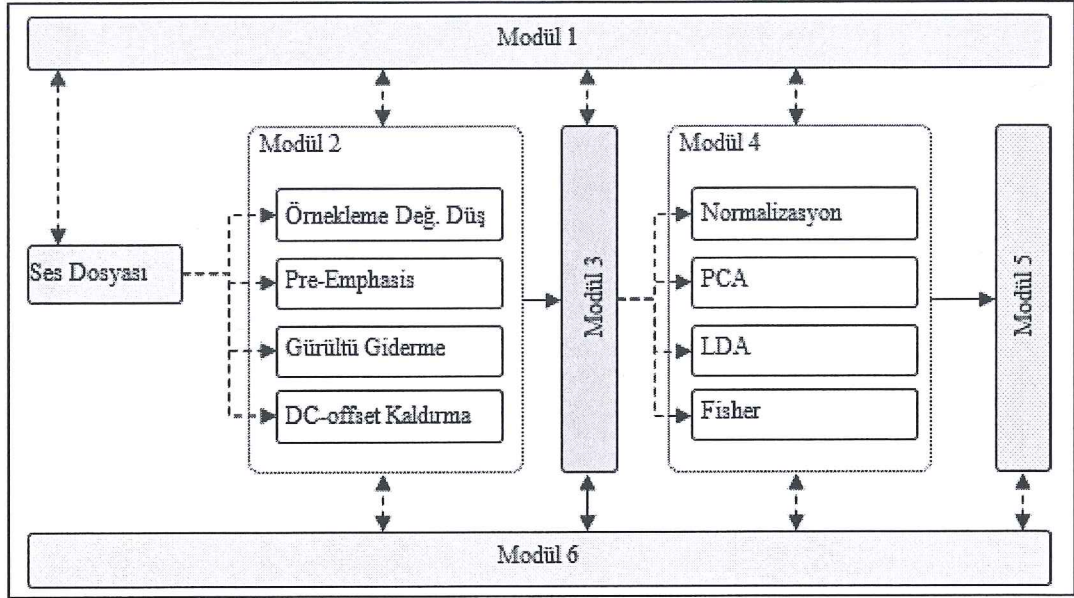
4.1. SPAC TASARIMI

SPAC konuşma işleme amaçlı MATLAB ortamında geliştirilmiş bir yazılımdır. Grafikselleşmiş ara yüzü ve sonuçların grafikselleşmiş gösterimi sayesinde kullanıcı dostu bir yapıya sahiptir. Grafikselleşmiş ara yüz sayesinde parametre tespit etmek diğer sistemlere göre SPAC üzerinde daha kolaydır. Modüler olarak tasarlanmış olup 6 modül içermektedir. Bunlar;

1. Modül 1: Konfigürasyon
2. Modül 2: Ön işleme
3. Modül 3: Öznelik çıkartımı,
4. Modül 4: Son işleme,
5. Modül 5: Sınıflandırma

6. Modül 6: Toplu dosya işleme

SPAC'a ait modüller yapı Şekil 4.1'de verilmiştir.



Şekil 4.1. SPAC modüler yapısı.

Modül 1 konuşma işleme için kullanılacak çerçeve boyutu, pencereleme ve örtüşme gibi parametrelerin ayarını içermektedir. Modül 2 pre-processing adımı olup öznelik çıkartımı öncesi sinyalin iyileştirilmesi için yöntemler içermektedir. Modül 3 öznelik çıkartımı adımıdır ve tek veya birden fazla dosya üzerinde kullanılabilir. Modül 4 elde edilen öznelikler üzerinde öznelik normalizasyonu ve öznelik seçimi için yöntemler içermektedir. Modül 5 üzerinden gerek SPAC ile elde edilen öznelik kümesi gerekse de harici bir dosya içerisindeki öznelikler kullanılarak sınıflandırma gerçekleştirilebilir. Modül 6 birden fazla dosya üzerinde konuşma işleme ve öznelik çıkartım işlemlerinin gerçekleştirilmesine olanak sağlar. SPAC modüller olarak tasarlandığı için her modül bağımsız olarak kullanılabilir. Örneğin, başka bir yazılım ile elde edilen öznelikler SPAC üzerinde sadece sınıflandırma modülü ile sınıflandırılabilir veya ses kalitesini artırmak için sadece ön işlem modülü kullanılabilir. SPAC ile örnek bir ses dosyası işlenmesine ait ekran görüntüsü Şekil 4.2'de verilmiştir.

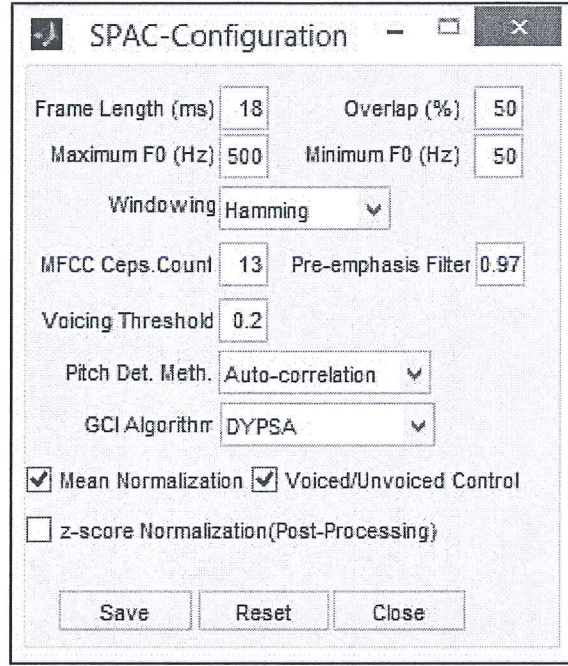


Şekil 4.2. SPAC ile örnek bir ses dosyasının işlenmesi.

Şekil 4.2’de örnek bir ses dosyasının SPAC ile işlenmesine ait ekran görüntüsü verilmiştir. Original Speech bölümü işlem yapılan ses dosyasını ve Analysis bölümü ses dosyasından öznitelik çıkartımı ve görselleştirilmesini içerir. SPAC ile her bir öznitelik şekilde de verildiği gibi Kullanıcı Grafik Arayüzü (GUI) aracılığıyla bağımsız olarak incelenebilir ve pre-processing uygulanabilir.

4.1.1. Modül 1: Konfigürasyon

Bu modül konuşma işleme ve öznitelik çıkartımı için kullanılacak ayarlamaları içermektedir. Modül 1’e ait ekran görüntüsü Şekil 4.3’de verilmiştir.



Şekil 4.3. Modül 1 grafiksel arayüzü.

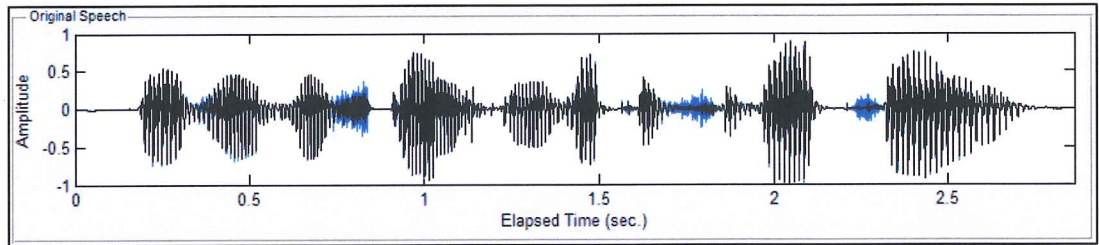
Konuşma küçük zaman aralıklarında kararlı kaldığı için konuşma sinyali çerçeveslere bölünerek işlem gerçekleştirilir. Frame Length parametresi ile konuşma işleme için kullanılacak çerçeve uzunluğu belirlemektedir. Genellikle 10-20 ms arasında olması tercih edilir [88]. Konuşma sinyalinin her bir çerçevesi işlenirken veri kaybı oluşmaması için bir sonraki çerçeve bir öncekinin belirli bir kısmını kapsar. Bu işleme örtüşme adı verilir. Overlap parametresi örtüşme oranını göstermektedir. Sinyal üzerindeki spektral sızmayı ve örtüşmeden dolayı oluşan kesişimi düzenlemek için pencereleme kullanılır. SPAC üzerinde pencereleme için Hamming, Hann, Bartlett, Gaussian ve Rectangular yöntemleri kullanılabilir. İnsanların konuşmaları belirli bir aralıkta F0 değerine sahiptir. Bu nedenle sinyal üzerindeki insan konuşmaları dışındaki frekansları dikkate almamak için konuşma işlemede kullanılacak frekans aralığı Maximum-Minimum F0 ile belirlenebilmektedir. MFCC Cepstrum Count, sinyal üzerinden MFCC özneliklerini tespit ederken kullanılacak MFCC katsayısını belirler. Pre-emphasis Filter, filtre için kullanılacak katsayı değerini belirler. Voicing Threshold, sesteki konuşma içeren ve içermeyen bölgelerin tespit edilmesi için kullanılacak enerji eşik değeridir. Pitch Detection Method, F0'ın tespit edilmesi için kullanılacak yöntemin seçilmesini sağlar. Pitch tespiti için otokorelasyon ve kepstrum olmak üzere iki yöntem sunulmuştur. Ses sinyalinin

gırtlakta oluşum şeklini tespit etmek için ters filtreleme kullanılmaktadır. Böylece sesin gırtlakta oluşum şeklinin elde edilmesi ile Glottal Closure Instant (GCI) yani gırtlığın kapanma anları tespit edilebilmektedir. GCI noktalarının tespit edilmesi için DYPSA [53] ve SE_VQ [138] olmak üzere iki yöntem SPAC içerisinde yer almaktadır. Modül üzerinde yer alan GCI Algorithm ile kullanılacak yöntem seçilmektedir. Buradan elde edilen GCI değerleri jitter parametresinin tespitinde kullanılmaktadır. Sinyalden sinyalin ortalama değeri çıkartılarak normalleştirme uygulanıp uygulanmayacağını Mean Normalization ile belirlenir. Voiced/Unvoiced Control, ses sinyali üzerinde konuşma içeren ve içermeyen bölümlerin dikkate alınıp alınmayacağını belirler. Çıkarılacak akustik özneliklerde bu seçim önem arz etmektedir. z-score normalization, SPAC ile elde edilen özneliklere normalizasyon uygulanıp uygulanmayacağını belirler.

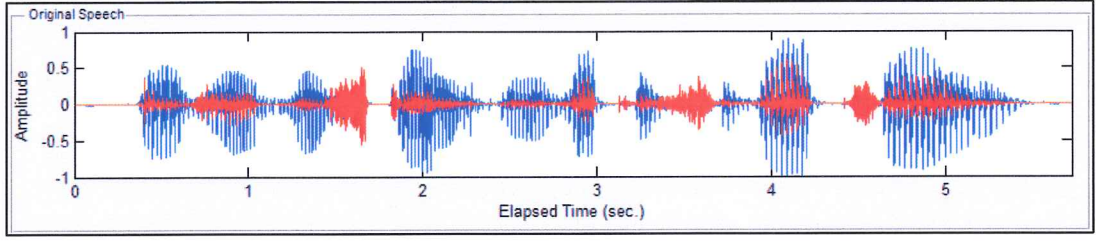
4.1.2. Modül 2: Ön İşleme

Ön işleme öznelik çıkartım algoritmalarının performansını iyileştirmek için öznelik çıkartımı öncesi konuşma sinyaline uygulanır. SPAC örnekleme değerini düşürme, pre-emphasis, gürültü giderme ve DC-offset kaldırma yöntemlerini ve bu yöntemlerin kombinasyonlarını içermektedir.

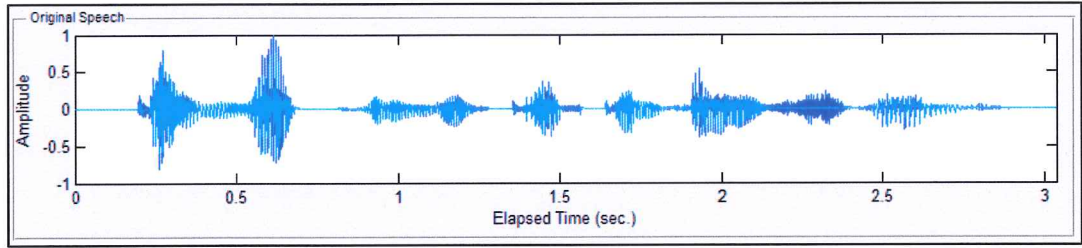
Örnek bir konuşma için ön işleme sonrası elde edilen görüntüler aşağıdaki şekillerde verilmiştir.



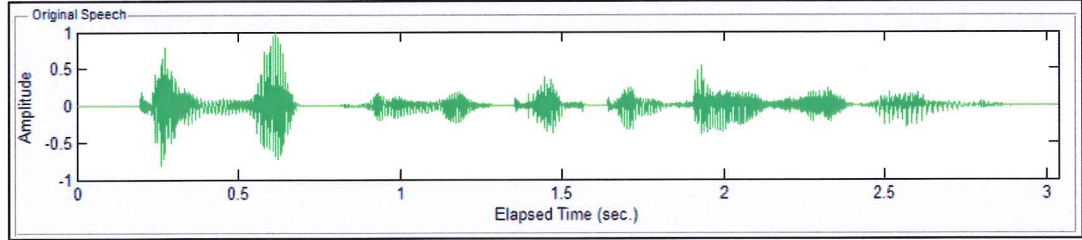
Şekil 4.4. Örnekleme değerini düşürme sonrası sinyaldeki değişim (mavi: orijinal sinyal, siyah: ön işleme sonrası).



Şekil 4.5. Pre-emphasis sonrası sinyaldeki değişim (mavi: orijinal sinyal, kırmızı: ön işleme sonrası).



Şekil 4.6. Gürültü giderme sonrası sinyaldeki değişim (mavi: orijinal sinyal, cyan: ön işleme sonrası).



Şekil 4.7. DC-offset kaldırma sonrası sinyaldeki değişim (mavi: orijinal sinyal, yeşil: ön işleme sonrası).

4.1.3. Modül 3: Öznitelik Çıkartımı

Öznitelik çıkartmaktaki temel amaç konuşmayı tanımlayıcı özniteliklerin sinyalden elde edilmesidir. SPAC aracılığı ile 9 ana başlık ve bunların istatistiksel varyasyonları ile birlikte yaklaşık 723 öznitelik elde edilebilmekte ve cepstrum count değerine göre değişmektedir. Çizelge 4.1'de SPAC ile elde edilebilecek öznitelikler verilmiştir.

Çizelge 4.1. SPAC ile elde edilebilecek öznelik öznelikler.

Öznelik	İstatistiksel Değerler	Öznelik Boyutu
Pitch (F0)	Ortalama, standart sapma,	18
Formant (F1, F2, F3 ve bandwidth)	maksimum, minimum, aralık, ortanca, Δ , Δ^2	108
Ses kalitesi (SNR, SNR Power Jitter, Shimmer)	SNR ve SNR Power (ortalama, Δ , Δ^2)	18
LPCC	Ortalama, standart sapma, maksimum, minimum, aralık, ortanca, Δ , Δ^2	Ceps. countX18
MFCC		Ceps. countX18
ZCR		18
Enerji		18
Konuşma (Konuşma oranı, duraksama oranı, duraksama sayısı)	Ortalama	3
Dalgacık (D1, D2, D3, D4, D5, A1)	Ortalama, standart sapma, maksimum, minimum, aralık, ortanca, L1 norm, L2 norm, Maksimum Norm, ortalama mutlak sapma, ortanca mutlak sapma, mod	72

Pitch tespiti için SPAC üzerinde otokorelasyon ve kepstrum olmak üzere 2 yöntem kullanılmıştır. Hangi yöntemin kullanılacağı Modül 1 üzerinden belirlenebilmektedir. Formant frekanslarının tespiti için LPC analizi kullanılmıştır. SNR tespiti için Matlab Signal Processing Toolbox kullanılmıştır. Jitter'in hesaplanmasında Modül 1'den seçilen GCI algoritması kullanılmaktadır. LPC hesaplanması için otokorelasyon kullanılmıştır. Dalgacık dönüşümü kullanılarak 5 seviye ayrıştırma yapılmıştır. Ayrıştırma sonucu her bir seviye için katsayılar hesaplanmıştır. Dalgacık katsayılarının tespitinde Matlab Wavelet Toolbox'dan yararlanılmıştır.

Birçok öznelik kümesinde kullanılan Delta (Δ) ve Delta-delta (Δ^2) ilgili öznelik için zaman türevinin eklenmesidir.

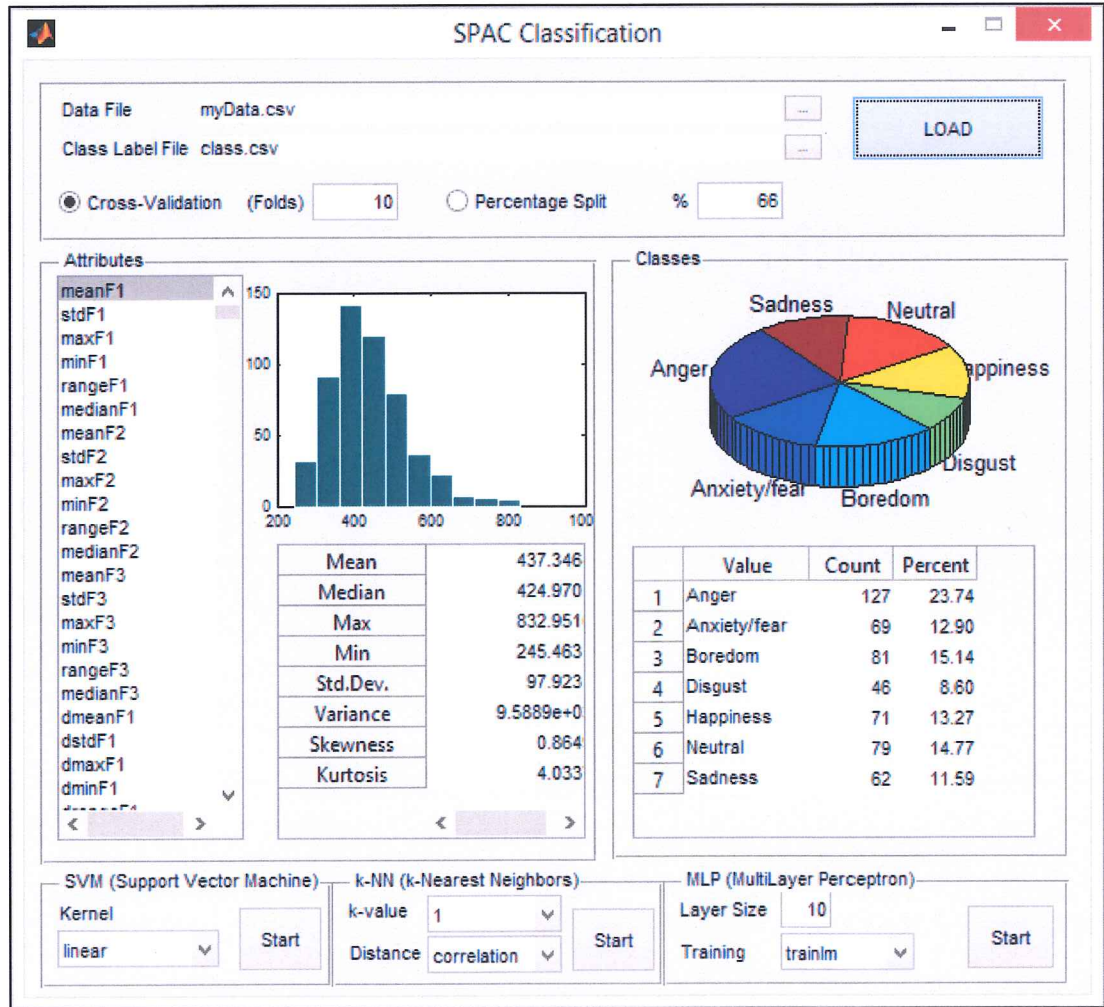
4.1.4. Modül 4: Son İşleme

Öznelik çıkartımı sonrası elde edilen öznelik kümesi üzerinde normalizasyon veya öznelik seçimi işlemleri Modül 4 üzerinden gerçekleştirilebilir. Normalizasyon, özneliklerin içerdiği değerler arasındaki birim farklılıklarını kaldırmak için

kullanılır. Öznitelik seçimi ise öznitelik kümesinin boyut indirgemesi için kullanılır. SPAC normalizasyon için z-score, öznitelik seçimi için ise PCA, LDA ve Fisher selection yöntemlerini içermektedir.

4.1.5. Modül 5: Sınıflandırma

Modül 5 çıkartılan öznitelikler üzerinden sınıflandırma yapmak için SVM, Çok Katmanlı Algılayıcı (MLP) ve k-NN sınıflandırıcıları içermektedir. Ancak bu sınıflandırıcılar gözetimli öğrenme olduğu için kullanılan verinin sınıflarının belirtilmesi gerekir. Bu işlem modül içerisinde yer alan Class Label File seçeneği ile gerçekleştirilir. Modül 5'e ait GUI Şekil 4.8'de verilmiştir.



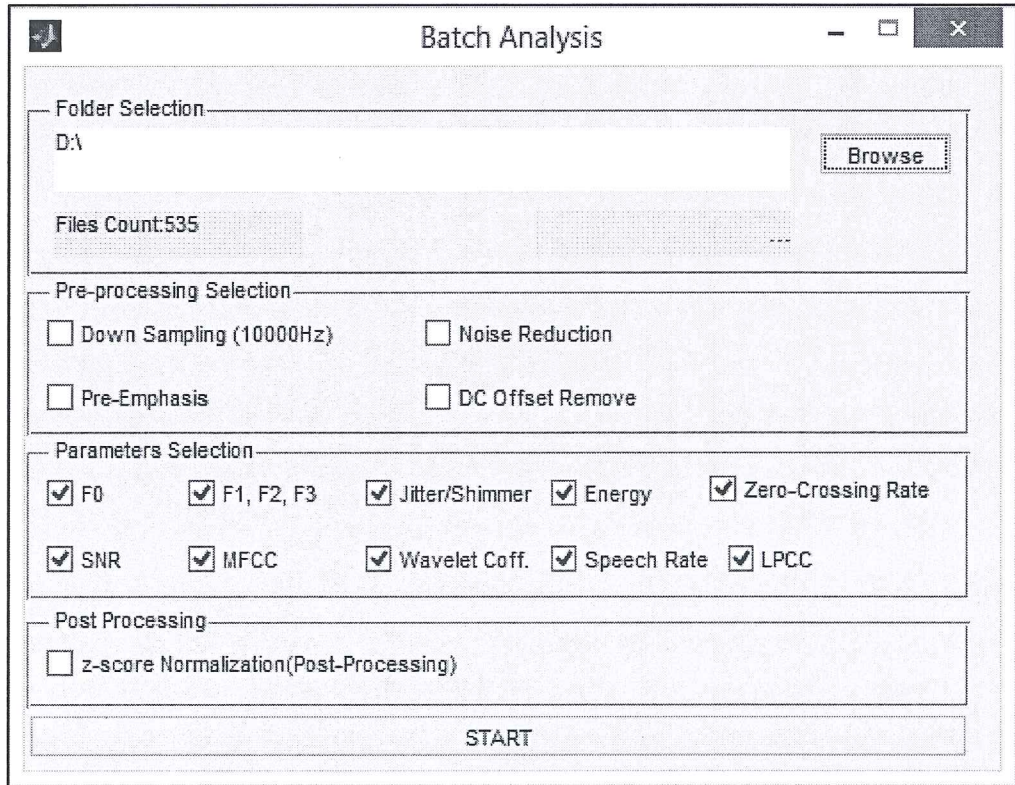
Şekil 4.8. Modül 5'e ait ekran görüntüsü.

Şekil 4.8'de verilen classification modülü üzerinden özniteliklere ait istatistiksel bilgiler ve sınıfların dağılımı takip edilebilmektedir.

Sonuçların doğru ve istatistiksel olarak anlamlı olduğundan emin olmak için sınıflandırma m-fold çapraz doğrulama (cross-validation) ile gerçekleştirilir. m-fold çapraz doğrulamada veri birbirinden bağımsız m adet alt gruplara ayrıştırılır. Sınıflandırıcı eğitim ve test işlemini m kez gerçekleştirir.

4.1.6. Modül 6: Toplu Dosya İşleme

Modül 6 birden fazla dosya üzerinde işlem yapmak için kullanılmaktadır. İçerdiği GUI sayesinde kullanıcı dostu bir ara yüze sahiptir ve diğer modüllerde gerçekleştirilen işlemlerin tamamı bu modül aracılığı birden fazla dosyaya uygulanabilmektedir. Modül 6'ya ait GUI Şekil 4.9'da verilmiştir.



Şekil 4.9. Modül 6'ya ait ekran görüntüsü.

GUI üzerinde yer alan Folder Selection paneli ile işlem yapılacak dosyaların bulunduğu klasör seçilir. Pre-processing Selection panelinden kullanılacak ön işleme yöntemleri, Parameters Selection panelinden ses dosyalarından elde edilecek öznitelikler ve Post-processing panelinden de normalizasyon yapılıp yapılmayacağı belirlenir. Analiz sonrası elde edilen öznitelikler mat dosyasında saklanır.

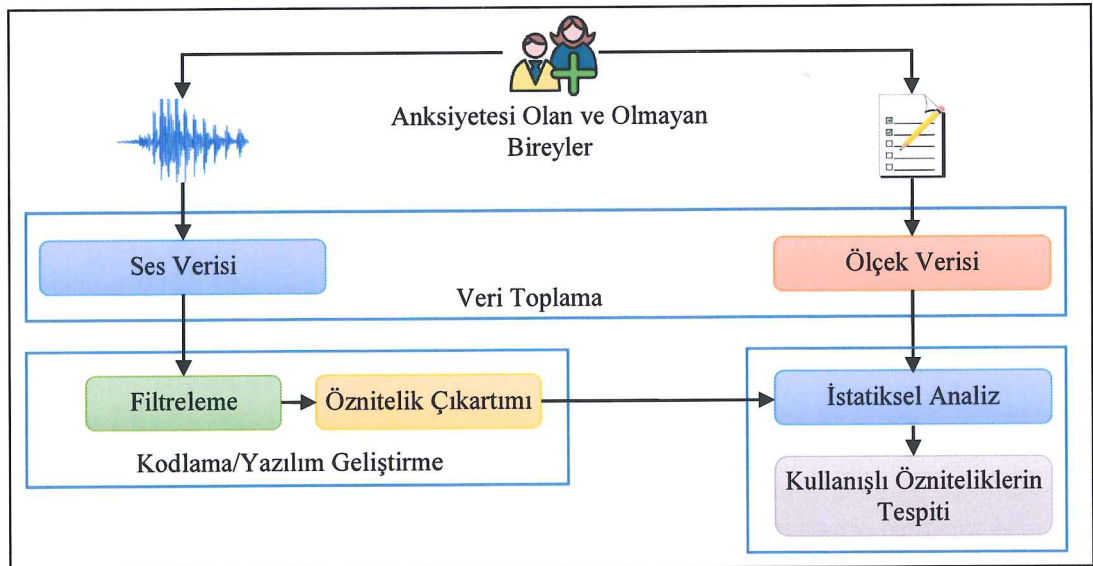
BÖLÜM 5

MATERYAL VE METOD

Bu bölümde tezde kullanılan veriler ve çalışma kapsamında kullanılan yöntemler hakkında bilgiler verilmiştir.

5.1. SİSTEMİN GENEL YAPISI

Bu çalışmanın temel amacı ses analizinde akustik parametrelerin tespiti ve bu parametrelerin anksiyete ilişkisini araştırmaktır. Çalışmanın birinci amacı olan ses analizi ve akustik parametre tespiti için MATLAB ortamında yazılım geliştirilmiştir. Anksiyetenin akustik parametrelerle ilişkisini tespit etmek için de toplanan gerçek verilerden hazırlanan yazılım ile öznelik çıkartımı yapılmış ve elde edilen sonuçlar istatistiksel olarak yorumlanmıştır. Sistemin genel yapısı Şekil 5.1’de verilmiştir.



5.2. ANKSİYÖZ KONUŞMA VERİSİNİN EDİNİMİ

Anksiyöz durum ile ilgili veri bulmadaki sınırlılığını aşmak için Gülhane Eğitim ve Araştırma Hastanesi Ruh Sağlığı ve Hastalıkları Anabilim Dalında 01.09.2015 ile 31.06.2016 tarihleri arasında toplanan gerçek ses kayıtları kullanılmıştır. Ses kayıtlarının toplanabilmesi için Gülhane Eğitim ve Araştırma Hastanesi'nden Prof. Dr. Ali DORUK başkanlığında 22 Ekim 2015 tarihli Ek Açıklamalar B'de verilen Etik Kurul izni alınmıştır.

Veri toplama için gerçekleştirilen çalışma kesitsel nitelikte bir olgu-kontrol çalışması olup ses kayıtları yatan ve ayaktan hastalardan elde edilmiştir. Olgulardan alınan gönüllü olur formu Ek Açıklamalar C'de verilmiştir.

Ses kayıtları toplanırken araştırmaya dahil olma, araştırmaya almama ve araştırmadan çekilme kriterleri aşağıda verilmiştir.

1. Araştırmaya dahil olma kriterleri

- 1.1. Yaygın anksiyete bozukluğu ve panik bozukluk tanısı konanlar.
- 1.2. Sağlıklı kontroller.
- 1.3. Gönüllü olanlar.
- 1.4. 18 – 65 yaş.
- 1.5. En az 8 yıllık temel öğrenimi olanlar.

2. Araştırmaya almama kriterleri

- 2.1. Yapısal ses bozukluğu olanlar.
- 2.2. Daha önce ses problemi nedeniyle tedavi görmüş olanlar.
- 2.3. Uygulama sırasında üst solunum yolu enfeksiyonu olanlar.
- 2.4. Psikotrop ilaçlar veya alkol-madde kullanımı olanlar.

3. Araştırmadan çekilme kriterleri

3.1. Denekler arařtırmadan neden belirterek ya da belirtmeksizin istedikleri zaman ayrılabilirler.

Veriler yaygın anksiyete bozukluęu, panik bozukluk ve saęlıklı kontrol grubu üzerinden elde edilmiřtir. Deneklerin anksiyete düzeyleri uluslararası standartlarda olan psikometrik ölçeklerle deęerlendirilmiřtir.

Panik bozukluk ve yaygın anksiyete bozukluęu anksiyetenin temel özneliklerini en iyi řekilde yansıtan bozukluklardandır. İstatistiki analizi aısından etkili olması için gruplar 20-30 kiřiden oluřmaktadır.

alıřmada kullanılan gönüllülerin sosyodemografik bilgilerini elde etmek için Ek Aıklamalar D’de verilen form kullanılmıřtır.

Olguların anksiyete düzeylerinin deęerlendirilmesi için ‘‘Hamilton Anksiyete Öleęi’’ ve ‘‘Beck Anksiyete Envanteri’’ kullanılmıřtır. Hamilton Anksiyete Öleęi [139] ve Beck Anksiyete Envanteri [139] Ek Aıklamalar F’de verilmiřtir.

Hamilton Anksiyete Öleęi, Hamilton tarafından 1959 yılında anksiyete düzeyini belirlemek ve semptom řiddetini deęerlendirmek için geliřtirilmiřtir. Hem zihinsel hem de bedensel belirtileri deęerlendiren 14 maddeden oluřmaktadır [140]. Öleęin Türke versiyonu Yazıcı vd. tarafından Türkiye’de geerli ve güvenilir olarak ifade edilmiřtir [141].

Beck Anksiyete Envanteri, 21 maddeden oluřmaktadır. Artan puanlar anksiyete semptomlarının yoęunluęunun řiddetini gösterir [142]. Öleęin Türke versiyonu Ulusoy vd. tarafından Türkiye’de geerli ve güvenilir olarak ifade edilmiřtir [143].

alıřmaya Ruhsal Bozuklukların Tanısal ve İstatistiksel El Kitabı Beřinci Versiyon (DSM-5) tanı ölçütlerine [144] göre yaygın anksiyete bozukluęu ve panik bozukluk tanısı konanlar alınmıřtır. Daha önce psikiyatrik tedavi görmemiř, halen psikiyatrik yakınması olmayan ve anksiyete bozukluęu řiddetini deęerlendiren ölçeklerden düşük puan alanlar (Beck Anksiyete Envanterinden 9 ve altında [145]), Hamilton

Anksiyete Ölçeğinden 14 ve altında puan alanlar [140]) kontrol grubu olarak belirlenmiştir. Olgulardan son 3 saat içinde sigara içmemeleri istenmiştir.

Hamilton Anksiyete Ölçeği ve Beck Anksiyete Envanteri sonuçlarına göre 23'ü anksiyete bozukluğu (Hasta grup) ve 20'si sağlıklı (Kontrol grubu) toplam 43 olgu çalışmaya dahil edilmiştir. Anksiyete bozukluğu olgularının 2'si panik bozukluk ve 21'i yaygın anksiyete bozukluğundan oluşmaktadır.

Çalışmada kullanılan ses kayıtlarının tamamı aynı oda ve ortamda Philips DVT3000 dijital ses kayıt cihazı ile wav formatında kaydedilmiştir. Olguların her birinden 3 adet ses kaydı alınmıştır. Bunlardan ilkinde olgu nötr durumda, diğer 2'sinde ise anksiyöz durumdadır. Nötr duygu durum için olgular herhangi bir duygusal kelime veya cümle içermeyen masallarla ilgili tanımlayıcı bir metni (NMO: Nötr Metin Okuma) okumuştur. Anksiyöz duygu içinse, Beck Anksiyete Envanteri temel alınarak ve kişilerin günlük hayatta karşılaşılabilecekleri durumları içeren bir metin olgular tarafından okunmuştur (EMO: Endişeli Metin Okuma). Bu metinlerin istenilen duyguları tetiklediği olgulardan gelen geribildirimler ve araştırmacıların gözlemleri ile olumlu olarak değerlendirilmiştir. Ayrıca, her bir olgu, son yaşadığı anksiyete atağını 3 dakika anlatmıştır (AY: Anksiyöz Yaşantı). Öncesinde neleri anlatabilecekleri konusunda görüşme yapılarak, kesintisiz anlatım sağlanmıştır. Panik bozukluk hastalarından yaşadıkları son panik ataklarını, yaygın anksiyete bozukluğu hastalarından endişe yaşadıkları bir durumlarını ve kontrol grubundan ise daha önce yaşadıkları her hangi bir stresli durumlarını (sınav heyecanı, topluluk önünde sunum yapma gibi) yaklaşık 3 dakika süresince anlatmaları istenmiştir. Endişe ve nötr durum için kullanılan metinler Ek Açıklamalar E'de verilmiştir.

5.3. KODLAMA VE YAZILIM GELİŞTİRME

Ses sinyallerinin incelendiği bu aşama, daha çok sinyal işleme ağırlıklıdır. Ses analizi konusunda yapılan çalışmalarda akustik ve linguistik öznitelikler ön plana çıkmaktadır. Akustik özniteliklerin bir kısmı duygusal içerikle ilgilenirken bir kısmı da sinyal kalitesi ile ilgilenmektedir. Duygusal içerikte seste kullanılan kelimelerin bir anlamı yoktur. Linguistic özniteliklerde ise kullanılan kelime ve kişi önem

taşımaktadır. Bu bilgiler doğrultusunda Bölüm 4’de ayrıntıları verilen SPAC isimli yazılım geliştirilmiştir.

Hazırlanan yazılım üzerinde ön işleme yöntemleri olarak çerçeveleme, pencereleme, pre-emphasis filtreleme, gürültü giderme, dc-offset kaldırma ve örnekleme değerini düşürme kullanılmıştır. Bu yöntemlerin ayrıntıları Bölüm 2’de verilmiştir.

SPAC ile bir ses dosyası üzerinden pitch, formant frekansları, sinyal gürültü oranı, jitter, shimmer, LPCC, MFCC, ZCR, enerji, konuşma oranı, duraksama oranı duraksama sayısı ve dalgacık katsayıları tespit edilebilmektedir. Öznitelik seçimi için PCA, LDA ve Fisher yöntemleri sınıflandırma için SVM, MLP ve k-NN kullanılmıştır.

5.3.1. Pitch Tespiti

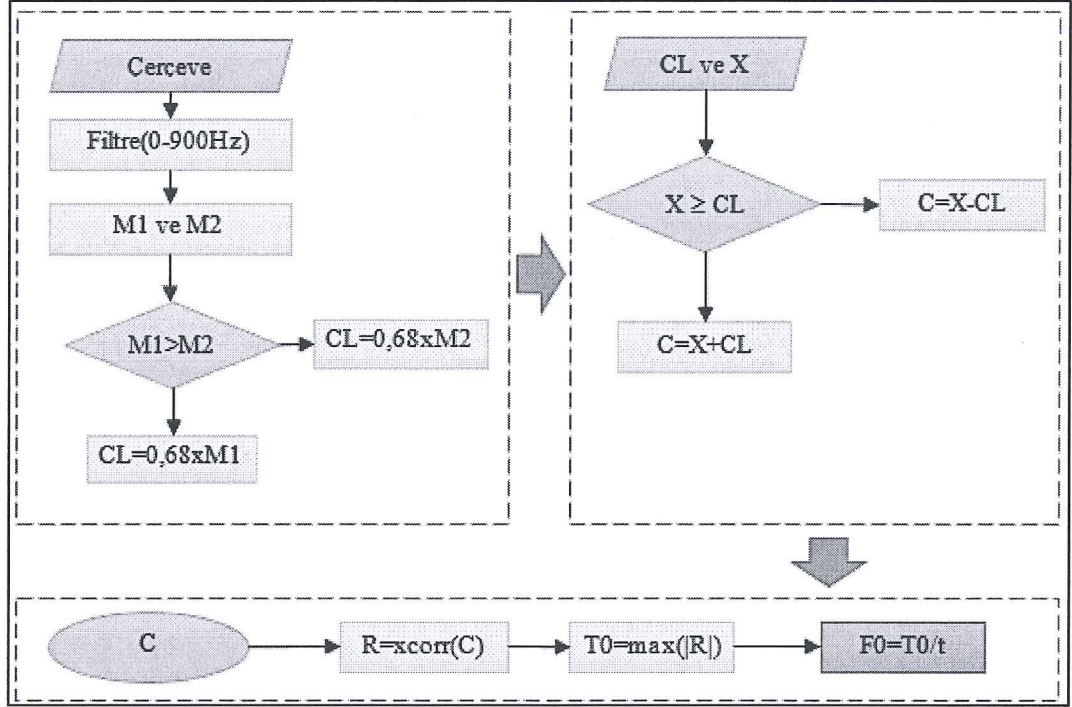
Çalışmada pitch tespiti için otokorelasyon ve kepstum olmak üzere iki yöntem kullanılmıştır.

İki sinyal arasındaki korelasyon onların benzerlik ölçüsüdür. Sinyal farklı zaman aralıklarında karşılaştırılarak benzerlikleri hesaplanır. Bir korelasyon işleminin sonucu iki sinyalin başlangıçları arasındaki zaman gecikmesinin bir fonksiyonu olarak benzerlik ölçüsüdür. Zaman gecikmesi arttıkça farklılıklar artarken sıfır zaman gecikmesi durumunda tam benzerlik görülebilir [131].

Otokorelasyon tabanlı yöntemler, orijinal sinyal ve kaydırılmış sinyalin noktasal çarpımını hesaplar. Sinyal harmonik bir sinyal ise otokorelasyon fonksiyonu temel frekansın katlarındaki pik değerlerine sahip olacaktır. Sınırlı pitch aralığına sahip konuşma tanıma uygulamalarında popüler bir yöntemdir. τ gecikme zamanına sahip bir sinyal otokorelasyon fonksiyonu $r(\tau)$ aşağıdaki şekilde (Eşitlik 5.1) tanımlanır [146].

$$r(\tau) = \frac{1}{N} \sum_{n=0}^{N-1} x[n]x[n + \tau] \quad (5.1)$$

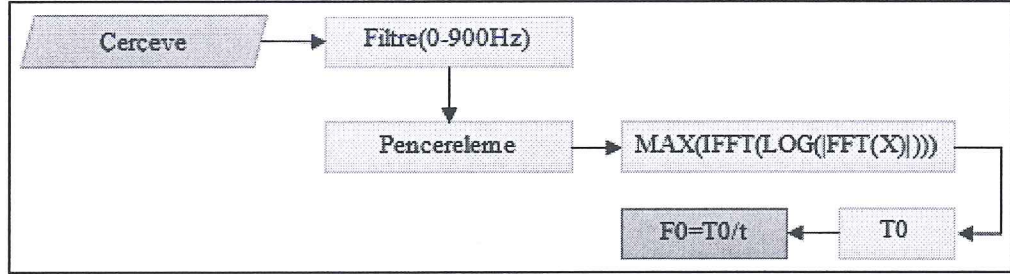
Otokorelasyon fonksiyonu $\tau=0$ için daima global maksimumdur. Temel frekans $1/\tau$ olarak hesaplanır. Şekil 5.2’de otokorelasyon tabanlı pitch tespiti için akış diyagramı verilmiştir.



Şekil 5.2. Otokorelasyon tabanlı pitch tespiti akış diyagramı.

Şekil 5.2’de verilen pitch tespit yöntemi konuşma sinyalinden alınan bir çerçeve ile algoritmayı başlatır. X üzerinde filtreleme yapılarak 900 Hz üzerindeki sinyaller silinir. Sinyalin baştan 1/3’lük kısmının mutlak değerinin maksimumu alınarak M1 ve sondan 1/3’lük kısmının mutlak değerinin maksimumu alınarak M2 değerleri hesaplanır. M1 değeri M2’den büyükse M2 değeri, değilse M1 değeri 0,68 ile çarpılarak CL (kırpma seviyesi) değeri hesaplanır. X sinyalinde CL değerinden büyük değerlerden CL çıkartılarak, küçük değerlerde ise CL değeri eklenerek sinyalin kırpma uygulanmış hali (C) elde edilmiş olur. C’ye çapraz korelasyon uygulandıktan sonra mutlak değerinin maksimumu alınarak temel periyot (T0) elde edilir. T0’değeri tespit edildiği bölgenin zaman değerine (t) bölünerek pitch (F0) elde edilir.

Kepstrum analizi bir giriş sinyalinin magnitud spektrumunun logaritmasının fourier dönüşümü olan spektral analiz biçimidir [131]. Şekil 5.3’de kepstum tabanlı pitch tespiti akış diyagramı verilmiştir.



Şekil 5.3. Kepstrum tabanlı pitch tespiti akış diyagramı.

Şekil 5.3’de verilen pitch tespit yöntemi konuşma sinyalinden alınan bir çerçeve ile algoritmayı başlatır. X üzerinde filtreleme yapılarak 900Hz üzerindeki sinyaller silinir. Daha sonra sırasıyla pencereleme, FFT, logaritma ve IFFT uygulanarak elde edilen sonucun en büyüğü temel periyod olan T0’ı oluşturur. T0’değeri tespit edildiği bölgenin zaman değerine (t) bölünerek F0 elde edilir.

5.3.2. Formant Frekanslarının Hesaplanması Ve LPCC

Formant frekanslarının hesaplanmasında LPC yöntemi kullanılmıştır. LPC, aşağıdaki eşitlikte görüldüğü gibi, konuşma sinyalinin n . örneğinin önceki p adet örneğin doğrusal kombinasyonu şeklinde ifade edilmesidir [147].

$$X_{LPC}[n] = \sum_{i=1}^p b_i x[n-i] \quad (5.2)$$

b_i ’ler çerçeve süresinde sabit olarak kabul edilir. Konuşmanın gerçek değeri ile öngörülen değeri arasındaki fark öngörü hatasıdır (Eşitlik 5.3).

$$e[n] = x[n] - x_{LPC}[n] = x[n] - \sum_{i=1}^p b_i x[n-i] \quad (5.3)$$

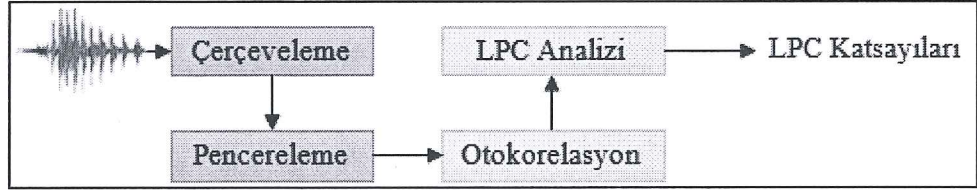
Hata sinyali ile birlikte LPC eşitliği Eşitlik 5.4'deki gibi olacaktır.

$$X_{LPC}[n] = \sum_{i=1}^p b_i x[n-i] + e[n] \quad (5.4)$$

Eşitlik 5.4'den z-dönüşümü ile transfer fonksiyonu hesaplanırsa sonlu uzunlukta dürtü tepkili bir süzgeç elde edilir (Eşitlik 5.5).

$$H(z) = \frac{E(z)}{S(z)} = 1 - \sum_{i=1}^p x^i z^i = A(z) \quad (5.5)$$

Süzgeç katsayıları b_i hatanın karesini minimum yapacak şekilde hesaplanır. $A(z)$ süzgeci ses yolunu (vocal tract) modeller. Böylece konuşma sinyali, incelendiği zaman aralığında, uyarım sinyaline ve süzgeç katsayılarına ayrıştırılmış olur. LPC blok diyagramı Şekil 5.4'de verilmiştir.



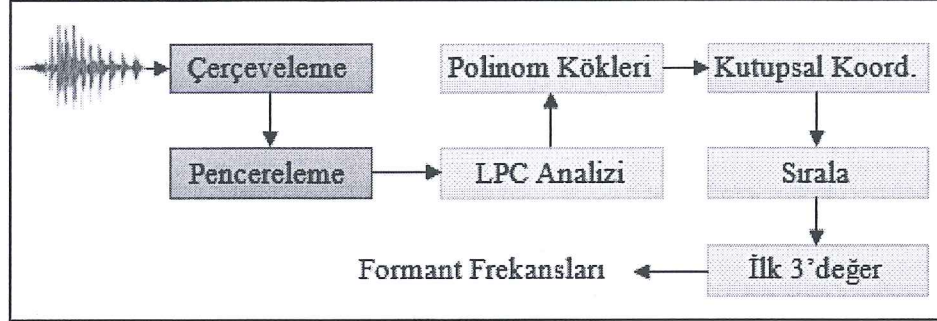
Şekil 5.4. LPC blok diyagramı [147].

Otokorelasyon hesaplamak için Eşitlik 5.6'da verilen denklem kullanılır.

$$X_l(\tau) = \sum_{n=0}^{N-1-\tau} x_l[n]x_l[n+\tau], \quad \tau = 0, 1, \dots, p \quad (5.6)$$

LPC analizinde $p+1$ otokorelasyondan LPC parametre kümesi hesaplanır. LPC analizine geçişte otokorelasyon yerine kovaryans analizi de yapılabilir, fakat yaygın olarak otokorelasyon kullanılmaktadır [147], [148]. LPCC katsayıları temel olarak LPC katsayılarının fourier dönüşümü ile kepsral katsayılarla dönüştürülmesi prensibine dayanır [147].

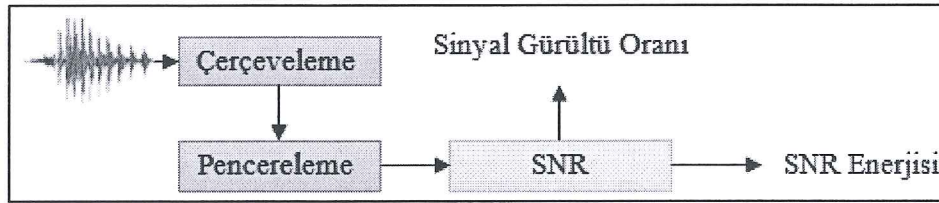
Formant frekanslarının LPC analizi ile tespiti Şekil 5.5’de verilen diyagrama göre gerçekleştirilmektedir.



Şekil 5.5. Formant frekanslarının tespiti.

5.3.3. Sinyal Gürültü Oranının Hesaplanması

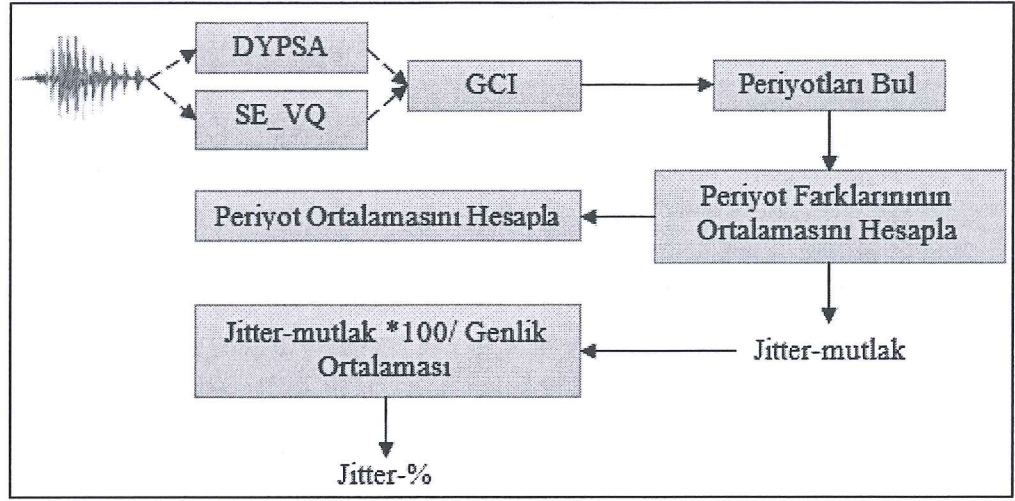
Sinyal gürültü oranı ve sinyal gürültü oranı enerjisini hesaplamak için Matlab Signal Processing Toolbox içerisinde yer alan *snr* fonksiyonu kullanılmıştır. Sinyal gürültü oranı enerjisi sinyalin harmonik olmayan bileşenlerinin enerjisidir. Şekil 5.6’da sinyal gürültü oranı ve enerjisinin hesaplanması için akış diyagramı verilmiştir.



Şekil 5.6. Sinyal gürültü oranı ve enerjisinin hesaplanması için akış diyagramı.

5.3.4. Jitter Ve Shimmer Değerlerinin Hesaplanması

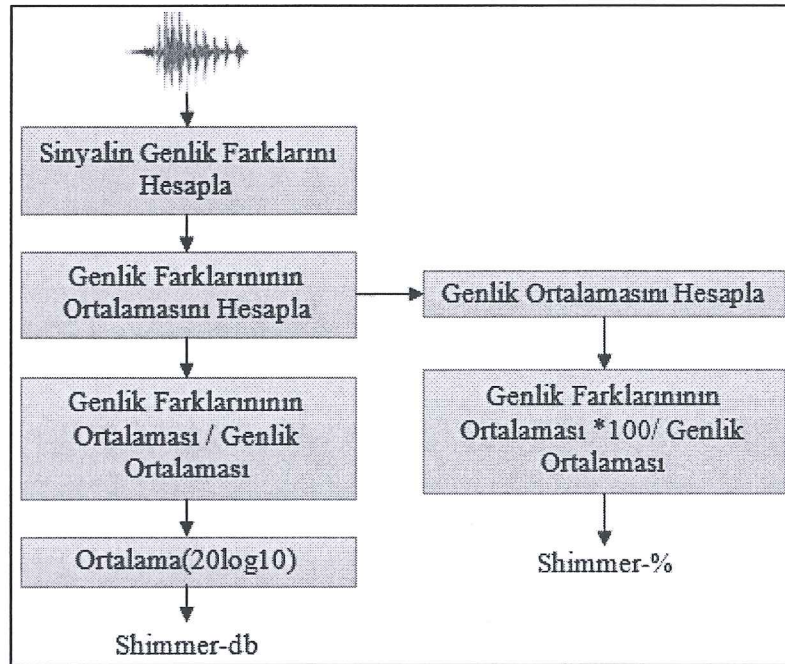
Jitter değerinin hesaplanması ters filtreleme ile ilgili olup bu işlem için DYPSA ve SE_VQ algoritmaları kullanılmıştır. Jitter ile ilgili jitter-mutlak ve jitter-% değerleri Şekil 5.7’de verilen diyagrama göre hesaplanmıştır.



Şekil 5.7. Jitter değerlerinin hesaplanması için akış diyagramı.

Yukarıdaki şekilde kullanılan GCI konuşma esnasında gırtlığın kapanma anlarını verirken, periyot gırtlığın kapanma anlarında elde edilen periyodu ifade etmektedir.

Shimmer değeri hesaplanırken sinyalin genlik değerleri kullanılmaktadır. Çalışmada shimmer-db ve shimmer-% olmak üzere shimmer ile ilgili iki değer hesaplanmıştır. Shimmer değerlerinin hesaplanması için kullanılan akış diyagramı Şekil 5.8’de verilmiştir.



Şekil 5.8. Shimmer değerlerinin hesaplanması için akış diyagramı.

Yukarıda şekilde kullanılan genlik farkı konuşma sinyalinin t ve $t+1$ zamanlarındaki genlik değerinin farkıdır.

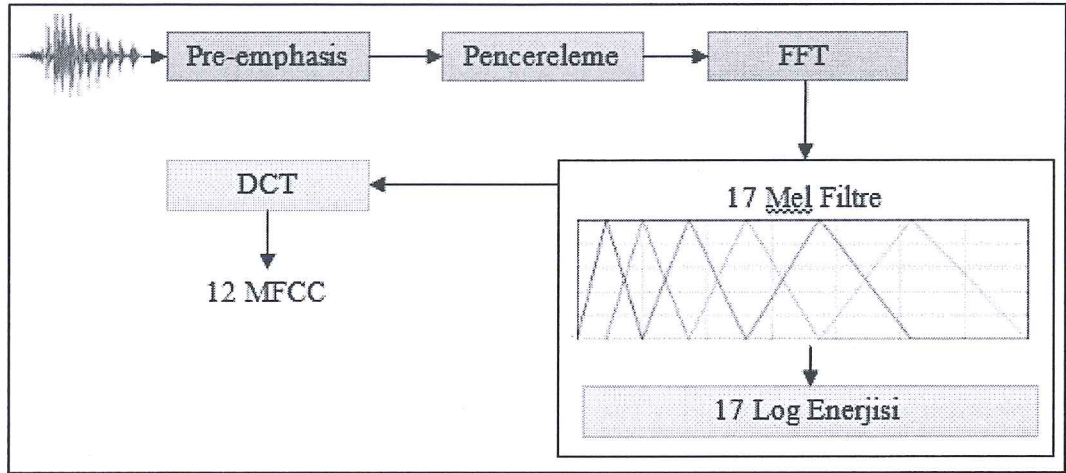
5.3.5. MFCC

Bir sinyal bant geçiş filtreleri ve bu filtrelerin enerjileri ile ifade edilir. Böylece sinyal spektrumunun tahmini sağlanır. MFCC'lerde filter bank eşit Mel ölçeği aralığında üçgen bir dizi filtreden oluşur. Mel birimi, insan kulağının algısal özelliğini taklit edecek şekilde tasarlanmış bir birimdir. Yapılan araştırmalar göstermiştir ki, ölçüler, 1 kHz'e kadar doğrusal, daha yüksek değerlerde ise logaritmiktir. Mel frekans ölçeği Eşitlik 5.7'de verilmiştir [38].

$$f_{MEL} = 2595 \log \left(1 + \frac{F_{Hz}}{700} \right) \quad (5.7)$$

Üçgen filtreler Mel frekans ölçeğinde doğrusal düzenlenmiş olduğundan, doğrusal ölçekte (Hz) bakıldığında filtreler birbirine yakın ve düşük frekansta dar bant genişliklerine sahiptir ve frekans ile aralık ve bant genişliği artmaktadır. Doğrusal olmayan bu konuşma analizi insan işitme sistemine dayanır [38].

Bu algısal spektrumu uygulamanın bir yolu mel skalasında uniform olarak yerleştirilmiş bir süzgeç öbeği kullanmaktır. Bu süzgeç öbeği % 50 oranında birbiri üstüne getirilen, üçgen bant geçiren ve aralıkları ve bant genişliği sabit mel frekansı aralığına bağlı bir banktır. 10 Hz'den 1000 Hz'e kadar lineer olarak 10 filtre atanır. Frekans her iki katına çıktığında logaritmik olarak 5 filtre bu frekans aralığına atanır. Süzgeç bant genişliği 3 dB noktası, mevcut dalga ile bir önceki veya bir sonraki dalganın ortasına gelecek şekilde seçilir. Genellikle 20 Süzgeç kullanılır. 17 üçgen Mel ölçekli filtreler ve 12 Ayrık Kosinüs Dönüşümü (DCT) katsayıları ile 12 boyutlu MFCC vektörü hesaplamak için kullanılan blok diyagram Şekil 5.9'da verilmiştir [38].



Şekil 5.9. MFCC akış diyagramı.

MFCC hesaplamak için konuşma sinyaline ilk olarak pre-emphasis filtre uygulanarak gırtlak ve dudak titreşimlerinin etkisi kaldırılır ve sonra küçük çerçevelerle pencerelenir (20 ms-30 ms). Her çerçevenin spektrum büyüklüğü hesaplandıktan sonra üçgen filtrelerin her birinin filtre cevabı ile çarpılır. Filtrelenmiş her spektrumun ortalamaları hesaplanır ve bunların logaritması her filtre bandındaki ortalama log-enerjiyi elde etmek için hesaplanır. N boyutlu MFCC 'yi elde etmek için seçilen ilk N katsayıları DCT ile hesaplanır [38].

5.3.6. Zero-Crossing Rate

ZCR, sinyaldeki işaret değişimlerinin oranını gösterir. Ses sinyalinin sıfırdan geçiş sayısı olarak bilinir. Sinüzoidal bir sinyalde her periyotta iki sıfırdan geçiş olduğu için sinyal sıklığı sıfırdan geçiş sayısının yarısı olarak hesaplanır. Farklı periyotlar ile çok sayıda sinüs dalgasından oluşan karmaşık sinyaller üzerinde iyi çalışmaz. Temel bir yöntemdir ve hesaplama maliyeti düşüktür. ZCR için hesaplama formülü Eşitlik 5.8'de verilmiştir [146]. Bir sinyalin frekansı ise Eşitlik 5.9'da verilmiştir.

$$zcr = \frac{1}{T-1} \sum_{t=1}^{T-1} \Pi\{s_t s_{t+1} < 0\} \quad (5.8)$$

Eşitlik 5.8'de s_t sinyalin t zamanındaki değeri ve $\Pi\{A\}$ eğer A argümanı sağlanırsa 1, sağlanmazsa 0.

$$F = \frac{zcr * F_s}{2} \quad (5.9)$$

5.3.8. Enerji Ve Konuşma İle İlgili Öznitelikler

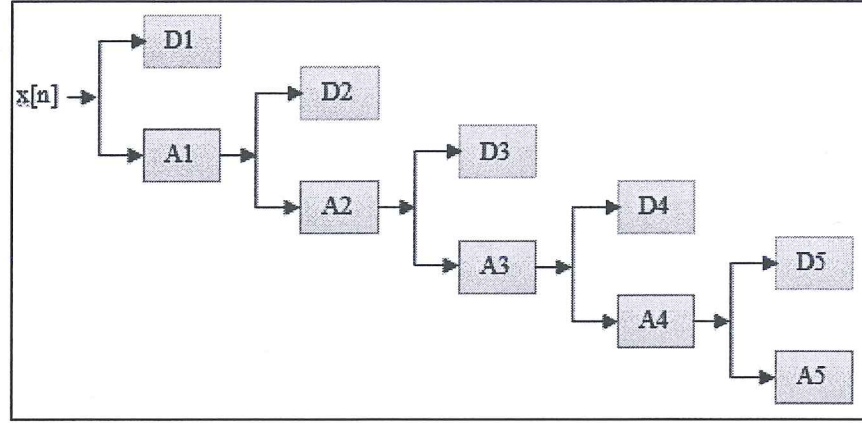
Konuşmanın içerdiği enerji zamana göre değişkendir ve enerji konuşmanın kısa süreli (short-time) bölgeleri ile ilgilidir. Herhangi bir konuşma esnasında konuşma sürekli gerçekleşmez ve bazen sessiz kalınır. Bu durum konuşma sinyalinin enerjisine yansımaktadır. Çalışmada enerji değerinin hesaplanması için short-time energy yöntemi kullanılmıştır. Elde enerji değerleri belirli bir eşik değerinin üstünde ise konuşma içeren bölge (sesli) değilse konuşma içermeyen bölge (sessiz) olarak belirlenmiştir. Bir $x[n]$ sinyali için toplam enerji formülü Eşitlik 5.10'da verilmiştir.

$$E = \sum_{n=-\infty}^{\infty} x^2[n] \quad (5.10)$$

Çalışmada enerji değeri her bir çerçeve için hesaplanmış ve bunların ortalaması alınarak tespit edilmiştir. Konuşma oranı hesaplanırken enerjisi eşik değerinde üstünde olan çerçeve sayısının tüm çerçeve sayısına oranı ve bekleme (sessizlik) oranı hesaplanırken enerjisi eşik değerinde altında olan çerçeve sayısının tüm çerçeve sayısına oranı alınmıştır.

5.3.9. Dalgacık Katsayıları

Dalgacık katsayıları hesaplanırken konuşma sinyalleri 5 seviye ayrıştırma yapılmış (D1, D2, D3, D4, D5, A1) ve her seviye için katsayılar (ortalama, standart sapma, maksimum, minimum, aralık, ortanca, L1 norm, L2 norm, maksimum norm, ortalama mutlak sapma, ortanca mutlak sapma, mod) hesaplanmıştır. Çalışmada kullanılan ayrışım basamakları Şekil 5.10'da verilmiştir.



Şekil 5.10. 5 seviye dalgacık dönüşümü.

5.3.10. Temel Bileşen Analizi (PCA)

PCA çok sayıda birbiri ile ilişkili değişkenler içeren veri setinin boyutlarını veri içerisinde varolan değişimlerin mümkün olduğunca korunarak daha az boyuta indirgenmesini sağlayan bir dönüşüm tekniğidir [149]. PCA yönteminde öznitelik vektörleri arasındaki ilişkiyi temsil eden kovaryans matrisinin özvektör ve özdeğer çarpımına eşit olduğunu varsayılır [95]. Buradan bulunan özvektörleri yeni temel bileşenler olarak kabul edilir ve verinin yeni bileşenleri hesaplanır. Temel bileşen sayısı örnek sayısına eşit ya da daha azdır. PCA, boyut azaltma için kullanıldığında sapma değeri az olan boyut silinir ve gerekiyorsa veri kendi boyutuna geri dönüştürülür. Gürültüye karşı düşük hassasiyet, bellek ve kapasite ihtiyaçlarının azalması, az boyutlu uzaylarda daha etkin çalışması PCA'nın temel avantajları arasında sıralanabilir [149]. PCA'nın işleme basamakları aşağıda verilmiştir [150]:

1. Her örnekten öznitelik vektörü elde edilir.
2. Öznitelik vektörlerinin ortalaması alınarak ortalama vektör elde edilir.
3. Ortalama vektöründen elde edilen her vektör çıkartılarak normalize vektörü elde edilir.
4. Normalize vektör ile normalize vektörün tersi çarpılarak kovaryans vektörü elde edilir.
5. Kovaryans vektörünün özvektörü elde edilir.
6. Normalize vektör ile özvektör çarpılarak kovaryans matrisinin özvektörü elde edilir.

7. Her örnek için normalize vektör ile kovaryans matrisinin özvektörünün tersi çarpılarak projeksiyon vektörü elde edilir.

5.3.11. Lineer Diskriminant Analizi (LDA)

LDA, sınıflar arası kovaryans matrisi ve sınıf içi kovaryans matrisinin doğrusal ayrılabilirliğini maksimize eden öznitelik indirgeme yöntemidir [151]. Temel amaç iki sınıfı en iyi ayıran özniteliklerin tespit edilmesidir. LDA'nın iki temel amacı vardır. Bunlardan birisi yüksek boyutlu öznitelik uzayında veri seyrekliği problemini gidermek amacıyla öznitelik uzayının boyutunu azaltmaktır. Diğeri ise sınıf ayrılabilirlik bilgisini maksimum yapmaktır [152].

5.3.12. Fisher Oranı

Fisher oranı [153] bireysel özniteliklere ait bilgiyi elde etmede sıklıkla kullanılan bir istatistiksel yöntemdir. Yöntem her bir sınıf için ayrı olmak üzere özniteliklere ait sayısal verilerin ortalaması ve yine her bir özniteliğin standart sapma değerleri bilgisini ölçüm metodunda kullanmaktadır. İki sınıf durumunda her bir x özelliği için fisher oranı Eşitlik 5.11'de verilmiştir [154].

$$FO(x) = \frac{(\mu_1 - \mu_2)^2}{\sigma_1^2 + \sigma_2^2} \quad (5.11)$$

Eşitlik 5.11'de bulunan μ_1 ve μ_2 sınıflara ait ortalama değer, σ_1 ve σ_2 sınıflara ait standart sapmadır. Temel amaç bir öznitelik iki sınıfa ne kadar çok ayırabiliyorsa o kadar iyidir. Bu yöntem ile büyük veri kümelerinde öznitelik alt kümesi elde etmek için gürültülü verilerin kaldırılması mümkün olabilmektedir [154].

Sınıf sayısı ikiden fazla ise Fisher ölçüt skoru Eşitlik 5.12'de verilmiştir [154].

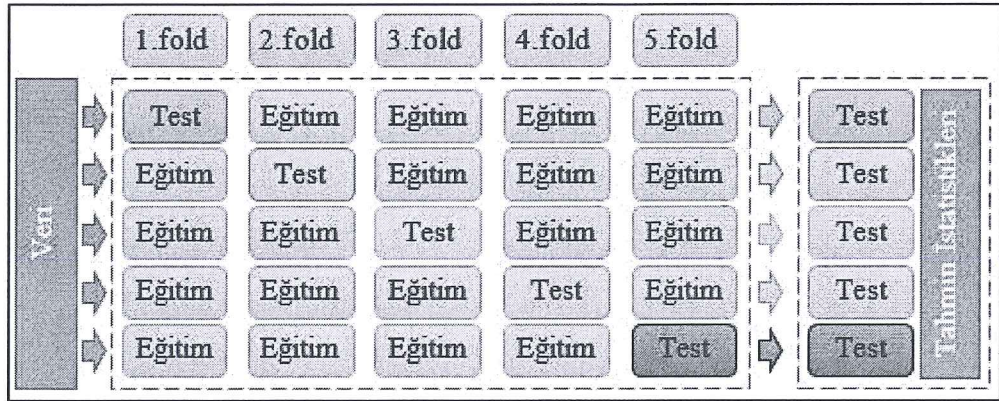
$$FO(x) = \frac{\left(\mu_1 - \sum_{i=2}^C \mu_i\right)^2 + \dots + \left(\mu_j - \sum_{\substack{i=1 \\ i \neq j}}^C \mu_i\right)^2 + \dots + \left(\mu_C - \sum_{i=1}^{C-1} \mu_i\right)^2}{\sum_{i=1}^C \sigma_i^2} \quad (5.12)$$

Eşitlik 5.12’de verilen C sınıf sayısıdır. Fisher ölçüt skoru ile her bir öznitelik kümesi için hesaplanan skor değerine göre öznitelikler sıralanır ve kullanılacak öznitelik sayısı araştırmacının tercihinine bağlıdır. Bu yöntemi karakterize eden öznitelikler arasındaki ilişkinin dikkate alınmamasıdır [154].

5.3.13. Çapraz Doğrulama

Çapraz doğrulama (cross-validation) öğrenme algoritmalarının eğitim ve test şeklinde ikiye ayrılarak doğrulanması ve değerlendirilmesi için kullanılan istatistiksel bir yöntemdir. Bu yöntemin kullandığı metodolojilerden birisi k-fold çapraz doğrulamadır. Burada veri eşit boyutlu k adet parçaya (fold) bölünür. Sonra, eğitim ve test için k adet iterasyon gerçekleştirilir. Her iterasyonda eğitim için $k-1$ fold kullanılırken test için bir fold kullanılır [132].

Veri madenciliğinde, bu yaklaşım performans tahmini ve model seçimi için standart bir prosedür olarak kullanılmaktadır. Çalışmamızda SPAC sınıflandırma modülünde verinin eğitim ve test olarak ayrılması için k-fold çapraz doğrulama ve yüzde bölme kullanılmıştır. Yüzde bölmede verinin belirlenen yüzdelik kısmı eğitim geriye kalan kısmı test için kullanılır. k-fold çapraz doğrulamada kullanılacak k sayısı varsayılan olarak 10 ayarlanmış olup kullanıcı tercihinine göre değiştirebilmektedir. 5-fold çapraz doğrulama için verinin dağılımı Şekil 5.11’de verilmiştir.



Şekil 5.11. 5-fold çapraz doğrulama için verinin dağılımı.

5.4. İSTATİSTİKSEL ANALİZ

Tez kapsamında istatistiksel analiz yöntemleri akustik parametrelerin varyasyonlarını hesaplama, öznelilik normalizasyonu, öznelilik seçimi, SPAC yazılımının testi ve anksiyetenin akustik parametrelerdeki etkilerinin değerlendirilmesinde kullanılmıştır.

Akustik parametrelerde ortalama, standart sapma, maksimum, minimum, aralık, ortanca, delta (Δ), delta-delta (Δ^2), L1 norm, L2 norm, ortalama mutlak sapma, medyan mutlak sapma ve mod (en sık değer) istatistiksel yöntemleri kullanılmıştır. Bu yöntemlere ait eşitlikler Eşitlik 5.13'de verilmiştir. Formüllerde kullanılan x öznelilik vektörünü ve N öznelilik vektörünün boyutunu ifade etmektedir.

Ortalama;

$$\mu = \frac{1}{N} \sum_{i=1}^N x_i = \frac{1}{N} (x_1 + \dots + x_N) \quad (5.13a)$$

Standart sapma;

$$\sigma = \sqrt{\frac{1}{N-1} \sum_{i=1}^N (x_i - \mu_x)^2} \quad (5.13b)$$

Aralık;

$$aralik = \max(x) - \min(x) \quad (5.13c)$$

Ortanca (verinin tek sayıda değer içermesi durumunda);

$$med = \left(\frac{N+1}{2} \right)^{th} \text{ terim} \quad (5.13d)$$

Ortanca (verinin çift sayıda değer içermesi durumunda);

$$med = \frac{\left(\frac{N}{2} \right)^{th} \text{ terim} + \left(\frac{N}{2} + 1 \right)^{th} \text{ terim}}{2} \quad (5.13e)$$

Delta ve Delta-delta;

$$x_{delta} = \frac{\sum_{w=1}^W w.(x(t+w) - x(t-w))}{2 \sum_{w=1}^W w^2} \quad (5.13f)$$

Eşitlik 5.13f'de yer alan x , mevcut öznitelik kümesi ve x_{delta} mevcut öznitelik kümesinden türetilen yeni bir öznitelik kümesidir. w , komşu pencere indeksi; W yarı-pencere uzunluğunu göstermektedir ve çalışmada 2 alınmıştır. Delta-delta öznitelik kümesini elde etmek için elde edilen delta öznitelik kümesine bir kez daha delta uygulaması yapılır.

L norm;

$$L_p = \sqrt[p]{\sum_{i=1}^N |x_i|^p} \quad (5.13g)$$

Örneğin L1 norm için p değeri 1'dir.

Ortalama mutlak sapma;

$$OMS = \frac{1}{N} \sum_{i=1}^N |x_i - \mu_x| \quad (5.13h)$$

Medyan mutlak sapma;

$$MMS = \frac{1}{N} \sum_{i=1}^N |x_i - med(x)| \quad (5.13i)$$

Öznitelik normalizasyonunda z-score normalizasyon kullanılmıştır. Bir x özelliği için z-score normalizasyon yöntemi Eşitlik 5.14'de verilmiştir [79].

$$x_n = \frac{x - \mu}{\sigma} \quad (5.14)$$

Eşitlik 5.14'de yer alan μ , x özelliğinin ortalaması ve σ standart sapmasıdır.

SPAC yazılımının test edilmesi için karesel ortalama hata (MSE) ve yüzde ortalama hata (MPE) yöntemleri kullanılmıştır. Bu yöntemlere ait formüller Eşitlik 5.15 ve Eşitlik 5.16'da verilmiştir [55].

$$E_{MSE} = \frac{1}{N} \sum_{i=1}^N |x_i - y_i|^2 \quad (5.15)$$

Eşitlik 5.15'de x_i test verisini içeren öznitelik vektörünün i . elemanı ve y_i standart veriyi içeren öznitelik vektörünün i . elemanıdır.

$$E_{MPE} = \frac{1}{N} \sum_{i=1}^N \frac{|x_i - y_i|}{|x_i|} \times 100 \quad (5.16)$$

Eşitlik 5.16'da x_i test verisini içeren öznitelik vektörünün i . elemanı ve y_i standart veriyi içeren öznitelik vektörünün i . elemanıdır.

Akustik parametreler anksiyete açısından değerlendirilirken tekrarlı ölçümlerde Friedman ve Wilcoxon testi, gruplar arası ilişkilerde Mann Whitney-U testi ve parametrelerin benzerlikleri incelenirken korelasyon analizi kullanılmıştır.

Friedman testi, birbiriyle ilişkili 2 veya daha fazla değişkene ait dağılımların arasındaki farklılığı incelemek için ve tekrarlı ölçümlerde ölçümler arasındaki ilişkinin araştırılması için kullanılır [155].

Wilcoxon testi ilişkili iki değişken arasındaki ilişkiyi incelerken hem gruplar arasındaki farkların yönünü hem de miktarını dikkate alır [155].

Mann-Whitney U Testi bağımsız iki değişken arasındaki ilişkinin ölçülmesi için kullanılır. Grupların medyanlarını karşılaştırır. Sürekli değişkenlerin, iki grup içerisinde değerlerini sıralı hale dönüştürür. Böylece, iki grup arasındaki sıralamanın farklı olup olmadığını değerlendirir [155].

Korelasyon bağımlı değişkenle bağımsız değişken veya değişkenler arasındaki ilişkinin gücünü gösteren bir katsayıdır. Korelasyon katsayısı, değişkenlerin yönü ve etkileşimlerin nasıl olduğu hakkında bilgi verir. Değişkenlerin birbiri arasındaki etkileşim var mı, varsa etkileşimin kuvvetli mi olduğu ve gözlem gruplarından birinin gözlem değerleri artarken diğerinin azalıyor mu, yoksa aynı yönde mi değerleri değişiyor olduğu korelasyon katsayısı ile gözlenebilir. Korelasyon katsayısı iki değişken arasındaki doğrusal ilişkinin ölçüsü olup, incelenen değişkenlerin birimlerinden bağımsızdır ve $-1 \leq r \leq 1$ arasındadır. Korelasyon katsayısının 0'a yaklaşması değişkenler arasında zayıf ilişkinin varlığını gösterir. Değişkenler birlikte artıyor veya azalıyorsa pozitif yönde, değişkenlerden biri artarken diğeri azalıyorsa ise negatif yönde bir ilişki vardır [156]. Korelasyon aralığına göre ilişki düzeyi aşağıda Çizelge 5.1'de verilmiştir.

Çizelge 5.1. Korelasyon aralıklarına göre ilişki düzeyleri [156].

Korelasyon Aralığı	İlişki Düzeyi
-0,25...0,00 ve 0,00...0,25	Çok Zayıf
-0,49...-0,26 ve 0,26...0,49	Zayıf
-0,69...-0,50 ve 0,50...0,69	Orta

Çizelge 5.1. (devam ediyor).

-0,89...-0,70 ve 0,70...0,89	Yüksek
-1,00...-0,90 ve 0,90...1,00	Çok Yüksek

İki değişken için korelasyon katsayısı Eşitlik 5.17’de verilen formül ile hesaplanır.

$$r = \frac{\sum xy - \frac{\sum x \sum y}{N}}{\sqrt{\left(\sum x^2 - \frac{(\sum x)^2}{N} \right) \left(\sum y^2 - \frac{(\sum y)^2}{N} \right)}} \quad (5.17)$$

5.5. SINIFLANDIRMA

Örüntü keşfi uygulamalarında en çok kullanılan yöntemlerden birisi olan sınıflandırma, bir veriyi daha önceden tanımlanmış sınıflara dağıtma tekniğidir. Hangi sınıflandırma yönteminin en iyi ya da en uygun olduğu kesin olarak söylenemez [157]. Örüntünün çeşidi, örüntüden elde edilen özneliklerin yapısı ve sayısı, işlem süresi ve karmaşıklığı gibi faktörlere göre tercih edilmesi gereken sınıflandırma yöntemi probleme göre farklılık gösterebilir. Tez kapsamında hazırlanan SPAC yazılımında sınıflandırma için SVM, k-NN ve MLP sınıflandırıcılar kullanılmıştır.

5.5.1. Destek Vektör Makineleri

SVM istatistiksel öğrenme teorisi temelli sınıflandırma algoritmasıdır. Başlangıçta iki sınıflı doğrusal verilerin sınıflandırılması için kullanılırken daha sonra çok sınıflı ve doğrusal olmayan verilerin sınıflandırılması için geliştirilmiştir.

SVM temel olarak iki sınıfı birbirinden ayırabilen en uygun hiper-düzlemin tanımlanması esasına dayanır [158].

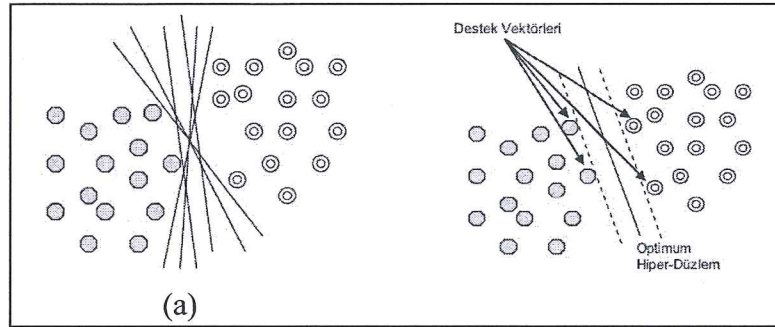
SVM’nin eğitimi için n sayıda örnekten oluşan eğitim verisinin $\{x_n, y_n\}$, $n = 1, \dots, k$ olduğu kabul edilirse, optimum hiper-düzleme ait eşitsizlikler aşağıdaki şekilde olur.

$$w \cdot x_n + b \geq +1 \quad \text{her } y = +1 \text{ için} \quad (5.18a)$$

$$w \cdot x_n + b \leq -1 \quad \text{her } y = -1 \text{ için} \quad (5.18b)$$

Burada $x \in R^N$ olup N-boyutlu bir uzayı, $y \in \{-1, +1\}$ ise sınıf etiketlerini, w ağırlık vektörünü (hiper-düzlemin normali) ve b eğilim değerini göstermektedir [159], [160].

Şekil 5.12 (a)'da görüldüğü gibi iki sınıflı verileri birbirinden ayırmak için birden fazla hiper-düzlem çizilebilir. SVM'nin amacı Şekil 5.12 (b)'de görüldüğü gibi sınıflara ait en yakın noktalar arasındaki uzaklığı maksimuma çıkaran hiper-düzlemi bulmaktır.



Şekil 5.12. a) İki sınıflı bir problem için oluşturulabilecek hiper-düzlemler, b) Optimum hiper-düzlem ve destek vektörleri [161].

Sonuç olarak, doğrusal olarak ayrılabilen iki sınıflı bir problem için karar fonksiyonu aşağıdaki şekilde yazılabilir [159], [160].

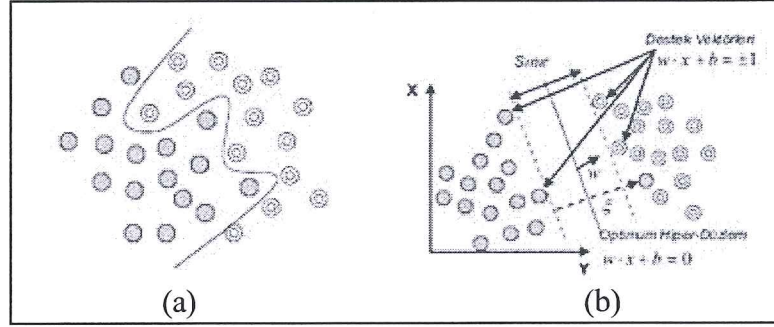
$$f(x) = \text{sign}\left(\sum_{n=1}^k \lambda_n y_n(x x_n) + b\right) \quad (5.19a)$$

Şekil 5.13 (a)'da görüldüğü gibi birçok durumda sınıfların doğrusal olarak ayrılması mümkün olmamaktadır. Bu tür durumlar Şekil 5.13 (b)'de görüldüğü gibi tanımlanan pozitif bir yapay değişken (ξ_n) ile çözülür. Doğrusal ayrılabilen sınıflar için kullanılan eşitler tekrar düzenlenirse aşağıdaki şekilde olur.

$$w \cdot x_n + b \geq +1 - \xi_n \quad \text{her } y = +1 \text{ için} \quad (5.18c)$$

$$w \cdot x_n + b \leq -1 + \xi_n \quad \text{her } y = -1 \text{ için} \quad (5.18d)$$

$\xi_n=0$ olması durumunda x_n örneği doğru sınıflandırılmış, $0 < \xi_n < 1$ olması durumunda doğru sınıflandırılmış ama aşırı düzlemler arasında yer alıyor, $\xi_n > 1$ ise yanlış sınıflandırılmış demektir.

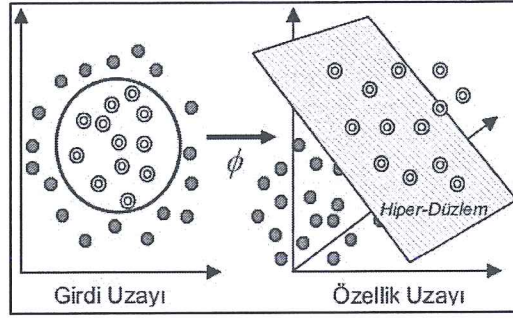


Şekil 5.13. a) Doğrusal olarak ayrılamayan sınıflar, b) Doğrusal ayrılamayan sınıflar için hiper düzlemin belirlenmesi [161].

SVM doğrusal olarak ayrılamayan verileri çekirdek fonksiyonları yardımıyla çok boyutlu düzleme taşıyarak sınıflandırma yapabilmektedir. Sonuç olarak doğrusal olarak ayrılamayan iki sınıflı bir problemin çözümü ile ilgili kural Eşitlik 5.19b'deki şekilde yazılabilir (φ : çekirdek fonksiyonu) [159,160]:

$$f(x) = \text{sign}\left(\sum_{n=1}^k \alpha_n y_n \varphi(x) \varphi(x_n) + b\right) \quad (5.19b)$$

SVM temel olarak eğitim ve sınıflandırma olmak üzere iki aşamada gerçekleşmektedir. Eğitim aşamasında sınıfları birbirinden ayıracak öznelikteki veriler öznelilik vektörü ile ifade edilir ve çekirdek fonksiyonları ile sınıfları ayıran hiper-düzlem tespit edilir.



Şekil 5.14. Çekirdek fonksiyonu ile verinin daha yüksek bir boyuta dönüştürülmesi [161].

Çekirdek fonksiyonları sınıflandırmanın daha yüksek boyutlarda yapılabilmesi için taşıma işleminde kullanılabilir. Uygulamaların büyük çoğunluğunda birçok çekirdek fonksiyonu iyi sonuçlar verir. Sık kullanılan çekirdek fonksiyonları doğrusal, polinom, radyal tabanlı ve sigmoid fonksiyonlarıdır [41].

Çekirdek fonksiyonları karşılaştırıldığında polinom ve radyal tabanlı çekirdeklerin daha sade ve anlaşılabilir olduğu ifade edilebilir. Matematiksel olarak basit görünse de, polinomun derecesindeki artış algoritmanın karmaşık bir hal almasına neden olmaktadır. Bu da hem işlem süresini önemli ölçüde artırmakta hem de bir noktadan sonra sınıflandırma doğruluğunu düşürmektedir. Buna karşın radyal tabanlı fonksiyonun çekirdek boyutundaki değişimlerin sınıflandırma performansına etkisinin daha az olduğu görülmüştür [162].

Doğrusal çekirdek, sınıflandırma işlemini doğrular çizerek tanımlar. Bu yüzden dikdörtgen ve kare olan şekillerin sınıflandırılmasında daha etkilidir. Radyal tabanlı çekirdek, sınıflandırma işlemini doğrusal olmayan örneği daha yüksek boyutlu bir uzaya taşıyarak gerçekleştirir. Polinom çekirdek, radyal tabanlı'ya göre daha fazla parametre içerir. Bu yüzden, radyal tabanlı çekirdeğin daha az sayısal zorlukları bulunmaktadır. Sonuç olarak, düşük derecelerde polinom çekirdek ile iyi sonuçlar alınabiliyorsa polinom çekirdek, sınıflandırmada kaçan kısımlar oluyorsa radyal tabanlı çekirdek kullanılmalıdır. Doğrusal çekirdek, öznelik sayısının çok fazla olduğu durumlarda daha iyi sonuçlar vermektedir.

SVM temel olarak iki sınıflı problemler için tasarlanmıştır. Çoklu sınıflandırmada SVM şekilde kullanılabilir. Bunlar [41];

1. Multi-Class Ranking SVM yaklaşımında tüm sınıfların sınıflandırılması için tek bir karar fonksiyonu kullanılır.
2. One-Against-Rest (One-vs-Rest) yaklaşımı bir sınıfın diğerlerinden ayrılmasını temel alır. Bu yaklaşımda k adet ikili sınıflandırıcı oluşturulur. i 'nci sınıflandırıcı, i 'nci sınıftaki pozitif etiketli tüm örnekler ile negatif etiketli diğer tüm örneklerle eğitilir. Sonuç en yüksek sınıflandırma başarısına karşılık gelen sınıftır.
3. One-Against-One (One-vs-One) yaklaşımı bir sınıfın diğer bir sınıftan ayrılmasını temel alır. Bu yaklaşımda k adet sınıflandırıcı için tanımlanabilecek tüm ikili sınıflandırıcılar tanımlanır. Sonuç elde edilirken her sınıflandırıcı belirlediği bir sınıf için oy verir ve en çok oy alan sınıfa göre sonuç belirlenir (voting).

Tez kapsamında hazırlanan SPAC üzerinde SVM sınıflandırıcı için one-vs-one yaklaşımı kullanılmıştır. SVM çekirdek yapısı olarak da doğrusal, karesel, radyal tabanlı, polinom ve mlp çekirdekleri kullanıcının tercihine göre kullanılabilir.

5.5.2. k-En Yakın Komşu

k-NN sınıflandırıcı örnek tabanlı öğrenme algoritması olup gözlemlerin uzaklıklarına dayalıdır. Önce eğitim kümesinin k yakın komşuları hesaplanır, sonra test verisinden bir örneğin k en yakın komşuluklarına olan benzerlikler komşulukların sınıflarına göre kümelendir. k-NN'in bir avantajı sınıflandırma kararının benzer nesnelere küçük komşuluklarına dayalı olmasından ötürü çoklu sınıflar için uygun olmasıdır. Yakınlıkla kasıt Öklid mesafesi gibi bir uzaklık ölçütü ile uzaklığı belirlemektir. Genel olarak en iyi k parametresinin seçimi veriye bağlı olup k 'nın büyük değerleri sınıflandırma üzerindeki gürültüyü azaltır, fakat sınıflar arasındaki sınırların belirginliğini en aza indirir. İyi bir k değerinin seçimi çapraz geçerlilik (cross-validation) gibi farklı sezgisel tekniklerle yapılabilir [41].

Yeni bir örneğin sınıfı belirlenirken diğer örneklerle uzaklığı uzaklık ölçütü hesaplanarak, yeni örneğe kendisine en yakın örneğin sınıf bilgisi verilebilir. Bu işlemde birden fazla örneğe ait uzaklıklar da kullanılabilir [117–119]. Kullanılabilecek uzaklık ölçütleri Çizelge 5.2’de verilmiştir.

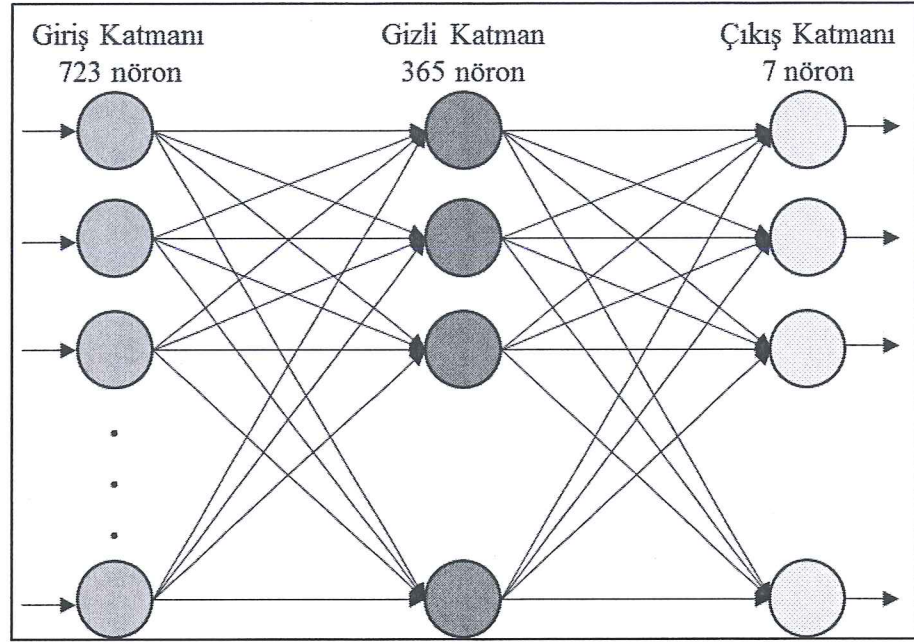
Çizelge 5.2. Uzaklık ölçütleri.

Uzaklık Ölçütü	Formül
Öklid	$d(x_s, x_t) = \sqrt{(x_{s_1} - x_{t_1})^2 + (x_{s_2} - x_{t_2})^2 + \dots + (x_{s_n} - x_{t_n})^2}$ (5.20a)
Mahalanobis [166]	$d(x_s, x_t) = \sqrt{(x_s - x_t)^T C^{-1} (x_s - x_t)}$ (5.20b)
Manhattan [167]	$d(x_s, x_t) = \sum_{i=1}^n x_{s_i} - x_{t_i} $ (5.20c)
Minkowski [163]	$d(x_s, x_t) = \sqrt[p]{\sum_{i=1}^n x_{s_i} - x_{t_i} ^p}$ (5.20d)
Kosinüs	$d(x_s, x_t) = 1 - \cos(x_s, x_t)$ (5.20e)
Korelasyon	$d(x_s, x_t) = 1 - \frac{C}{\sigma_{x_s} \sigma_{x_t}}$ (5.20f)
Tabloda verilen x_s ve x_t uzaklıkları ölçülecek iki örneği, iki örneğin C kovaryans matrisini, σ_{x_s} , x_s örneğinin, σ_{x_t} , x_t örneğinin standart sapmasını göstermektedir.	

Tez kapsamında hazırlanan SPAC üzerinde k-NN sınıflandırıcı için kullanılacak uzaklık ölçütü için korelasyon, öklid ve kosinüs seçenekleri ve k değeri için 1 ile 10 arasında seçenek sunulmuştur.

5.5.3. Çok Katmanlı Algılayıcı

MLP birden fazla gizli katmana sahip olan ve genellikle ileri yönlü (feed-forward) bağlantılı bir yapay sinir ağı modelidir. MLP en çok kullanılan YSA türüdür [41]. Aşağıdaki şekilde giriş, gizli ve çıkış katmanından oluşan bir MLP modeli verilmiştir.



Şekil 5.15. EMO-DB verisetinde sınıflandırma için kullanılan MLP modeli.

Giriş katmanı, giriş verilerinin ağı sunulduğu katmandır. Bu katmandaki nöron sayısı giriş veri sayısı kadardır. Bu katmanda giriş verileri ile ilgili bir işlem yapılmadan veriler gizli katmana geçiş yapılır. Gizli katman şekilde görüldüğü gibi 2 veya daha fazla katmandan oluşabilir. Gizli katman ağ ile ilgili temel işlevlerin yapıldığı katmandır. Gizli katman sayısı ve katmandaki nöron sayısı MLP modelinin kullanılacağı probleme göre değişiklik göstermekle birlikte sabit bir sayı değildir. Bu katmanın temel görevi giriş katmanından alınan veriyi belirlenen fonksiyonla işleyerek verinin bir sonraki katmana iletilmesidir. Çıkış katmanı gizli katmandan aldığı veriyi ağı kullandığı fonksiyonla işleyerek çıktıyı oluşturur. Çıkış katmanında kullanılan nöron sayısı ağı sunulan verinin çıkışı kadardır. Örneğin, 10 öznitelik kümesi ve 3 sınıf içeren bir veri için giriş nöronu sayısı 10, çıkış nöronu sayısı 3 olup gizli katmandaki nöron sayısı problemin içeriğine göre değişmektedir.

Tez kapsamında hazırlanan SPAC üzerinde MLP sınıflandırıcı için ileri yönlü yapı kullanılmıştır. Kullanılacak katman sayısı ve eğitim yöntemi kullanıcı tarafından belirlenebilmektedir. Eğitim yöntemi olarak bayesian regularization, Levenberg-Marquardt, conjugate gradient ve gradient descent tabanlı geri yayılım (backpropagation) algoritmaları kullanılabilir.

BÖLÜM 6

BULGULAR

6.1. SPAC ÖZELLİK KÜMESİNİN BAŞARI ANALİZİ

SPAC ile elde edilen özniteliklerin doğruluğunu değerlendirmek için iki yöntem kullanılmıştır. Bunlardan ilkinde SPAC öznitelik kümesi Praat ve OpenSMILE öznitelik kümeleri ile karşılaştırılmıştır. Diğer yöntemde ise SPAC ile çıkartılan özniteliklerin duyguları sınıflandırma başarısı test edilmiştir.

6.1.1. Test Verisi

SPAC testi için Berlin Database of Emotional Speech (EMO-DB), EMOVA, eNTERFACE05 ve Surrey Audio-Visual Expressed Emotion (SAVEE) hazır veritabanları kullanılmıştır.

EMO-DB veritabanı 7 farklı duygunun (öfke, korku, mutluluk, nötr, üzüntü, sıkıntı, tiksinti) aktörler tarafından Almanca dilinde simüle edilmesi ile oluşturulmuştur. İçerdiği ses kayıtları 16 kHz örnekleme frekansına sahip olup 16 bit mono'dur [72]. EMOVA veritabanı 6 farklı aktör tarafından İtalyanca 14 cümle okunarak 7 farklı duygunun (öfke, korku, sevinç, nötr, üzüntü, sürpriz, tiksinti) simüle edilmesi ile oluşturulmuştur. İçerdiği ses kayıtları 48 kHz örnekleme frekansına sahip olup 16 bit stereo'dur [168]. eNTERFACE'05 duygusal durumların görsel ve işitsel olarak ifade edildiği veritabanıdır. 14 farklı ülkeden 42 katılımcı ile oluşturulmuştur (Belçika, Türkiye, Fransa, İspanya, Yunanistan, İtalya, Avusturya, Küba, Slovakya, Brezilya, ABD, Hırvatistan, Kanada, Rusya). Katılımcılar 6 farklı duygu durum (öfke, tiksinti, korku, mutluluk, hüzn, sürpriz) içeren İngilizce dilinde 6 farklı cümle okumuştur [77]. SAVEE veritabanı da eNTERFACE'05 gibi duygusal durumların görsel ve işitsel olarak ifadesini içerir. İngilizce dilinde 4 erkek konuşmacının 7 (öfke, korku,

mutluluk, nötr, üzüntü, sürpriz, tiksinti) farklı duygu durum içeren 14 farklı cümleyi İngilizce dilinde okumuştur [169].

Testi verisi olarak üç veritabanında ortak 4 duygu için (öfke, korku, mutluluk/sevinç, üzüntü) 5'er örnek alınarak 20 konuşma seçilmiştir. Farklı verisetlerinin kullanım amacı dil/söylem bilgilerine ve katılımcıya göre oluşacak farklılıkları ortadan kaldırmaktır.

6.1.2. Test Yöntemleri

SPAC doğruluğunu değerlendirmede öznitelik kümeleri karşılaştırılırken MSE ve MPE yöntemleri kullanılmıştır. Her iki yöntemde de standart veri Praat ve OpenSMILE ile test verisi SPAC ile elde edilmiştir.

SPAC'ın doğruluğunu değerlendirmek için kullanılan diğer yöntemde ise SPAC, Praat ve OpenSMILE ile elde edilen öznitelik kümelerinin sınıflandırma performansları EMO-DB veritabanı üzerinde karşılaştırılmıştır. Duygusal durum tespitinde çoğunlukla kullanılan SVM, k-NN ve MLP sınıflandırıcılar kullanılmıştır.

6.1.2.1. MSE Ve MPE Sonuçları

OpenSMILE, Praat ve SPAC ile 20 veriden elde edilen özelliklerin değerlendirilmesi için MSE ve MPE kullanılmış ve sonuçlar Çizelge 6.1'de verilmiştir.

Çizelge 6.1. MSE ve MPE sonuçları.

Öznitelik	OpenSMILE-Praat		OpenSMILE-SPAC		Praat-SPAC	
	Ort. MSE	Ort. MPE	Ort. MSE	Ort. MPE	Ort. MSE	Ort. MPE
Pitch	11,494	% 4,2	2,195	%1,5	0,027	% 0,2
Formant	-	-	-	-	20,156	% 5,7
Ses Kalitesi	-	-	-	-	0,021	% 2,9
LPCC	-	-	-	-	0,001	% 8,7
MFCC	14,594	% 70,3	0,231	%30,4	0,029	% 9,7
ZCR	0,021	% 5,6	0,018	%5,2	0,078	% 11,4
Enerji	0,253	% 13,1	0,383	%16,2	0,014	% 2,7
Konuşma	-	-	-	-	0,299	% 5,3

Çizelge 6.1’de verilen sonuçlar incelendiğinde SPAC ile mevcut yazılımlar arasında farklılıklar vardır ve farklılık derecesi öznitelik kümesine göre değişiklik göstermektedir. Pitch öznitelik kümesi incelendiğinde OpenSMILE ve Praat arasında % 4,2 farklılık oluşurken OpenSMILE-SPAC arasında % 1,5 ve Praat-SPAC arasında % 0,2 farklılık oluşmaktadır. SPAC öznitelik kümeleri diğer yazılımların sonuçlarına göre çok büyük farklılıklar içermediği için geçerliliği kabul edilebilir düzeydedir. Genel anlamda SPAC ile elde edilen öznitelik kümeleri Praat’a daha yakındır.

6.1.2.2. Sınıflandırma Sonuçları

SPAC doğruluğunu değerlendirmek için kullanılan bir diğer yöntem SPAC ile elde edilen özniteliklerin sınıflandırmada kullanılmasıdır. Bu amaçla OpenSMILE, Praat ve SPAC ile EMO-DB veritabanının içerdiği 535 ses dosyasından öznitelik çıkartımı gerçekleştirilmiştir. Üç yazılım üzerinden elde edilen öznitelik kümelerinin başarısı SVM, k-NN ve MLP sınıflandırıcılar ile analiz edilmiştir. SVM için linear çekirdek türü ve k-NN için euclidean uzaklık ölçümü kullanılmıştır. MLP için epoch 500, öğrenme oranı 0,3 ve gizli katman nöron sayısı (öznitelik sayısı + sınıf sayısı)/2 kadar kullanılmıştır.

SPAC ile elde edilen 723 öznitelik SVM, k-NN ve MLP sınıflandırıcılar ile sınıflandırılmış ve Çizelge 6.2’de sonuçlar verilmiştir. Diğerleri öznitelik kümesi F0, formants, ses kalitesi, ZCR, enerji ve konuşma özniteliklerini içermektedir.

Çizelge 6.2. SPAC öznitelik kümelerinin sınıflandırma başarısı.

Öznitelik Kümesi	Öznitelik Boyutu	Sınıflandırma Başarısı		
		SVM	k-NN	MLP
MFCC	234	% 69,5	% 58,7	% 74,1
LPCC	234	% 67,5	% 59,4	% 71,3
Dalgacık	72	% 62,3	% 49,2	% 60,3
Diğerleri	183	% 68,6	% 51,3	% 72,3
Tümü	723	% 82,1	% 65,4	% 85,1

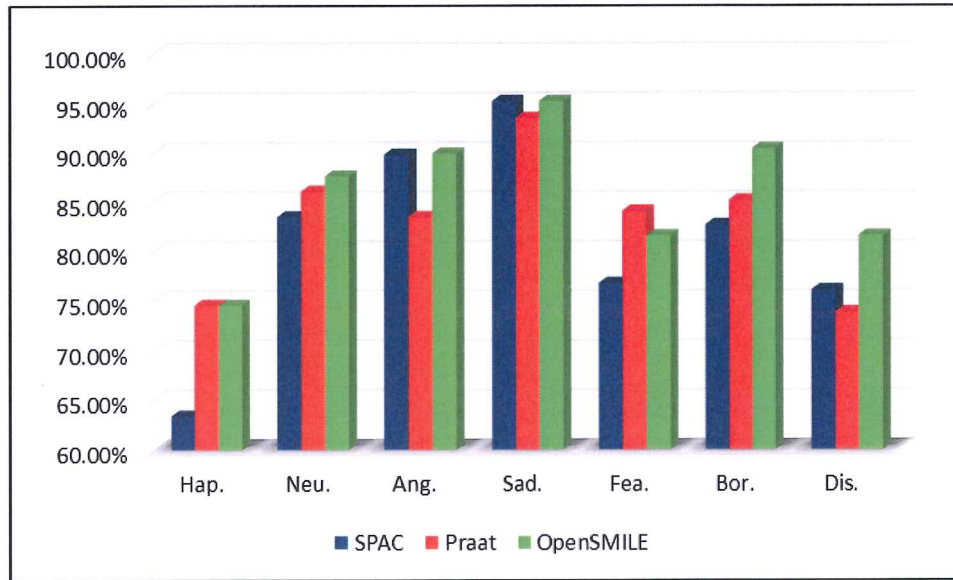
Çizelge 6.2 incelendiğinde tek bir öznitelik kümesinin sınıflandırma başarısını önemli oranda düşürdüğü ve birden fazla öznitelik kümesinin kullanımının

sınıflandırma başarısını artırdığı görülmektedir. Sınıflandırma başarısını en fazla MFCC öznitelik kümesi artırmaktadır ve bunu sırasıyla diğerleri, LPCC ve dalgacık katsayıları takip etmektedir. Tüm öznitelik kümesindeki sonuçlar incelendiğinde SPAC öznitelik kümesi üç sınıflandırıcıda da kabul edilebilir seviyede başarı elde etmiştir.

Bir diğer sınıflandırma analizinde SPAC'ın sınıflandırma başarısı OpenSMILE ve Praat öznitelik kümelerinin sınıflandırma başarısı ile karşılaştırılmış ve genel başarı sonuçları Çizelge 6.3'de sınıflandırma bazlı sonuçlar Şekil 6.1'de verilmiştir.

Çizelge 6.3. SPAC, Praat, OpenSMILE öznitelik kümeleri üzerinde genel başarı.

Araç	Öznitelik Boyutu	Sınıflandırma Baş.		
		SVM	k-NN	MLP
SPAC	723 (F0, F1, F2, F3, jitter, shimmer, ZCR, enerji, MFCC, LPCC, dalgacık katsayıları)	%82,1	%65,4	%85,1
Praat	123 (F0, F1, F2, F3, jitter, shimmer, HNR, enerji, MFCC, LPCC)	%83,4	%68,8	%85,2
OpenSMILE	1583 (Loudness, MFCC, logMelFreqBand, lspFreq, F0, jitter, shimmer)	%85,2	%61,4	%87,6



Şekil 6.1. SPAC, Praat ve OpenSMILE öznitelik kümelerinde SVM sınıf bazlı başarı.

Çizelge 6.3 ve Şekil 6.1'e göre SPAC ile elde edilen başarı % 1 - % 3 farklılıkla diğer yazılımlarla elde edilen sonuçlara yakındır. Ayrıca SPAC her sınıf için sınıflandırma gerçekleştirmiştir. Bu durum SPAC ile elde edilen öznitelik kümelerinin geçerliliğini göstermektedir.

6.2. OLGULARIN SOSYODEMOGRAFİK ÖZELLİKLERİ

DeneySEL çalışmalarda olguların sosyodemografik özelliklerinin benzerlik göstermesi gerekmektedir. Aksi durumda elde edilen sonuçların güvenilirliği tartışmaya açıktır. Bu nedenle, çalışmaya katılan olgulara ait veriler normal dağılım göstermediği için olguların benzerlikleri Mann-Whitney U testi ile incelenmiştir. Çalışmaya katılan gönüllülere ait tanımlayıcı istatistikler Çizelge 6.4'de ve istatistiksel analiz sonuçları Çizelge 6.5'de verilmiştir.

Çizelge 6.4. Olguların sosyodemografik verileri.

Demografik Özellik	Hasta Grubu				Kontrol Grubu			
	%	Min	Mak	Ort. ± SS	%	Min	Mak	Ort. ± SS
Erkek	26	-	-	-	55	-	-	-
Bayan	74	-	-	-	45	-	-	-
Yaş (yıl)	-	17	50	35,2 ± 8,9	-	23	55	34,1 ± 10,2
Öğrenim (yıl)	-	8	18	13,4 ± 2,8	-	8	19	15,2 ± 3,0
Boy (cm)	-	150	185	165,4 ± 9,6	-	158	185	170,2 ± 7,9
Kilo (kg)	-	54	165	76,1 ± 13,4	-	48	96	70,6 ± 11,8
BMI	-	19,4	48,2	27,5 ± 5,9	-	17,8	32,1	24,3 ± 3,3
Evli	65	-	-	-	60	-	-	-
Bekar	35	-	-	-	40	-	-	-
Sigara Kullanan	22	-	-	-	40	-	-	-
Sigara Kullanmayan	78	-	-	-	60	-	-	-
Beck Ank. Envanteri	-	12	53	31,0 ± 12,4	-	0	7	4,0 ± 2,5
Hamilton Ank. Ölçeği	-	10	52	26,4 ± 11,2	-	0	6	2,2 ± 1,7

Ort.: Ortalama, SS: Standart Sapma

Çizelge 6.5. Olguların sosyodemografik verilerinin karşılaştırması.

Demografik Özellik	MWU		
	U	z	p
Cinsiyet	163,5	-1,912	0,056
Yaş (yıl)	203,0	-0,658	0,510
Öğrenim (yıl)	139,5	-2,231	0,026*
Boy (cm)	151,5	-1,916	0,055
Kilo (kg)	222,0	-0,195	0,834
BMI	151,0	-1,924	0,054
Medeni Durumu	218,0	-0,349	0,727
Sigara Kullanımı	188,0	-1,285	0,199
Beck Ank. Envanteri	0,0	-5,608	0,000*
Hamilton Ank. Ölçeği	0,0	-5,610	0,000*

* $p \leq 0,05$, BMI: Beden kitle indeksi

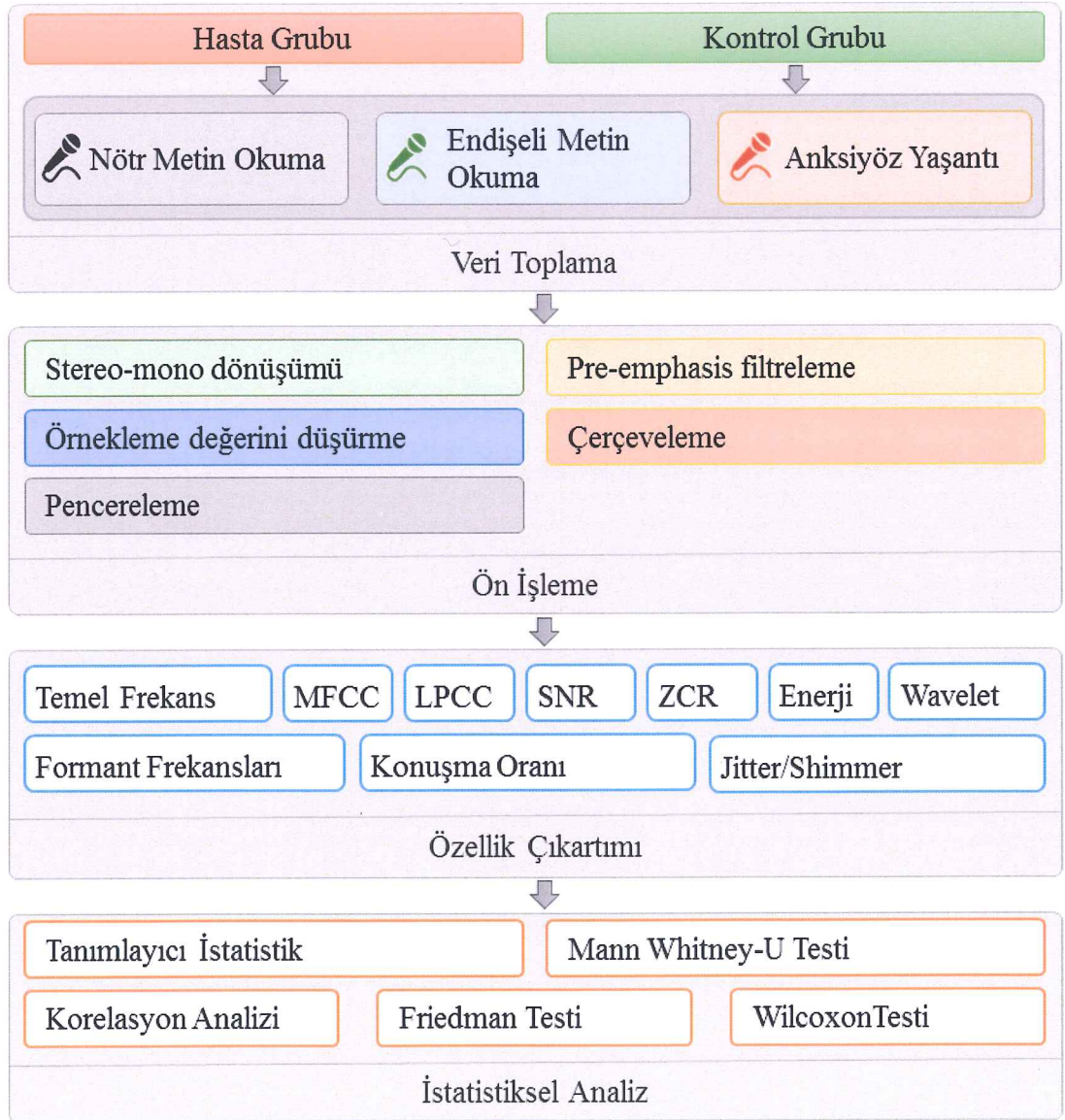
Çizelge 6.5'deki sonuçlar incelendiğinde, çalışmaya katılan olguların öğrenim düzeyleri dışındaki sosyodemografik verileri gruplar arasında benzer bulunmuştur ($p > 0,05$). Kontrol grubunun eğitim düzeyi hasta gruptan daha yüksektir ($p \leq 0,05$). Genel olarak olguların sosyodemografik özellikleri benzerlik gösterdiği için olgulara ait veriler analizlerde kullanılabilir.

6.3. AKUSTİK PARAMETRELERİN ANKSİYETE İLE İLİŞKİSİ

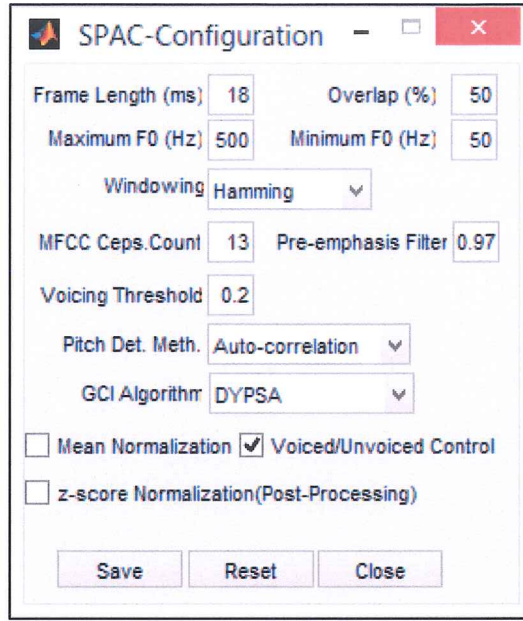
Bölüm 5'de ayrıntıları verilen anksiyöz konuşma verisinin toplanması sonrası 23'ü hasta ve 20'si kontrol grubunda olmak üzere toplam 43 olguya ait ses kayıtları elde edilmiştir. Her bir olgunun nötr metin okuma, endişe metni okuma ve anksiyöz yaşantı olmak üzere 3 durum için ses kaydı alındığı için toplam ses kaydı sayısı 129'dur. Ses kayıtlarının toplanması ve sonraki sürece ait akış diyagramı Şekil 6.2'de verilmiştir.

Veri toplama aşaması sonrası elde edilen her bir ses kaydı ön işleme aşamasına tabi tutularak hem iş yükü azaltılmış hem de öznelik çıkartımı için hazır hale getirilmiştir. Ses kayıtları stereo olduğu için öznelik çıkartımı öncesi monoya dönüştürülmüştür. Ayrıca pre-emphasis filtreleme uygulanarak konuşmalardaki vurgular güçlendirilmiş ve örnekleme değeri düşürülerek öznelik çıkartımı

aşamasının iş yükü azaltılmıştır. Çerçeveleme ile ses kayıtları 18 ms'lik parçalara %50 örtüşme ile bölünmüş ve pencereleme ile her bir çerçevedeki süreksizlikler kaldırılmıştır. SPAC ile parametreler elde edilirken kullanılan yapılandırma Şekil 6.3'de verilmiştir.



Şekil 6.2. Deneysel çalışma akış diyagramı.



Şekil 6.3. Öznitelik çıkartımı aşamasında kullanılan ön tanımlı ayarlar.

Öznitelik çıkartımı aşamasında konuşma sinyali çerçevelere bölünmüş her bir çerçeveden akustik parametreler elde edilmiştir. Böylece her bir dosya için bir akustik parametrenin çerçeve sayısı kadar değeri bulunmaktadır. Bu değerler tanımlayıcı istatistikler (ortalama ve standart sapma) kullanılarak her bir ses kaydı için bire düşürülmüştür. Çalışmada kullanılan parametre sayısı 122 olup kullanılan parametreler Çizelge 6.6’da verilmiştir.

Çizelge 6.6. Kullanılan akustik parametreler.

Akustik Parametre	Tanımlayıcı İstatistik	
	Ortalama	Standart Sapma
F0, F1, F2 ve F3	✓	✓
F1, F2 ve F3 bandwidh	✓	
MFCC1...MFCC13	✓	
LPCC1...LPCC13	✓	
Jitta, Jitt, Shimmer ve ShimmerDB	✓	
ZCR	✓	✓
SNR ve SNR power	✓	
Enerji	✓	✓
Konuşma Oranı	✓	
Duraksama Oranı	✓	
Duraksama Sayısı	✓	
Dalgacık Katsayıları (D1...D5, A1)	Ortalama, medyan, mod, maks, min, range, std.sp., mad, mad2, norm1, norm2, norminf	

Öznitelik çıkartımı sonrası elde edilen akustik parametrelerin istatistiksel analizi için IBM Statistics 20 yazılımı kullanılmıştır. Sonuçlar, %95 anlamlılık ($p \leq 0,05$) düzeyinde değerlendirilmiştir. Çalışmada kullanılan veriler 2 grup üzerinde gerçekleştirilmiş tekrarlı ölçümleri (Nötr, Endişe ve Anksiyete) içermektedir. Bu nedenle tekrarlı ölçümler arasındaki ilişkiyi analiz için Friedman testi kullanılmış ve gruplar arasında farklılık gösteren parametreler tespit edilmiştir. Farklılığın hangi grupta gerçekleştiğini tespit için Wilcoxon testi kullanılmıştır. Hasta ve sağlam olgular arasındaki farklılıkları analiz için Mann Whitney-U testi ve parametreler arası ilişki için korelasyon analizi kullanılmıştır.

Gerçekleştirilen çalışmanın araştırma soruları;

1. Elde edilen akustik parametrelerden benzerlik gösterenler var mı?
2. Çalışma grupları ve duygu durumları üzerinde duyguları birbirinden ayırt etmek için etkin olan parametreler hangileridir?
3. Endişe ve anksiyetenin etkilediği parametrelerin değişimi nasıl?
4. Hasta grubu ve kontrol grubu arasında farklılık var mı?
5. Hasta ve kontrol grupları duygularını sese yansıtmada farklılık içeriyor mu?

6.3.1. Benzerlik Gösteren Parametreler

Çalışmada gerçekleştirilecek olan analiz sayısını ve iş yükünü azaltmak için benzerlik gösteren parametreler tek bir parametre altında toplanmıştır. Benzerlik gösteren parametreleri tespit etmek için akustik parametrelerin duygu durumuna göre almış olduğu değerlerin değişimleri (nötr-endişe, nötr-anksiyete, endişe-anksiyete) korelasyon yöntemi ile tespit edilmiştir. Korelasyon katsayısı 0,95 ve yukarı olan parametreler benzer kabul edilmiş ve sadece bir tanesi kullanılmıştır. Benzerlik gösteren parametreler Çizelge 6.7'de verilmiştir.

Çizelge 6.7. Korelasyon analizi sonrası benzer olduğu kabul edilen parametreler.

Parametre	Benzer Olduğu Kabul Edilen Parametreler	Korelasyon Katsayısı
modeD1	minD1	1,000
rangeD1	norminfD1	0,989
stdD1	norm2D1, mad2D1	0,966 – 0,974
rangeD2	maxD2, norminfD2	0,972 – 0,970
madD2	madD1	0,975
norm1D2	norm1D1	0,967
stdD2	norm2D2, mad2D2	0,972 – 0,985
modeD3	minD3	1,000
rangeD3	maxD3, norminfD3	0,965 – 0,959
stdD3	norm2D3, mad2D3	0,977 – 0,987
modeD4	minD4	1,000
rangeD4	maxD4, norminfD4	0,977 – 0,980
stdD4	norm2D4, mad2D4	0,976 – 0,986
modeD5	minD5	1,000
rangeD5	maxD5, norminfD5	0,978 – 0,981
stdD5	norm2D5, mad2D5	0,981 – 0,990
modeA1	minA1	1,000
rangeA1	maxA1, norminfA1	0,986 – 0,985
stdA1	madA1, mad2A1	0,957 – 0,994
norm1A1	norm2A1	0,975

Benzerlik tespiti sonuçlarına göre sadece dalgacık katsayılarında benzerlik gösteren parametreler tespit edilmiştir. Bundan sonraki analizlerde benzer olduğu kabul edilen parametreler analizden çıkartılıp sadece bir tanesi kullanılacaktır. Benzerlik analizi sonrasında 122 olan parametre sayısı 91'e düşürülmüştür.

6.3.2. Duygular Üzerinde Etkin Olan Akustik Parametreler

Çalışmada kullanılan veriler nötr, endişe ve anksiyete olmak üzere üç tekrarlı ölçüm içermektedir. Bu nedenle hasta, kontrol ve her iki grup üzerinde etkin olan parametrelerin tespit edilmesi için Friedman testi kullanılmıştır.

Friedman testi için hipotezler;

1. H0: Hasta, kontrol ve her iki grupta duygular arasında fark yoktur. ($p>0,05$)
2. H1: Hasta, kontrol ve her iki grupta duygular arasında fark vardır. ($p\leq 0,05$)

Friedman testi sonrası elde edilen sonuçlar Çizelge 6.8’de verilmiştir.

Çizelge 6.8. Friedman testi sonuçları.

Parametre	Önem Düzeyi (p)			Parametre	Önem Düzeyi (p)		
	Hasta	Kontrol	Tümü		Hasta	Kontrol	Tümü
F0_ort	0,000*	0,002*	0,000*	minD2	0,015*	0,558	0,033*
F0_std	0,000*	0,264	0,000*	rangeD2	0,287	0,174	0,380
F1_ort	0,000*	0,000*	0,000*	stdD2	0,086	0,558	0,216
F1_std	0,000*	0,006*	0,000*	madD2	0,247	0,338	0,458
F2_ort	0,000*	0,000*	0,000*	norm1D2	0,522	0,004*	0,430
F2_std	0,000*	0,205	0,000*	meanD3	0,043*	0,558	0,040*
F3_ort	0,000*	0,076	0,000*	medianD3	0,142	0,174	0,031*
F3_std	0,212	0,005*	0,058	modeD3	0,705	0,920	0,755
F1_bw	0,861	0,472	0,911	rangeD3	0,951	0,920	0,882
F2_bw	0,043*	0,174	0,008*	stdD3	0,705	0,368	0,969
F3_bw	0,001*	0,264	0,168	madD3	0,350	0,558	0,417
MFCC1	0,449	0,205	0,140	norm1D3	0,705	0,013*	0,048*
MFCC2	0,001*	0,105	0,000*	meanD4	0,522	0,779	0,552
MFCC3	0,003*	0,338	0,002*	medianD4	0,861	0,717	0,666
MFCC4	0,000*	0,009*	0,000*	modeD4	0,350	0,558	0,216
MFCC5	0,006*	0,920	0,021*	rangeD4	0,522	0,368	0,261
MFCC6	0,387	0,558	0,380	stdD4	0,819	0,779	0,969
MFCC7	0,048*	0,039*	0,005*	madD4	0,522	0,125	0,140
MFCC8	0,035*	0,042*	0,013*	norm1D4	0,819	0,028*	0,140
MFCC9	0,049*	0,006*	0,001*	meanD5	0,074	0,779	0,315
MFCC10	0,387	0,717	0,458	medianD5	0,004*	0,717	0,026*
MFCC11	0,035*	0,076	0,003*	modeD5	0,035*	0,558	0,102
MFCC12	0,043*	0,076	0,005*	rangeD5	0,212	0,920	0,417
MFCC13	0,024*	0,039*	0,013*	stdD5	0,259	0,338	0,216
LPCC1	0,000*	0,047*	0,000*	madD5	0,116	0,097	0,016*
LPCC2	0,259	0,076	0,030*	norm1D5	0,549	0,001*	0,034*
LPCC3	0,212	0,205	0,048*	meanA1	0,041*	0,125	0,005*
LPCC4	0,951	0,558	0,666	medianA1	0,157	0,717	0,261
LPCC5	0,008*	0,046*	0,001*	modeA1	0,035*	0,920	0,168
LPCC6	0,017*	0,029*	0,013*	rangeA1	0,026*	0,558	0,123
LPCC7	0,000*	0,039*	0,000*	stdA1	0,142	0,920	0,216
LPCC8	0,000*	0,032*	0,000*	norm1A1	0,350	0,000*	0,031*
LPCC9	0,086	0,174	0,030*	jitta	0,001*	0,028*	0,000*
LPCC10	0,522	0,338	0,666	jitt	0,001*	0,013*	0,000*
LPCC11	0,259	0,205	0,969	shimmer	0,017*	0,558	0,016*
LPCC12	0,116	0,717	0,519	shimmerDB	0,142	0,205	0,033*
LPCC13	0,116	1,000	0,261	meanZCR	0,951	0,000*	0,019*
meanD1	0,157	0,205	0,058	stdZCR	0,001*	0,004*	0,000*
medianD1	0,048*	0,028*	0,168	meanSNR	0,350	0,174	0,140
modeD1	0,001*	0,264	0,001*	meanSNRpw	0,522	0,046*	0,168
maxD1	0,165	0,264	0,093	meanEng	0,549	0,779	0,755

Çizelge 6.8. (devam ediyor).

rangeD1	0,002*	0,049*	0,000*	stdEng	0,522	0,472	0,417
stdD1	0,015*	0,049*	0,002*	spRate	0,951	0,000*	0,040*
meanD2	0,024*	0,097	0,666	pauseRate	0,951	0,000*	0,040*
medianD2	0,047*	0,205	0,026*	numberPause	0,350	0,000*	0,190
modeD2	0,026*	0,033*	0,055				

*p≤0,05 (H1)

Friedman testine göre üç duygu durumunun değişiminden (nötr-endişe, nötr-anksiyete, endişe-anksiyete) herhangi birisi anlamlı bulunursa ilgili parametre anlamlı olarak kabul edilmelidir. Ancak anlamlı değişimin hangi ilişkide olduğu hakkında bilgi vermemektedir. Bu nedenle, H1'i sağlayan parametrelerin hangi grupta (nötr-endişe, anksiyete-endişe, nötr-anksiyete) farklılık oluşturduğunu bulmak için Bonferroni düzeltmeli Wilcoxon işaretli sıra testi veya işaret testi kullanılabilir. Duygular arasında anlamlı olanları tespit etmek için wilcoxon testi kullanılmış ve elde edilen sonuçlar Çizelge 6.9'da verilmiştir.

Wilcoxon testi için hipotezler;

1. H0: Duygular arası fark anlamsızdır. (p>0,05)
2. H1: Duygular arası fark anlamlıdır. (p≤0,05)

Çizelge 6.9. Wilcoxon testi sonuçları.

Parametre	Hasta Grubu			Kontrol Grubu			Tüm Grup		
	Önem Düzeyi (p)			Önem Düzeyi (p)			Önem Düzeyi (p)		
	N-E	N-A	E-A	N-E	N-A	E-A	N-E	N-A	E-A
F0_ort	0,000*	0,000*	0,601	0,433	0,002*	0,012*	0,000*	0,000*	0,112
F0_std	0,001*	0,000*	0,010*	0,272	0,638	0,388	0,015*	0,000*	0,011*
F1_ort	0,000*	0,000*	0,263	0,050	0,002*	0,002*	0,000*	0,204	0,001*
F1_std	0,218	0,000*	0,000*	0,019*	0,388	0,071	0,695	0,000*	0,000*
F2_ort	0,000*	0,000*	0,009*	0,340	0,010*	0,023*	0,000*	0,000*	0,000*
F2_std	0,000*	0,006*	0,351	0,084	0,875	0,050	0,000*	0,027*	0,026*
F3_ort	0,000*	0,000*	0,478	0,023*	0,158	0,814	0,000*	0,000*	0,751
F3_std	0,117	0,881	0,351	0,638	0,004*	0,019*	0,390	0,050	0,023*
F2_bw	0,010*	0,086	0,351	0,010*	0,347	0,754	0,000*	0,043*	0,322
F3_bw	0,040*	0,004*	0,145	0,480	0,388	0,814	0,197	0,052	0,340
MFCC2	0,028*	0,004*	0,014*	0,583	0,015*	0,041*	0,018*	0,000*	0,001*
MFCC3	0,001*	0,052	0,108	0,136	1,000	0,388	0,000*	0,112	0,092
MFCC4	0,000*	0,001*	0,218	0,003*	0,388	0,117	0,000*	0,002*	0,059
MFCC5	0,001*	0,052	0,313	0,308	0,583	0,937	0,001*	0,059	0,575
MFCC7	0,002*	0,526	0,100	0,015*	0,754	0,015*	0,000*	0,779	0,005*
MFCC8	0,970	0,017*	0,028*	0,049*	0,040*	0,084	0,695	0,003*	0,008*

Çizelge 6.9. (devam ediyor).

MFCC9	0,033*	0,017*	0,550	0,049*	0,034*	0,070	0,681	0,016*	0,082
MFCC11	0,007*	0,681	0,023*	0,028*	0,388	0,347	0,000*	0,411	0,021*
MFCC12	0,019*	0,351	0,019*	0,480	0,182	0,050	0,025*	0,161	0,004*
MFCC13	0,526	0,010*	0,025*	0,209	0,695	0,050	0,217	0,210	0,003*
LPCC1	0,001*	0,001*	0,057	0,347	0,034*	0,117	0,000*	0,000*	0,018*
LPCC2	0,135	0,057	0,794	0,158	0,239	0,937	0,043*	0,032*	0,793
LPCC3	0,433	0,502	0,263	0,272	0,182	0,308	0,178	0,191	0,161
LPCC5	0,002*	0,823	0,019*	0,875	0,071	0,041*	0,009*	0,239	0,002*
LPCC6	0,073	0,004*	0,037*	0,049*	0,047*	0,530	0,015*	0,007*	0,172
LPCC7	0,000*	0,001*	0,145	0,003*	0,049*	0,136	0,000*	0,002*	0,032*
LPCC8	0,000*	0,062	0,006*	0,012*	0,480	0,136	0,000*	0,064	0,003*
LPCC9	0,023*	0,218	0,296	0,308	0,937	0,099	0,012*	0,304	0,064
medianD1	0,045*	0,032*	0,455	0,005*	0,048*	0,433	0,100	0,278	0,985
modeD1	0,001*	0,021*	0,709	0,272	0,071	0,272	0,000*	0,003*	0,654
rangeD1	0,001*	0,044*	0,881	0,347	0,099	0,239	0,001*	0,005*	0,466
stdD1	0,002*	0,033*	0,351	0,028*	0,023*	0,239	0,001*	0,003*	0,881
meanD2	0,093	0,008*	0,411	0,060	0,084	0,388	0,654	0,286	0,751
medianD2	0,550	0,007*	0,108	0,023*	0,347	0,638	0,043*	0,005*	0,331
modeD2	0,005*	0,030*	0,601	0,027*	0,019*	0,347	0,026*	0,019*	0,837
minD2	0,002*	0,030*	0,627	0,754	0,272	0,347	0,018*	0,018*	0,793
normlD2	0,681	0,126	0,478	0,388	0,005*	0,012*	0,822	0,525	0,360
meanD3	0,126	0,526	0,057	0,937	0,875	0,695	0,239	0,513	0,096
medianD3	0,108	0,296	0,015*	0,583	0,209	0,084	0,116	0,096	0,003*
normlD3	0,232	0,550	0,794	0,028*	0,005*	0,050	0,020*	0,012*	0,246
normlD4	0,263	0,852	0,940	0,028*	0,010*	0,050	0,036*	0,040*	0,178
medianD5	0,006*	0,370	0,014*	0,875	0,388	0,209	0,014*	0,224	0,006*
modeD5	0,009*	0,073	0,204	0,754	0,272	0,754	0,018*	0,032*	0,501
madD5	0,117	0,601	0,011*	0,937	0,041*	0,019*	0,224	0,064	0,000*
normlD5	0,332	0,601	0,601	0,272	0,005*	0,006*	0,793	0,112	0,024*
meanA1	0,005*	0,279	0,737	0,084	1,000	0,239	0,001*	0,350	0,278
modeA1	0,044*	0,145	0,940	0,814	0,530	0,272	0,089	0,360	0,614
rangeA1	0,048*	0,049*	0,970	0,937	0,308	0,272	0,140	0,627	0,550
normlA1	0,086	0,279	0,765	0,583	0,002*	0,002*	0,262	0,262	0,023*
jitta	0,049*	0,001*	0,051	0,239	0,084	0,023*	0,313	0,000*	0,004*
jitt	0,167	0,001*	0,009*	0,099	0,117	0,023*	0,926	0,000*	0,001*
shimmer	0,179	0,011*	0,108	1,000	0,117	0,433	0,313	0,003*	0,059
shimmerDB	0,332	0,012*	0,167	0,875	0,117	0,117	0,501	0,003*	0,042*
meanZCR	0,370	0,526	0,526	0,814	0,002*	0,002*	0,360	0,020*	0,023*
stdZCR	0,002*	0,008*	0,370	0,012*	0,004*	0,182	0,000*	0,000*	0,911
meanSNRpw	0,351	0,247	0,940	0,182	0,754	0,084	0,079	0,454	0,295
spRate	0,881	0,575	0,926	0,060	0,002*	0,002*	0,360	0,067	0,035*
pauseRate	0,881	0,575	0,926	0,060	0,002*	0,002*	0,360	0,067	0,035*
numberPause	0,896	0,322	0,422	0,033*	0,002*	0,002*	0,313	0,184	0,221

* $p \leq 0,05$ (H1), N-E: Nötr ve endişe arasındaki ilişki, N-A: Nötr ve anksiyete arasındaki ilişki, E-A: Endişe ve anksiyete arasındaki ilişki.

Wilcoxon testi sonuçlarına göre anksiyete için hasta grubunda 30, kontrol grubunda 24 ve her iki grupta 30 parametre H1'i sağlamakta olup anksiyete tespiti için kullanılabilir. Endişe için hasta grubunda 34, kontrol grubunda 19 ve her iki grupta 33 parametre H1'i sağlamakta olup endişe tespiti için kullanılabilir. Anlamlı parametrelerdeki bu farklılık kullanılan çalışma grupları, endişe ve anksiyetenin

tetiklenmesinden kaynaklanmaktadır. Çünkü hasta grubu hasta denekleri, kontrol grubu sağlıklı denekleri içermektedir. Ayrıca, endişe'nin tetiklenmesi için metin okutulmuş ve anksiyetenin tetiklenmesi için de deneklerin doğal durumları kullanılmıştır. Oluşturulacak bir karar destek sisteminde sisteme ses girişi yapılan kişinin hasta mı yoksa sağlıklı mı olduğu bilinmediği için ve sistemin temel amacının bu ayrımı yapabilmesi olduğu için her iki grupta anlamlı olan parametrelerin dikkate alınması önemlidir.

6.3.3. Duygu Durumları Üzerinde Etkin Olan Parametrelerin Değişim Yönü Ve Şiddeti

Anksiyöz durumunda parametrelerin değişim şiddeti ve yönünü tespit etmek için nötr durumdan endişe ve anksiyete durumuna geçişte parametrelerin ortalama değerlerindeki değişim oranı kullanılmıştır. Hasta olgular gerçek anksiyöz durumu içermektedir. Bu nedenle sadece hasta grup üzerinde EMO ve AY durumunda etkili olan parametreler kullanılmıştır. Elde edilen sonuçlar Çizelge 6.10'da verilmiştir.

Çizelge 6.10. Parametrelerin duygulara göre değişim yönü ve şiddeti.

Parametre	EMO	AY	Parametre	EMO	AY	Parametre	EMO	AY
F0_ort	↑	↑	MFCC8	×	↓	meanD2	×	↓
F0_std	↑	↑	MFCC9	↘	↓	medianD2	×	↓
F1_ort	↘	↘	MFCC11	↓	×	modeD2	↑	↑
F1_std	×	↑	MFCC12	↓	×	minD2	↑	↑
F2_ort	↗	↗	LPCC1	↘	↘	medianD5	↑	×
F2_std	↑	↑	LPCC5	↑	×	modeD5	↑	×
F3_ort	↗	↗	LPCC6	×	↑	meanA1	↑	×
F2_bw	↑	×	LPCC7	↓	↓	modeA1	↑	×
F3_bw	↗	↑	LPCC8	↓	×	rangeA1	↑	×
MFCC2	↘	↓	LPCC9	↓	×	jitta	↗	↗
MFCC3	↓	×	medianD1	↓	↑	jitt	×	↗
MFCC4	↓	↓	modeD1	↑	↑	shimmer	×	↗
MFCC5	↓	×	rangeD1	↑	↑	shimmerDB	×	↗
MFCC7	↑	×	stdD1	↑	↑	stdZCR	↑	↑

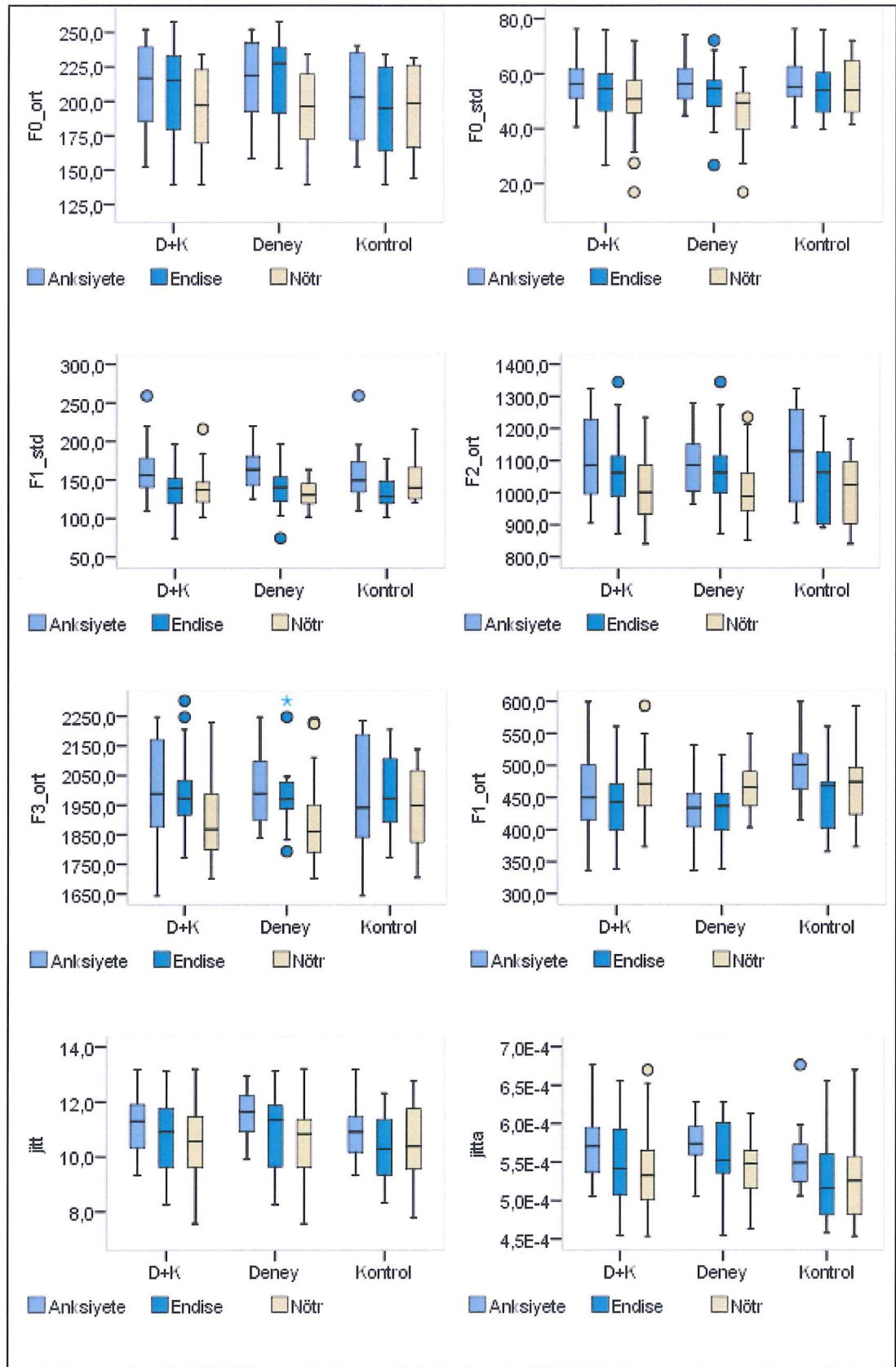
↓: Çok azalma(< -%10), ↑: Çok artma(>%10), ↗: Artma(%0...%10), ↘: Azalma(%0...-%10), ×: Etkisiz

Çizelge 6.10’da elde edilen sonuçlara göre anksiyöz durumda literatürde kullanılan F0, F1, jitter, shimmer gibi genel parametrelere ek olarak MFCC, LPCC ve dalgacık katsayıları da önemli oranda değişiklik göstermektedir. Ayrıca, metin okuyarak anksiyetenin tetiklendiği endişe ile doğal konuşma sonucu elde edilen anksiyete ‘yi etkileyen parametrelerin değişim şiddeti farklılık gösterse de değişim yönü aynıdır. Örneğin, MFCC2 parametresi endişe durumunda az bir düşüş gösterirken doğal anksiyete durumunda fazla düşüş göstermiştir. Ana parametreler (F0, F1, F2, F3, jitta ve ZCR) dikkate alındığında, EMO ile AY arasında benzer parametrelerde değişim görülmüştür. Bununla birlikte, EMO ile daha fazla sayıda parametrede değişim olmuştur. Çizelge 6.9 ve Çizelge 6.10 incelendiğinde anksiyete durumundan etkilenen parametrelerin öncelik sıralaması Çizelge 6.11’de verilmiştir. Öncelik sıralaması belirlenirken nötr durumdan endişe ve anksiyete durumuna geçişte parametrelerin ortalama değerlerindeki değişim oranı kullanılmıştır.

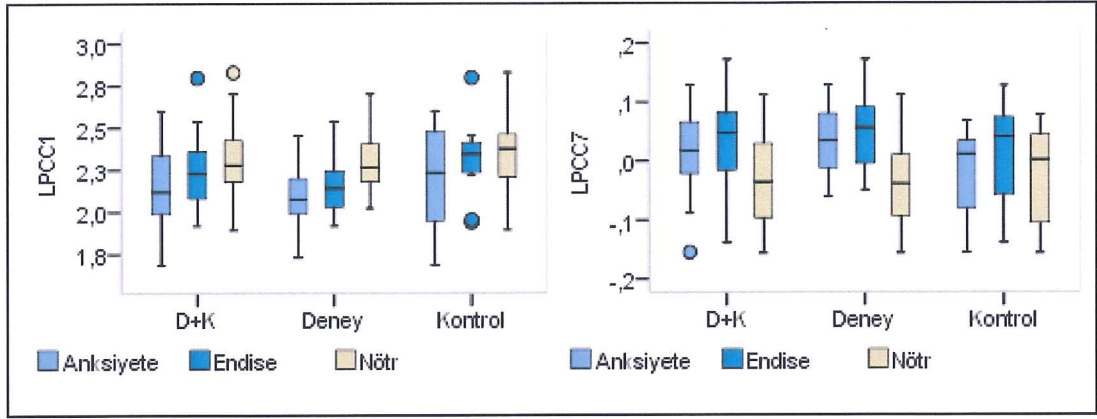
Çizelge 6.11. Anksiyetenin etkilediği parametrelerin etkilenme derecesine göre sıralaması.

Sıra	Parametre	Sıra	Parametre	Sıra	Parametre	Sıra	Parametre
1	F0_ort	12	LPCC6	23	modeD1	34	modeA1
2	F0_std	13	F3_bw	24	minD2	35	LPCC9
3	F1_std	14	MFCC2	25	modeD2	36	rangeA1
4	F2_ort	15	F2_std	26	medianD1	37	meanA1
5	F3_ort	16	medianD2	27	stdD1	38	MFCC12
6	F1_ort	17	stdZCR	28	rangeD1	39	medianD5
7	jitt	18	meanD2	29	MFCC5	40	MFCC7
8	jitta	19	shimmer	30	MFCC3	41	MFCC11
9	LPCC1	20	shimmerDB	31	LPCC8	42	LPCC5
10	LPCC7	21	MFCC9	32	modeD5		
11	MFCC4	22	MFCC8	33	F2_bw		

Anlamli bulunan ilk 10 parametrenin içerdigi deęerlere göre grup bazli daęılımı Şekil 6.3’de verilmiştir.



Şekil 6.4. Anksiyetenin etkilediği en önemli 10 parametrenin içerdiği değerlere göre grup bazlı dağılım grafiği.



Şekil 6.3. (devam ediyor).

6.3.4. Hasta Ve Sağlam Olguları Ayırt Etmede Etkin Olan Parametreler

Mann Whitney-U testi ile sağlıklı ve hasta ayırımı gerçekleştirilen parametreler incelenmiştir. Analiz gerçekleştirilirken her bir grupta yer alan üç duygunun farkları (nötr-endişe, nötr-anksiyete, endişe-anksiyete) kullanılmıştır. Analizler grup bazlı gerçekleştirilmiştir. Friedman testi sonrası anlamlı bulunan parametreler kullanılmıştır. Elde edilen sonuçlar aşağıdaki Çizelge 6.12’de verilmiştir.

Çizelge 6.12. Hasta ve sağlam olguları ayırt etmede etkin parametreler.

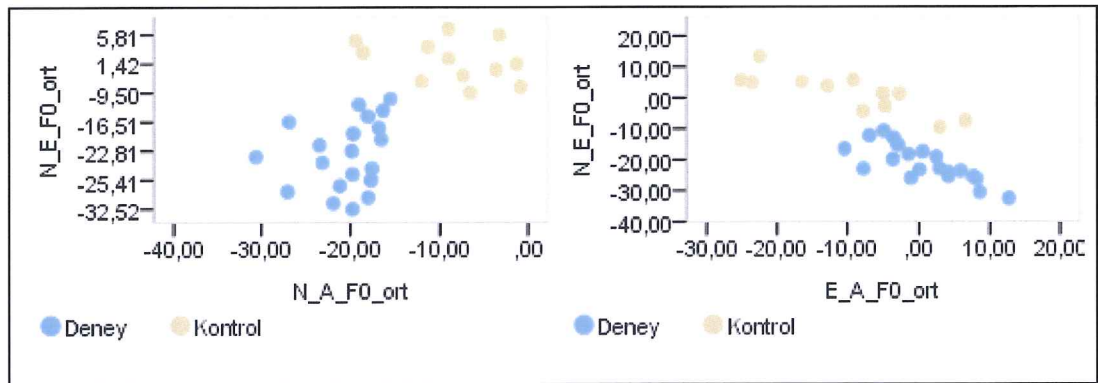
Parametre	Önem Düzeyi (p)			Parametre	Önem Düzeyi (p)		
	N-E	N-A	E-A		N-E	N-A	E-A
F0_ort	0,000*	0,000*	0,004*	rangeD1	0,047*	0,586	0,243
F0_std	0,002*	0,020*	0,640	stdD1	0,036*	0,047*	0,726
F1_ort	0,213	0,000*	0,000*	meanD2	0,024*	0,003*	0,213
F1_std	0,020*	0,020*	0,533	medianD2	0,161	0,414	0,350
F2_ort	0,669	0,043*	0,036*	modeD2	0,029*	0,586	0,330
F2_std	0,199	0,024*	0,150	minD2	0,027*	0,414	0,350
F3_ort	0,073	0,173	0,414	norm1D2	0,371	0,002*	0,014*
F3_std	0,312	0,024*	0,161	meanD3	0,640	0,876	0,533
F2_bw	0,508	0,669	0,846	medianD3	0,846	0,559	0,533
F3_bw	0,073	0,018*	0,330	norm1D3	0,312	0,006*	0,043*
MFCC2	0,111	0,640	0,697	norm1D4	0,330	0,013*	0,056
MFCC3	0,062	0,161	0,846	medianD5	0,094	0,785	0,414
MFCC4	0,043*	0,047*	0,436	modeD5	0,119	0,846	0,350
MFCC5	0,073	0,697	0,350	madD5	0,330	0,119	0,150
MFCC7	0,613	0,484	0,276	norm1D5	0,199	0,006*	0,013*
MFCC8	0,938	0,755	0,969	meanA1	0,293	0,414	0,392
MFCC9	0,533	0,087	0,199	modeA1	0,392	0,150	0,508

Çizelge 6.12. (devam ediyor).

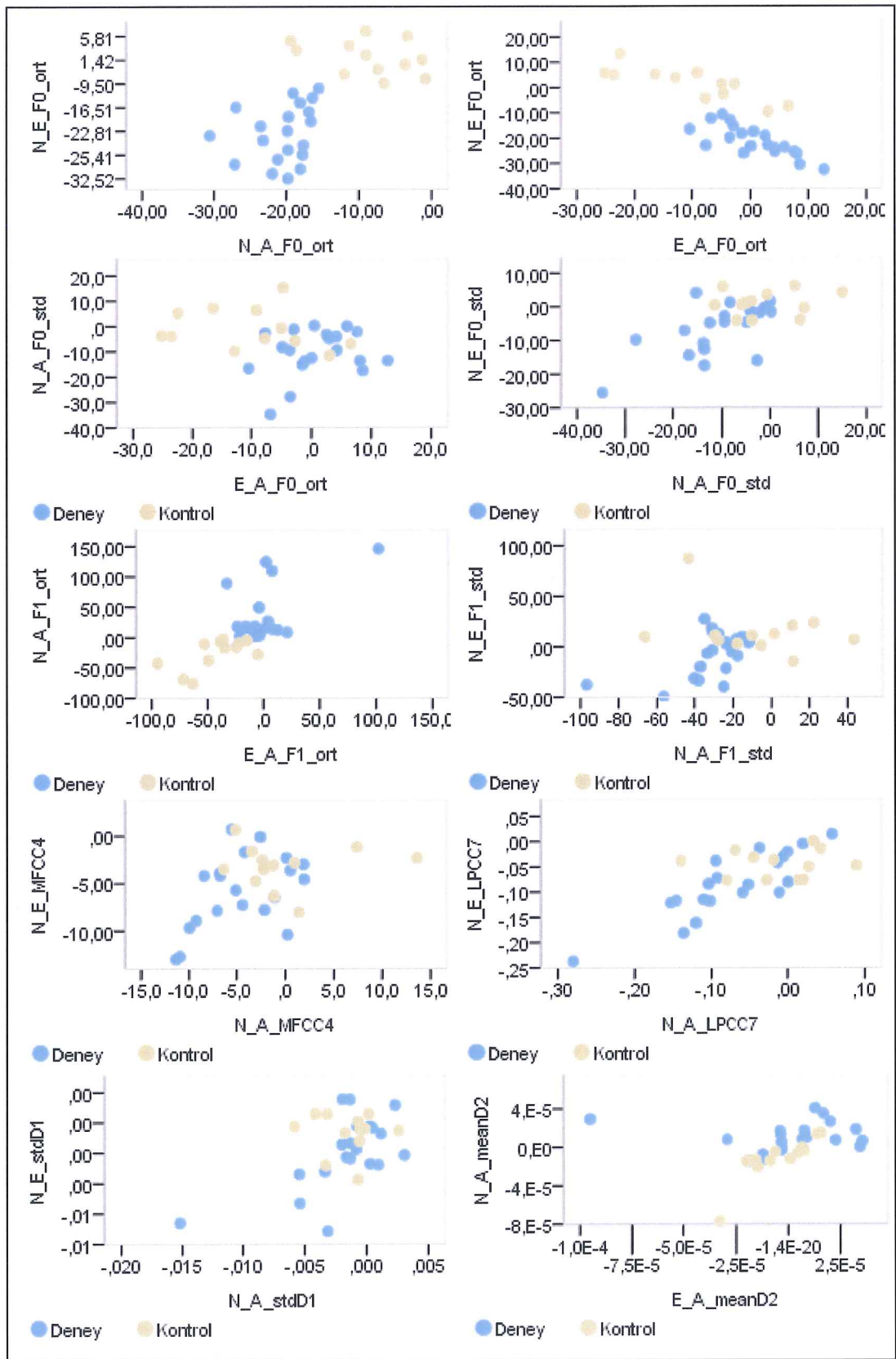
MFCC11	0,586	0,697	0,508	rangeA1	0,228	0,094	0,276
MFCC12	0,484	0,846	0,669	norm1A1	0,021*	0,051	0,181
MFCC13	0,613	0,213	0,755	jitta	0,056	0,907	0,392
LPCC1	0,005*	0,559	0,508	jitt	0,024*	0,330	0,460
LPCC2	0,815	0,969	0,938	shimmer	0,484	0,293	0,697
LPCC3	1,000	0,726	1,000	shimmerDB	0,436	0,613	0,755
LPCC5	0,027*	0,056	0,484	meanZCR	0,755	0,102	0,087
LPCC6	0,559	0,129	0,139	stdZCR	0,043*	0,697	0,139
LPCC7	0,047*	0,032*	0,228	SNRpow	0,508	0,119	0,330
LPCC8	0,161	0,436	0,697	spRate	0,199	0,001*	0,004*
LPCC9	0,484	0,371	0,613	pauseRate	0,199	0,001*	0,004*
medianD1	0,392	0,484	0,186	numPause	0,220	0,001*	0,004*
modeD1	0,043*	0,815	0,243				

* $p \leq 0,05$, N-E: Nötr ve endişe farkı, N-A: Nötr ve anksiyete farkı, E-A: Endişe ve anksiyete farkı

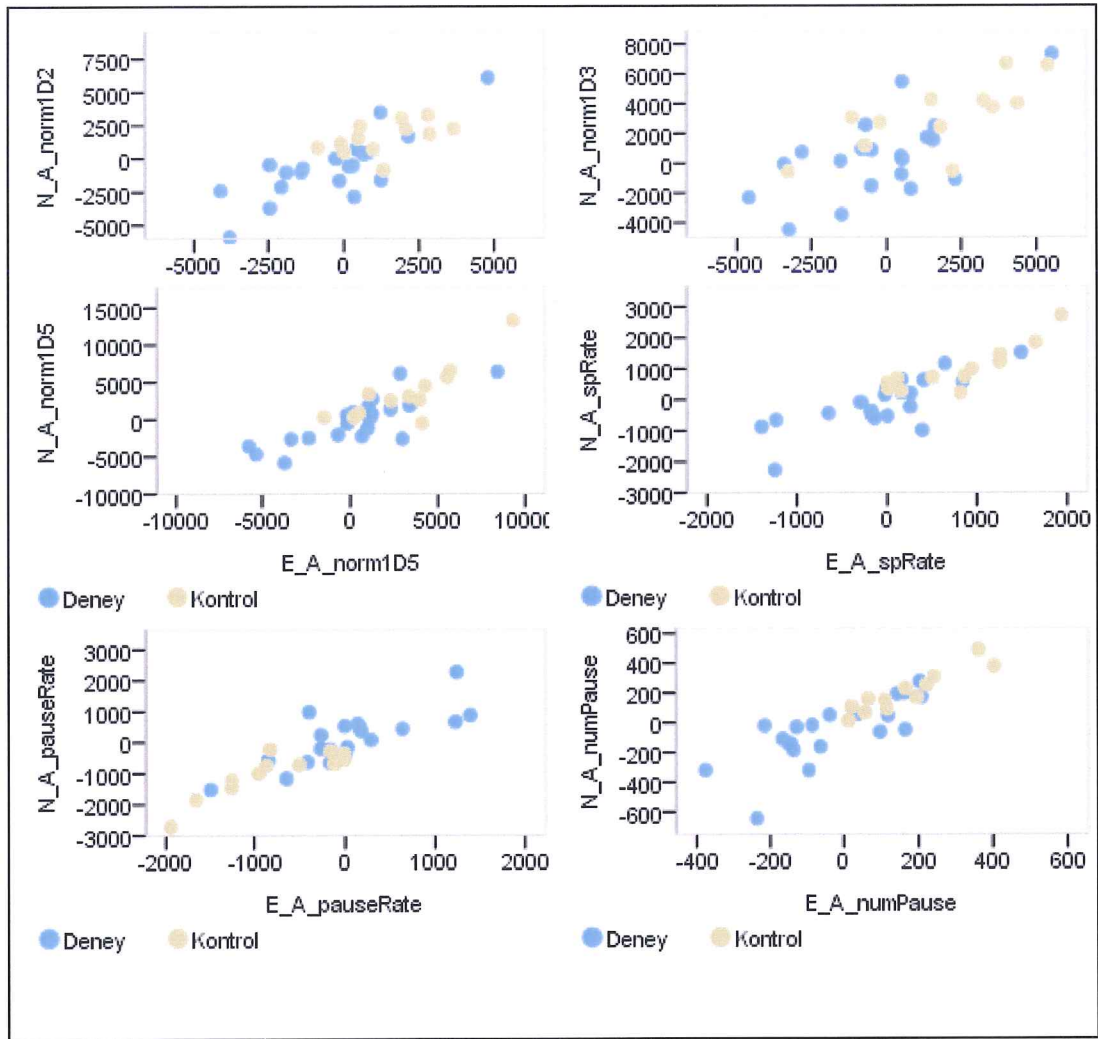
Sonuçlar incelendiğinde bazı parametreler 3 durum için de önemliken bazıları sadece 1 veya 2 durum için önemlidir. Bunun anlamı; örneğin nötr-endişe için önemli ve diğer 2 durum için önemsiz olan LPCC1 parametresi nötr ve endişe duygusu üzerinden sağlıklı ve hasta grubu tespit edebilirken diğer 2 durum üzerinden tespit edememektedir. Bundan dolayı 2 durum için anlamlı olan parametreler sağlıklı ve hasta ayırımı gerçekleştirebilmektedir. Tablo incelendiğinde hasta ve sağlıklı grupları birbirinden ayırmada etkin olan parametreler; F0_ort, F0_std, F1_ort, F1_std, F2_ort, MFCC4, LPCC7, stdD1, meanD2, norm1D2, norm1D3, norm1D5, spRate, pauseRate, numPause. Anlamlı olan parametrelerin grupları ayırt etme gücü Şekil 6.4’de verilmiştir.



Şekil 6.4. Parametrelerin hasta ve sağlam denekleri ayırt etme gücünün grup bazlı grafiksel gösterimi.



Şekil 6.4. (devam ediyor).



Şekil 6.4. (devam ediyor).

Elde edilen önem dereceleri ve grafiksel dağılımlar incelendiğinde hasta ve kontrol grubunu ayırt etmede başarılı olan parametreler önem derecesine göre sırasıyla F0_ort, F1_ort, spRate, pauseRate, numPause, norm1D2, norm1D1, norm1D5, F0_std, meanD2, F1_std, norm1D3, F2_ort, LPCC7, stdD1, MFCC4 'dür.

6.3.5. Hasta Ve Kontrol Gruplarının Duygularını Sese Yansıtmadaki Farklılıkları

Çalışmada kullanılan iki grup için duygu durumları aynı ortam, aynı kayıt cihazı ve aynı metin için elde edilmiştir. Ayrıca, olgular benzer sosyodemografik verilere sahiptir. Bundan dolayı hasta ve kontrol grubuna ait verilerin toplanmasında ki farklılıklar duygu bazlı (hasta nötr-kontrol nötr, hasta endişe-kontrol endişe, hasta

anksiyete-kontrol anksiyete) incelenmiştir. Analizler Mann Whitney-U testi ile yapılmış ve elde edilen sonuçlar Çizelge 6.13’de verilmiştir.

Çizelge 6.13. Çalışma gruplarının duygularını sese yansıtmadaki farklılıkları.

Parametre	Önem Düzeyi (p)			Parametre	Önem Düzeyi (p)		
	N-N	E-E	A-A		N-N	E-E	A-A
F0_ort	0,697	0,062	0,392	norminfD2	0,199	0,969	0,330
F0_std	0,129	0,907	0,846	meanD3	0,371	0,508	0,484
F1_ort	0,876	0,259	*0,004	medianD3	0,259	0,586	0,613
F1_std	0,087	0,559	0,330	modeD3	0,213	0,785	0,067
F2_ort	0,876	0,613	0,726	maxD3	0,276	0,330	0,228
F2_std	0,815	0,213	0,080	minD3	0,213	0,785	0,067
F3_ort	0,312	0,846	0,559	rangeD3	0,293	0,586	0,119
F3_std	0,613	0,350	0,161	stdD3	0,073	0,213	0,056
F1_bw	0,392	0,350	0,907	madD3	0,139	0,371	0,173
F2_bw	0,213	0,087	0,161	mad2D3	0,067	0,228	*0,036
F3_bw	*0,011	0,161	0,697	norm1D3	0,080	0,199	0,199
MFCC1	0,484	0,640	0,371	norminfD3	0,293	0,508	0,129
MFCC2	0,161	0,056	0,213	meanD4	0,161	1,000	0,815
MFCC3	1,000	0,392	0,484	medianD4	0,414	0,199	0,969
MFCC4	0,508	0,846	0,392	modeD4	0,613	0,228	0,371
MFCC5	0,484	0,938	0,559	maxD4	0,276	0,199	0,414
MFCC6	0,846	0,726	0,293	minD4	0,613	0,228	0,371
MFCC7	0,186	0,330	0,161	rangeD4	0,533	0,259	0,371
MFCC8	0,484	0,414	0,276	stdD4	0,129	0,293	0,111
MFCC9	0,938	0,613	0,213	madD4	0,243	0,293	0,533
MFCC10	0,484	0,243	0,161	mad2D4	0,161	0,276	0,161
MFCC11	0,508	0,586	0,815	norm1D4	0,087	0,150	0,094
MFCC12	0,228	0,213	0,119	norm2D4	0,102	0,150	0,640
MFCC13	0,640	0,907	0,969	norminfD4	0,392	0,228	0,484
LPCC1	0,559	*0,020	0,276	meanD5	*0,004	0,726	0,815
LPCC2	0,330	0,460	0,150	medianD5	0,938	0,259	0,392
LPCC3	0,697	0,907	0,755	modeD5	0,199	0,508	0,259
LPCC4	0,228	0,436	0,199	maxD5	0,111	0,414	0,414
LPCC5	0,533	0,907	0,586	rangeD5	0,129	0,508	0,350
LPCC6	0,640	0,846	0,243	stdD5	0,161	0,228	0,371
LPCC7	0,785	0,312	0,052	madD5	0,213	0,259	0,815
LPCC8	0,938	0,213	0,392	mad2D5	0,161	0,293	0,508
LPCC9	0,969	0,846	0,755	norm1D5	0,173	0,213	*0,029
LPCC10	0,559	0,969	0,371	norminfD5	0,213	0,533	0,293
LPCC11	0,876	0,094	0,161	meanA1	0,350	0,150	*0,036
LPCC12	0,640	0,533	0,938	medianA1	0,876	0,436	0,640

Çizelge 6.13. (devam ediyor).

LPCC13	0,460	0,969	0,293	modeA1	0,119	0,533	0,392
meanD1	0,259	0,199	0,508	maxA1	0,186	0,697	0,613
medianD1	0,199	0,846	0,186	rangeA1	0,173	0,640	0,559
modeD1	0,414	0,259	0,907	stdA1	0,111	0,815	0,969
maxD1	0,697	0,392	0,259	madA1	0,143	0,293	0,640
rangeD1	0,508	0,293	0,726	mad2A1	0,087	0,640	0,907
stdD1	0,640	0,533	0,697	jitta	0,350	0,119	0,094
madD1	0,586	0,907	0,938	jitt	0,697	0,228	0,102
mad2D1	0,697	0,876	0,640	shimmer	0,938	0,669	0,755
norm1D1	0,484	0,697	*0,010	shimmerDB	0,392	0,785	0,846
meanD2	0,094	0,726	*0,004	meanZCR	0,243	0,371	0,330
medianD2	0,414	0,484	0,056	stdZCR	0,484	0,073	0,484
modeD2	0,087	0,697	0,436	meanSNR	*0,036	0,080	0,293
maxD2	0,139	0,846	0,259	meanSNRpw	0,350	0,414	0,640
minD2	0,056	0,726	0,312	meanEng	0,080	0,073	0,102
rangeD2	0,129	0,907	0,199	stdEng	0,150	0,129	0,111
stdD2	0,392	1,000	0,228	spRate	0,094	0,213	*0,039
madD2	0,559	0,907	0,907	pauseRate	0,094	0,213	*0,039
norm1D2	0,533	0,755	*0,022	numberPause	0,586	0,741	*0,007
norm2D2	0,436	0,755	0,392				

* $p \leq 0,05$

Önem derecesi $p \leq 0,05$ değerinden küçük olan parametreler aynı duygu için hasta ve kontrol grubunda toplanan verilerin farklı olduğunu göstermektedir. Başka bir ifadeyle grupların farklı olduğunu göstermektedir. İstedığımız durum gruplardaki aynı duyguların benzerliğinin tespiti olduğu için $p \leq 0,05$ şartını sağlamayan parametreler gruplarda aynı etkiye sahiptir. Tabloda verilen sonuçlar incelendiğinde sadece sınırlı sayıda parametre önem düzeyine sahiptir. Ayrıca, aynı anda iki duygu üzerinde de önem derecesini sağlayan parametre bulunmamaktadır. Bu nedenle, hasta ve kontrol grubu duyguları seslerine başarılı bir şekilde yansıtmış olup gruplar arasında farklılık bulunmamaktadır.

BÖLÜM 7

SONUÇLAR VE ÖNERİLER

İnsanların konuşmalarından duygusal durumlarının tespiti insan makine etkileşimi çalışmalarında önemli bir yer almaktadır. Duygusal durumların tespit edilebilmesi için konuşma, sinyal işleme yöntemleri ile sayısallaştırılmakta ve akustik analiz ile konuşmaya ait çeşitli akustik parametreler tespit edilmektedir. Elde edilen her akustik parametre her duygu üzerinde etkin bir özelliğe sahip değildir. Bu çalışmada, esas olarak anksiyöz duygu ile ilişkili akustik parametrelerin saptanması amaçlanmıştır. Bu amaca ulaşmak için gerçek veriler kullanılmış ve bu verilerin akustik analizi hazırlanan SPAC isimli yazılım ile gerçekleştirilmiştir.

SPAC modüler yapısı, geniş öznitelik kümesi, kullanıcı dostu arayüzü, ön işleme, son işleme ve sınıflandırma modülleri ile mevcut yazılımlardaki eksiklikleri gidermektedir. Ayrıca, dalgacık katsayıları, delta ve delta-delta öznitelik kümelerini de içermektedir. SPAC yazılımının geçerlilik testi literatürde kabul görmüş hazır araçlarla karşılaştırılarak gerçekleştirilmiş ve elde edilen sonuçlar SPAC ile öznitelik kümesinin kullanılabilirliğini göstermiştir.

Duygusal durum üzerine yapılan çalışmalardaki en büyük sınırlılık duygusal durumun tam olarak tetiklendiği verilerin edinilebilmesidir. Ayrıca, mevcut çalışmaların büyük çoğunluğu anksiyeteyi korku ailesinde kabul etmiş ve çalışmalarına korkuyu da dahil etmiştir. Sadece anksiyöz durumu kullanan çalışmalar literatürde yer almakla birlikte kullandığı akustik parametre sayısı sınırlıdır. Bu bağlamda, çalışma yeni bir veri kümesi sunarak biyomedikal sinyal işleme ve duygu tanıma konusunda çalışma yapanların ilgisini çekmesi beklenmektedir.

Çalışmada anksiyete bozukluğu olan hastalar ile sağlıklı kontrol grubunun nötr ve anksiyöz olmak üzere 2 duyguya ilişkin 3 ayrı ses kaydı (NMO, EMO ve AY) alınmıştır. Çalışmada toplam 43 olgu ve 129 ses kaydı yer almaktadır.

Anksiyete duygusunun tetiklenmesinde anksiyete şiddetini ölçmek için kullanılan ölçeklerden yararlanarak oluşturulan endişe içerikli metin okuma yöntemi ve bir anksiyete atağını anlatma yöntemi uygulanmıştır. Duygunun tetiklenmesi için metin okutmanın da anksiyeteyi ortaya çıkartmada etkili olduğu görülmüştür. Ayrıca, endişe ve anksiyeteden etkilenen bazı parametrelerin değişim miktarı farklı olsa da değişim yönü aynıdır.

Yapılan çalışma daha önceki çalışmalarda elde edilen sonuçların uygulanabilirliğini ve geçerliliğini göstermesine ek olarak anksiyöz durumda farklı parametrelerde de değişim tespit edilmiştir.

Duygular üzerinde etkili olan parametrelerin değişim yönü ve şiddetinin incelenmesi sonucunda, anksiyetenin F0, F1, jitter, shimmer gibi genel parametrelere ek olarak MFCC, LPCC ve dalgacık katsayılarını da önemli oranda etkilediği tespit edilmiştir. Bu parametrelerdeki değişimler incelendiğinde özellikle dalgacık katsayılarında önemli derecede artış gözlemlenmiştir. Anksiyeteden en çok etkilenen parametre literatüre paralel olarak F0'ın ortalaması olarak tespit edilmiştir. Çalışmada kullanılan toplam 122 parametrenin 42 tanesi anksiyöz durumda değişkenlik göstermektedir.

F0 glotik döngünün hızı ile ilişkili olduğu için duygusal uyarılma durumunda değişim gösterecektir. Bu değişimin yönü duyguya göre değişmekle birlikte anksiyete durumunda subglottal basınç ve F0 değeri artmaktadır.

Her formant kendi merkez frekansı ve bant genişliği ile karakterize edilir ve duygu hakkında önemli bilgiler içerir. Örneğin, insanlar stres ve depresyon altında sesli sesleri üretmez ve aynı durum nötr duygu durumunda da gerçekleşir. Sesteki bu değişim formant bant genişliklerinde farklılıklara neden olur. Anksiyöz durum

formant frekanslarında da deęişim oluřturmuřtur. Yani, anksiyete durumunda konuřmada sesli harflerin dil vuruřu azalmaktadır.

MFCC ve LPCC parametreleri spektral öznitelik olup sinyalin yapısı ile ilgilidir. MFCC insan iřitme sistemini modeller. Anksiyöz durumda MFCC parametrelerindeki deęişim anksiyetenin sese yansıdığını ve insanların konuřmayı dinlerken bu durumu algılayabileceklerini göstermektedir.

Dięer önemli bir ayrıntı ise jitter ve shimmer parametrelerinin deęerleri ses bozukluęu olan hastalarda artmaktadır. Hasta grubunda bu artış görölürken kontrol grubunda deęişim olmamıřtır. Sonuçlarda elde edilen jitter ve shimmer deęerindeki artış anksiyöz durumda vokal kordların düzensiz olduęunu ve ses řiddetinin düzensiz olduęunu göstermektedir.

Enerji duyguların uyarılma düzeyi hakkında bilgiler içerir ve ses řiddeti uygulanarak elde edilir. Neřeli duygu durumlarında konuřma enerjisinde artış beklenmektedir. Ancak olumsuz duygu durumlarında konuřma enerjisi nötr duruma yakın olmakta hatta nötr durumunda altına düşmektedir. Bu nedenle anksiyöz durumda konuřma enerjisinde deęişim beklenmemektedir.

Konuřma esnasında içinde bulunulan duygu durumuna göre konuřmada duraksamalar yařanabilmektedir. Ancak, yapılan çalıřmada hasta olguların konuřurken duraksamalarında bir deęişim olmazken kontrol grubunda deęişim gözlemlenmiřtir. Yani, konuřmadaki duraksamalar anksiyöz durum hakkında bilgi vermemektedir.

Dalgacık katsayıları hesaplanırken konuřma sinyali alçak ve yüksek frekans bantlarına ayrıřtırılmaktadır. Anksiyöz durumda frekans ile ilgili parametrelerde deęişim oluřtuęu için dalgacık katsayılarında da deęişim gözlemlenmesi olaęandır.

Çalıřmanın en önemli sınırlılıęı olgu sayısının görece az olmasıdır. Ayrıca, kadın ve erkek oranının eřit olmaması da önemli bir etken olarak düşünölebilir. Geçmiş çalıřmalarda kadın ve erkek sesleri arasında karakteristik farklılıklar bildirilmiřtir.

Bir diđer sınırlılık ise SPAC bir toolbox olarak hazırlandıđı için MATLAB ortamına ihtiya duymaktadır.

Gelecek alıřmalar daha fazla sayıda olgu ile ve kadın erkek oranlarının da kıyaslanması tarzında yapılması uygun olacaktır. Ayrıca, olguların yař ve kilo bilgilerinin de sonuçlar üzerindeki etkisi araştırılabilir.

Elde edilen parametrelerin sadece anksiyeteye özgü olup olmadığı, depresyon gibi diđer duygularda farklılık olup olmadığı araştırılmalıdır.

SPAC geliştirilip platform bağımsız hale getirilip gerek zamanlı analiz yeteneđi eklenerek uzmanlara karar vermede yardımcı olarak kullanılabilir. Ayrıca, SPAC gömülü bir sisteme dönüřtürülerek biyomedikal cihaz dünyasına yeni bir cihaz olarak dahil edilebilir.

KAYNAKLAR

1. Koç, A., "Septum deviasyonlu hastaların septoplasti operasyonu öncesi ve sonrası akustik ses analizi ile değerlendirilmesi", Uzmanlık Tezi, *Sağlık Bakanlığı Şişli Etfal Hastanesi*, İstanbul, (2008).
2. Okur, E. ve Kılıç, M. A., "CSL ve Dr. Speech ile ölçülen temel frekans ve pertürbasyon değerlerinin karşılaştırılması", *KBB İhtisas Dergisi*, 8: 152–157 (2001).
3. Sataloff, R. T., "Treatment of Voice Disorders", *Plural Publishing*, San Diego, (2005).
4. Ververidis, D. and Kotropoulos, C., "Emotional speech recognition: Resources, features, and methods", *Speech Communication*, 48 (9): 1162–1181 (2006).
5. Chul, M. L. and Narayanan, S. S., "Toward detecting emotions in spoken dialogs", *IEEE Transactions on Speech and Audio Processing*, 13 (2): 293–303 (2005).
6. Aydın, Ö., "Yapay sinir ağlarını kullanarak bir ses tanıma sistemi geliştirilmesi", Yüksek Lisans Tezi, *Trakya Üniversitesi Fen Bilimleri Enstitüsü*, Edirne, (2005).
7. Sarıca, S., "Ses analizinde kullanılan akustik parametreler", Tıpta Uzmanlık Tezi, *Kahramanmaraş Sütçü İmam Üniversitesi Tıp Fakültesi*, Kahramanmaraş, (2012).
8. Honda, M., "Human speech production mechanisms", *NTT Technical Review*, 1 (2): 24–29 (2003).
9. Gick, B., Wilson, I. and Derrick, D., "Articulatory Phonetics", *John Wiley & Sons*, West Sussex, (2012).
10. Sarımehtetoğlu, E. A., "Vokal nodülü olan pediatrik olgularda ses kalitesinin akustik ses analizi ve pediatric voice handicap index ile değerlendirilmesi", Yüksek Lisans Tezi, *Gazi Üniversitesi Sağlık Bilimleri Enstitüsü*, Ankara, (2012).
11. Proakis, J. G. and Manolakis, D. G., "Digital Signal Processing Principles Algorithms and Applications", *Prentice-Hall Inc.*, New Jersey, 1033 (1996).
12. Kuo, S. M., Lee, B. H., and Tian, W., "Real-Time Digital Signal Processing: Fundamentals, Implementations and Applications 2nd ed.", *John Wiley & Sons*, West Sussex, (2013).

13. Smith, S. W., "The Scientist and Engineer's Guide to Digital Signal Processing 2nd ed.", *California Technical Publishing*, San Diego, California, 688 (1999).
14. Roberts, M. J., "Signals and Systems: Analysis Using Transform Methods and MATLAB 2nd ed.", *McGraw-Hill*, New York, (2011).
15. Orfanidis, S. J., "Introduction to Signal Processing", *Prentice-Hall Inc.*, New Jersey, (2010).
16. Jones, D. L., "Digital Signal Processing: A User's Guide", *Rice University*, Texas, (2008).
17. Weeks, M., "Digital Signal Processing Using MATLAB & Wavelets", *Jones & Bartlett Learning*, Massachusetts, (2010).
18. Hayes, M. H., "Schaum's Outline of Theory and Problems of Digital Signal Processing", *McGraw-Hill*, New York, (1999).
19. Erdoğan, H., "Mühendislik yapılarındaki dinamik davranışların jeodezik ölçmelerle belirlenmesi", Doktora Tezi, *Yıldız Teknik Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul, (2006).
20. İnternet: Secrets of Home Theater and High Fidelity, "Up-Sampling, Aliasing, Filtering, and Ringing: A Clarification of Terminology", <http://hometheaterhifi.com/technical/technical-reviews/up-sampling-aliasing-filtering-ringing-a-clarification-of-terminology/> (2016).
21. Ertürk, S., "Sayısal İşaret İşleme", *Birsen Yayınevi*, İstanbul, (2002).
22. Lee, J. J., Lee, S. M., Kim, I. Y., Min, H. K., and Hong, S. H., "Comparison between short time Fourier and wavelet transform for feature extraction of heart sound", *TENCON 99. Proceedings Of The IEEE Region 10 Conference*, Korea, 21547–1550 (1999).
23. Karadağ, K. and Özerdem, M. S., "Classification of ECoG patterns related to finger movements with wavelet based SVM methods", *22nd Signal Processing and Communications Applications Conference (SIU)*, 2174–2177 (2014).
24. Tawade, L. and Warpe, H., "Detection of epilepsy disorder using discrete wavelet transforms using MATLABs", *International Journal of Advanced Science and Technology*, 28 (3): 17–24 (2011).
25. Jiang, X., Zhang, L., Zhao, Q., and Albayrak, S., "ECG arrhythmias recognition system based on independent component analysis feature extraction", *TENCON 2006. 2006 IEEE Region 10 Conference*, Hong Kong, 1–4 (2006).
26. Misiti, M., Misiti, Y., Oppenheim, G., and Poggi, J. M., "Wavelet Toolbox MATLAB user's guide Version 4", *Mathworks Inc*, Massachusetts, (2008).

27. Mintzer, F., "Filters for distortion-free two-band multirate filter banks", *Acoustics, Speech And Signal Processing, IEEE Transactions On*, 33 (3): 626–630 (1985).
28. Smith, M. and Barnwell T. P., "A procedure for designing exact reconstruction filter banks for tree-structured subband coders", *Acoustics, Speech, And Signal Processing, IEEE International Conference On ICASSP'84.*, California, 9421–424 (1984).
29. Rizon, M., "Discrete wavelet transform based classification of human emotions using electroencephalogram signals", *American Journal Of Applied Sciences*, 7 (7): 878 (2010).
30. Murugappan, M., Nagarajan, R., and Yaacob, S., "Discrete Wavelet Transform Based Selection of Salient EEG Frequency Band for Assessing Human Emotions", *INTECH Open Access Publisher*, Croatia, (2011).
31. Murugappan, M., Rizon, M., Nagarajan, R., and Yaacob, S., "Inferring of human emotional states using multichannel EEG", *European Journal Of Scientific Research*, 48 (2): 281–299 (2010)..
32. Scherer, K. R., Scherer, K. R., and Ekman, P., "On the nature and function of emotion: A component process approach", *Approaches to Emotion*, 2: 293-317 (1984).
33. Tompkins, S., "Affect Imagery Consciousness: Volume I: The Positive Affects", *Springer Publishing Company*, New York, (1962).
34. Arnold, M. B., "Emotion and Personality", *Columbia University Press*, New York, 430 (1960).
35. Scherer, K. R., "Emotion as a process: Function, origin and regulation", *Social Science Information/sur Les Sciences Sociales*, 21 (4/5): 555-570 (1982).
36. Bower, G. H., "Mood and memory", *American Psychologist*, 36 (2): 129-148 (1981).
37. Mowrer, O., "Learning theory and behavior", *Learning Theory and Behavior*, *John Wiley & Sons Inc*, West Sussex, 555 (1960).
38. Sethu, V., "Automatic emotion recognition: An investigation of acoustic and prosodic parameters", *Doktora Tezi, The University of New South Wales Electrical Engineering & Telecommunications*, Sydney, (2009).
39. Damasio, A. R., "A Second Chance for Emotion", *Cognitive Neuroscience of Emotion*, Oxford, 12–23 (2000).
40. Ekman, P., "Facial expressions of emotion: New findings, new questions", *Psychological Science*, 3 (1): 34–38 (1992).

41. Polat, G. ve Altun, H., "Ses öznelik vektörlerinin duygusal durum sınıflandırılmasında kullanımı", *Elektrik-Elektronik-Bilgisayar Mühendisliği Sempozyumu ELECO 2006*, Bursa, 30–34 (2006).
42. Fónagy, I., "Emotions, voice and music", *Research Aspects On Singing*, 3 (3): 51–79 (1981).
43. Bou-Ghazale, S. E. and Hansen, J. H., "A comparative study of traditional and newly proposed features for recognition of speech under stress", *Speech And Audio Processing, IEEE Transactions On*, 8 (4): 429–442 (2000).
44. Loizou, P., "Colea: A MATLAB software-tool for Speech Analysis", *University Of Arkansas*, Arkansas, (2003).
45. Boersma, P. and Weenink, D., "Praat, a system for doing phonetics by computer", *Glott International*, 5 (9): 341–345 (2002).
46. Eyben, F., Wöllmer, M., and Schuller, B., "Opensmile: The munich versatile and fast open-source audio feature extractor", *Proceedings of the 18th ACM International Conference on Multimedia*, Frenze, 1459–1462 (2010).
47. Eyben, F., Wollmer, M., and Schuller, B., "OpenEAR—introducing the Munich open-source emotion and affect recognition toolkit", *3rd International Conference on Affective Computing and Intelligent Interaction And Workshops*, Amsterdam, 1–6 (2009).
48. İnternet: Ent & Speech, "Computerized Speech Lab, Kay Elemetrics", http://www.kaypentax.com/index.php?option=com_product&controller=product&Itemid=3&cid%5B%5D=11&task=pro_details (2015)..
49. Slaney, M., "Auditory toolbox: A matlab toolbox for auditory modelling work", *Apple Technical Report*, Kaliforniya, 45 (1994).
50. İnternet: Speech and Voice Assessment Tools, "lingWAVES", <https://www.wevosys.com/products/lingwaves/lingwaves.html> (2016).
51. İnternet: SIL International, "Speech Analyzer", <http://www-01.sil.org/computing/sa/index.htm> (2016).
52. " İnternet: LSTA Lab., "Welcome to Talkbox Documentation — Talkbox v0.1 Documentation", http://www.ar.media.kyoto-u.ac.jp/members/david/software/talkbox/talkbox_doc/index.html (2016).
53. İnternet: Imperial College London, "Voicebox: Speech processing toolbox for matlab", <http://www.ee.ic.ac.uk/hp/staff/dmb/voicebox/voicebox.html> (2016).
54. Hansen, J. H., Bou-Ghazale, S. E., Sarikaya, R. and Pellom, B., "Getting started with SUSAS: A speech under simulated and actual stress database", *Eurospeech*, 97 (4): 1743–1746 (1997).

55. Li, X., "SPeECH Feature Toolbox (SPEFT) Design and Emotional Speech Feature Extraction", Doktora Tezi, *Marquette University Faculty of Graduate School*, Milwaukee, (2007).
56. Laukka, P., Linnman, C., Ahs, F., Pissioti, A., Frans, Ö., Faria, V., Michelgaard, A., Appel, L., Fredrikson, M. and Furmark, T., "In a nervous voice: Acoustic analysis and perception of anxiety in social phobics' speech", *Journal Of Nonverbal Behavior*, 32 (4): 195–214 (2008).
57. Smith, C. A. and Lazarus, R. S., "Emotion and adaptation", *Handbook of Personality: Theory and Research Guilford*, New York, 609–637 (1990).
58. Murray, I. R. and Arnott, J. L., "Toward the simulation of emotion in synthetic speech: A review of the literature on human vocal emotion", *The Journal of the Acoustical Society of America*, 93 (2): 1097–1108 (1993).
59. Ruiz, R., Absil, E., Harmegnies, B., Legros, C. and Poch, D., "Time-and spectrum-related variabilities in stressed speech under laboratory and real conditions", *Speech Communication*, 20 (1): 111–129 (1996).
60. Drioli, C., Tisato, G., Cosi, P. and Tesser, F., "Emotions and voice quality: experiments with sinusoidal modeling", *ISCA Tutorial and Research Workshop on Voice Quality: Functions, Analysis and Synthesis*, Geneva, (2003).
61. Diamond, G. M., Rochman, D., and Amir, O., "Arousing primary vulnerable emotions in the context of unresolved anger: Speaking about versus speaking to", *Journal of Counseling Psychology*, 57 (4): 402 (2010).
62. Goberman, A. M., Hughes, S., and Haydock, T., "Acoustic characteristics of public speaking: Anxiety and practice effects", *Speech Communication*, 53 (6): 867–876 (2011).
63. Weeks, J. W., Lee, C.-Y., Reilly, A. R., Howell, A. N., France, C., Kowalsky, J. M., and Bush, A., "'The Sound of Fear': Assessing vocal fundamental frequency as a physiological indicator of social anxiety disorder", *Journal of Anxiety Disorders*, 26 (8): 811–822 (2012).
64. Protopapas, A. and Lieberman, P., "Fundamental frequency of phonation and perceived emotional stress", *The Journal of the Acoustical Society of America*, 101 (4): 2267–2277 (1997).
65. Hagenaaars, M. A. and van Minnen, A., "The effect of fear on paralinguistic aspects of speech in patients with panic disorder with agoraphobia", *Journal of Anxiety Disorders*, 19 (5): 521–537 (2005).
66. Fuller, B. F., Horii, Y., and Conner, D. A., "Validity and reliability of nonverbal voice measures as indicators of stressor-provoked anxiety", *Research In Nursing & Health*, 15 (5): 379–389 (1992).

67. Siegman, A. W., "The telltale voice: Nonverbal messages of verbal communication", *Nonverbal Messages of Verbal Communication*, 2nd ed. Ed., *Lawrence Erlbaum Associates*, England, 351–433 (1987).
68. Juslin, P. N. and Laukka, P., "Communication of emotions in vocal expression and music performance: Different channels, same code?", *Psychological Bulletin*, 129 (5): 770 (2003).
69. McLoughlin, I., "Applied Speech and Audio Processing: With Matlab Examples", *Cambridge University Press*, (2009).
70. Grimm, M., Kroschel, K., and Narayanan, S., "The Vera am Mittag German audio-visual emotional speech database", *Multimedia and Expo, 2008 IEEE International Conference On*, 865–868 (2008).
71. Douglas-Cowie, E., Cowie, R., Sneddon, I., Cox, C., Lowry, O., Mcrorie, M., Martin, J.-C., Devillers, L., Abrilian, S., Batliner, A., and others, "The HUMAINE database: addressing the collection and annotation of naturalistic and induced emotional data", *Affective Computing and Intelligent Interaction*, *Springer*, 488–500 (2007).
72. Burkhardt, F., Paeschke, A., Rolfes, M., Sendlmeier, W. F., and Weiss, B., "A database of German emotional speech.", *Interspeech*, 51517–1520 (2005).
73. Steidl, S., "Automatic Classification of Emotion Related User States in Spontaneous Children's Speech", *University of Erlangen-Nuremberg Germany*, (2009).
74. Busso, C., Bulut, M., Lee, C.-C., Kazemzadeh, A., Mower, E., Kim, S., Chang, J. N., Lee, S., and Narayanan, S. S., "IEMOCAP: Interactive emotional dyadic motion capture database", *Language Resources and Evaluation*, 42 (4): 335–359 (2008).
75. McKeown, G., Valstar, M. F., Cowie, R., and Pantic, M., "The SEMAINE corpus of emotionally coloured character interactions", *Multimedia and Expo (ICME), 2010 IEEE International Conference On*, 1079–1084 (2010).
76. Engberg, I. S. and Hansen, A. V., "Documentation of the danish emotional speech database des", *Internal AAU Report, Center For Person Kommunikation, Denmark*, 22 (1996).
77. Martin, O., Kotsia, I., Macq, B., and Pitas, I., "The eNTERFACE'05 audio-visual emotion database", *Data Engineering Workshops, 2006. Proceedings. 22nd International Conference On*, 8–8 (2006).
78. Fu, L., Mao, X., and Chen, L., "Speaker independent emotion recognition based on svm/hmms fusion system", *Audio, Language and Image Processing, 2008. ICALIP 2008. International Conference On*, 61–65 (2008).

79. El Ayadi, M., Kamel, M. S., and Karray, F., "Survey on speech emotion recognition: Features, classification schemes, and databases", *Pattern Recognition*, 44 (3): 572–587 (2011).
80. Anolli, L., Lei Wang, Mantovani, F., and De Toni, A., "The Voice of Emotion in Chinese and Italian Young Adults", *Journal of Cross-Cultural Psychology*, 39 (5): 565–598 (2008).
81. Clavel, C., Vasilescu, I., Devillers, L., Richard, G., and Ehrette, T., "Fear-type emotion recognition for future audio-based surveillance systems", *Speech Communication*, 50 (6): 487–503 (2008).
82. Pell, M. D., Paulmann, S., Dara, C., Alasseri, A., and Kotz, S. A., "Factors in the recognition of vocally expressed emotions: A comparison of four languages", *Journal of Phonetics*, 37 (4): 417–435 (2009).
83. López-Cózar, R., Silovsky, J., and Kroul, M., "Enhancement of emotion detection in spoken dialogue systems by combining several information sources", *Speech Communication*, 53 (9–10): 1210–1228 (2011).
84. Coutinho, E. and Dibben, N., "Psychoacoustic cues to emotion in speech prosody and music", *Cognition & Emotion*, 27 (4): 658–684 (2013).
85. Oflazoglu, C. and Yildirim, S., "Recognizing emotion from Turkish speech using acoustic features", *EURASIP Journal on Audio, Speech, and Music Processing*, 2013 (1): 1–11 (2013).
86. Aljaafreh, A. and Dong, L., "An evaluation of feature extraction methods for vehicle classification based on acoustic signals", *Networking, Sensing and Control (ICNSC), 2010 International Conference On*, 570–575 (2010).
87. "DC Bias", https://en.wikipedia.org/w/index.php?title=DC_bias&oldid=700415584 (2016).
88. Rabiner, L. R. and Schafer, R. W., "Digital Processing of Speech Signals", *Prentice Hall*, (1978).
89. Picone, J. W., "Signal modeling techniques in speech recognition", *Proceedings of the IEEE*, 81 (9): 1215–1247 (1993).
90. Kesarkar, M. P., "Feature extraction for speech recognition", *M. Tech. Credit Seminar Report, Electronic Systems Group, EE. Dept, IIT Bombay*, 1–11 (2003).
91. Vaseghi, S. V., "Advanced Digital Signal Processing and Noise Reduction", *John Wiley & Sons*, (2008).
92. Villanueva-Luna, A. E., Flores-Gil, A., Jaramillo-Nuñez, A., Ortiz-Lima, C. M., Sanchez-Lucero, D., Aguilar-Soto, J. G., and May-Alarcon, M., "De-Noiseing

Audio Signals Using MATLAB Wavelets Toolbox", *INTECH Open Access Publisher*, (2011).

93. Virtanen, T., Singh, R., and Raj, B., "Techniques for Noise Robustness in Automatic Speech Recognition", *John Wiley & Sons*, (2012).
94. Tarnag, W., Chen, Y.-Y., Li, C.-L., Hsie, K.-R., and Chen, M., "Applications of support vector machines on smart phone systems for emotional speech recognition", *World Academy of Science, Engineering And Technology*, 72106–113 (2010).
95. Chen, L., Mao, X., Xue, Y., and Cheng, L. L., "Speech emotion recognition: Features and classification models", *Digital Signal Processing*, 22 (6): 1154–1160 (2012).
96. Grimm, M., Kroschel, K., Mower, E., and Narayanan, S., "Primitives-based evaluation and estimation of emotions in speech", *Speech Communication*, 49 (10–11): 787–800 (2007).
97. Hoque, M. E., Yeasin, M., and Louwerse, M. M., "Robust recognition of emotion from speech", *Intelligent Virtual Agents*, 42–53 (2006).
98. Altun, H. and Polat, G., "Boosting selection of speech related features to improve performance of multi-class SVMs in emotion detection", *Expert Systems with Applications*, 36 (4): 8197–8203 (2009).
99. Batliner, A., Steidl, S., Schuller, B., Seppi, D., Vogt, T., Wagner, J., Devillers, L., Vidrascu, L., Aharonson, V., Kessous, L., and Amir, N., "Whodunnit – Searching for the most important feature types signalling emotion-related user states in speech", *Computer Speech & Language*, 25 (1): 4–28 (2011).
100. Polzehl, T., Schmitt, A., Metze, F., and Wagner, M., "Anger recognition in speech using acoustic and linguistic cues", *Speech Communication*, 53 (9–10): 1198–1209 (2011).
101. Lee, C.-C., Mower, E., Busso, C., Lee, S., and Narayanan, S., "Emotion recognition using a hierarchical binary decision tree approach", *Speech Communication*, 53 (9–10): 1162–1171 (2011).
102. Pérez-Espinosa, H., Reyes-García, C. A., and Villaseñor-Pineda, L., "Acoustic feature selection and classification of emotions in speech using a 3D continuous emotion model", *Biomedical Signal Processing and Control*, 7 (1): 79–87 (2012).
103. Chen, L., Mao, X., Wei, P., Xue, Y., and Ishizuka, M., "Mandarin emotion recognition combining acoustic and emotional point information", *Applied Intelligence*, 37 (4): 602–612 (2012).

104. Truong, K. P., van Leeuwen, D. A., and de Jong, F. M. G., "Speech-based recognition of self-reported and observed emotion in a dimensional space", *Speech Communication*, 54 (9): 1049–1063 (2012).
105. Sezgin, M. C., Gunesel, B., and Kurt, G. K., "Perceptual audio features for emotion detection", *EURASIP Journal on Audio, Speech, And Music Processing*, 2012 (1): 1–21 (2012).
106. Mariooryad, S. and Busso, C., "Exploring Cross-Modality Affective Reactions for Audiovisual Emotion Recognition", *IEEE Transactions on Affective Computing*, 4 (2): 183–196 (2013).
107. Planet, S. and Iriondo, I., "Children's Emotion Recognition from Spontaneous Speech Using a Reduced Set of Acoustic and Linguistic Features", *Cognitive Computation*, 5 (4): 526–532 (2013).
108. Busso, C., Mariooryad, S., Metallinou, A., and Narayanan, S., "Iterative Feature Normalization Scheme for Automatic Emotion Detection from Speech", *IEEE Transactions on Affective Computing*, 4 (4): 386–397 (2013).
109. Ramakrishnan, S. and El Emary, I. M. M., "Speech emotion recognition approaches in human computer interaction", *Telecommunication Systems*, 52 (3): 1467–1478 (2013).
110. Mariooryad, S. and Busso, C., "Compensating for speaker or lexical variabilities in speech for emotion recognition", *Speech Communication*, 57:1–12 (2014).
111. Milton, A. and Tamil Selvi, S., "Class-specific multiple classifiers scheme to recognize emotions from speech signals", *Computer Speech & Language*, 28 (3): 727–742 (2014).
112. Jianhua Tao, Yongguo Kang, and Aijun Li, "Prosody conversion from neutral speech to emotional speech", *IEEE Transactions on Audio, Speech and Language Processing*, 14 (4): 1145–1154 (2006).
113. Roh, Y.-W., Kim, D.-J., Lee, W.-S., and Hong, K.-S., "Novel acoustic features for speech emotion recognition", *Science in China Series E: Technological Sciences*, 52 (7): 1838–1848 (2009).
114. Kandali, A. B., Routray, A., and Basu, T. K., "Vocal emotion recognition in five native languages of Assam using new wavelet features", *International Journal of Speech Technology*, 12 (1): 1–13 (2009).
115. Kockmann, M., Burget, L., and "Honza" Černocký, J., "Application of speaker- and language identification state-of-the-art techniques for emotion recognition", *Speech Communication*, 53 (9–10): 1172–1185 (2011).

116. Jia, J., Zhang, S., Meng, F., Wang, Y., and Cai, L., "Emotional audio-visual speech synthesis based on PAD", *IEEE Transactions on Audio, Speech, And Language Processing*, 19 (3): 570–582 (2011).
117. Albornoz, E. M., Milone, D. H., and Rufiner, H. L., "Spoken emotion recognition using hierarchical classifiers", *Computer Speech & Language*, 25 (3): 556–570 (2011).
118. Huang, C., Liang, R., Wang, Q., Xi, J., Zha, C., and Zhao, L., "Practical speech emotion recognition based on online learning: From acted data to elicited data", *Mathematical Problems in Engineering*, 2013:1–9 (2013).
119. Gharavian, D., Sheikhan, M., and Ashoftedel, F., "Emotion recognition improvement using normalized formant supplementary features by hybrid of DTW-MLP-GMM model", *Neural Computing and Applications*, 22 (6): 1181–1191 (2013).
120. Lanjewar, R. B., and Chaudhari, D. S., "Comparative analysis of speech emotion recognition system using different classifiers on berlin emotional speech database", *International Journal of Electrical and Electronics Engineering Research (IJEEER)*, 3(5): 145-156 (2013).
121. Zao, L., Cavalcante, D., and Coelho, R., "Time-frequency feature and AMS-GMM mask for acoustic emotion classification", *IEEE Signal Processing Letters*, 21 (5): 620–624 (2014).
122. Siegert, I., Philippou-Hübner, D., Hartmann, K., Böck, R., and Wendemuth, A., "Investigation of speaker group-dependent modelling for recognition of affective states from speech", *Cognitive Computation*, 6 (4): 892–913 (2014).
123. Huang, Y., Zhang, G., Li, Y., and Wu, A., "Improved emotion recognition with novel task-oriented wavelet packet features", *Intelligent Computing Theory, Springer*, 706–714 (2014).
124. Song, M., You, M., Li, N., and Chen, C., "A robust multimodal approach for emotion recognition", *Neurocomputing*, 71 (10–12): 1913–1920 (2008).
125. San-Segundo, R., Cordoba, R., Ferreiros, J., Macias-Guarasa, J., Montero, J. M., Fernández, F., D'haro, L. F., Barra, R., and Barra, R., "Speech technology at home: enhanced interfaces for people with disabilities", *Intelligent Automation & Soft Computing*, 15 (4): 647–666 (2009).
126. Ntalampiras, S. and Fakotakis, N., "Modeling the temporal evolution of acoustic parameters for speech emotion recognition", *IEEE Transactions on Affective Computing*, 3 (1): 116–125 (2012).
127. Zupan, B., Neumann, D., Babbage, D. R., and Willer, B., "The importance of vocal affect to bimodal processing of emotion: implications for individuals

- with traumatic brain injury", *Journal of Communication Disorders*, 42 (1): 1–17 (2009).
128. Steidl, S., Batliner, A., Nöth, E., and Hornegger, J., "Quantification of segmentation and F0 errors and their effect on emotion recognition", *Text, Speech and Dialogue*, 525–534 (2008).
 129. Chiriacescu, I., "Automatic emotion analysis based on speech", *TU Delft, Delft University of Technology*, (2009).
 130. Vidrascu, L. and Devillers, L., "Detection of real-life emotions in call centers.", *INTERSPEECH*, 20051841–1844 (2005).
 131. Gerhard, D., "Pitch extraction and fundamental frequency: History and current techniques", *Regina: Department of Computer Science, University of Regina*, (2003).
 132. Escalona Mena, M., "Emotion recognition from speech signals", Master Thesis, *Faculty of Electrical Engineering, University of Ljubljana*, (2012).
 133. Rezaei, N. and Salehi, A., "An introduction to speech sciences (acoustic analysis of speech)", *Iranian Rehabilitation Journal*, 4 (4): 5–14 (2006).
 134. Farrus, M. and Hernando, J., "Using jitter and shimmer in speaker verification", *Signal Processing, IET*, 3 (4): 247–257 (2009).
 135. Deshmukh, O., Espy-Wilson, C. Y., Salomon, A., and Singh, J., "Use of temporal information: Detection of periodicity, aperiodicity, and pitch in speech", *Speech and Audio Processing, IEEE Transactions On*, 13 (5): 776–786 (2005).
 136. Cannizzaro, M., Harel, B., Reilly, N., Chappell, P., and Snyder, P. J., "Voice acoustical measurement of the severity of major depression", *Brain and Cognition*, 56 (1): 30–35 (2004).
 137. Rapcan, V., D'Arcy, S., Yeap, S., Afzal, N., Thakore, J., and Reilly, R. B., "Acoustic and temporal analysis of speech: A potential biomarker for schizophrenia", *Medical Engineering & Physics*, 32 (9): 1074–1079 (2010).
 138. Kane, J. and Gobl, C., "Evaluation of glottal closure instant detection in a range of voice qualities", *Speech Communication*, 55 (2): 295–314 (2013).
 139. İnternet: Bilişsel Davranış, "Bilişsel Davranışçı Psikoterapiler Derneği", <http://bilisseldavraniscisi.org> (2016).
 140. Hamilton, M. A. X., "The assessment of anxiety states by rating", *British Journal of Medical Psychology*, 32 (1): 50–55 (1959).

141. Yazıcı, M. K., Demir, B., Tanriverdi, N., Karaağaoğlu, E., and Yolac, P., "Hamilton anxiety rating scale: interrater reliability and validity study", *Türk Psikiyatri Derg*, 9114–117 (1998).
142. Beck, A. T., Epstein, N., Brown, G., and Steer, R. A., "An inventory for measuring clinical anxiety: psychometric properties.", *Journal Of Consulting And Clinical Psychology*, 56 (6): 893 (1988).
143. Ulusoy, M., Sahin, N. H., and Erkmen, H., "Turkish version of the Beck Anxiety Inventory: psychometric properties", *Journal of Cognitive Psychotherapy*, 12 (2): 163–172 (1998).
144. "Home | APA DSM-5", <http://www.dsm5.org/Pages/Default.aspx> (2016).
145. Beck, A. T., and Steer, R. A., "Beck anxiety inventory manual. San Antonio, Tx: The psychological corporation", *Harcourt Brace & Company*, (1993).
146. Cai, W., "Analysis of acoustic feature extraction algorithms in noisy environments", *University of Rochester*, (2013).
147. Karasartova, S., "Metinden bağımsız konuşmacı tanıma sistemlerinin incelenmesi ve gerçekleştirilmesi", Yüksek Lisans Tezi, *Ankara Üniversitesi Fen Bilimleri Enstitüsü*, Ankara, (2011).
148. Rabiner, L., and Juang, B.-H., "Fundamentals of speech recognition", *Prentice-Hall*, New Jersey, (1993).
149. Yazar, I., Yavuz, H. S., ve Çay, M. A., "Temel bileşen analizi yönteminin ve bazı klasik ve robust uyarlamalarının yüz tanıma uygulamaları", *Eskişehir Osmangazi Üniversitesi Mühendislik Mimarlık Fakültesi Dergisi*, 22 (1): 49–63 (2009).
150. Baykal, N., "Speech emotion recognition using auditory models", Yüksek Lisans Tezi, *Orta Doğu Teknik Üniversitesi Enformatik Enstitüsü*, Ankara, (2013).
151. Schuller, B., Batliner, A., Steidl, S., and Seppi, D., "Recognising realistic emotions and affect in speech: State of the art and lessons learnt from the first challenge", *Speech Communication*, 53 (9–10): 1062–1087 (2011).
152. Zolnay, A., and Haeb-Umbach, U. D.-I. R., "Acoustic feature combination for speech recognition", Doctoral Dissertation, *RWTH Aachen University*, (2006).
153. Duda, R. O., and Hart, P. E., "Pattern Classification and science analysis", *Wiley-Interscience*, New York, (1973).
154. Fröhlich, H., Chapelle, O., and Schölkopf, B., "Feature selection for support vector machines by means of genetic algorithm", *Tools With Artificial*

Intelligence, 2003. Proceedings. 15th IEEE International Conference On, 142–148 (2003).

155. Karagöz, Y., "Nonparametrik tekniklerin güç ve etkinlikleri", *Elektronik Sosyal Bilimler Dergisi*, 9 (33): 18–40 (2010).
156. Ersöz, F., "IBM SPSS İle İstatistiksel Veri Analizi", *Sage Yayıncılık*, Ankara, (2014).
157. Baldi, P., Brunak, S., Chauvin, Y., Andersen, C. A., and Nielsen, H., "Assessing the accuracy of prediction algorithms for classification: an overview", *Bioinformatics*, 16 (5): 412–424 (2000).
158. Vapnik, V., "The Nature of Statistical Learning Theory", 2nd edition. Ed., *Springer-Verlag*, New York, USA, (2000).
159. Osuna, E., "Support vector machines: Training and applications", Doctoral Dissertation, *Massachusetts Institute of Technology*, (1998).
160. Özseven, T., Ertürkler, M., Nurmhammed, M., Gül, M., and Harputluoğlu, M., "Quantifying the necrotic areas on liver tissues using support vector machine (SVM) algorithm and Gabor filters", *Innovations In Intelligent Systems And Applications (INISTA), 2012 International Symposium On*, 1–5 (2012).
161. Kavzoğlu, T., ve Çölkesen, İ., "Destek vektör makineleri ile uydu görüntülerinin sınıflandırılmasında kernel fonksiyonlarının etkilerinin incelenmesi", *Harita Dergisi*, 144 (7): 73–82 (2010).
162. Hsu, C. W., Chang, C. C., and Lin, C. J., "A practical guide to support vector classification Technical Report", *National Taiwan University Department of Computer Science*, Taipei, 1-16 (2003).
163. Şakar, B. E., "Çoklu ses kayıtları içeren bir parkinson konuşma veri kümesinin toplanması ve analizi", Doktora Tezi, *İstanbul Üniversitesi Fen Bilimleri Enstitüsü*, İstanbul, (2014).
164. Hastie, T., Tibshirani, R., Friedman, J., and Franklin, J., "The elements of statistical learning: data mining, inference and prediction", *The Mathematical Intelligencer*, 27 (2): 83–85 (2005).
165. Michie, D., Spiegelhalter, D. J., and Taylor, C. C., "Machine Learning, Neural and Statistical Classification", *Ellis Horwood Upper Saddle River*, New Jersey, (1994).
166. Xiang, S., Nie, F., and Zhang, C., "Learning a mahalanobis distance metric for data clustering and classification", *Pattern Recognition*, 41 (12): 3600–3612 (2008).

167. Garcia, V., Debreuve, E., and Barlaud, M., "Fast k nearest neighbor search using GPU", *Computer Vision and Pattern Recognition Workshops, 2008. CVPRW'08. IEEE Computer Society Conference On*, 1–6 (2008).
168. Costantini, G., Iaderola, I., Paoloni, A., and Todisco, M., "EMOVO corpus: An italian emotional speech database.", *LREC*, 3501–3504 (2014).
169. Jackson, P., Haq, S., and Edge, J. D., "Audio-visual feature selection and reduction for emotion classification", *In Proc. Int'l Conf. On Auditory-Visual Speech Processing*, 185–190 (2008).

EK AÇIKLAMALAR A.

DFT UYGULAMASI MATLAB KODLARI

```

clear all;
close all;
[x,fs]=audioread('ex2.wav');
x = (x(:,1)+x(:,2))/2; % stereo olan ses monoya çevrildi
figure, plot(x);
title('Orjinal Sinyal'); xlabel('Time (s)'); ylabel('Amplitude');
axis([0 length(x), min(x) max(x)])

X1=fft(x);
figure, plot(abs(X1));
title('Orjinal Sinyal-FFT'); xlabel('Frequency (kHz)');
ylabel('Amplitude');
axis([0 length(X1)/2, 0 max(abs(X1))])
%----- Tekrar Örnekleme
x_resample=resample(x,10000,fs); %resample yapıldı
figure, plot(x_resample);
title('Tekrar Örneklenmiş Sinyal'); xlabel('Time (s)');
ylabel('Amplitude');
axis([0 length(x_resample), min(x_resample) max(x_resample)])

X2=fft(x_resample);
figure, plot(abs(X2));
title('Tekrar Örneklenmiş Sinyal-FFT'); xlabel('Frequency (kHz)');
ylabel('Amplitude');
axis([0 length(X2)/2, 0 max(abs(X2))])
%----- Pencereleme
windowHanning = window(@hann, length(x));
x_pen = windowHanning.*x;
figure, plot(x_pen);
title('Pencerelenmiş Sinyal'); xlabel('Time (s)');
ylabel('Amplitude');
axis([0 length(x_pen), min(x_pen) max(x_pen)])

X3=fft(x_pen);
figure, plot(abs(X3));
title('Pencerelenmiş Sinyal-FFT'); xlabel('Frequency (kHz)');
ylabel('Amplitude');
axis([0 length(X3)/2, 0 max(abs(X3))])

%----- Alçak Geçiren FIR filtreleme
N = 37; % Order
Fc = 3000; % Cutoff Frequency
flag = 'scale'; % Sampling Flag
win = hamming(N+1);
b = fir1(N, Fc/(fs/2), 'low', win, flag);
x_filtre = filter(b,1,x);
figure, plot(x_filtre);
title('Lowpass-FIR Sinyal'); xlabel('Time (s)');
ylabel('Amplitude');
axis([0 length(x_filtre), min(x_filtre) max(x_filtre)])

X4=fft(x_filtre);
figure, plot(abs(X4));
title('Filtrelenmiş Sinyal-FFT'); xlabel('Frequency (kHz)');
ylabel('Amplitude');
axis([0 length(X4)/2, 0 max(abs(X4))])

```

Şekil Ek A.1. DFT uygulaması MATLAB kodları.

EK AÇIKLAMALAR B.

ETİK KURUL İZİNİ

HİZMETE ÖZEL

HİZMETE ÖZEL

T.C.
GENELKURMAY BAŞKANLIĞI
GÜLHANE ASKERİ TIP AKADEMİSİ KOMUTANLIĞI
ANKARA



KLN.ARŞ.ETK.KRL.: 50687469-1491-197 -15/SEK.1677

28 Ekim 2015

KONU : GATA (KAEK 2015-66) Doç.Tbp.Alb.
Ali DORUK'a Ait Kurul Raporu

PSIKİYATRİ ANABİLİM DALI BAŞKANLIĞINA

İLGİ: GATA Y 66-2 (A) Gülhane Askeri Tıp Akademisi Klinik Araştırmalar Etik Kurulu Yönergesi.

1. İlgili Yönerge gereğince değerlendirilmek üzere GATA Klinik Araştırmalar Etik Kuruluna sunulan Doç.Tbp.Alb. Ali DORUK'a ait GATA KA EK-2015-66 kayıt numaralı ve **"Anksiyete Bozukluğunda Sesin Akustik Parametreler Açısından İncelenmesi"** başlıklı araştırma projesi 22 Ekim 2015 tarihinde KA EK 7'inci toplantısında incelenmiş olup GATA Klinik Araştırmalar Etik Kurulunun ilgili kararı EK-A'dadır.

2. GATA Y 66-2 (A) Klinik Araştırma Etik Kurulu yönerge gereği;

- Araştırma süresince kullanılacak olan gönüllüler tarafından imzalanmış "Bilgilendirilmiş Gönüllü Olur Formları'nın" saklanması ve talep edildiğinde KA EK'a sunulmasını,
- Çalışmanın tamamlanmasını müteakiben KA EK'na bilgi verilmesini,
- Çalışmanın "Primer Sonlanım Noktası'na" gelmeden sonlandırılması durumunda; sonlandırma nedeni ve tarihi hakkında GATA Komutanlığı Klinik Araştırma Etik Kurul Başkanlığına bilgi verilmesini, rica ederim.

Mustafa GEREK
Profesör Hava Tabip Tuğgeneral
GATA Klinik Araştırmalar
Etik Kurul Başkanı

EKİ :
EK-A (2 Adet Klinik Araştırmalar Etik Kurulu
Karar Formu)

HİZMETE ÖZEL
Bağlantı Noktası Svl.Me.G.ÜNAL (TEL:2293)


HİZMETE ÖZEL

Şekil Ek B.1. Etik kurul izni.

EK AÇIKLAMALAR C.

GÖNÜLLÜ OLUR FORMU

Çizelge Ek C.1. Asgari bilgilendirilmiş gönüllü olur formu örneği.

 <p>TC Sağlık Bakanlığı Türkiye İktisadi ve Sosyal Reform Bakanlığı</p>	<p>ASGARİ BİLGİLENDİRİLMİŞ GÖNÜLLÜ OLUR FORMU ÖRNEĞİ</p>	Doküman Adı: KADB-F.23-R.00
		Yayın Tarihi: 18.04.2013
		Sayfa No: 120/153
		Onaylayan: Daire Başkanı
<p>1. Anksiyete bozukluklarında ses yapısı hakkında yeni bir araştırma yapmaktayız. Bu araştırma bir tez çalışmasıdır. Araştırmanın ismi "Anksiyete Bozukluğunda Sesin Akustik Parametreler Açısından İncelenmesi" dir.</p> <p>2. Bu araştırmayı yapmak istememizin nedeni, anksiyetenin ses üzerindeki etkisini inceleyerek tanı ve tedavi süreçlerine katkı sağlamaktır.</p> <p>3. 4. Bu araştırma yapılırken size herhangi bir tedavi yöntemi uygulanmayacaktır. Araştırma sonrası psikiyatrik hastalığınızın gerektirdiği rutin tedaviler yapılacaktır.</p> <p>5. Size hastalığınızla ilgili tanısız ve hastalığınızın şiddetini gösteren ölçekler uygulanacaktır. Bu ölçekler sizin hastalığınızda hastalığın seyrini ve şiddetini göstermek amaçlı uygulanan ölçeklerdir. Yine izniniz doğrultusunda bu çalışmayı yapabilmek için görüşme esnasında ses kaydınız alınacak ve ses kaydınız uygun programlarla analiz edilecektir.</p> <p>6. Gönüllülük esasına göre araştırmaya dahil olmayı kabul ettiğiniz takdirde size sorulan sorulara doğru cevap verme, ölçekleri eksiksiz ve doğru biçimde doldurma, yapılan uyarı ve bilgilendirmelerine riayet etme sorumluluğuna sahip olacaksınız. Ayrıca, istediğiniz takdirde çalışmadan çekilme hakkına sahipsiniz, fakat araştırmacıları zor durumda bırakmamak için araştırmadan çekileceğinizi önceden bildirmeniz uygun olacaktır.</p> <p>7. Çalışmamızda herhangi bir deneysel kısım mevcut değildir.</p> <p>8. Araştırma esnasında size zarar verecek herhangi bir risk bulunmamaktadır.</p> <p>9. Bu araştırmanın size doğrudan bir faydası bulunmamaktadır. Sizin hastalığınızla ilgili bilimsel veriler elde edilmesi ve ileride bu bilgiler ışığında hastalara faydalı olunması amaçlanmaktadır.</p> <p>10. Size hastalığınızla ilgili rutin tedaviler uygulanacak olup alternatif bir tanı-tedavi yöntemi uygulanmayacaktır.</p> <p>11. Araştırmaya katıldığınız için size tazminat ödenmeyecek ve farklı bir tedavi hizmeti sağlanmayacaktır.</p> <p>12. Araştırma esnasında size ulaşım ya da yemek masrafı ile ilgili bir ödeme yapılmayacaktır.</p> <p>13. Bu çalışma, gönüllülük esasına bağlıdır, çalışmaya katılmayı reddedebilirsiniz. Bu araştırmaya katılmak tamamen isteğe bağlıdır ve reddettiğiniz takdirde size uygulanan tedavide herhangi bir değişiklik olmayacaktır. Yine çalışmanın herhangi bir aşamasında onayınızı çekmek hakkına da sahipsiniz.</p> <p>14. Sizinle ilgili tıbbi bilgiler gizli tutulacak, ancak, çalışmanın kalitesini denetleyen görevliler, etik kurullar ya da resmi makamlar tarafından gereği halinde erişilebilecektir. Bu formu imzalamanız halinde bu erişime izin vermiş olacaksınız.</p> <p>15. İlgili mevzuat gereğince kimliğinizi ortaya çıkaracak kayıtlar gizli tutulacak, kamuoyuna açıklanamayacak, araştırma sonuçlarının yayımlanması halinde dahi kimliğiniz gizli kalacaktır.</p> <p>16. Araştırma konusuyla ilgili ve araştırmaya katılmaya devam etme isteğinizi etkileyebilecek yeni bilgiler elde edildiğinde zamanında bilgilendirileceksiniz.</p> <p>17. Araştırma hakkında, kendi haklarınız hakkında veya araştırmayla ilgili herhangi bir yan etki yaşammanız halinde daha fazla bilgi temin edebilmek için temasa geçebileceğiniz kişi ve günün 24 saatinde erişebileceğiniz telefon numarası; Prof.Dr.Ali DORK: 0312 304 4509 ve 0541 885 4767 dir.</p> <p>18. Araştırmaya dahil edilme kriterlerini karşılamıyor olmanız durumunda veya araştırmaya</p>		

katılmaktan vazgeçmeniz halinde arařtırmaya katılımınız son bulacaktır.

19. Arařtırma kapsamında size birkaç ölçek uygulanacak ve ses kaydı yapılacak olup, bunun için 2 saatlik bir süre öngörülmektedir.

20. Arařtırmaya yaklaşık olarak 100 gönüllünün katılımı beklenmektedir.

21. Sizden alınacak ses kaydı incelenerek anksiyetenin ses üzerine olan etkileri arařtırılacaktır.

22. Yapılacak inceleme GATA Psikiyatri Ana Bilim Dalında yapılacak olup, kayıtlarınız başka bir merkeze ya da yurtdışına gönderilmeyecektir.

(Katılımcının/Hastanın Beyanı)

23. "Bilgilendirilmiş Gönüllü Olur Formundaki tüm açıklamaları okudum. Bana, yukarıda konusu ve amacı belirtilen arařtırma ile ilgili yazılı ve sözlü açıklama ařağıda adı belirtilen arařtırmacılar tarafından yapıldı. Arařtırmaya gönüllü olarak katıldığımı, istediğim zaman gerekçeli veya gerekçesiz olarak arařtırmadan ayrılabileceğimi biliyorum."

24. "Söz konusu arařtırmaya, hiçbir baskı ve zorlama olmaksızın kendi rızamla katılmayı kabul ediyorum."

29. "Anksiyete Bozukluğunda Sesin Akustik Parametreler Açısından İncelenmesi" arařtırması kapsamında alınan ses örneklerimin; sadece bu arařtırmada kullanılmasına izin veriyorum."

25. Katılımcı Adı, soyadı:

Adres-Telefon:

Tarih-İmza:

26. Katılımcı ile görüşen arařtırmacı Adı, soyadı:

Adres-Telefon:

Tarih-İmza

27. Görüşme tanığı Adı, soyadı:

Adres-Telefon:

Tarih-İmza

EK AÇIKLAMALAR D.

SOSYODEMOGRAFİK VERİ FORMU

Çizelge Ek D.1. Sosyodemografik veri formu.

SOSYODEMOGRAFİK ÖZELLİKLER	
Tarih: .../.../201...	
Adı Soyadı	
TC Kimlik numarası	
Yaşı	
Cinsiyet	
Boy (cm) ve ağırlık (kg)	
Adres	
Telefon	
Eğitim durumu (yıl olarak)	
Ekonomik durumunuz	Düşük () Orta () İyi ()
Medeni durum	Evli () Bekar ()
Ses tellerini etkileyen hastalık	Evet () Hayır ()
Solunum sistemini etkileyen hastalık (kulak burun boğaz, göğüs)	Evet () Hayır ()
Sigara kullanımı	Evet () Hayır () Miktarı (1 günde):
Alkol kullanıyor musunuz?	Evet () Hayır () Miktarı:
Yasal olmayan madde kullanıyor musunuz?	Evet () Hayır () İsmi, miktarı:
İlaç kullanımı	Evet () Hayır () Kullandığı ilaçlar:
Tanı:	

EK AÇIKLAMALAR E.

KULLANILAN METİNLER

Hiçbir şeye yetişemiyorum. Yaptığım bir şey yok, ama hiçbir şeyi yetiştiremiyorum.

Endişelerim beni bunaltıyor, sıkıyor. Pek çok durum beni endişelendirir. Çocuklar, iş, ev işleri, faturalar, sağlık... Bu kadar endişelenmemem gerektiğinin farkındayım, ama kendimi kontrol edemiyorum, kendime engel olamıyorum.

Özellikle baskı altında olduğumda çok fazla endişeleniyorum. Endişe veren düşüncelerimi aklımdan uzaklaştırıyorum. Bir işi bitirir bitirmez, yapmak zorunda olduğum diğer şeyler aklıma geliyor ve endişeleniyorum. Sıkılıyorum.

Bir sorun hakkında yapabileceğim bir şey olmadığında daha fazla endişe ediyorum. İşkence gibi. Hayatımı felç ediyor. Durup dururken bunaldığımı, endişelendiğimi fark ediyorum.

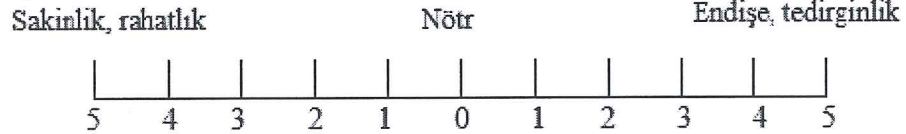
Sevdiklerimden ayrılmaya dayanamıyorum. Çocuklar ya da diğer yakınlarım başka bir yerde kaldıklarında içim içimi yiyor. Birinin başına bir şey geleceği korkusu, onlara zarar geleceği düşünceleri beni bunaltıyor. Yakınlarımla kaçırılacağı, kaza geçireceği, hastalanacağı, kanser olacağı bunlardan bazıları.

Uykuya dalmak zor, bir o tarafa bir bu tarafa dötüp duruyorum. Sıkı sık uyanıyorum. Sabahları kalkınca pestil gibi oluyorum. Sanki hiç uyumadım, hiç dinlenmedim. Uykuda hastalanırım, kalbim veya nefesim durar diye tedirgin oluyorum. Kalp çarpıntısı, terleme, mide bulantısı gibi vücudumda ufak bir değişiklik olması beni tedirgin ediyor. Acaba yine hastalanıyor muyum diye kaygılanıyorum.

Kapalı havalar, kapalı yerler beni sıkıyor. Yerimde duramıyorum. Moralistim düşüyor.

Okuduğumuz bazı yazılar bizde sakinlik, rahatlama, iç ferahlığı hissine, bazıları endişe, tedirginlik, sıkıntı hissine yola açarken bazı yazılar ise bizde hiç bir his (nötr) uyandırmaz.

Bu okuduğunuz metin sizde nasıl bir hissine yol açtı? Bunu aşağıdaki ölçek üzerinde işaretleyiniz.



Şekil Ek E.1. Endişe duygu durumu için kullanılan metin.

Masal Nedir?

Masallar; yaratıcısı bilinmeyen ve ağızdan ağıza sürüp gelen, olağanüstü, düş ürünü olaylarla örülmüş bir anlatı türüdür. Masallar, olağanüstü olabildikleri gibi gerçekçi de olabilir, yani gerçek kişilere, gerçek olaylara dayandıkları izlenimini veren masallar da vardır. Ama her iki masal çeşidinde de düş ürünü, uydurma oluş ana niteliktir. Bu ise biçimde ve anlatımda kendini gösterir.

Daha masalın başındaki tekerlemeyle dinleyici masala hazırlanır. Olayın uydurma olduğuna dikkat çekilir. Anlatım genellikle kısa ve yoğundur. Olaylar önemlerine göre sıralanarak aktarılır ve miş'li geçmiş, şimdiki zaman ya da geniş zaman rivayeti kullanılır.

Masallar, konularına göre şu ana çeşitlere ayrılır:

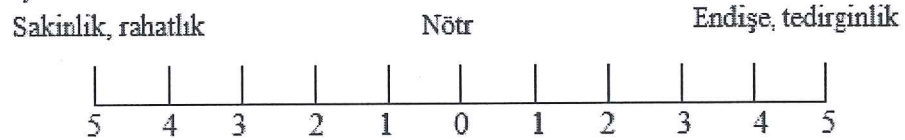
Hayvan masalları, asil masallar, olağanüstü masallar, gerçekçi masallar, güldürücü öyküler, nükteli fıkralar, yalanlamalar, zincirlemeli masallar.

Kahramanları hayvanlardan seçilen ve daha çok öğretici nitelik taşıyan hayvan masalları kısa olur, başlama tekerlemeleri de yoktur. Olağanüstü masalların kahramanları da insanlarla cinler, periler, devler gibi doğa dışı varlıklardır. Gerçekçi masallarda bu kişiler pek değişmez. Ama temel konu halk arasında seçilmiş tiplerin maceralarıdır. Güldürücü masallar ise, fıkra deyimiyle nitelenir. Yalanlamalı masalların ana niteliği yalana dayanmalarındır. Zincirlemeli masallar da kuruluşlarındaki özelliklerle ötekilerden ayrılır. Küçük, önemsiz olayların birbiri ardına bağlanmasıyla oluşturulmuşlardır.

Masallar sözlü halk edebiyatı türleri içinde ülkeden ülkeye, çağdan çağa en çok yayılan türlerdir.

Okuduğumuz bazı yazılar bizde sakinlik, rahatlama, iç ferahlığı hissine, bazıları endişe, tedirginlik, sıkıntı hissine yola açarken bazı yazılar ise bizde hiç bir his (nötr) uyandırmaz.

Bu okuduğumuz metin sizde nasıl bir hisse yol açtı? Bunu aşağıdaki ölçek üzerinde işaretleyiniz.



Şekil Ek E.2. Nötr duygu durumu için kullanılan metin.

EK AÇIKLAMALAR F.

ANKSİYETE ŞİDDETİNİ BELİRLEME ÖLÇEKLERİ

Çizelge Ek F.1. Hamilton anksiyete değerlendirme ölçeği.

HAMILTON ANKSİYETE DEĞERLENDİRME ÖLÇEĞİ/HAM-A					
Hasta adı soyadı:		Tarih			
Görüşmeyi yapan hekim:					
Aşağıdaki sorular için; HAYIR:0 puan, EVET: Şiddetine göre 1-4 puan					
	0	1	2	3	4
1.ANKSİYETELİ DUYGUDURUM Geçtiğimiz hafta ruh haliniz nasıldı? Kendinizi endişeli hissettiniz mi? Kötü bir şey olacak beklentisi oldu mu? Korkulacak bir şey olduğu hissine kapıldığınız veya çabuk parladığınız oluyor mu?					
2. GERGİNLİK Gerginlik, kolayca yorulma, sık sık irkilme, titremeler, yerinde duramama veya bir türlü gevşeyememe					
3.KORKULAR Karanlıktan, tanımadığınız insanlardan, yalnız kalmaktan, trafikten, kalabalıktan					
4.UYKUSUZLUK Geçtiğimiz hafta uykuya dalmakta zorlanma, huzursuz, rahatsız uyuma, dinlenememe, sabah yorgun ve halsiz hissetme, kötü rüya veya kabus görme					
5.ZİHİNSEL Her zamankine göre hafızada değişiklik, dikkatini toplamada zorluk(TV izlerken, kitap okurken), hatırlamada zorluk					
6. ÇOKKÜN DUYGUDURUM İlgisizlik-zevk almama, geçen hafta üzgün, ümitsiz, ağlamaklı hissetme, günün belli saatlerinde kendini daha kötü hissetme (sabahları)					
7. BEDENSEL (KAS) Ağrı, sızı, kaslarda seyirme, tutulma, diş gıcırdatma, sesin titremesi, kaslarda kasılma veya sıçramalar					
8.SOMATİK (DUYUSAL) Kulak çınlaması, vızıltı, tıkırtı veya gürleme gibi sesler görme bulanıklığı, ateş basması veya üşüme, takatsizlik, iğnelenmeler					
9. KALP-DAMAR BELİRTİLERİ Göğüs ağrısı, kalpte tekleme hissi, nabızda hızlanma, kalp atış hızında artış hissetme, damarlarda atma, bayılacak gibi olma					
10.SOLUNUM BELİRTİLERİ Göğüste baskı veya sıkışma; boğuluyor gibi hissetme, her zamankinden fazla iç çekme, nefes alırken zorlanma, güçlükle nefes alma, nefesin yetmemesi					
11.MİDE-BARSAK Mide, karında ağrı veya rahatsızlık, yutkunma zorluğu, fazla yellenme, midede ekşime, yanma, guruldama, karında şişkinlik ve gerilme hissi, mide bulantısı, barsaklarda gevşeklik veya ishal, kabızlık, kilo kaybı					
12.ÜREME-BOŞALTMA Sık idrara çıkma, adet gecikmesi veya kesilme, aşırı adet kanaması, cinsel bakımdan soğukluk, tepkisizlik, erken boşalma, sertleşme sorunu					
13.OTONOM BELİRTİLER Ağız kuruluğu, her zamankinden daha fazla terleme, yüz kızarması, solgunluk, sersemlik, baş dönmesi, gerginlik, tüylerin diken diken olması, baş ağrısı					
14.GÖRÜŞME SIRASINDAKİ DAVRANIŞLAR Sürekli kıpırdanma, yerinde duramama veya gezinme, ellerde titreme, çatık kaşlar, gergin yüz ifadesi, iç çekme veya sık sık soluk alıp verme, yüzde solukluk, şişme, geğirme, canlı refleksler, göz bebeklerinin büyümesi, gözlerin dışa doğru çıkması					
TOPLAM PUAN (0-5: Anksiyete yok, 6-14: Minör, 15-: Majör Anksiyete):					

Çizelge Ek F.2. Beck anksiyete envanteri.

BAE	
Hastanın adı soyadı:.....	Tarih:.....
Aşağıda insanların kaygılı ya da endişeli oldukları zamanlarda yaşadıkları bazı belirtiler verilmiştir. Lütfen her maddeyi dikkatle okuyunuz. Daha sonra, her maddedeki belirtinin BUGÜN DAHİL SON 1 HAFTADIR sizi ne kadar rahatsız ettiğini yandakine uygun yere (x) işareti koyarak belirleyiniz.	

	(0) Hiç	(1) Hafif derecede	(2) Orta derecede	(3) İleri derecede
1. Bedeninizin herhangi bir yerinde uyuşma veya karıncalanma				
2. Sıcak/ ateş basmaları				
3. Bacaklarda halsizlik, titreme				
4. Gevşeyememe				
5. Çok kötü şeyler olacak korkusu				
6. Baş dönmesi veya sersemlik				
7. Kalp çarpıntısı				
8. Dengeyi kaybetme duygusu				
9. Dehşete kapılma				
10. Sinirlilik				
11. Boğuluyormuş gibi olma duygusu				
12. Ellerde titreme				
13. Titreklilik				
14. Kontrolü kaybetme korkusu				
15. Nefes almada güçlük				
16. Ölüm korkusu				
17. Korkuya kapılma				
18. Midede hazımsızlık ya da rahatsızlık hissi				
19. Baygınlık				
20. Yüzün kızarması				
21. Terleme (sıcaklığa bağlı olmayan)				

ÖZGEÇMİŞ

Turgut ÖZSEVEN 1981 yılında Turhal'da doğdu; ilk ve orta öğrenimini aynı şehirde tamamladı. Tokat Teknik Lise Bilgisayar Programcılığı Bölümü'nden mezun oldu. 1999 yılında Kocaeli Üniversitesi Teknik Eğitim Fakültesi Elektronik ve Bilgisayar Eğitimi Bölümü'nde öğrenime başlayıp 2003 yılında mezun oldu. Eğitim süresince BİLKON Bilişim ve Kontrol Sistemleri A.Ş. 'de teknik personel olarak çalıştı. 2004 yılında öğretmen olarak Turhal Endüstri Meslek Lisesi'nde göreve başladı. 2008 yılında Gaziosmanpaşa Üniversitesi Turhal Meslek Yüksekokulu'nda öğretim görevlisi olarak göreve başladı. 2009 yılında Karabük Üniversitesi Fen Bilimleri Enstitüsü Bilgisayar Mühendisliği Anabilim Dalı'nda başlamış olduğu yüksek lisans programını 2011 yılında tamamladı. Turhal Meslek Yüksekokulu'nda başladığı görevine halen aynı yerde devam etmektedir.

ADRES BİLGİLERİ

Adres : Gaziosmanpaşa Üniversitesi
Turhal Meslek Yüksekokulu
Turhal / TOKAT

Tel : (533) 224 5325

E-posta : turgutozseven@gmail.com