

**T.C.
İSTANBUL AYDIN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**



**İŞİTME KAYIPLI BİREYLERİN DERİN ÖĞRENME TABANLI SES
DUYGU ANALİZİ**

YÜKSEK LİSANS TEZİ

Büşra ULUDAĞ

Odyoloji Anabilim Dalı

Odyoloji Programı

Ekim, 2020

**T.C.
İSTANBUL AYDIN ÜNİVERSİTESİ
LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ**



**İŞİTME KAYIPLI BİREYLERİN DERİN ÖĞRENME TABANLI SES
DUYGU ANALİZİ**

YÜKSEK LİSANS TEZİ

Büşra ULUDAĞ
(Y1816.070029)

Odyoloji Anabilim Dalı
Odyoloji Programı

Tez Danışmanı: Prof. Dr. B. Özlem Konukseven

Ekim, 2020

ONUR SÖZÜ

Yüksek Lisans tezi olarak sunduđum “İşitme Kayıplı Bireylerin Derin Öğrenme Tabanlı Ses Duygu Analizi” adlı çalışmanın, tezin proje safhasından sonuçlanmasına kadarki bütün süreçlerde bilimsel ahlak ve geleneklere aykırı düşecek bir yardıma başvurulmaksızın yazıldığını ve yararlandığım eserlerin Bibliyografya’da gösterilenlerden oluştuđunu, bunlara atıf yapılarak yararlanılmış olduğunu belirtir ve onurumla beyan ederim. (13/10/2020)

Büşra ULUDAĞ

ÖNSÖZ

Lisans eğitimimin son yılında, yüksek lisans eğitimim boyunca ve akademisyenliğe adım attığım günden bugüne ışığını, sevgisini ve desteğini her zaman hissettiğim, değerli bilgi ve tecrübeleri ile yolumu aydınlatan tez danışmanım Prof. Dr. Özlem Konukseven' e,

Desteğini hiçbir zaman esirgemeyen kıymetli hocam Dr. Öğr. Üyesi İnci Adalı'ya,

Çalışmamız boyunca yanımızda olan derin bilgi, birikim ve tecrübeleriyle bize yol gösteren değerli hocam Prof. Dr. Tahir Çetin Akıncı'ya,

Tez sürecimde sıkıştığım her konuda bana destek olan, gösterdiği anlayış ve hoşgöründen dolayı değerli hocam Prof. Dr. Ragıp Kutay Karaca'ya,

Tez çalışmama katkıda bulunan arkadaşlarıma ve Odyoloji bölümü öğrencilerimize,

Her zaman olduğu gibi desteklerini hiçbir zaman esirgemeyen, emeklerini ödeyemeyeceğim annem Zübeyde Uludağ'a, babam Hayati Uludağ'a, ablam Bahar Uludağ'a ve minik kardeşim Beyza Uludağ'a,

Sonsuz teşekkürler.

Ekim, 2020

Büşra ULUDAĞ

İŞİTME KAYIPLI BİREYLERİN DERİN ÖĞRENME TABANLI SES DUYGU ANALİZİ

ÖZET

Amaç: İşitme kayıplı bireylerin, hedef duyguyu tanıma ve ifade etme becerileri düşüktür. Bu çalışmanın amacı, koklear implant kullanan bireylerde duyguları sözel ifade etme becerisini derin öğrenme ile değerlendirmek ve koklear implant kullanan bireylerin eğitim öncesi ve eğitim sonrası derin öğrenme ile ses duygu analizleri arasındaki farkı ortaya koymaktır.

Gereç ve Yöntem: Çalışma, veri setinin oluşturulması, derin öğrenme algoritmasının oluşturulması/ eğitilmesi ve prototipin canlıya geçirilmesi (klinik uygulama) olmak üzere üç aşamada gerçekleştirilmiştir. Derin öğrenme algoritmasının eğitilmesi için hazırlanan veri seti cümleleri nötr, mutluluk, hüzün, öfke, korku ve şaşkınlık olmak üzere altı farklı duyguda Drama ve Oyunculuk Bölümü'nde öğrenim gören 20 kadın ve 20 erkek öğrenci tarafından seslendirilmiştir. Derin öğrenme eğitimi tamamlandıktan sonra 18- 40 yaş aralığında, 20 normal işiten ve 20 koklear implant kullanan olmak üzere toplamda 40 birey çalışmaya dahil edilmiştir. Koklear implant kullanan bireyler resim ve animasyon olmak üzere iki ayrı duygu eğitimi uygulandıktan sonra cümleleri tekrar seslendirmiştir. Derin öğrenme yöntemi kullanılarak, eğitim öncesi ve eğitim sonrası ses duygu analizleri karşılaştırılmıştır.

Bulgular: Derin öğrenme eğitimi tamamlandıktan sonra test aşamasında doğruluk oranı %93 elde edilmiştir. Derin öğrenme ile ses duygu analizi doğruluğu, normal işiten bireylerde %94.3, koklear implant kullanan bireylerde eğitim öncesi %68.25, koklear implant kullanan bireylerde eğitim sonrası %91.3 olarak elde edilmiştir.

Sonuç: Koklear implant kullanan bireylerin ses duygu analizi doğruluk oranı, normal işiten bireylerin ses duygu analizi doğruluk oranına göre daha düşük ve koklear implant kullanan bireylerin eğitim sonrası ses duygu analizi doğruluk oranı

daha yüksek elde edilmiştir. Bu çalışma, işitme kayıplı bireyler için ilk olması ile literatüre yenilikçi bir katkı sağlayacağı öngörülmektedir.

Anahtar Kelimeler: derin öğrenme, duygu, koklear implant, yapay zekâ, ses duygu analizi

DEEP LEARNING BASED VOICE EMOTION ANALYSIS OF INDIVIDUALS WITH HEARING LOSS

ABSTRACT

Aim: Individuals with hearing loss have low ability to recognize and express the target emotion. The aim of this study is to evaluate the verbal expression of emotions in individuals using cochlear implants with deep learning and to reveal the difference between pre-and post-training deep learning and sound emotion analysis of individuals using cochlear implants.

Material Metod: The study was carried out in three stages: creating the data set, creating / training the deep learning algorithm, and moving the prototype to life (clinical practice). The dataset sentences prepared for the training of the deep learning algorithm were voiced by 20 female and 20 male students studying in the Drama and Acting Department in six different emotions: neutral, happiness, sadness, anger, fear, and surprise. After the deep learning training was completed, a total of 40 individuals, including 20 normal hearing and 20 cochlear implant users, between the ages of 18-40 were included in the study. Individuals using cochlear implants vocalized the sentences again after applying two separate emotional training, namely painting and animation. Using the deep learning method, pre-training and post-training sound emotion analysis were compared.

Results: After the deep learning training was completed, the accuracy rate was 93% in the test phase. Sound emotion analysis accuracy with deep learning was achieved as 94.3% in individuals with normal hearing, 68.25% before training in individuals using cochlear implants, and 91.3% after training in individuals using cochlear implants.

Conclusion: The accuracy rate of sound emotion analysis of individuals using cochlear implants was lower than the accuracy rate of sound emotion analysis of normal hearing individuals, and the accuracy rate of sound emotion analysis after training of individuals using cochlear implants was higher. It is expected that this

study will make an innovative contribution to the literature, being the first for individuals with hearing loss.

Key Words: deep learning, emotion, cochlear implant, artificial intelligence, sound emotion analysis

İÇİNDEKİLER

	Sayfa
ONUR SÖZÜ	v
ÖNSÖZ.....	vii
ÖZET.....	ix
ABSTRACT	xi
İÇİNDEKİLER	xiii
KISALTMALAR	xvii
ÇİZELGE LİSTESİ.....	xix
ŞEKİLLER LİSTESİ.....	xxi
I. GİRİŞ..	1
A. Çalışma Konusu.....	1
B. Çalışmanın Amacı	2
C. Literatür Özeti.....	3
D. Hipotezler	8
II. GENEL BİLGİLER.....	9
A. Duygu	9
1. Duygu Tanımı	9
2. Duygu Teorileri.....	10
a. İki Faktörlü Duygu Teorisi	10
b. Gelişimsel Duygu Teorisi.....	11
c. Evrim Duygu Teorileri	12
3. Duygu Boyutları.....	14

B. İşitme Kaybı.....	16
1. İletim Tipi İşitme Kaybı.....	18
2. Sensörinöral Tip İşitme Kaybı	18
3. Mikst Tip İşitme Kaybı	19
C. Koklear İmplant	20
1. Koklear İmplant Tarihçesi.....	22
2. Koklear İmplant Kullanıcılarında Dil ve Konuşma Gelişimi	23
3. Tek Taraflı ve Bilateral Koklear İmplant.....	24
4. Koklear İmplant Kullanan Bireylerde Ses Duygu Tanıma ve İfade Etme ...	25
D. Geçmişten Günümüze Ses Duygu Analizi.....	27
E. Ses.....	29
1. Ses Nedir?	29
2. Ses Dalgalarının Temel Özellikleri	30
3. Ses Dalgalarının Karakteristik Özellikleri	31
F. Yapay Zekâ	32
1. Makine Öğrenmesi	33
2. Yapay Sinir Ağları	34
3. Derin Öğrenme.....	36
a. Convolutional Neural Networks	38
III. GEREÇ VE YÖNTEM.....	47
A. Çalışmanın Yürütüldüğü Birim	47
B. Çalışmanın Örnekleme	47
C. Katılımcıların Özellikleri	48
D. Veri Toplama Araçları ve Çalışma Planı	48
1. Veri Setinin Oluşturulması.....	49
a. Derin Öğrenme Eğitim Seti İçin Cümlelerin Hazırlanması.....	49

b. Veri Seti Ses Kayıtlarının Etiketlenmesi.....	51
2. Derin Öğrenme Algoritması Oluşturulması ve Eğitilmesi.....	53
a. Derin Öğrenme Algoritmasının Aşamaları	54
3. Prototipin Canlıya Geçirilmesi (Klinik Uygulama).....	59
a. Normal İşiten Bireylerin Test Aşamaları	59
b. İşitme Kayıplı Bireylerin Test Aşamaları.....	61
c. Normal İşiten ve İşitme Kayıplı Bireylerin Ses Kayıtlarının Etiketlenmesi.....	64
IV. BULGULAR.....	67
A. Derin Öğrenme Algoritması Sınıflandırma Bulguları	67
1. Nötr Duygu Analizi.....	67
2. Mutluluk Duygu Analizi	70
3. Hüzün Duygu Analizi	72
4. Öfke Duygu Analizi	75
5. Korku Duygu Analizi.....	77
6. Şaşkınlık Duygu Analizi	80
B. Derin Öğrenme Doğruluğu	85
C. Normal İşiten ve Koklear İmplant Kullanan Bireylerin Sosyodemografik Özellikleri.....	86
D. Normal İşiten ve Koklear İmplant Kullanan Bireylerin Derin Öğrenme ile Ses Duygu Analizi Bulguları	87
V. TARTIŞMA	89
VI. SONUÇ VE ÖNERİLER.....	97
VII.KAYNAKÇA	99
EKLER.....	135
ÖZGEÇMİŞ.....	143

KISALTMALAR

ABC	: Airplane Behaviour Corpus
ACE	: Advanced Combination Encoder
ANN	: Artificial Neural Networks
AVIC	: The Audiovisual Interest Corpus
BLSTM	: Bidirectional Long Short-Term Memory
CIS	: Continuous Interleaved Sampling
CNN	: Convolutional Neural Networks
CRBM	: Convolutional Restricted Boltzmann Machine
CRF	: Conditional Random Field
DCNN	: Deep Convolutional Neural Networks
DCT	: Discrete Cosine Transform
DES	: Danish Emotional Speech
DNN	: Deep Neural Network
DTPM	: Discriminant Temporal Pyramid Matching
ELM	: Extreme Learning Machine
EMO-DB	: Berlin Emotional Speech Database
FFT	: Fast Fourier Transform
GEMEP	: Cenevre Multimodal Duygu Tanımlama
GerDa	: Generalized Discriminant Analysis
HMM	: Hidden Markov Model
IADS	: International Affective Digitized Sounds
IAPS	: International Affective Picture System
IEMOCAP	: Interactive Emotional Dyadic Motion Capture

LSTM	: Long Short-Term Memory
MFCC	: Mel-Frequency Cepstral Coefficients
PSO-FF	: Particle Swarm Optimization- Firefly
RBM	: Restricted Boltzmann Machines
ReLU	: Rectified Linear Unit
RNN	: Recurrent Neural Networks
SAL	: Sensitive Artificial Listener
SSO	: Saf Ses Ortalaması
SUSAS	: The Speech Under Simulated and Actual Stress
SVM	: Support Vector Machine
Tanh	: Hiperbolik Tanjant
TED-LIUM	: TED l'Université du Maine
VAE	: Variational Autoencoders
VAM	: Vera-Am-Mittag
WHO	: Dünya Sağlık Örgütü
YSA	: Yapay Sinir Ağları

ÇİZELGE LİSTESİ

	Sayfa
Çizelge 1. Konuşma Duygu Tanıma Derin Öğrenme Modelleri.....	5
Çizelge 2. Makine Öğrenmesi Türleri	34
Çizelge 3. Derin Öğrenme Yöntemleri, Kullanıldığı Alanlar ve Örnek Uygulamalar	38
Çizelge 4. Çalışmaya Dahil Edilen Bireylerin Sayısı.....	48
Çizelge 5. Oluşturulan Cümleler	49
Çizelge 6. Eğitim Aşaması / Cümle Numaraları	52
Çizelge 7. Eğitim Aşaması / Duyguları İfade Eden Numaralar	53
Çizelge 8. CNN Mimarisi.....	59
Çizelge 9. Goodman İşitme Kaybı Sınıflandırması.....	60
Çizelge 10. Test Aşaması/ Cümle Numaraları	65
Çizelge 11. Test Aşaması/ Duyguları İfade Eden Numaralar	65
Çizelge 12. Koklear İmplant Kullanan Bireylere Ait Özellikler.....	86

ŞEKİLLER LİSTESİ

	Sayfa
Şekil 1. İki Faktörlü Duygu Teorisi'nin Şematik Diyagramı.....	10
Şekil 2. Evrim Teorisinin Şematik Diyagramı	12
Şekil 3. Altı Temel Duygunun Yüz İfadeleri	14
Şekil 4. Değerlik ve Uyarılma Modeli.	15
Şekil 5. Normal İşitmeye Ait Örnek Odyogram.....	17
Şekil 6. Periferik ve Santral İşitme Sistemi Lezyon Yerine Göre İşitme Kaybı Sınıflandırılması.....	17
Şekil 7. İletim Tipi İşitme Kaybı Örnek Odyogram.....	18
Şekil 8. Sensörinöral Tip İşitme Kaybı Örnek Odyogram	19
Şekil 9. Mikst Tip İşitme Kaybı Örnek Odyogram	20
Şekil 10. Koklear İmplant Sisteminin Bileşenleri	21
Şekil 11. Ses Dalgası	30
Şekil 12. Yapay Zeka, Makine Öğrenmesi ve Derin Öğrenme Kronolojisi.....	33
Şekil 13. Nöron Mimarisi	35
Şekil 14. Yapay Sinir Ağı (YSA) ve Derin Öğrenme (DL) Farkı.....	37
Şekil 15. İlkel CNN Modeli: Lenet 5 Örneği.	40
Şekil 16. Geliştirilmiş CNN Modeli.....	41
Şekil 17. Ses Sinyali ve Mel Spektrograma Dönüştürülmüş Hali.....	42
Şekil 18. Mel Ölçeği.....	43
Şekil 19. Tanh Fonksiyonu (-1,1) ve Sigmoid Fonksiyonu (0,1).....	44
Şekil 20. Audacity Ekran Görüntüsü.....	51

Şekil 21.	Normal İşiten Bireylerin Ses Kayıtlarının Etiketlenmesi.....	52
Şekil 22.	Mel Spektrogram Şeması	55
Şekil 23.	Hamming ve Hanning Fonksiyonu Farkı	56
Şekil 24.	Maksimum ve Ortalama Havuzlama Yöntemleri.....	58
Şekil 25.	Dropout Uygulanmadan Önce (A) ve Dropout Uygulandıktan Sonra CNN.....	58
Şekil 26.	Altı Farklı Duygu İçin Kadın ve Erkek Yüz İfadeleri.....	62
Şekil 27.	Animasyon Modülü Eğitim Materyalinden Bir Kesit	63
Şekil 28.	Normal İşiten Bireylerin Ses Kayıtlarının Etiketlenmesi.....	64
Şekil 29.	Koklear İmplant Kullanan Bireylerin Ses Kayıtlarının Etiketlenmesi	64
Şekil 30.	‘Çatal Elinden Düştü’ Cümlesine Ait Nötr Duygu Ham Ses İşareti	67
Şekil 31.	‘Çatal Elinden Düştü’ Cümlesine Ait Nötr Duygu Histogramı	68
Şekil 32.	‘Çatal Elinden Düştü’ Cümlesine Ait Nötr Duygu Rose Analizi.....	68
Şekil 33.	‘Çatal Elinden Düştü’ Cümlesine Ait Nötr Duygusu Ks Yoğunluk Analizi.....	69
Şekil 34.	‘Çatal Elinden Düştü’ Cümlesine Ait Nötr Duygusu Normal Dağılım İçin Olasılık Grafiği.....	69
Şekil 35.	‘Çatal Elinden Düştü’ Cümlesine Ait Nötr Duygusu İçin Makine Öğrenmesi Analizi.....	70
Şekil 36.	‘Çatal Elinden Düştü’ Cümlesine Ait Mutluluk Duygusu Ham Ses İşareti.....	70
Şekil 37.	‘Çatal Elinden Düştü’ Cümlesine Ait Mutluluk Duygusu Histogramı .	71
Şekil 38.	‘Çatal Elinden Düştü’ Cümlesine Ait Mutluluk Duygusu Rose Analizi.	71
Şekil 39.	‘Çatal Elinden Düştü’ Cümlesine Ait Mutluluk Duygusu Ks Yoğunluk Analizi.....	72
Şekil 40.	‘Çatal Elinden Düştü’ Cümlesi Mutluluk Duygusuna Ait Normal Dağılım İçin Olasılık Grafiği	72
Şekil 41.	‘Çatal Elinden Düştü’ Cümlesine Ait Hüzün Duygusu Ham Ses İşareti	73

Şekil 42.	‘Çatal Elinden Düştü’ Cümlesine Ait Hüzün Duygusu Histogramı	73
Şekil 43.	‘Çatal Elinden Düştü’ Cümlesine Ait Hüzün Duygusu Rose Analizi.....	74
Şekil 44.	‘Çatal Elinden Düştü’ Cümlesine Ait Hüzün Duygusu Ks Yoğunluk Analizi.....	74
Şekil 45.	‘Çatal Elinden Düştü’ Cümlesi Hüzün Duygusuna Ait Normal Dağılım İçin Olasılık Grafiği	75
Şekil 46.	‘Çatal Elinden Düştü’ Cümlesine Ait Öfke Duygusu Ham Ses İşareti...	75
Şekil 47.	‘Çatal Elinden Düştü’ Cümlesine Ait Öfke Duygusu Histogramı	76
Şekil 48.	‘Çatal Elinden Düştü’ cümlesine ait öfke duygusu Rose analizi	76
Şekil 49.	‘Çatal Elinden Düştü’ Cümlesine Ait Öfke Duygusu Ks Yoğunluk Analizi.....	77
Şekil 50.	‘Çatal Elinden Düştü’ Cümlesi Öfke Duygusuna Ait Normal Dağılım İçin Olasılık Grafiği	77
Şekil 51.	‘Çatal Elinden Düştü’ Cümlesine Ait Korku Duygusu Ham Ses İşareti	78
Şekil 52.	‘Çatal Elinden Düştü’ Cümlesine Ait Korku Duygusu Histogramı.....	78
Şekil 53.	‘Çatal Elinden Düştü’ Cümlesine Ait Korku Duygusu Rose Analizi	79
Şekil 54.	‘Çatal Elinden Düştü’ Cümlesine Ait Korku Duygusu Ks Yoğunluk Analizi.....	79
Şekil 55.	‘Çatal Elinden Düştü’ Cümlesi Korku Duygusuna Ait Normal Dağılım İçin Olasılık Grafiği	80
Şekil 56.	‘Çatal Elinden Düştü’ Cümlesine Ait Şaşkınlık Duygusu Ham Ses İşareti.	80
Şekil 57.	‘Çatal Elinden Düştü’ Cümlesine Ait Şaşkınlık Duygusu Histogramı ...	81
Şekil 58.	‘Çatal Elinden Düştü’ Cümlesine Ait Şaşkınlık Duygusu Rose Analizi	81
Şekil 59.	Çatal Elinden Düştü Cümlesi Şaşkınlık Duygusuna Ait Ks Yoğunluk Analizi.....	82
Şekil 60.	‘Çatal Elinden Düştü’ Cümlesi Şaşkınlık Duygusuna Ait Normal Dağılım İçin Olasılık Grafiği	82
Şekil 61.	Altı Farklı Duyguya Ait Derin Öğrenme Analizleri	84

Şekil 62. Derin Öğrenme Modeli Doğruluk Grafiği.....	85
Şekil 63. Derin Öğrenme Modeli Kayıp Grafiği	85
Şekil 64. Normal İşiten ve Koklear İmplant Kullanan Bireylerin Derin Öğrenme ile Ses Duygu Analizi Doğruluđu	87

I. GİRİŞ

A. Çalışma Konusu

Konuşma sadece dilsel içeriği değil aynı zamanda konuşmacının cinsiyet, yaş, aksan ve duygusal durumu ile ilgili bilgileri de iletmektedir. İletişimin önemli boyutlarından biri akustik ipuçları kullanarak konuşmacının duygusal durumunun tanınmasıdır. Konuşmacıların yüz ifadeleri, duygusal durumun güçlü bir göstergesi olsa da ses duygu tanıma, özellikle görsel ipuçlarının olmadığı telefon görüşmeleri gibi sadece işitsel iletişimin söz konusu olduğu durumlarda önemli bir bileşendir. Konuşmacının duygusal durumuna işaret eden prozodik ipuçlarının algılanması, sonraki aşama olan duygusal durumun sözel ifade edilmesi açısından önemlidir (Fu & Nogaki, 2005). İleri ve çok ileri derecede işitme kaybı yalnızca işitmenin işlevsel bant genişliğini azaltmakla kalmaz, aynı zamanda spektral ve temporal çözünürlüğünü de sınırlamaktadır. Bu durum işitme kayıplı bireylerin, ince akustik özellikleri algılanmasını ve buna bağlı olarak duygusal prozodinin üretilmesini zorlaştırmaktadır (Moore & Carlyon, 2005; Luo, Fu & Galvin, 2007).

Derin öğrenme, makine öğrenmesinin alt dalıdır. Derin öğrenme, doğrusal olmayan katmanları kullanarak öznetelik çıkarma ve sınıflandırma yapabilen çok katmanlı bir sinir ağıdır. Ardışık katmanlar, önceki katmanlardaki çıktıyı alıp girdi olarak kullanmaktadır (Zhang, Quan & Ren, 2016) . Derin öğrenme, kodlanmış kurallar ile öğrenmek yerine ses, resim, video verilerinden otomatik olarak öğrenebilme özelliği sayesinde makine öğrenmesi yöntemlerinden farklıdır. Derin öğrenme algoritmasının veri büyüklüğüne bağlı olarak tahmin doğruluğu artmaktadır (Buduma & Locascio, 2017). Günümüzde derin öğrenme, hastalıkların erken tanısında, medikal görüntüleme, beyin doku sınıflandırmasında, nöral hücre sınıflandırmasında ve tümör sınıflandırması gibi sağlık alanlarında da etkinliğini arttırmıştır (Umut, Yılmaz & Dikmen, 2019).

Literatürde işitme kaybı ile duygu algısı ve ses duygu ifade etme arasındaki ilişkiyi inceleyen çok sayıda çalışma yer almaktadır (Most & Michaelis, 2012; Cannon & Chatterjee, 2019; Christensen ve ark., 2019). Çalışmamızda işitme kayıplı bireylerde (18- 40 yaş) ses duygu analizi, literatürdeki çalışmalardan farklı olarak derin öğrenme algoritması kullanılarak yapılmıştır. Ek olarak koklear implant kullanan bireylere duygu eğitimi uygulanarak eğitim öncesi ve eğitim sonrası duyguları sözel ifade etme becerileri karşılaştırılmıştır.

B. Çalışmanın Amacı

Koklear implant kullanıcıları, harmonik perde algısı için gerekli olan spektro-zamansal ayrıntının azalması sebebiyle soru-ifade tezatlığı, ses duygu ve sözcüksel ton tanımayı algılamada yetersiz kalırlar (Zeng, 2002; Peng ve ark., 2004; Luo & Fu, 2004; Wei ve ark., 2004; Luo, Fu & Galvin, 2007; Chatterjee & Peng, 2008). Bu sınırlama, koklear implant kullanan bireylerin duygu algısı ve duyguları ifade etme becerisini olumsuz etkilemektedir (Kong ve ark., 2004; Xu ve ark., 2009; Chatterjee ve ark., 2015). Koklear implant kullanan bireylerin duyguları algılama ve sözel ifade etme becerileri ile ilgili çalışmalar literatürde yer almaktadır. Ancak ülkemizde ve dünyada işitme kayıplı bireylerin ses duygu analizini derin öğrenme gibi objektif bir şekilde analiz eden çalışmalara rastlanılmamıştır. Bu çalışmanın amacı koklear implant kullanan bireylerde duyguları sözel ifade etme becerisini derin öğrenme teknolojisi ile değerlendirmektir.

İşitsel rehabilitasyon programının, koklear implant kullanan bireylerin iletişim becerileri üzerindeki etkinliği bilinmektedir. Bu sayede koklear implant kullanan bireylerin, işitsel algı ve konuşma üretimi performansı iyileşmektedir (DesJardin & Eisenberg, 2007). Literatürde ses duygu üretimine yönelik eğitim seanslarının işitsel rehabilitasyonda önemi ile ilgili çalışmalar yetersizdir. Bu bağlamda çalışmanın ikincil amacı koklear implant kullanan bireylerin eğitim öncesi ve eğitim sonrası ses duygu analizleri arasındaki farkı ortaya koymaktır.

Bu doğrultuda;

- Derin öğrenme algoritması nötr, mutluluk, hüznün, korku, öfke ve şaşkınlık duygularında eğitilmiştir.

- Koklear implant kullanan bireylerin eğitim öncesi ve eğitim sonrası duyguları sözel ifade etme becerisi derin öğrenme algoritması ile analiz edilmiştir.
- Normal işiten bireylerin duyguları sözel ifade etme becerisi derin öğrenme algoritması ile analiz edilmiştir.

C. Literatür Özeti

Koklear implantlar, işitme sinirini doğrudan elektrik akımı yoluyla uyarmak için elektrotlar kullanarak, ileri ve çok ileri düzeyde işitme kaybı olan bireylerde konuşma algılamayı sağlayan cerrahi olarak yerleştirilen işitsel protezlerdir. Koklear implant kullanıcılarının sessiz ortamlarda konuşmayı anlama becerisi yüksektir (Barrett ve ark., 2020). Bununla birlikte, koklear implant aracılığıyla iletilen konuşma sinyalinin spektro- temporal ayrıntıda eksik olması, özellikle gürültülü ortamlar veya çoklu konuşma durumları gibi karmaşık dinleme senaryolarında konuşmanın algılanması zorlaştırmaktadır (Shannon ve ark., 2004; Pisoni, 2005; Başkent ve ark., 2016; Plack, 2018).

Duygusal prozodi, konuşmacının duygusal durumu hakkında bilgi aktarır ve dinleyicinin duyguyu tanıma becerisine katkıda bulunur. Normal işiten bireylere kıyasla, koklear implant kullanıcıları hedef duyguyu tanıma ve ifade etme görevlerinde başarısız olmaktadır (Pereira, 2000; Luo ve ark., 2007; Hopyan-Misakyan ve ark., 2009; Nakata ve ark., 2012; Volkova ve ark., 2013; Gilbers ve ark., 2015; Chatterjee ve ark., 2015). Everhardt ve ark (2020), koklear implantın tonal olmayan dillerde linguistik ve duygusal prozodinin tanımlanmasını nasıl etkilediğini değerlendirmek amacıyla metaanaliz çalışması yapmışlardır. Metaanaliz, koklear implantların linguistik ve duygusal prozodi üzerinde olumsuz etkiye sahip olduğunu ortaya çıkarmıştır (Everhardt ve ark., 2020). Koklear implant kullanıcılarının karşılaştığı zorluklar, melodi algısı/ üretimi (Gfeller ve ark., 2002; Kong ve ark., 2004; McDermott, 2004; Limb & Roy, 2014) ve nörofizyolojik düzeyde müzikal duyguların algısında (Deroche ve ark., 2019) da gözlenmektedir (Barrett ve ark., 2020).

Bugüne kadar yapılan çalışmalar, koklear implant kullanan çocukların ses duygu tanıma ve ses duygu ifade etme görevinde, normal işiten akranlarından önemli ölçüde zayıf kaldığını göstermektedir (Most & Aviner, 2009; Ketelaar ve ark., 2012; Volkova ve ark., 2013; Wang ve ark., 2013). Nakata ve ark. (2012), 5-13 yaş arası Japonca konuşan çocuk koklear implant kullanıcılarının, duyguları sözel taklit etme becerilerinin, normal işiten çocuklara göre anlamlı derecede düşük olduğunu göstermiştir. Bu durum, koklear implant kullanan çocukların duyguları algılamadaki yetersizlikleriyle ilişkili bulunmuştur. Benzer şekilde, koklear implant kullanan çocukların (5- 7 yaş arası), normal işiten çocuklara göre mutlu ve üzgün duygu ifadelerini sözel taklit etmede yetersiz kaldığı belirtilmiştir (Wang ve ark., 2013). Normal işitmeye sahip okul çağındaki çocukların çoğu, yetişkinlere benzer perde ayrımcılığına sahiptir (Deroche ve ark., 2012). Ses duygu tanıma, normal işiten çocuklarda 5 yaşında mümkündür ve zamanla gelişmeye devam etmektedir (Tonks ve ark., 2007; Sauter ve ark., 2013). İşitme kayıplı çocukların ise duygu tanıma eksiklikleri vardır ve bunların bir kısmı eğitimle hafifletilebilmektedir (Dyck & Denver, 2003; Dyck ve ark., 2004). Yapılan çalışmalar, çocuk koklear implant kullanıcılarının, konuşma tonlamasını algılama ve ifade etmede zayıflıklar gösterdiğini bildirmiştir. (Peng ve ark., 2004; Peng ve ark. (2008); Chatterjee ve ark., 2015).

Prozodik özellikler, konuşma duygusunu tanıma için ayırt edici özellik taşımaktadır (Zeng ve ark., 2008). En yaygın kullanılan prozodik özellikler temel frekans, enerji ve süredir. Temel frekans olan F0, konuşmanın ritmik ve tonal özelliklerini vermektedir. Hacim veya yoğunluk olarak adlandırılan konuşma sinyalinin enerjisi, konuşma sinyallerinin zaman içindeki genlik değişimini yansıtmaktadır. Araştırmacılar, öfke, mutluluk veya şaşkınlık gibi yüksek uyarılma duygularının artan enerji verdiğini, iğrenme ve üzüntünün ise azalmış enerji ile sonuçlandığını öne sürmektedir (Lin ve ark., 2012). Süre; harf, kelime ve benzer yapıları oluşturmak için geçen süreyi ifade etmektedir. Prozodik özellikler ile konuşmanın duygu içeriği arasında ilişki bulunmaktadır. Örneğin, öfke, korku, kaygı ve mutluluk gibi üst düzey uyarılma duygularının üretimi sırasında ortalama F0, F0 değişkenliği ve ses şiddeti artmaktadır. Üzüntü gibi düşük seviyeli uyarılmada ise doğal konuşmaya kıyasla ortalama F0, F0

değişkenliği ve ses yoğunluğu azalmaktadır (Frick, 1985; Bachorowski, 1999). Öfkeyi ifade etme süresi, üzüntü ifade etme süresinden daha kısadır (Rao ve ark., 2013; Akçay & Oğuz, 2020).

Konuşma içerdiği duygular aracılığıyla bilgiyi aktarmanın doğrudan yoludur. İnsan sesini ve yüz ifadelerini analiz ederek duyguların otomatik olarak tanınması, son yıllarda çok sayıda çalışmaya konu olmuştur (Rao ve ark., 2013). Otomatik duygu tanıma sistemlerinin pek çok alanda farklı amaçlarla kullanılabilmesi, bu konudaki çalışma sayısının önemli ölçüde artmasına neden olmuştur (Aouani & Ayed, 2020). Derin öğrenme, yeni bir yapay zekâ tekniğidir ve tıbbi uygulamaları son yıllarda büyük ilgi uyandırmaktadır (LeCun, Bengio & Hinton, 2015). Bilgiyi işlemek için insan beynindeki nöron katmanlarını taklit etmek ve bilgisayarların programlanmadan öğrenmesini sağlamak için tasarlanmıştır (Ting ve ark., 2018). Konuşma duygu tanıma mimarilerinin çoğu, CNN (Convolutional Neural Networks), LSTM (Long Short Term Memory), RNN (Recurrent Neural Networks) veya bunların kombinasyonudur (Trigeorgis ve ark., 2016; Niu ve ark., 2018). Konuşma duygu tanımanın ana hedeflerinden biri, model eğitimi için kullanılabilir önemli özelliklerin belirlenmesidir (Issa, Demirci & Yazıcı, 2020). CNN, iki boyutlu şekilleri tanımlamak için özel olarak tasarlanmış çok katmanlı bir algılayıcıdır. CNN modeli, uyarlanabilir özellik çıkarma özelliklerinden dolayı, ses sinyallerinden duygu tanıma için kullanılmaktadır (Zhang, Quan & Ren, 2016).

Çizelge 1. Konuşma Duygu Tanıma Derin Öğrenme Modelleri

Konuşma Duygu Tanıma Literatürü	Yıl	Derin Öğrenme Modelleri	Sınıflandırmada Kullanılan Duygular	Kullanılan Duygu Veri Seti	Doğruluk
Lin & Wei	2005	HMM (Hidden Markov Model) SVM (Support Vector Machine)	Öfke, mutluluk, üzüntü, şaşkın, nötr	DES (Danish Emotional Speech)	%99.5 %88.9
Chavhan ve ark.	2010	SVM	Öfke, mutluluk, üzüntü, korku, nötr	EMO-DB (Berlin Emotional Speech Database)	%93.75

Çizelge 1. (devamı) Konuşma Duygu Tanıma Derin Öğrenme Modelleri

Konuşma Duygu Tanıma Literatürü	Yıl	Derin Öğrenme Modelleri	Sınıflandırmada Kullanılan Duygular	Kullanılan Duygu Veri Seti	Doğruluk
Stuhlsatz ve ark.	2011	RBM (Restricted Boltzmann Machines) + Generalized Discriminant Analysis (GerDa)	Öfke, mutluluk, üzüntü, korku, nötr, tiksinti, can sıkıntısı	EMO-DB	%85.6
			Öfke, mutluluk, üzüntü, korku, tiksinti, şaşkın	eNTERFAC E	%72.4
			Agresif, mutluluk, sersem, gergin, nötr, yorgun	ABC (Airplane Behaviour Corpus)	%61.5
			Yüksek stres, orta düzeyde stres, nötr, bağırma	SUSAS (The Speech Under Simulated and Actual Stress)	%56.5
			Spontan konuşma ve doğal duygular	AVIC (The Audiovisual Interest Corpus)	%79.1
			Öfke, mutluluk, üzüntü, şaşkın ve nötr	DES	%60.1
			Spontan konuşma ve doğal duygular	SAL (Sensitive Artificial Listener)	%34.3
Wöllmer ve ark.	2013	LSTM (Long Short-Term Memory) + BLSTM (Bidirectional Long Short-Term Memory)	Öfke, çaresizlik, mutluluk, nötr, şaşkınlık	SmartKom	%59.5
			Değerlik, uyarılma ve baskınlık temelinde duygular	VAM (Vera-Am-Mittag)	%68.0
Han, Yu & Tashev	2014	DNN (Deep Neural Network) + ANN + ELM (Extreme Learning Machine)	Duygusal içeriği olan görsel-ışitsel veri kümesi	SEMAINE	%71.2
			Heyecanlı, hayal kırıklığı, mutluluk, nötr, şaşkın	IEMOCAP (Interactive Emotional Dyadic Motion Capture)	%54.3

Çizelge 1. (devamı) Konuşma Duygu Tanıma Derin Öğrenme Modelleri

Konuşma Duygu Tanıma Literatürü	Yıl	Derin Öğrenme Modelleri	Sınıflandırmada Kullanılan Duygular	Kullanılan Duygu Veri Seti	Doğruluk	
Amer ve ark.	2014	CRF (Conditional Random Field) + CRBM (Convolutional Restricted Boltzmann Machine)	Duygusal içeriği olan görsel- işitsel veri kümesi	AVEC	%69.2	
				VAM SPD	%66.3 %75.6	
Shaw, Vardhan & Saxena	2016	ANN (Artificial Neural Networks)	Öfke, mutluluk, üzüntü, nötr	Kendi oluşturdukları kayıt	%86.87	
Bertero & Fung	2017	CNN	TED konuşmaları/ Öfke, mutluluk, üzüntü	TED-LIUM (TED l'Université du Maine)	%66.1	
Zhang ve ark.	2017	DCNN AlexNet (Deep Convolutional Neural Networks) + DTPM (Discriminant Temporal Pyramid Matching) + SVM	Öfke, mutluluk, üzüntü, korku, nötr, tiksinti, can sıkıntısı	Öfke, mutluluk, üzüntü, korku, RML	%87.31	
				Öfke, mutluluk, üzüntü, korku, tiksinti, şaşkınlık	%75.34	
				Öfke, mutluluk, üzüntü, korku, tiksinti, şaşkınlık	eNTERFACE	%79.25
				Öfke, mutluluk, üzüntü, tiksinti, korku, şaşkınlık, can sıkıntısı, küçümseme	BAUM-1	%44.61
Harar, Burget & Dutta	2017	DCNN/10	Öfke, nötr, üzüntü	EMO-DB	%96.97	
Latif ve ark.	2017	VAE (Variational Autoencoders) + LSTM	Nötr, mutluluk, üzüntü ve öfke	IEMOCAP	%64.93	
Darekar & Dhande	2018	ANN + PSO-FF (Particle Swarm Optimization-Firefly)	Öfke, mutluluk, üzüntü, şaşkın, korku, nötr	Benchmark Database Marathi Database	%88.7 %80.0	
Eskimez, Duan & Heinzelman	2018	CNN + VAE	Öfke, heyecanlı, hayal kırıklığı, nötr, üzüntü	IEMOCAP	%48.54	
Zhao, Mao & Chen	2019	DCNN, LSTM	Öfke, heyecanlı, hayal kırıklığı, mutluluk, nötr, üzüntü	EMO-DB	%95.33	
				Öfke, heyecanlı, hayal kırıklığı, mutluluk, nötr, üzüntü	IEMOCAP	%86.16
Li, Zhao & Kawakara	2019	CNN + BLSTM + ATTN	Nötr, mutluluk, üzüntü ve öfke	IEMOCAP	%82.8	
Tiwari ve ark.	2020	DNN + Generative	Öfke, mutluluk, üzüntü, nötr	EMO-DB	%82.73	
				Öfke, mutluluk, üzüntü, nötr	IEMOCAP	%62.74

D. Hipotezler

H0: Normal işiten bireylerdeki ses duygu analizi bulguları ile koklear implant kullanan bireylerdeki ses duygu analizi bulguları arasında anlamlı bir fark yoktur.

H1: Normal işiten bireylerdeki ses duygu analizi bulguları ile koklear implant kullanan bireylerdeki ses duygu analizi bulguları arasında anlamlı bir fark vardır.

H0: Koklear implant kullanan bireylerin eğitim öncesi ses duygu analizi bulguları ile eğitim sonrası ses duygu analizi bulguları arasında anlamlı bir fark yoktur.

H1: Koklear implant kullanan bireylerin eğitim öncesi ses duygu analizi bulguları ile eğitim sonrası ses duygu analizi bulguları arasında anlamlı bir fark vardır.

II. GENEL BİLGİLER

A. Duygu

1. Duygu Tanımı

Duygular insanları ilişkilerde ve sosyal gruplarda bir arada tutan, ilişkiler arasındaki öncelikleri belirlemeye yardımcı olan, kişilerin çevre ile ilişkisinde tutumunu bildiren, sinir sistemi, dikkat süreci, karar verme, iletişim ve davranış yönetimi fonksiyonel evriminin önemli bir parçasıdır (Niedenthal & Ric, 2017).

Türk Dil Kurumu (TDK), ‘duygu’ kelimesini “1. Duyularla algılama, his 2. Belirli nesne, olay veya bireylerin insanın iç dünyasında uyandırdığı izlenim. 3. Önsezi. 4. Nesnelere veya olayları ahlaki ve estetik yönden değerlendirme yeteneği. 5. Kendine özgü bir ruhsal hareket ve hareketlilik.” olarak tanımlamaktadır.

Keltner ve Gross (1999) duyguları, “belirli fiziksel, sosyal zorluklara ve fırsatlara yanıt olarak ortaya çıkan epizodik, nispeten kısa vadeli, biyolojik temelli algı, deneyim, fizyoloji, eylem ve iletişim kalıpları” olarak tanımlamaktadır. Duygular işlevsel olarak insanların, kendilerine doğru şeyleri çekmelerini ve kendilerini yanlış şeylerden korumalarına yardımcı olmaktadır. Başka bir ifadeyle, “Duygular, deneyim akışı boyunca anlam oluşturan ve taşıyan bir tür radar ve hızlı tepki sistemidir. Duygular, deneyimi değerlendirdiğimiz ve durumlara göre hareket etmeye hazırladığımız araçlardır.” (Cole, Martin & Dennis, 2004).

Günümüzde birçok duygu bilimci, duyguların insan hayatı için gerekli ve vazgeçilmez olduğuna inanmaktadır (Damasio, 1999; Fredrickson, 2001; Keltner & Haidt, 2001; Barrett & Salovey, 2002; Mehu & Dunbar, 2008; Niedenthal & Brauer, 2012).

2. Duygu Teorileri

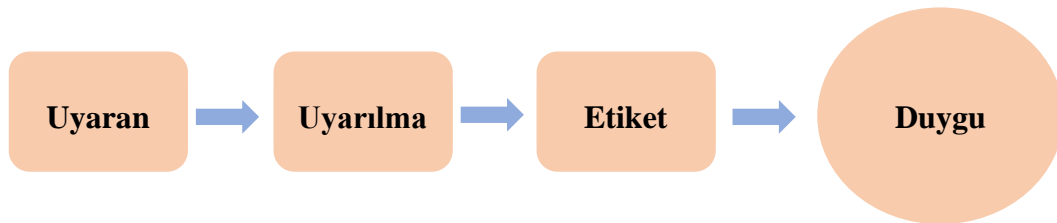
Geçmişten günümüze kadar duygular pek çok araştırmannın konusu olmuştur. Belirgin olarak nitelendirilen duygular, her kültürde az ya da çok olarak karşılık bulsa dahi evrensel bir denklik bulunmamaktadır. Duygu teorileri, duyguların tam olarak nasıl oluştuğuna dair test edilebilir ifadelerdir. Farklı teori gruplarının belirli fikirlerini karşılaştırmak için aşağıdaki üç maddenin cevaplarının alınması gerekmektedir (Niedenthal & Ric, 2017).

- 1) Bir duygunun öncülleri (onlara neden olan şey)
- 2) Biyolojik değerler (doğuştan gelen duygusal kapasiteler)
- 3) Duygusal deneyimin entegrasyonu (duygu bileşenlerinin birbirine nasıl uyum sağladığı)

Duygunun ortaya çıkması ile ilgili ortaya konan kuramlardan bazıları; Bilişsel Kuram, Canon-Bard Kuramı, James-Lange Kuramı, Arnold-Linsey Kuramı, ve Sosyo-Biyolojik Kuramdır. Deneylerle desteklenmesi ve gündelik hayattan örnekler ile açıklanması sebebiyle ‘Bilişsel Kuram’ bilim dünyasında kabul görmüştür. Stanley Schachter tarafından ‘Bilişsel Kuram’, “Bedenimizde olup biten fizyolojik değişiklikler için çevremizde bulunan uyarıcılar çerçevesinde anlamlı bir tanımlama yaparız.” şeklinde açıklanmıştır (Schachter ve Singer, Akt. Cüceloğlu, 1991: 162).

a. İki Faktörlü Duygu Teorisi

Schachter ve Singer (1962) tarafından ortaya konan iki faktörlü duygu teorisine göre duygular neredeyse her durumdan kaynaklanan otonomik uyarılma ve mevcut durum açısından deneyimi tanımlayan bir etiketın ortaya çıkmasıyla oluşmaktadır. Teori Şekil 1'de gösterilmiştir.



Şekil 1. İki Faktörlü Duygu Teorisi'nin Şematik Diyagramı

Kaynak: Schachter & Singer, 1962

İki faktörlü duygu teorisine göre, karanlık, sessiz bir sokakta karşınızdaki size doğru koşarak gelen adam ile aydınlık, kalabalık bir sokakta karşınızdaki size doğru koşarak gelen adamın yarattığı uyarılma hissi farklıdır. İlk durumda korku hissi, ikinci durumda merak ve endişe hâkim olmaktadır. Uyarılma, uyarılmanın olduğu koşulların doğasından etkilenmekte ve duygu ortaya çıkmaktadır (Schachter & Singer, 1962).

b. Gelişimsel Duygu Teorisi

Çağdaş duygu bilimciler, duygular için uyarılma ve etiketlerden daha fazlası olduğuna inanmaktadır. Gelişimsel duygu teorileri hem duygusal gelişimin organizmada kısmen önceden programlandığını hem de duyguların tüm yönlerinin çocuğun geliştiği ortama duyarlı olduğunu kabul etmektedir. Yani teoriye göre duygular kültürel kurallar ve normlar da dahil olmak üzere sosyal deneyimi birleştirmektedir (Saarni ve ark., 2008). Gelişim teorilerinde hem doğa (genler) hem de yetiştirme (deneyim) duygusal yaşamın tam olarak farklılaşmasını ve detaylandırılmasını mümkün kılmaktadır (Pollak, 2013).

Duyguya ilişkin gelişimsel bir perspektif, duygu bileşenlerinin bebeklerde ve çocuklarda ilk gözlemlendiği yaşları belgelemekten ve yaşam döngüsü boyunca seyrini çizmekten daha fazlasını içermektedir. Tam bir gelişim teorisi, duygu bileşenlerinin ortaya çıkmasını diğer birçok beceri ve kapasitenin katkısıyla ilişkilendirmektedir (Saarni, 2008). Bunlar görsel ve motor sistemlerinin olgunlaşmasını ve kendini bir düşünce nesnesi olarak temsil edebilme yeteneği gibi karmaşık bilişsel kapasitelerin elde edilmesini içermektedir. Örneğin, korkunun davranışsal ve fizyolojik bileşenleri, tehlikeyi algılayacak ve kaçacak kadar olgun yanıt sistemlerine sahip bir vücuda sahip olmaktır. Benzer şekilde, kıskançlık ve gurur gibi duygular, öz değerlendirme ve sosyal karşılaştırma bilişsel yeteneklerini gerektirmektedir (Lewis, 2007).

Duygu ve duygusal yetkinliklerin (empati gibi) gelişimi için önemli bir sosyal bağlam, bir bakıcıyla ilk bağlanma ilişkisidir. Bir bebeğin bakıcısı ihtiyaçlarına yanıt verirse bebek, dünyanın güvenli bir yer olduğunu ve diğer insanların güvenebileceğini öğrenmektedir. İhtiyaçların karşılanması deneyimi, duygu için nöral ve endokrin sistemlerin sağlıklı gelişimini de desteklemektedir (Pollak, 2012).

c. Evrim Duygu Teorileri

Evrimsel duygu teorileri Charles Darwin'in yazılarına dayanmaktadır (Darwin, 1872; Darwin 1998). Darwin, doğal seçim yoluyla genel evrim teorisini desteklemek için duyguların yüz ifadesini kullanmak istemiştir. Hayvan türleri arasında ifadenin sürekliliğinin belgelenmesi nispeten daha kolay olduğu için, teorisini geliştirmede hayvanları kullanmıştır. Darwin hayvanların duygusal görüntülerini incelemiş ve insanlarla karşılaştırmıştır. Darwin, çalışmalarını sonucunda bir dizi yüz ifadesinin türler arası sürekliliği ve evrenselliği gösterdiğine inanmıştır. Bu tür gözlemlere dayanarak, duygusal ifadelerin nedenleri hakkında hipotezler geliştirmiştir (Niedenthal & Ric, 2017). Teori Şekil 2'de gösterilmiştir.



Şekil 2. Evrim Teorisinin Şematik Diyagramı

Kaynak: Schachter & Singer, 1962

Darwin yüz ifadelerinin problemle karşı karşıya kalınan ilk etapta ortaya çıkan jest ve yararlı alışkanlıklar olduğunu düşünmüştür. Örneğin, kontaminasyona maruz kalma (örneğin zehir, çürüyen yiyecekler) ölümle sonuçlanabilecek bir durumken, potansiyel olarak zararlı maddelerle daha fazla temastan korunmamızı sağlayan iğrenme duygusu ağızdan bir şeyleri çıkarmayı sağlamaktadır (Susskind ve ark., 2008).

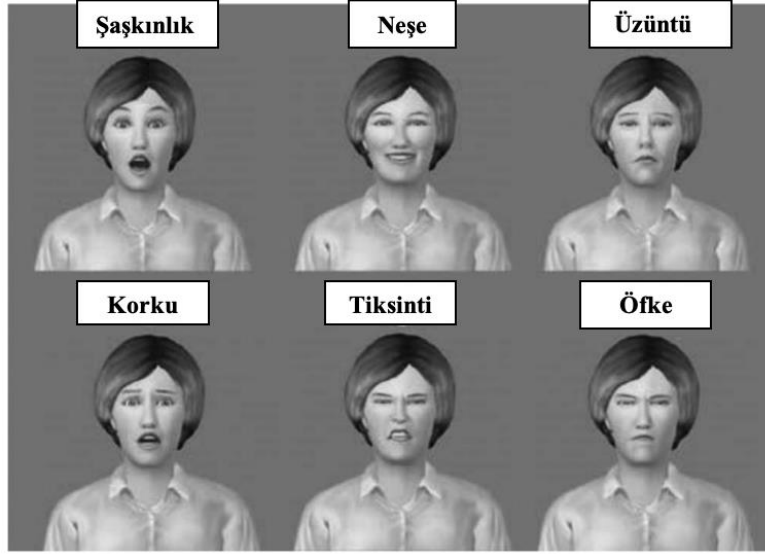
Darwin ayrıca yüz ifadesinin iletişimsel bir işlevinin olduğunu düşünmektedir. Başka bir kişinin bir olay karşısında, onunla başa çıkmak için hangi davranışta bulunacağı hakkında yüz ifadesinin önemli bilgiler sağladığını öne sürmüştür. Darwin, bu konu hakkında herhangi bir bilimsel çalışma yapmasa dahi birçok bilim adamı tarafından çalışılmıştır (Buck, 1983; Buck ve ark., 1992; Ekman & Friesen, 1971, 1975; Horstmann, 2003; Keltner & Kring, 1998; Marsh & Ambady, 2007). Bu görüşe göre, her temel duygu için ayrı bir program vardır.

Etki programları tüm insanlar için (doğuştan gelenler) benzer şekilde gelişmesine rağmen, bireysel deneyimler yoluyla kazanılan bilgileri içerecek şekilde değişebilmektedir (Ekman & Cordaro, 2011).

Silvan Tomkins (1962, 1963), Carroll Izard (1977, 2007) ve ardından Paul Ekman (1992), evrim teorisi perspektifinden duyguların biyolojik temelleri hakkında teori geliştirerek, temel duygular adını verdikleri küçük bir duygu kümesinin varlığını onaylamışlardır. Temel duygular, yeni uyarı sinyalleri karşısında hızlı, düşünmeden ve otomatik olarak ortaya çıkan doğuştan gelen sinirsel ve bedensel durumlardır (Izard, 1977; Johnson-Laird & Oatley, 1992; MacLean, 1993; Ekman, 1994; Panksepp, 1998; Buck, 1999; Ohman & Mineka, 2001). Duygu sınıflandırma konusunda evrensel bir görüş bulunmamaktadır. Bu bağlamda teorisyenler duyguları temel olarak sınıflandırmak için test edilebilir kriterler önermişlerdir. Bu kriterler:

- Yüz ifadesi ve daha fazlasını içeren evrensel ifadeler (örneğin, ses),
- Farklı fizyoloji,
- Primatlarda varlığı,
- Otomatik veya çevre tarafından değerlendirilmesi.

Temel duygu olarak nitelendirilen kavram, bedensel ve yüz ifadesiyle tanımlanabilmesi ve vücut fizyolojisinde belirli bir imzaya sahip olmasıdır. Bilim insanları bu özellikleri henüz tüm duygular için test edemedikleri için, temel duygu seti henüz oluşturulamamıştır. Ancak mevcut araştırmalara dayanarak korku, tiksinti, öfke, şaşkınlık, neşe ve üzüntü duyguları temel olarak adlandırılmaktadır (Ekman, Friesen & Ellsworth, 1972). Bu altı temel duygu, farklı yüz ifadeleriyle Şekil 3' de gösterilmiştir.



Şekil 3. Altı Temel Duygunun Yüz İfadeleri

Kaynak: Niedenthal & Ric, 2017

Aşağılama duygusu yedinci temel duygu olarak dahil edilmiştir (Ekman, 1999). 2000'li yılların başlarında, araştırmacılar utanç ve gurur gibi diğer duyguların evrensel ifadesi için kanıt bulmuşlardır (Tracy & Robins, 2004). Tiksinti ve şaşkınlığın duygu olarak sayılmaması gerektiğini savunan araştırmacılar bulunmaktadır (Ortony & Turner, 1990; Panksepp, 2007). Görüldüğü üzere bilim ilerledikçe sözde temel duyguların listesi değişmektedir.

3. Duygu Boyutları

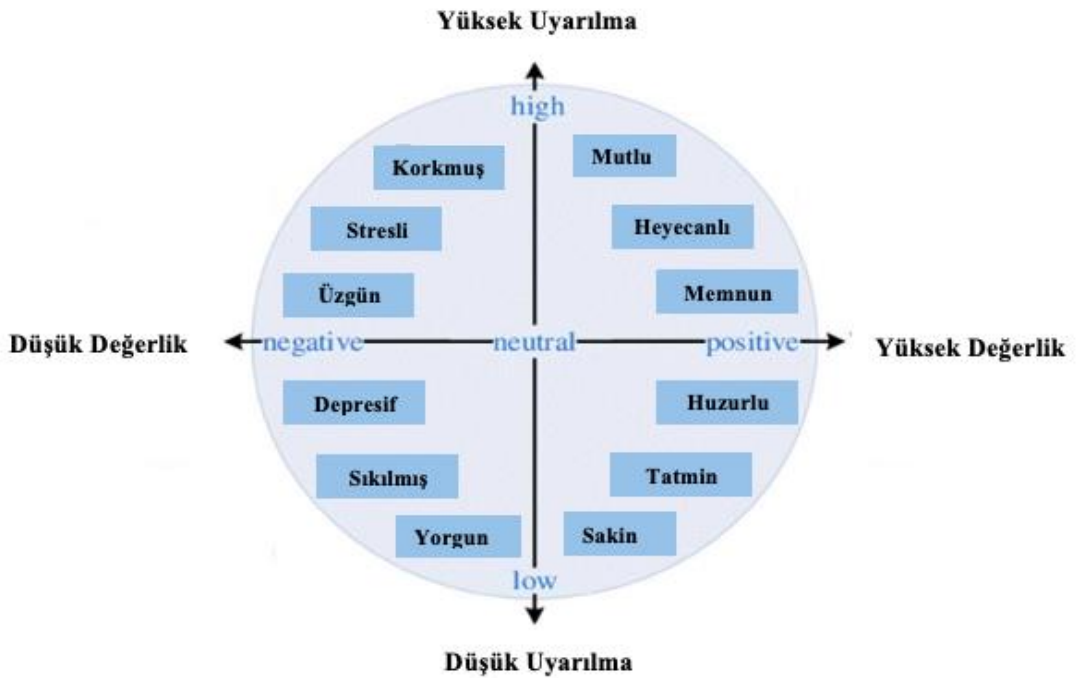
Duygu, beyin tarafından koordine edilen, organizmalar için dışsal ya da içsel olaylara karşı tepkiyi kolaylaştıran, kısa ve güçlü davranıştır. Duygu terimi, daha çok birkaç dakikadan birkaç saate kadar süren anlık ya da sınırlı bir süre içerisinde yaşanan durumlar için kullanılmaktadır (Davidson, Scherer & Goldsmith, 2002).

Tarihsel olarak, duyguları incelemeye yönelik ayırık ve boyutsal olmak üzere iki ana yaklaşım vardır. Ayırık yaklaşım, duygusal durumları birbirinden ayıran özelliklere odaklanırken (Ekman, 1992) boyutsal yaklaşım duyguları, duygusal durumların altında yatan önceden belirlenmiş özelliklere göre tanımlamaktadır (Russell, 1980).

Duyguyla ilgili pek çok boyut olmasına rağmen, öznel duygu durumlarının en yaygın atıfta bulunulan dört boyutu aktivasyon, değerlik, güç ve yoğunluktur

(Smith & Ellsworth, 1985). Aktivasyon, alçaktan yükseğe değişen enerji hissini ifade etmektedir (Gosselin ve ark., 2007, Sammler ve ark., 2007). Değerlik, bir olayın, nesnenin veya durumun içsel değerlendirmesiyle ilgilidir ve olumludan olumsuzu doğru değişmektedir (Schubert, 1999; Dalla Bella ve ark., 2001). Güç, bir bireyin evrensel olarak belirli bir duygu ile özdeşleştiği güçlülük veya güçsüzlük derecesini tanımlamak için kullanılan bir boyuttur. Olumlu duygular neredeyse her zaman yüksek düzeyde kontrol veya hakimiyet oluşturmaktadır. Öfkenin yüksek bir güç derecesine sahip olması ve korkunun düşük bir güç derecesine sahip olması sebebiyle güç korku ve öfke gibi olumsuz duyguları ayırt etmede kullanılmaktadır. Son olarak, duygusal yoğunluk, hissedilen duygunun derecesini ölçmek için kullanılmaktadır (Jiam, Caldwell & Deroche, 2017).

Değerlik-uyarılma modeli, duyguya değerlik ve uyarılmanın iki ayrılabilir boyutu olarak yaklaşmaktadır (Şekil 4) (Russell, 1980). Bu iki boyut yaygın olarak ses duygu çalışmalarında kullanılmaktadır ve duygunun psikofizyolojik bileşenlerinin çoğunu temsil etmektedir. Bu nedenle bazı araştırmacılar duygu teorisini yalnızca değerlik ve uyarılma olmak üzere iki bileşene indirgemektedir (Bachorowski, 1999).



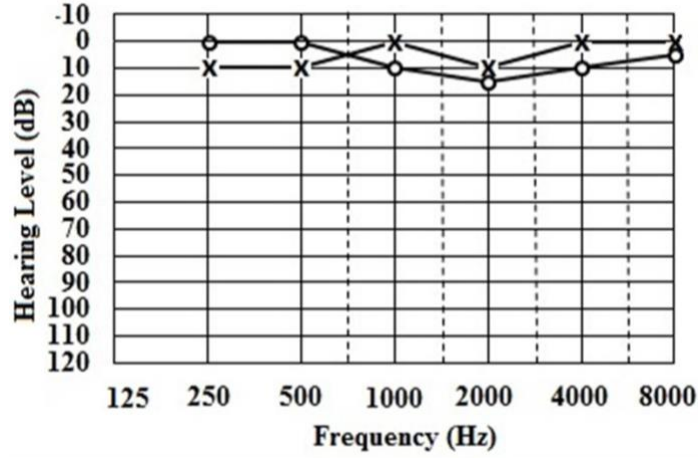
Şekil 4. Değerlik ve Uyarılma Modeli.

B. İşitme Kaybı

İşitme kaybı, en yaygın sensörinöral eksikliklerden biridir. Sosyal akademik, bilişsel, duygusal gelişim olmak üzere genel yaşam kalitesini büyük ölçüde etkilemektedir (Cunningham ve ark., 2003). 34 milyonu çocuk olmak üzere dünya çapında yaklaşık 466 milyon insanı etkilemektedir (WHO, 2020). 2030 yılına kadar bu sayının yaklaşık 630 milyon kişiye ulaşması, 2050 yılına kadar ise 900 milyondan fazla insanda işitme kaybı olacağı öngörülmektedir (WHO, 2020).

İşitme kaybı, tek taraflı veya çift taraflı olarak sesi kısmen veya tamamen işitememektir. İşitme kaybına genetik, yaş, gürültüye maruz kalma, hastalık, kimyasallar ve fiziksel travma gibi birçok faktör neden olabilmektedir. İşitme kaybı her yaş grubunu etkileyerek, çocuklarda konuşma ve öğrenmeyi geciktirmekte, yetişkinlerde sosyal ve mesleki sorunlara neden olabilmektedir. İşitme kaybının yaygınlığı yüksek sesli müziğe maruz kalma ile ilişkili olarak ergenlerde ve genç yetişkinlerde artmaktadır. Yaşa bağlı işitme kaybı ise özellikle arka plan gürültüsü varlığında bulunulan ortamlarda sesleri ayırt etme yeteneğini tehlikeye atmaktadır. Erken teşhis ve müdahaleler ile, işitme kayıplı bireylerin çoğu tedavi edilmektedir (Alshuaib, Al-Kandari & Hasan, 2015).

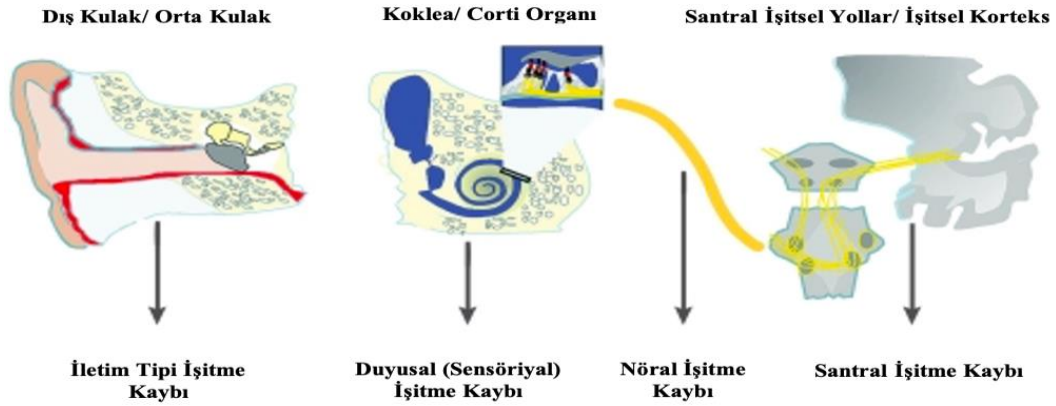
Dış kulak ortamdan sesi toplayarak orta kulağa iletmede, orta kulak ses titreşimlerinin iç kulağa yükseltilecek şekilde transfer etmede, koklea gelen mekanik enerjiyi elektrokimyasal enerjiye dönüştürmede görev almaktadır. (Belgin & Çalışkan, 2004). Periferik işitme organlarından daha üst merkezlere çıkıldığında, sesin lokalizasyon ve lateralizasyonu, işitsel ayırt etme, işitsel patern tanıma, temporal maskeleyme, temporal çözünürlük, temporal birleştirme, temporal sıralama gibi daha kompleks analizler yapılmaktadır.



Şekil 5. Normal İşitmeye Ait Örnek Odyogram

Kaynak: Alshuaib, Al-Kandari & Hasan, 2015

İşitme sistemi genel olarak periferik ve santral işitme sistemi olarak ikiye ayrılmaktadır. Periferik işitme kaybı, farklı tip ve derecelerde ortaya çıkarak işitsel bilgilerin beyindeki işitme merkezlerine iletilmesine engel olmaktadır. Periferik işitme sistemindeki hasarın yeri, işitme kaybının iletim, sensörinöral veya mikst olarak sınıflandırılmasını sağlamaktadır (Hall, 2014; Kramer & Brown, 2019). İşitme sisteminin periferik ve santral komponentleri dahil olmak üzere lezyon yerine göre sınıflandırılması Şekil 6'da gösterilmiştir.

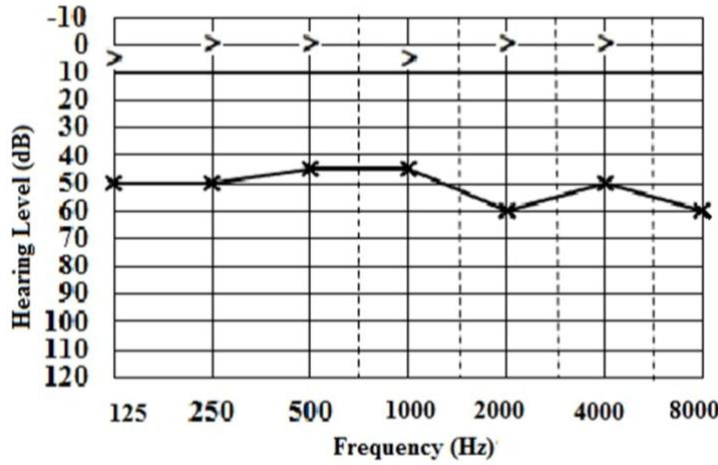


Şekil 6. Periferik ve Santral İşitme Sistemi Lezyon Yerine Göre İşitme Kaybı Sınıflandırılması

Kaynak: Zahnert, 2011

1. İletim Tipi İşitme Kaybı

İletim tipi işitme kaybı, normal bir iç kulak fonksiyonuna sahipken genellikle dış ve / veya orta kulakta bulunan işlev bozukluğu ile ilişkilidir. İletim tipi işitme kaybı hem çocuklarda hem de yetişkinlerde işitme bozukluğunun en yaygın nedenidir. İnsidansı çocuklarda önemli ölçüde yüksektir. İletim tipi işitme kaybında, iç kulak normal olarak işlev görür, ancak ses titreşimlerinin kulak kanalından, kulak zarından veya orta kulaktaki kemiklerden geçmesi engellenir. Hava yolu ile iletilen sinyallere kıyasla kemik yolu ile iletilen sinyallerde daha iyi işitme eşiklerinin olmasıyla karakterize edilen bir işitme kaybı türüdür. İletim tipi işitme kaybı olan hastalar, vibratörle kafatasına sunulan kemikle iletilen sesleri hava yolu ile sunulan seslerden daha iyi eşiklerle algılamaktadır. İletim tipi işitme kaybı genellikle hafif ila orta dereceli ve çoğu durumda tek taraflı olarak görülmektedir. İletim tipi işitme kaybının çoğu, nispeten küçük tıbbi veya cerrahi tedavilerle düzeltilebilmektedir (Afolabi ve ark., 2012).



Şekil 7. İletim Tipi İşitme Kaybı Örnek Odyogram

Kaynak: Alshuaib, Al-Kandari & Hasan, 2015

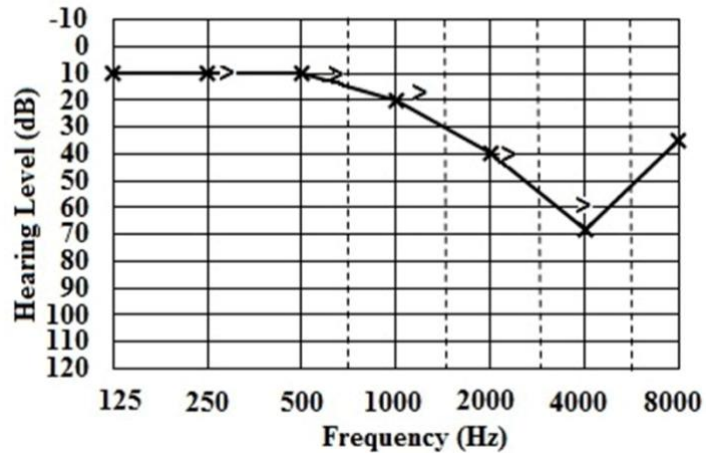
2. Sensörinöral Tip İşitme Kaybı

Sensörinöral tip işitme kaybı, kokleada, 8. kranial sinirde veya beyinde oluşan hasarın sonucu olarak ortaya çıkan bir işitme kaybıdır. Sensörinöral tip işitme kaybında, hava yolu ve kemik yolu iletim eşikleri benzerlik göstermektedir.

Corti organı, perilenfatik titreşimleri nöral sinyallere dönüştüren işlevsel bir birimdir. Baziler membranın titreşimi, stereosilyaların mekanik olarak sapmasına

ve bunun sonucunda depolarizasyona neden olmaktadır. Dış tüylü hücrelerin hücre duvarlarındaki motor proteinler mekano-elektriksel transdüksiyon (MET) işleminin, hücreler boyunca bir reseptör akımı oluşturmasını sağlamaktadır (Flock & Orman, 1983; Zenner, 1997). Duyusal (Sensöriyal) işitme kaybı, tanımı gereği Corti organının işlev bozukluğuna bağlıdır. Sorun genellikle dış tüylü hücrelerde bulunur ve bu gibi durumlarda kemik iletim eşiğinde 50 dB' e kadar düşme, yüksek sese tahammülsüzlük, frekans distorsiyonu ve buna bağlı olarak konuşmayı anlama yeteneğinde bozulmalar olarak kendini göstermektedir (Zahnert, 2011).

Nöral işitme kaybı, koklear siniri etkileyen hastalıklara bağlı her tür işitme bozukluğunu içermektedir. Nöral işitme kaybının yaygın nedenleri menenjiyom, akustik nörinom, kordoma, kondrosarkom gibi tümörler, ve petröz kemiğin kolesterol granülomu, kolesteatom gibi enflamatuar yıkımıdır (Murugasu, Yong & Yoon, 2004; Zahnert, 2011). Akustik nörinom (vestibüler schwannom), yılda 100.000 kişide 1,74 insidansla nöral işitme kaybının en yaygın nedenidir (Tos ve ark., 2004).

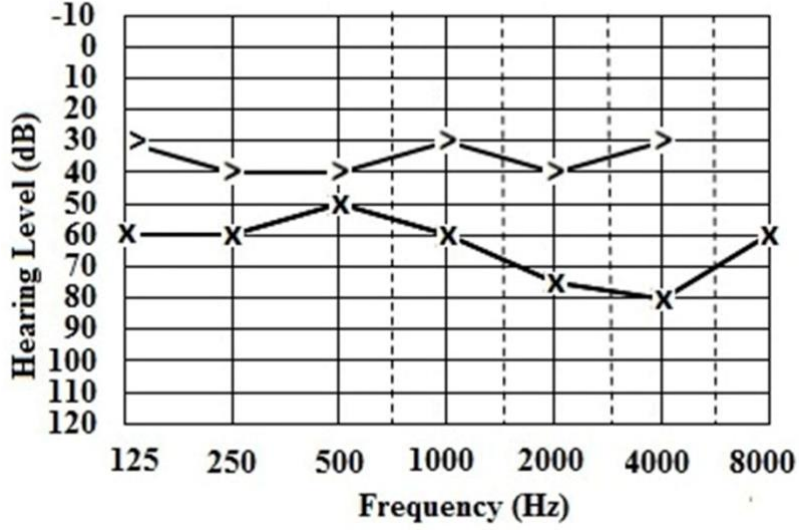


Şekil 8. Sensörinöral Tip İşitme Kaybı Örnek Odyogram

Kaynak: Alshuaib, Al-Kandari & Hasan, 2015

3. Mikst Tip İşitme Kaybı

Mikst tip işitme kaybında, aynı tarafta iletim ve sensörinöral işitme kaybı kombinasyonu vardır. Kemik yolu işitme eşiğinin düşmesi ile beraber, hava yolu ve kemik yolu işitme eşiği arasındaki farkın 10 dB'den fazla olması mikst tip işitme kaybını düşündürmektedir.



Şekil 9. Mikst Tip İşitme Kaybı Örnek Odyogram

Kaynak: Alshuaib, Al-Kandari & Hasan, 2015

C. Koklear İmplant

Koklear implantlar evrensel olarak ileri ve çok ileri derecede sensörinöral işitme kayıplı olan çocuk ve yetişkinlere işitsel girdi sağlamak amacıyla tasarlanan ve cerrahi olarak yerleştirilen biyomedikal bir cihazdır (Dillon, 2012; Wolfe, 2019).

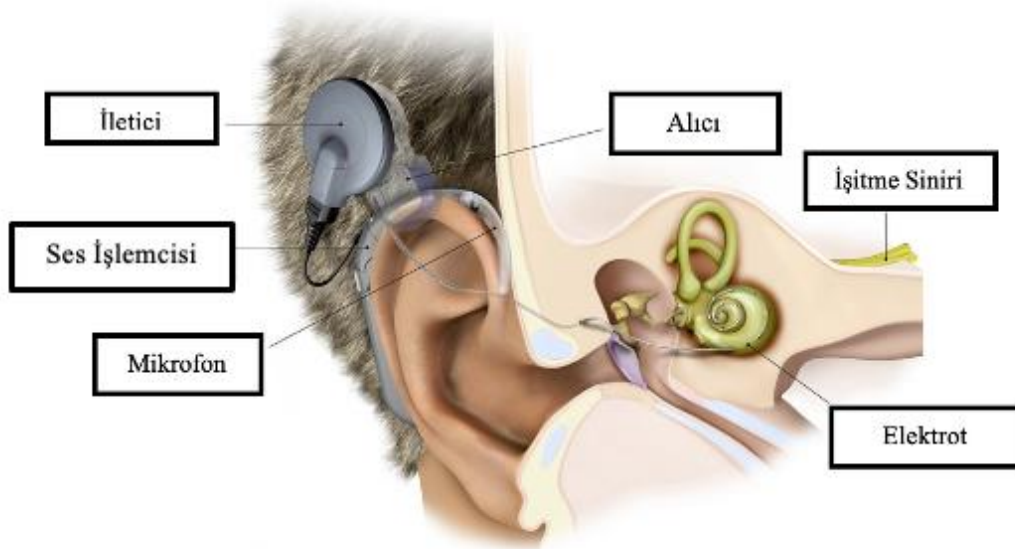
Günümüzde koklear implantlar, ileri/ çok ileri derecede işitme kaybı veya total işitme kayıplı bireylerde konuşma bilgilerini algılamada önemli düzeyde başarı elde etmektedir. Dünya çapında 300.000'den fazla işitme kayıplı birey koklear implant kullanmaktadır (NICDD, 2013).

Koklear implantlar akustik basınç dalgasını koklea içine yerleştirilen elektrotlara ileterek doğrudan işitsel nöronları uyaran ve seslerin algılanmasını sağlayan işitsel protezlerdir (Huawei & Chai, 2019). Akustik siniri doğrudan uyararak kulağın normal ses iletim mekanizmasını atlarlar. Harici bir mikrofon sesi yakalar, işlemciye iletir ve ardından sesi, deri altındaki bir alıcıyla kablosuz olarak iletişim kuran vericiye gönderir. Alıcı daha sonra sesi, koklea içindeki işitsel sinir liflerine cerrahi olarak bitişik olarak yerleştirilen bir elektrot dizisi tarafından taşınan elektriksel uyarılara dönüştürür. Sinyal, işitme kayıplı bireylerde merkezi işitme sistemi içinde işitme korteksine iletilerek hastaların işitmesini sağlar (Yawn ve ark., 2015).

Koklear implantların işitme cihazlarından farkı, sesin sadece amplifikasyonunu sağlamaktan ziyade iç kulağın hasarlı veya çalışmayan kısımlarını telafi etmesidir (Niparko, 2004). Koklear implantların başarılı bir şekilde uygulanmasının temel koşulu, işlevsel işitme siniri, patolojik olmayan merkezi işitme yolları ve elektrotların yerleştirilmesi için anatomik olarak gelişmiş bir kokleadır. Bu nedenle, ciddi koklear malformasyon, işitme siniri eksikliği ve akut süperatif otitis, koklear implantasyon için kontrendikasyondur (Eğilmez & Kalcıoğlu, 2015).

Koklear implantların temel bileşenleri:

- (1) sesi ortamdan alan bir veya daha fazla mikrofon,
- (2) akustik sinyali işleyerek bir dizi elektrik sinyaline dönüştüren sinyal işlemcisi,
- (3) elektrik sinyallerini deri üzerinden implante edilmiş elektrot dizisine ileten bir iletim sistemi,
- (4) koklea içine gömülü bir elektrot veya elektrot dizisi.



Şekil 10. Koklear İmplant Sisteminin Bileşenleri

Kaynak: (<http://hearingsolutions7.com/cochlear-implant/>) (Erişim Tarihi: 22.08.2020)

1. Koklear İmplant Tarihi

İşitme sisteminin elektriksel uyarımı ilk olarak Alessandro Volta tarafından kulak kanalına yerleştirdiği iki metal sondanın kutupları pile bağlanarak sağlanmış ve elektriksel uyarımın işitsel hissi uyandırabildiğinin ilk göstergesi olmuştur (Volta, 1800). André Djourno ve Charles Eyriès (1957), insan işitme sinirini elektrikle doğrudan uyararak için ilk deneyi gerçekleştirmiştir. Total işitme kaybı olup ameliyat olan iki hasta net işitsel algı bildirmiştir. Dr. William House ve John Doyle (1961) sağır olan iki hastayı altın yalıtımlı elektrotlar ile implante etmiştir. Bu hastalar, elektriksel bir uyarı ile perde değişikliklerini tespit edebildiklerini bildirmişlerdir. Simmons ve ark. (1964), bir hastanın vestibül ve doğrudan koklea modiolusuna bir elektrot implante etmişlerdir. Hasta, süre ve tonalitedeki değişiklikleri tespit edebilmiştir.

İmplant teknolojisi 1970'lerden 1990'lara kadar belirgin bir şekilde ilerlemiş ve 1972'den 1980'lerin ortalarına kadar implantasyon ameliyatlarının sayısı 1000'i geçmiştir (Huawei & Chai, 2019). İlk tek kanallı koklear implant 1972'de implante edilmiştir. 1980'lerin ortalarında, erken dönemde implante edilen tek kanallı koklear implant konuşma algısını önemli ölçüde artırmıştır. Ardından farklı bir hastaya takılan koklear implant, fonksiyonel bir kokleada iç tüylü hücrelerin akustik uyarımı ile üretilen normal sinirsel aktivite seviyesine ulaşamamıştır.

Bilger ve ark. (1975), Amerikan Ulusal Sağlık Enstitüleri (NIH) tarafından dünyanın ilk tek elektrotlu koklear implant alıcıları grubunda, konuşma performansını değerlendirmek üzere görevlendirilmiştir. Araştırma raporu, tek elektrotlu koklear implantın çevresel sesi tanımlamak için yararlı bilgiler sağlayabilmesine rağmen, açık set konuşma tanıma sağlayamadığını göstermiştir. 1984 yılında 3M / House tek elektrotlu koklear implant, ticari kullanım için Amerikan Gıda ve İlaç Dairesi (FDA) tarafından onaylanmıştır (Fretz & Fravel, 1985). İlk çok kanallı koklear implant sistemi aynı yıl Cochlear Şirketi tarafından geliştirilmiştir. Büyük ölçekli klinik araştırmalar, çok elektrotlu cihazın gerçekten de tek elektrotlu cihaza kıyasla daha üstün ses algılaması sağladığı sonucuna varmıştır (Loizou, 1999). Çok elektrotlu koklear implantın piyasaya sürülmesi, tek elektrotlu koklear implantın ticari olarak tercih edilmemesine yol açmıştır. Geçmişten günümüze, birçok araştırmacı koklear implant sisteminde kullanılan

tasarım ve konuşma işleme stratejilerini iyileştirmeye çalışmıştır. 1990'larda, konuşma işleme teknolojisindeki ve minyatürleştirmedeki büyük gelişmeler, koklear implantların hastaların yaşam kalitesini önemli ölçüde iyileştirdiğini göstermiş ve yaygın olarak tıpta kullanılmaya başlanmıştır. Son 10 yılda mikroelektronik alanındaki gelişmeler ve sinyal işleme tekniklerindeki ilerlemeler, koklear implant sistemlerinde önemli iyileştirmelere yol açmıştır. Önemli ilerleme kaydedilmiş olmasına rağmen, koklear implantlar başka zorluklar da ortaya koymaktadır. Koklear implant aracılığıyla sessiz bir ortamda konuşma tanıma plato seviyesine ulaşmıştır. Araştırmacılar günümüzde yeni koklear implant modellerini ve hasta kullanımını değerlendirirken, mevcut implant kullanıcıları için zorlu görevler olan gürültüde konuşma tanıma, müzik algısı ve tonal dili anlamayı geliştirmeye odaklanmıştır (Huawei & Chai, 2019).

2. Koklear İmplant Kullanıcılarında Dil ve Konuşma Gelişimi

Bebek ve çocuklarda, işitsel öğrenmenin gerçekleşmesi için beynin sese ayrıntılı erişmesi gerekmektedir. İşitme kaybıyla ilgili başlıca sorun, işitsel yeteneklerin hiç olmaması veya azalması ile beyindeki işitme merkezlerine ulaşan ses ve bilginin engellenmesidir (Musiek, 2009; Fitzpatrick & Doucet, 2013; Kral & Lenarz, 2015). Beyindeki işitme merkezlerinin nöral gelişimi ve organizasyonu, duyuşal girdi ve kapsamlı işitme deneyimini gerektirmektedir (Sharma ve ark., 2004; Sharma, Dorman & Kral, 2005; Kilgard ve ark., 2007; Estabrooks, 2012; Chermak & Musiek, 2014; Kral ve ark., 2016;).

Koklear implant kullanan bireylerin, gerekli işitsel rehabilitasyon eğitimini almaları durumunda, ameliyat öncesi işitme cihazlı performanslarına kıyasla konuşma algısı ve konuşma üretiminin tüm alanlarında önemli gelişmeler gösterdikleri ifade edilmiştir (Zwolan ve ark., 2004; Waltzman, 2005; Bradham & Houston, 2015; Lovett ve ark., 2015; Geers ve ark., 2017; Ching ve ark., 2018; Percy-Smith, 2018).

Dil edinme yeteneği uygun akustik uyarım koşullarında öğrenilmektedir. Bu yetenek, sinir sisteminin belirli girdileri işleme yeterliliğini geliştirmek için dış akustik uyarıma ihtiyaç duyduğu hassas dönemlerde gerçekleşmektedir (Gatliffe, 2010). Konuşmayı öğrenmek için en hassas dönemin yaşamın ilk 5 yılı olduğu ve bu erken gelişim döneminde işitme kaybı olduğu takdirde dil gelişiminin geri

dönüşü olmayacak bir şekilde etkileneceği belgelenmiştir (Dale, 1976; Crystal, 1987; Ruben, 1997). Erken yaşta koklear implantın sağlanması, kritik dönemde çocuğa sunulan işitsel bilgi miktarını en üst düzeye çıkarıp, işitsel yoksunluğun çocuğun konuşma ve dil becerilerinin gelişimi üzerindeki etkilerini azaltmaktadır (Zwolan ve ark., 2004).

Genel olarak, koklear implantasyondan en çok yararlanan bireyler 12 aylık veya daha erken implante edilmiş bebeklerdir. Ek bir anahtar faktör, çocuğun implantı kullandığı sürenin uzunluğudur. Cihazlarını 3 yıldan fazla süredir her gün kullanan çocuklar dinleme becerilerinde önemli gelişme göstermektedirler (Cole & Flexer, 2019). Genel olarak, implantasyonu takiben iyi olmayan sonuçlar, ek engeller, geç implantasyon, zayıf cihaz programlama, sınırlı kullanım süresi, koklear sinir aplazisi, bakteriyel menenjit ve ciddi koklear malformasyonlar ile ilişkili olma eğilimindedir (Wolfe, 2018). Bununla birlikte, koklear implantların nihai etkinliği, büyük ölçüde, implantasyonu izleyen işitsel rehabilitasyon eğitiminin türü ve derecesi ile belirlenmektedir (Zwolan ve ark., 2004; Waltzman, 2005; Tobey, 2010; Bradham & Houston, 2015; Geers ve ark., 2017; Smith, Wolfe & Dettman, 2018). İmplant edilen bebek ve çocukların, sözlü iletişimi vurgulayan bir ortamda yetiştirilmesi, dinleme becerilerini sistematik olarak geliştirmesi gerekmektedir.

Araştırmalar, önemli bir değişkenin bebeğin veya çocuğun implantasyondaki yaşı olduğunu belirtmektedir (Dettman ve ark., 2016; Ching ve ark., 2018). Beynin nöroplastisite kapasitesinin en yüksek olduğu küçük yaşlarda implante edilen çocuklar, uygun işitsel temelli müdahale alırlarsa en yüksek işitsel performansı gösterebilmektedir. İmplant kullanan bireylere, koklear implant yoluyla algıladıkları işitsel bilgiyi yorumlama ve bunlardan anlam çıkarma öğretilmektedir. Böylece bireyler, koklear implanttan en iyi şekilde yararlanabilmektedir (Cole & Flexer, 2019).

3. Tek Taraflı ve Bilateral Koklear İmplant

Başlangıçta koklear implant sadece tek taraflı olarak uygulanmaktaydı ve bunun üzerine hastaların tek taraflı veya çift taraflı implantlarla daha iyi performans gösterip göstermeyeceği sorgulanmaya başlanmıştır. Bilateral implantasyonun özellikle gürültüde konuşma algısında önemli derecede

iyileşmeler gösterdiği ortaya konmuştur (Venail ve ark., 2008). Ek olarak, bilateral implant hastaları, tek bir implantla performanslarına kıyasla, lokalizasyonda önemli bir gelişme göstermektedir (Dunn ve ark., 2012). Cerrahi zamanlamayla ilgili olarak, bilateral implantasyon sırayla veya aynı anda yerleştirilebilir. Bununla birlikte, eş zamanlı bilateral implantasyon yapılan çocuklar, sırayla implante edilen çocuklara kıyasla konuşma tanımada daha iyi performans göstermektedir (Lammers ve ark., 2014).

İki taraflı işitme cihazlarında olduğu gibi, iki taraflı koklear implantasyon zor dinleme durumlarına yardımcı olarak dil, konuşma, eğitim ve psikososyal gelişimi optimize etmektedir. Çalışmalar, bilateral koklear implantasyonun bireylere önemli faydalar sağladığını göstermiştir (Sarant ve ark., 2014; Dorman & Gifford, 2017; Grimes, 2017). Lovett ve ark., (2015), her iki kulakta saf ses ortalamasının (SSO) cihazsız 80 dB HL veya daha kötü olan çocukların, bilateral koklear implant adayları olarak kabul edilmesi gerektiğini savunmaktadır. Ekip iş birliği, ikinci bir implant için adaylığı belirlemede ve her çocuk için fonksiyonel sonuçları en üst düzeye çıkarmak için uygun takibi düzenlemede kritik öneme sahiptir. İkinci koklear implantın işitme kayıplı birey için uygunluğu, çocuğun yaşı, ilk koklear implant veya işitme cihazı ile işitsel beceri gelişimi, ilk implantın ne zaman yapıldığı gibi birçok bireysel faktöre bağlıdır (Cole & Flexer, 2019).

İmplantasyondan önce çocuğun işitme yetenekleri (işitme cihazı veya ilk koklear implant ile) ne kadar iyi olursa, çocuk ikinci implantın yeni elektroakustik girdisini kullanmayı o kadar çabuk öğrenmektedir. İşitsel girdiye alışkın olan merkezi korteks, işitsel olarak kodlanmış yeni girdinin kaynağını daha hızlı tanıyarak anlamlandırmaktadır. Çocuğun ikinci implantasyonda yaşı ne kadar büyükse, rehabilitasyon süreci güçleşerek adaptasyon zaman almaktadır. Birinci ve ikinci implant sırasında çocuk ne kadar küçükse, ikinci cihazın adaptasyonu için gereken süre o kadar kısa olur (Litovsky, 2010).

4. Koklear İmplant Kullanan Bireylerde Ses Duygu Tanıma ve İfade Etme

Konuşma, koklear implantın zamansal kodlaması yoluyla kullanıcılara iletilmesine rağmen, ses perde bilgisi iyi bir şekilde sunulmamaktadır (Faulkner ve ark., 2000; Geurts & Wouters, 2001). Zamansal açıdan bakıldığında, F0 / F2,

F0 / F1 / F2 ve SPEAK (Spectral Peak) gibi konuşma kodlama stratejileri, 250 Hz' in üzerindeki frekanslarda yalnızca darbe dizisi hızlarının iletilmesine izin vermektedir. Koklear implant kullanıcıları, ACE (Advanced Combination Encoder), CIS (Continuous Interleaved Sampling) ve HiResolution gibi yüksek uyarım hızı kullanılan konuşma kodlama stratejilerinde dahi, uyarım hızına bağlı olarak 300 Hz'nin üzerindeki ses perdesi değişikliklerini tanımak için geçici zarf ipuçlarını kullanamazlar. Bu, insan işitme sisteminin zamansal ipuçlarını kullanmadaki genel sınırlamalarından kaynaklanmaktadır (Peng, Tomblin & Turner, 2008). Bu durumda koklear implant kullanıcıları, özellikle 300 Hz' in üzerindeki geçici zarf bilgilerine dayalı olarak frekans farklılıklarını ayırt etme becerilerinde oldukça yetersizdir. Bu sebeple çocuk ve kadın konuşmacılar gibi belirli konuşmacıların ses perdesi bilgilerine yalnızca sınırlı erişime sahiptir, çünkü bu kişilerin doğal ses perdesi aralıkları genellikle 300 Hz' i aşmaktadır (Shannon, 1983).

Koklear implant kullanıcıları, bazılar membranın tonotopik organizasyonu tarafından sağlanan yer ipuçlarından tam olarak faydalanamamaktadır. Yani, spektral bilgiyi iletmek için kullanılan elektrotların sayısı, özellikle F0 aralığında sınırlıdır. Spektral bilgiler F0 algılarını desteklese dahi koklear implant kullanan bireylerin, yetersiz spektral çözünürlük nedeniyle ses perde bilgisi sınırlanmaktadır (Green, Faulkner & Rosen, 2002; Green, Faulkner & Rosen, 2004). Osberger ve ark. (1991), iki yıldır koklear implant kullanan okul çağındaki çocukların tonlama ve diğer prozodik konuşma bileşenlerinin üretiminde veya algılamasında iyi performans sergilemediğini göstermiştir. Peng ve ark. (2007), pediatrik koklear implant kullanıcılarının, 7 yıllık koklear implant kullanımında dahi tutarlı bir şekilde uygun tonlamaya sahip sorular üretmediklerini bildirmiştir.

İşitme kayıplı çocuklarla yapılan araştırmalar, normal işiten akranlarına kıyasla ses duygu tanımada iyi performans göstermediklerini ortaya koymaktadır (Dyck ve ark., 2004; Nakata, Trehub & Kanda, 2012; Chatterjee ve ark., 2015). Wiefferink ve ark (2013), işitme kaybına bağlı olarak ifade edici dilin gelişmemesinin, duygularla ilgili kavramsal bilginin azalmasına yol açabileceğini ve bunun da duygu tanıma üzerinde olumsuz bir etkiye yol açabileceğini öne sürmüştür.

Hopyan-Misakyan ve ark. (2009), koklear implant kullanan 7 yařındaki çocukların normal işiten akranlarının yanı sıra görsel duygu tanımada iyi performans sergilediklerini ancak ses duygu tanımada önemli ölçüde daha düşük puan aldıklarını göstermiştir. Görsel duygu tanıma genel olarak, işitme kayıplı çocuklarda ses duygu tanımadan daha hızlı gelişmektedir, bu da görsel duygu ipuçlarının algısal olarak daha belirgin olduğunu veya ses duygu ipuçlarından daha kolay kategorize edildiğini göstermektedir. İşitme kayıplı çocuklar için, spektro- temporal olarak bozulmuş işitsel girdinin telafisi olarak görsel duygu ipuçlarına daha fazla güvenmek, günlük yaşamda duygu tanıma genellikle çok modlu olduğu için etkili bir strateji olabilmektedir (Nagels ve ark., 2020).

Duygu tanıma arařtırmalarının çoğunun odak noktası yüz ifadeleri üzerine olmasına rağmen, arařtırmalar çocukların ve yetişkinlerin diđer ipuçlarını anlama kapasitelerini incelemiş ve konuşma prozodisinin diđer alanlardaki ifadeleri anlamadan daha sonra gelişebileceğini bulmuştur. Okul öncesi çocuklar yüz ifadesi ve postür ipuçlarından mutluluk, üzüntü, öfke ve korkuyu fark edebilmektedir. Nelson ve Russell (2011) tarafından yapılan bir çalışmada üç-beş yařındaki çocuklar ses olarak sunulan üzüntü duygusu hariç, mutluluk, öfke ve korku duygularını tanımada yetersiz kalmıştır. Yalnızca sesle sunulan duygular yetişkinler tarafından iyi tanınırken, üç-beş yařındaki çocukların sesle sunulan duyguları tanınması zayıftır ve çoğu çocuk doğru etiketi seçmekte başarısız olmaktadır. (Nelson & Russell, 2011; Vidas, Dingle & Nelson, 2018). Benzer şekilde yüz ifadesinden duyguların tanınması 11 yařına kadar yetişkinlik düzeyine ulaşırken, ses duygu tanıma bu yařtan sonra gelişmeye devam etmektedir (Chronaki ve ark., 2015; Vidas, Dingle & Nelson, 2018).

D. Geçmişten Günümüze Ses Duygu Analizi

İnsan konuşması sadece dil bilgisini değil aynı zamanda duygusal mesajları da iletmektedir. Duygusal ifadenin sözel iletişimdeki önemi ve dinleyici üzerindeki güçlü etkisi tarih boyunca vurgulanmıştır. Duygusal olarak ifade edici konuşmanın stratejik kullanımına yönelik somut önerilerle birlikte konunun sistematik incelemelerine Batı felsefesinde Hitabet Sanatı ile ilgili Yunan ve Roma klavuzlarında (örneğin, Aristoteles, Cicero, Quintilian) rastlanmıştır. (Kennedy, 1972).

Yüz ve seste yansıtılan duygu ifadesine olan ilgi, Spencer, Bell ve özellikle Darwin'in katkılarıyla, modern evrimsel biyolojinin ortaya çıkmasıyla 19. yüzyılda ateşlenmiştir (Darwin,1872). Vokal duyguya daha teori odaklı bir yaklaşım, insan duygularının muhtemelen hayvan duygularından evrimleştiğini ve muhtemelen hayatta kalma değerleri için seçim baskısı ile şekillenmiş olduklarını öne süren Darwin (1871, 1872) kadar eskilere dayanmaktadır. Morton (1977), birçok kuş ve memelinin seslerinde tutarlı bir eğilim gözlemlemiştir. Örneğin düşmanca davranıldığında sert, nispeten daha düşük frekanslı seslerin çıkması, itaatkâr, yatıştırıcı olduğunda daha yüksek frekanslı, daha saf ton benzeri seslerin çıkmasıdır. Morton, bu modeli, vücut büyüklüğü ile ilgili evrimin varsayımsal seçim baskısına dayanan “motivasyonel-yapısal kurallar” olarak adlandırarak açıklamaktadır. Daha büyük bir hayvanın daha küçük bir hayvana karşı fiziksel bir çatışma kazanması muhtemel olduğundan, hayvanların rakiplerini korkutmak için mümkün olduğunca büyük görünme stratejisini benimsemeleri için güçlü bir seçim baskısı vardır (Morton, 1977). Böylece saldırgan saçlarını veya tüylerini diker, kuyruğunu veya kuyruk tüylerini kaldırır ve omzunun daha büyük görünmesini sağlar (Ohala, 1984; Hauser, 1993, 1997; Chuenwattanapranithi ve ark., 2008). Hatta bazı hayvanlar kamburluk veya kalın yüz kılı gibi kalıcı boyutta işaretler geliştirmişlerdir (Ohala, 1984). Ancak Morton (1977) tarafından öne sürüldüğü gibi hayvanlar sadece vücut boyutlarını büyütme için görsel sinyaller kullanmakla kalmaz, aynı zamanda aynı etkiyi elde etmek için F0 ve ses kalitesi gibi akustik sinyalleri de kullanmaktadırlar. Saldırganlığın hayvan sesine yansıttığı kaba nitelik, Darwin (1872) zamanında kaydedilmiştir. Aksine, itaatkâr bir hayvan, tehdit olmadığı durumda uysallığı ifade etmek için tam tersini yapmaktadır. Kulakları, kuyruğu, saçları veya tüyleri düzleştirip, genellikle yüksek frekanslı ses üretmektedir (Morton, 1977; Ohal, 1984; Hauser, 1997). Tehdit olmadığını göstermenin yanı sıra, yüksek frekanslı ve saf tonlu sesler, ebeveynlerin bakım ve koruma içgüdüsünü ortaya koyan bebeklerin seslerini de yansıtmaktadır (Morton, 1977).

Duygunun ses üzerindeki etkisinin ampirik araştırması, 20. yüzyılın başında psikiyatristlerin yeni geliştirilen elektroakustik analiz yöntemleriyle duygusal rahatsızlıkları teşhis etmeye çalışmasıyla başlamıştır (Isserlin, 1925). Telefon ve radyonun icadı ve hızlı yayılması, iletişimde konuşmadaki vokal ipuçları

aracılığıyla konuşmacı özelliklerine bilimsel ilginin artmasına da yol açmıştır (Pear, 1931; Herzog, 1933; Allport & Cantril, 1934). Bununla birlikte, sistematik araştırmalar, psikiyatristlerin duygusal durumları sesli ifade yoluyla teşhis etme konusundaki ilgilerini yeniledikleri 1960'larda başlamıştır (Moses, 1954; Alpert ve ark., 1963; Ostwald, 1964; Hargreaves ve ark., 1965). Sözlü iletişim araştırmacıları, farklı bedensel kanalların duygu sinyallerini taşıma kapasitesini araştırmıştır (Knapp, 1972; Harper ve ark., 1978; Scherer, 1982; Feldman & Rime, 1991). Dilbilimciler ve özellikle fonetisyenler konuşmada pragmatik bilginin önemini keşfetmiştir (Pittenger ve ark., 1960; Mahl & Schulze, 1964; Caffi & Janney, 1994). Akustik sinyal işlemede uzmanlaşan mühendisler ve fonetisyenler, duyguların ses üzerindeki etkilerini incelemek amacıyla daha gelişmiş teknolojiden yararlanmaya başlamıştır (Lieberman & Michaels, 1962; Williams & Stevens, 1969, 1972).

2000'li yılların başlarında, konuşma teknolojisi uygulamaları için konuşma iletişimi modelleri geliştirme çabalarında olan konuşma bilimcileri ve mühendisler, konuşmanın pragmatik ve paralinguistik yönlerini göz ardı ederek konuşmacı tutumlarına ve duygularına daha fazla ilgi göstermeye başlamıştır (Scherer, 2003). Kelime anlamları duygusal olarak nötr olsa dahi konuşmanın duygusal içeriği, dinleyiciler tarafından anlaşılmaktadır. Alanın en yaygın uygulaması, mümkün olduğunca çok sayıda akustik parametreyi incelemek ve bunların çoklu duygularla olan ilişkilerini ölçmeyi amaçlamıştır [Murray & Arnott, 1993; Scherer, 2003; Shami & Verhelst, 2007].

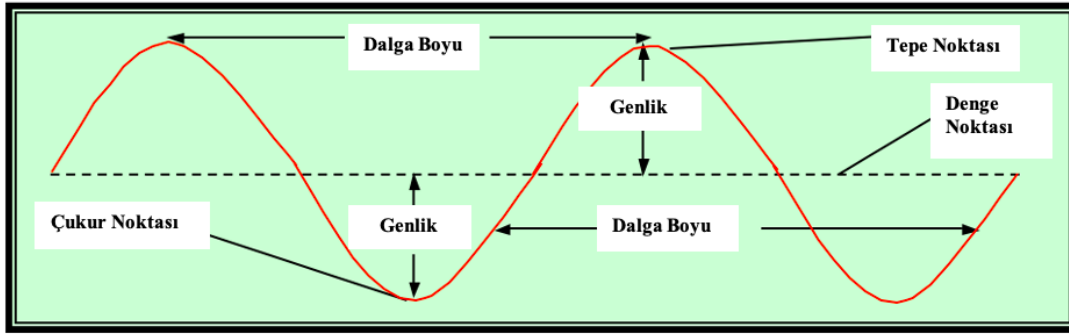
E. Ses

1. Ses Nedir?

Ses, titreşim kaynaklı oluşan ve sıkışma- seyrelme dalgaları halinde yayılan fiziki bir enerjidir (Rezaei & Salehi, 2006). Bu enerji, insanda ses tellerinin titreşmesi ile ortaya çıkarak ses organlarında şekillenerek kompleks bir fonksiyon olarak konuşma veya şarkı söyleme şeklinde ortaya çıkmaktadır (Story, 2015). Konuşma ağız, farenks, larenks, akciğer, diyafram, boyun ve karın kaslarının koordine aktivasyonunu gerektiren kompleks bir dizi aşamayı barındırmaktadır (Sataloff ve ark., 2007).

İşitsel uyarılar, transvers temporal gyrus (Heschl gyrus) merkezinde (Broadman (Br.) 41) sonlanarak sol hemisferde bulunan Wernicke alanına (Br. 22) transfer edilmektedir. Tanıma, anlama, değerlendirme ve yorum gibi işlemler bu merkezde gerçekleşmektedir. Uyarılar, fasciculus arcuatus aracılığıyla frontal lobda bulunan Brocca Merkezine (Br. 44-45) iletilerek, motor cevaplar uyarıya göre düzenlenmektedir (Gerçeker, Yorulmaz & Ural, 2000). İnsan gelişiminde ses biliş yetenekleri, konuşma algısından daha önce ortaya çıkmıştır. Fonem ayırmacılığı bebeklerde yaklaşık 2 aylıkken ve sözcüksel- anlamsal işlem doğumdan 12 ila 14 ay sonra ortaya çıkarken (Friederici, 2005), bebekler halihazırda gelişmiş ses algılama yetenekleri göstermektedir.

Ses dalgaları boyuna ilerleyen mekanik bir dalgadır. Sesin yayılabilmesi için maddesel bir ortama ihtiyacı vardır. Katı, sıvı ve gaz ortamlarında ses dalgaları ortamın özelliğine bağlı olarak farklı hızlarda yayılmaktadır. En hızlı katı ortamlarda yayılan ses dalgaları, en yavaş olarak gaz ortamlarda yayılır. Frekans, dalga boyu, periyot ve genlik ses dalgalarının temel fiziksel parametreleridir (Yücel, 2013). Şekil 11’de ses dalgası gösterilmiştir.



Şekil 11. Ses Dalgası

Kaynak: Lapp, 2003

2. Ses Dalgalarının Temel Özellikleri

Sesin frekansı bir saniye de tamamlanan döngü sayısı olarak ifade edilmektedir. Frekansın birimi Hertz (Hz)'dir. Düşük frekanslar saniyede daha az döngü ile karakterize edilirken, saniyedeki birçok döngü daha yüksek frekanslı bir sesi belirtmektedir (Cole & Flexer, 2019). Frekans değeri ortamdan bağımsız olarak sadece sesin kaynağına bağlı olarak değişmektedir (Johnston, 2002).

Birbiri ardına gelen iki dalga tepesi ya da iki dalga çukuru arasındaki mesafe dalga boyu olarak adlandırılmaktadır. Dalga boyu “ λ ” sembolü ile gösterilir ve birimi metredir. Dalga boyu arttıkça, bir saniye de tamamlanan döngü sayısı azalacağı için frekans değeri azalmaktadır. Yani dalga boyu ile frekans ters orantılıdır (Lapp, 2003).

Periyot ard arda iki dalga tepesi ya da iki dalga çukurunun oluşması için geçen süredir. Periyot “T” ile gösterilir ve birimi saniyedir. Periyot, dalga boyu ile doğru orantılı, frekans ile ters orantılıdır (Gettys, Keller & Skove, 1995).

Salınım hareketi yapan ses dalgaları, kendinden sonra gelen taneciklere enerji aktardıktan sonra minimum enerjiye sahip olduğu denge noktasına geri dönmektedir. Denge noktasından maksimum uzaklık genlik olarak adlandırılmaktadır. Genlik “A” ile gösterilir ve birimi metredir. Ses dalgasının genliği arttığı zaman sesin şiddeti de artmaktadır (Johnston, 2002).

3. Ses Dalgalarının Karakteristik Özellikleri

Sesin fiziksel ve psikoakustik bileşenlerinin olması sebebiyle ses algısını açıklamaya çalışmak karmaşıktır. Ses fiziğinin kaynağı, bir nesnenin titreşimi sonucu oluşan basınç değişimleridir. Bu tür değişiklikler insan dış kulağı tarafından algılanmakta, timpanik membran ve orta kulak kemikçikleri aracılığıyla oval pencereye iletilmekte ve güçlendirilmektedir. Kokleada baziler membranın hareketiyle başlar ve insan beyninin her iki yarım küresine ulaşmak için koklear çekirdekler ve santral işitme yolları boyunca ilerlemektedir (Iakovides ve ark., 2004).

Ses yalnızca konuşmayı iletmekle kalmaz, aynı zamanda konuşmacının cinsiyeti, kimliği ve duygusal durumu hakkında bilgi sağlamaktadır (Imaizumi ve ark., 1997; Agus ve ark., 2010). Bireylere özgü olarak değişen sesin karakteristik özelliklerine perde, tını ve gürlük örnek verilebilir. Sesin fiziksel parametreleri olan frekans ve ses şiddeti ile sesin öznel parametreleri olan perde ve gürlük kavramları ilişkilendirilmektedir. Sesin tınısı ise fiziksel parametrelerle ilişkilendirilmemektedir (Sethares, 2005).

Farklı ses kaynaklarından aynı frekansta elde edilen seslerin farklı olarak algılanması sesin tınısını ifade etmektedir. Pratt ve Doak (1976), tınıyı bir dinleyicinin perde, ses yüksekliği ve süreden başka herhangi bir kriteri kullanarak

iki sesin farklı olduğunu anlayabildiği işitsel algı niteliği olarak tanımlamaktadır. Sesin tınısı subjektif bir kavramdır.

Fiziksel karşılığı frekans olan perde kavramı, sesin kalınlığı ve inceliğini ifade eden algısal bir terimdir (Kılıç, 2002). Sesin frekansı arttığında ses tizleşip, sesin perdesi yükselmekte; sesin frekansı azaldığında ise ses pesleşip, sesin perdesi düşmektedir (Yücel, 2013).

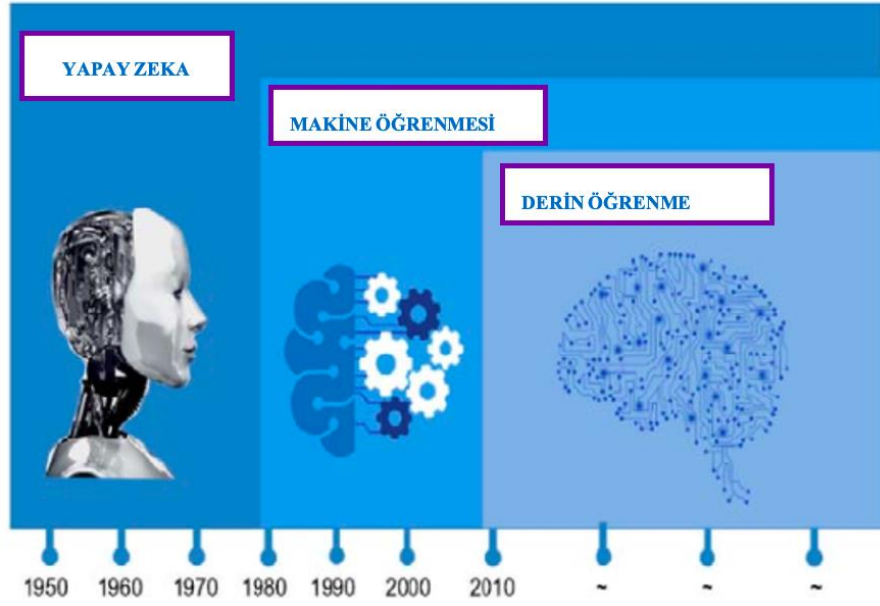
Fiziksel karşılığı ses şiddeti olan gürlük kavramı, sesin miktarı ve düzeyini ifade eden algısal bir terimdir. Sesin şiddetinin ve titreşim genliğinin artması gürlüğün de artmasına sebep olmaktadır (Lee, 2007). Gürlük algısı öznel bir kavram olduğu için, aynı şiddete ve farklı frekansa sahip olan seslerin gürlüğü farklı algılanabilmektedir.

F. Yapay Zekâ

İnsan beyninin çalışma prensibini anlamaya yönelik çalışmaların yapılmaya başlanması ile ortaya çıkan yapay zekâ, beynin algılama ve karar verme mekanizmasını matematiksel ifadelerle kodlamaya yönelik çalışmalar ile geliştirilmeye devam edilmektedir. Yapay zekâ, günlük sosyal yaşamı, ekonomik faaliyetleri ve sağlık sektörünü destekleyen önemli bir teknolojidir (Lu ve ark., 2018).

Yapay zekâ fikri, M. Minsky, A. Newell ve H. Simon' un katılımıyla Dartmouth College' da 1956 yılında düzenlenen konferans ile hayatımıza girmiştir. 1956 yılında ilk defa John McCarthy tarafından 'yapay zekâ' terimi kullanılmıştır (Russell & Norvig, 2009).

Günümüzde yapay zekanın en çok kullanıldığı alanlardan bazıları doğal dil oluşturma, konuşma algılama, sanal araçlar (müşteri hizmetleri, akıllı ev sistemleri), makine öğrenimi platformları (tahmin ve sınıflandırma içeren uygulamalar), yapay zekâ optimizasyonlu uygulamalar, karar yönetimi, derin öğrenme platformları, biyometrik, robotik süreç otomasyonu, metin analizidir (Tan, 2019).



Şekil 12. Yapay Zeka, Makine Öğrenmesi ve Derin Öğrenme Kronolojisi

Kaynak: Kayaalp & Süzen, 2018

1. Makine Öğrenmesi

Yaşamı pek çok alanda kolaylaştıran bilgisayarlara teknolojinin imkân sağladıkları dahilinde insan gibi düşünebilme özelliğini entegre etmek makine öğrenimini sağlamıştır. Makine öğrenmesi, yapay zekanın bir alt kümesi olarak kabul edilir. Makine öğrenmesi deneyimlerden öğrenerek performansını iyileştirir. Süreçlerin ve kaynak tahsisinin optimize edilmesine yardımcı olma konusunda umut veren bir alandır (Bini, 2018).

Genel olarak, makine öğrenimi denetimli öğrenme, denetimsiz öğrenme ve pekiştirmeli öğrenme olmak üzere üç alt alana ayrılmaktadır (Adam & Smith, 2008). Girdileri ve istenen çıktıları olan denetimli öğrenmenin, etiketli verilerle eğitimi gerekmektedir. Denetimli öğrenmenin aksine, denetimsiz öğrenme etiketli eğitim verisi gerektirmez ve hedef çıktı olmadan yalnızca girdiler ile öğrenmektedir. Pekiştirmeli öğrenme, dış çevre ile etkileşimler yoluyla alınan geribildirimden öğrenmeyi sağlamaktadır (Bekkerman ve ark., 2002; Jones, 2014; Qiu ve ark., 2016).

Çizelge 2. Makine Öğrenmesi Türleri

Öğrenme Türleri	Veri İşleme Görevleri	Sınıflandırma Modeli
Denetimli Öğrenme	Sınıflandırma/Regresyon/ Tahmin	Hesaplamalı sınıflandırıcılar İstatistiksel sınıflandırıcılar Bağlantılı sınıflandırıcılar
Denetimsiz Öğrenme	Kümeleme / Tahmin	Parametrik Parametrik Olmayan
Pekiştirmeli Öğrenme	Karar verme	Model içermeyen Model tabanlı

Kaynak: Qiu ve ark., 2016

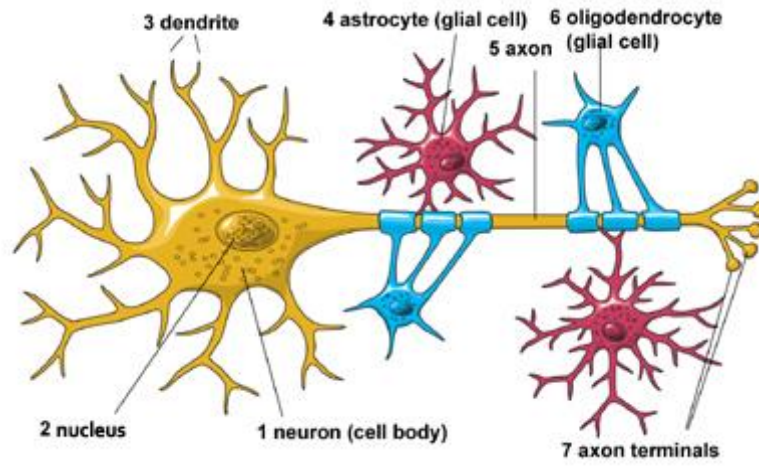
Veri seti ne kadar büyük olursa, makine öğrenmesi için görev o kadar kolay olmaktadır. Ancak görevin zorluğu aynı zamanda veri setindeki verilere atanan etiketlerin doğruluğuna da bağlıdır. Makine öğrenimi, doğru tahmin algoritmaları tasarımı ile gerçekleşmektedir (Mohri, Rostamizadeh & Dikmen). Sağlık alanında erken tanılamaya yönelik yapılan çalışmalar incelendiğinde en çok kullanılan yöntem makine öğrenmesidir (Kaya, Yılmaz & Dikmen, 2019).

Sağlık alanına yönelik uygulamalar öne sürülen hipotezler üzerinden ilerlemektedir. Makine öğrenimi, verilerde önceden gizlenmiş eğilimleri ortaya çıkararak hipotez üretimini geliştirme potansiyeline sahiptir ve bu nedenle hem bireysel olarak hasta düzeyinde hem de sistem düzeyinde önemli etki potansiyeline sahiptir (Panch, Szolovits & Atun, 2018). Makine öğrenimi, verilerin dağıtımı ile ilgili ön varsayımlara dayanmayan yöntemleri kullanarak mevcut istatistiksel teknikleri genişletir ve verilerde hipotezleri formüle etmek için kullanılabilir örüntüler bulabilmektedir. Bu nedenle, makine öğrenimi modellerinin yorumlanması zor olsa da, çok daha fazla değişkeni dahil ederek, daha karmaşık durumlarda sonuçlar üretebilmektedir (Henglin ve ark., 2017; Beam & Kohane, 2018).

2. Yapay Sinir Ağları

İnsan beyni yapısında bulunan sinir hücreleri, öğrenme, hafıza, düşünme ve hatırlama gibi bilişsel davranışların temelini oluşturur. Nöronlar, elektriksel uyarılar ve kimyasal sinyaller aracılığıyla beyinde bulunan farklı alanlar ve sinir sistemi arasında bilgi iletimini sağlar (Ağaoğlu ve ark., 2017).

Nöronların hücre gövdesi, akson ve dendrit olmak üzere üç temel bölümü vardır. Hücre gövdesinin içinde, hücrenin faaliyetlerini kontrol eden ve hücrenin genetik materyalini içeren bir çekirdek bulunur. Akson, dendritlerle bağlantı kurarak, aldıkları bilgileri diğer hücrelere iletmektedir. Dendritler, hücre için mesajları toplamada görevlidir (Kutlu, 2010). Nöronlar, komşu nöronların aksonları ve dendritleri arasındaki sinaps adı verilen küçük bir boşluk boyunca nörotransmitterler adı verilen kimyasallar göndererek birbirleriyle iletişim kurarlar (Duman, 2017). Bilişsel ve motor becerilerin tümü nöronlar ve nöronların destek hücreleri olan astrositler ve oligodendrositler adı verilen glial hücreler olmadan imkansızdır (Avcı & Yağbasan, 2008).



Şekil 13. Nöron Mimarisi

1- Nöron (Hücre Gövdesi), 2- Nücleus (Çekirdek), 3- Dendrit, 4- Astrosit (Glial Hücre), 5- Akson, 6- Oligodendrosit, 7- Akson terminalleri

Kaynak: <https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Life-and-Death-Neuron> (Erişim Tarihi: 09.09.2020)

Son yıllarda, yapay sinir ağları (YSA), insan beyindeki bilgi işleme ve yorumlama tekniğinden esinlenilerek karmaşık biyomedikal verileri işlemek için umut verici bir araç olarak önerilmiştir (Schmidhuber, 2015). Verileri öğrenme, depolama ve birbirleri arasında ilişkiyi ortaya koyma kapasitesi bulunan YSA, simüle edilen sinir hücrelerindeki nöronların birbirleri ile ağ oluşturması temeline dayanır. YSA, örnekler üzerinden öğrendiği bilgiyi sinir hücrelerinin arasındaki bağlantı aralıkları olarak depolar (Şen, 2004).

YSA' nın karmaşık fonksiyonlar için iyi bir yaklaşım sağladığı ve hesaplama açısından verimli olduğu bilinmektedir (Hornik, Stinchcombe & White, 1989). Ek olarak, YSA' da hesaplama genellikle hesaplamayı büyük ölçüde paralel hale getirerek verimliliği artıran bir grafik işlemci birimi (GPU) kullanılarak gerçekleştirilir (Lee ve ark., 2020). GPU, başlangıçta görüntülerle ilgili hesaplama görevlerini yerine getirmek ve görüntüleme aygıtlarına çıktılar üretmek için tasarlanmıştır. Modern GPU'lar çok sayıda hesaplama çekirdeğine sahip olduğu ve hızlı hesaplama yapabilmesinden dolayı yapay sinir ağlarını eğitmek için kullanılmaktadır (Yu, Beam & Kohane, 2018). Bu avantajlardan yararlanarak YSA, eğri uydurma ve ters problemler dahil olmak üzere bir dizi veri işleme görevi için uygulanmış ve büyük boyutlu verileri geleneksel yöntemlerden önemli ölçüde daha kısa işleme süresinde işleme yeteneğini göstermiştir (Jin ve ark., 2017; Yoon ve ark., 2018; Cohen, Zhu & Rosen, 2018; Lee ve ark., 2018).

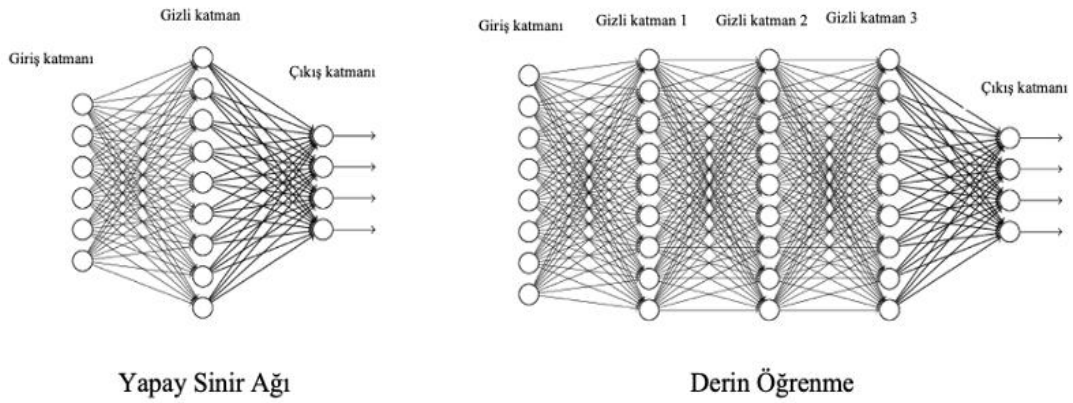
YSA'da öğrenme yöntemleri, denetimli (supervised) ve denetimsiz (unsupervised) olarak ikiye ayrılır (Çuhadar & Kayacan, 2005). Günümüzde yaygın olarak denetimli öğrenme yöntemleri kullanılmaktadır. Denetimli öğrenme yönteminde eğitme işlemi uygulanmaktadır. Girdi sonucunda yapay sinir ağının çıktısı ile beklenen çıktının karşılaştırılması esasına dayanır. Yapay sinir ağlarının insan katkısı olmadan öğrenebileceğinin göstergesi olan denetimsiz öğrenme yöntemleri üzerinde çalışmalar halen devam etmektedir (Anderson & McNeill, 1992).

3. Derin Öğrenme

Derin öğrenme; bilgisayarların, deneyimleri öğrenerek ve kaydederek kavramlar ile ilgili dünyayı daha iyi algılamalarını sağlayan makine öğrenimi olarak tanımlanmaktadır (Gu ve ark., 2016). Yine derin öğrenme, insan beyninin son derece karmaşık problemler için gözlem, analiz, öğrenme ve karar verme yeteneklerini taklit etmeyi amaçlayan, hataya yer vermeyen ve büyük miktarda denetimsiz veri kullanan bir makine öğrenmesi olarak kabul edilmektedir (Najafabadi ve ark., 2015).

21. yüzyıldan itibaren makine öğrenmesi çeşitlerinden biri olan derin öğrenme kavramı ile, karmaşık ve zor problemlerin çözümü için insan sinir

hücrelerini taklit edebilen yapay sinir ağları kullanılmaktadır. Yapay sinir ağları ile gizli katman sayısının beşi geçtiği durumlarda karmaşa artmaktadır (Özcan, 2004). Derin öğrenme ana hatlarıyla yapay sinir ağlarına benzese dahi en büyük fark katman sayısının fazla olması ve ezberlemeyi önlemede kullanılan ‘iletim sönümü’ (Drop Out) yöntemidir (Srivastava ve ark., 2014). Bu yöntem sayesinde eğitim esnasında ağdaki bazı düğümler gelişmiş bir şekilde kaldırılarak ezberlemenin önüne geçilmektedir. Derin öğrenme mimarileri ise birçok gizli katmana sahip yapay sinir ağlarına dayanır ve birçok parametreyi içinde barındırmaktadır (Korotcov ve ark., 2017). Bu sayede derin öğrenme ile yapay sinir ağlarına göre daha az ön işlem yapılmaktadır.



Şekil 14. Yapay Sinir Ağı (YSA) ve Derin Öğrenme (DL) Farkı.

Derin öğrenme, birbirini takip eden katmanlardan dolayı bu ismi almıştır. Geliştirilen farklı modellerdeki katman sayıları modellerin katman derinliğini oluşturmaktadır. Derin Öğrenme modelleri, yüzlerce birbirini takip eden katmanlar içerebilmektedir. Makine öğrenme algoritmaları ise genelde bir veya iki katmandan oluşmaktadır (Chollet, 2017). Derin öğrenme yöntemleri makine öğrenmesinden farklı olarak matematiksel olarak kodlanmış kurallar yerine ses, resim, metin ve video gibi veri simgelerinden otomatik olarak öğrenebilmektedir (Kaya & Yılmaz, 2019).

Konuşma tanıma için derin öğrenme kullanımındaki artış 1990'lı yıllara dayanmaktadır. Hata Derin öğrenmenin konuşma tanımada kullanımının artmasıyla birlikte hata oranlarında önemli derecede büyük düşüşler elde edilmiştir. Sesten duygu tanıma ile ilgili yapılan çalışmalar derin öğrenme kullanımı ile son yıllarda artış göstermiştir (Deng ve ark., 2010; Dahl ve ark., 2010; Seide, Li & Yu, 2011).

Literatür, derin öğrenme yöntemi kullanılarak yapılan birçok çalışmada başarılı sonuçlar elde edildiğini ortaya koymaktadır. CNN, derin öğrenmenin özelleşmiş bir mimarisi olarak özellikle görüntü işleme konusunda oldukça başarılı bir yöntemdir (Cengil & Çınar, 2016). Veriler büyüdükçe, derin öğrenme, özellikle artan işlem gücü ve grafik işlemcilerdeki gelişmelerle, büyük ölçekli veri kümeleri için tahmine dayalı analitik çözümleri sağlamada çok önemli bir rol oynamaya başlamıştır (Chen & Lin, 2014). Derin öğrenme yöntemlerinin kullanım alanları Çizelge 3’de gösterilmiştir.

Çizelge 3. Derin Öğrenme Yöntemleri, Kullanıldığı Alanlar ve Örnek Uygulamalar

Derin Öğrenme Yöntemleri	Derin Öğrenme Yöntemlerinin Kullanıldığı Alanlar	Derin Öğrenme Uygulamaları Örnekleri
Boltzmann machine	Sınıflandırma ve eniyileme	Konu modelleme Boyut indirgeme
Deep Belief Networks	Sınıflandırma, robotik ve bilgisayar görüşü	İnsan aktivitesi tanıma
Deep Feedforward Network	Regresyon, sınıflandırma, tahmin, robotik ve bilgisayar görüşü	Otomatik dil tanıma
Convolutional Neural Network	Regresyon, sınıflandırma, robotik ve bilgisayar görüşü	Görsel doküman analizi Duygu sınıflandırma
Recurrent Neural Network	Regresyon, sınıflandırma, tahmin, robotik ve bilgisayar görüşü	Çeviri yapma El yazısı tanıma Konuşma tanıma
Long Short Term Memory	Zaman serisi verilerini sınıflandırmak	Resimlerden otomatik başlık çıkarma Düzensiz dillerde öğrenme

Kaynak: Kaya, Yılmaz & Dikmen, 2019.

a. Convolutional Neural Networks

Beyin görsel sistemi, karmaşık fiziksel dünyada anlamlı modeller ortaya çıkarmak için tasarlanmıştır (James, 1890). Hayvanların hayatta kalma, üreme olasılığı yiyecekleri ve tehlike işaretlerini fark ettiklerinde daha yüksektir. Bununla birlikte nesnelere görselliği, konum, poz, kontrast, arka plan, ön plan ve diğer birçok faktör açısından birbirlerinden büyük ölçüde farklıdır ve bu özellikleri düşük seviyeli görüntü özelliklerinden ayırt etmek kolay değildir (Pinto ve ark., 2008). Primatlarda görsel sistem, bir görüntünün orijinal piksel değerlerini üst düzey özelliklerin daha açık olduğu yeni bir iç sembole dönüştürerek bu sorunu çözmektedir (DiCarlo ve ark., 2012). Bu, her görüntüyü

bir dizi doğrusal olmayan özellik olarak kodlayan bir algoritma olarak modellenenmektedir (Hung ve ark., 2005; Majaj ve ark., 2015; Nayebi ve ark., 2018). Primat görsel korteksinde niceliksel olarak doğru kodlama modellerinden optimize edilerek CNN geliştirilmiştir (Yamins ve ark., 2014, Khaligh-Razavi & Kriegeskorte, 2014, Güçlü ve Van Gerven, 2015).

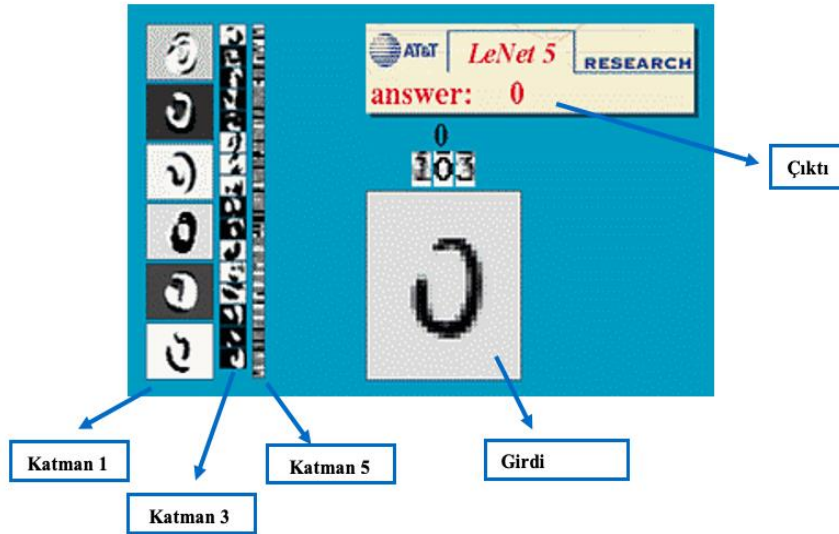
“Convolutional Neural Network” adı ağın, evrişim (convolution) adı verilen bir matematiksel işlemi kullanması sebebiyle verilmiştir. Evrişim özel bir çeşit doğrusal işlemdir. Evrişimsel ağların katmanlarından en az birinde genel matris çarpımı yerine doğrusal işlem kullanılmaktadır (Şeker, Diri & Balık, 2017).

CNN 1980’li yılların başında LeCun ve arkadaşları (1990) tarafından sayıları algılamak için kullanılan bir sistem olan LeNet-5 mimarisi ile tanıtılmıştır. CNN, özel bir tür çok katmanlı sinir ağıdır. Diğer tüm sinir ağıları gibi, geri yayılma algoritmasının bir versiyonu ile eğitilmektedir. CNN, minimum ön işleme ile doğrudan piksel görüntülerinden görsel desenleri tanımak için tasarlanmıştır. El yazısı karakterleri gibi aşırı değişkenliğe sahip desenleri ve basit geometrik dönüşümlere, bozulmalara dayanıklı olan desenleri tanıyabilmektedir. LeNet-5, el yazısı ve makine baskılı karakter tanıma için tasarlanmış yapay sinir ağıdır (LeCun, 2015).

Giriş katmanı, konvolüsyonel katmanlar, havuzlama ve çıktı katmanlarından oluşan CNN, bir önceki katmandan devraldığı girdiler üzerinde daha yüksek düzeyli özelliklerin çıkarılmasını sağlamaktadır (Başbuğ, 2019). Öznitelik haritası elde etmek için konvolüsyonel katmanda giriş verisi filtrelenir. Yapay sinir ağlarının daha hızlı öğrenbilmesi için havuzlama katmanında öznitelik haritaları örneklenir. Her bir verinin bir önceki katmandan gelen tüm girişlerine uygun olarak her bir veri için çıkış değerinin üretildiği katman çıkış katmandır. Her bir katmanda öznitelik çıkarma işlemi yapılarak düşük seviyeli katmanlardan yüksek seviyeli katmanlara kadar etkili bir öğrenme gerçekleştirilmektedir (Çiğdem & Çırak, 2019). Konvolüsyon katmanı sayısı arttıkça daha yüksek düzeyde özelliklerin çıkarılması mümkün olmaktadır. Günümüzde CNN, girdi katmanında ses, görüntü ve video gibi veriler kullanılarak yapılan sinyal işleme çalışmalarında tercih edilmektedir (Valenti ve ark., 2017; Eghbal-Zadeh, 2016).

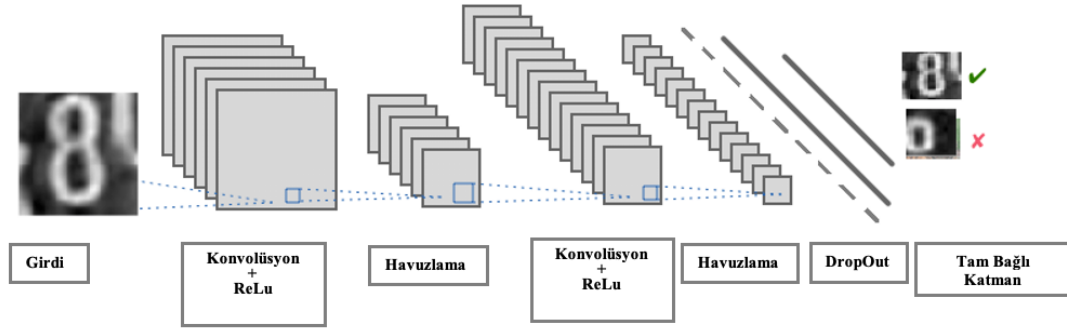
Derin öğrenme mimarilerinden biri olan CNN özellikle veri ön işleme gereksinimlerini en aza indirmek için önerilmiştir. CNN'in ağ eğitimi parametrelerini azaltmasının üç ana yolu vardır. Bunlar yerel bağlantı, ağırlık paylaşımı ve havuzlama yöntemleridir (Xin ve ark., 2018). CNN'lerin en önemli özellikleri, ham görüntü üzerinde çalışabilmeleri ve eğitim setlerinden yararlı özellikleri öğrenebilmeleridir. Böylece eğitim sürecini basitleştirip, görüntü modellerinin tanımlanmasını kolaylaştırmaktadırlar (Yu, Beam & Kohane, 2018).

LeNet mimarisi ile tasarlanan CNN, AlexNet, VGGNet (VGG16, VGG19), ZFNet, GoogLeNet, Microsoft ResNet, DenseNet, CaffeNet, Shallow ConvNet vb. mimariler ile halen geliştirilmeye devam edilmektedir (Adit ve ark., 2020; Blalock ve ark., 2020; Esmailpour, Cardinal & Koerich, 2020; Islam ve ark., 2020). Katman sayılarını arttırma, farklı boyutlarda filtreler uygulama, ardışık yapıyı değiştirme, havuzlama işlemi uygulama ve hiper parametreleri iyileştirme gibi düzenlemeler yeni CNN mimarilerin geliştirilmesini ve test başarı oranının artmasını sağlamaktadır (Akin & Duman, 2019). Şekil 15' de ilkel bir CNN modeli, Şekil 16'da daha kompleks bir CNN modeli örneği verilmiştir.



Şekil 15. İlkel CNN Modeli: Lenet 5 Örneği.

Kaynak: LeCun, 2015



Şekil 16. Geliştirilmiş CNN Modeli

Kaynak: Tipari, 2018

Konuşma ve ses işleme uygulamalarında, veriler genellikle sinyallerin dinamik bilgilerini yakalayamayan sıkıştırılmış gösterimler hesaplanarak işlenmektedir. Son yıllarda ses ve konuşma analizi için, CNN'ler ve tekrarlayan sinir ağları (RNN) gibi derin öğrenme yöntemlerini kullanan çalışmaların sayısı artmıştır (Purwins ve ark., 2019).

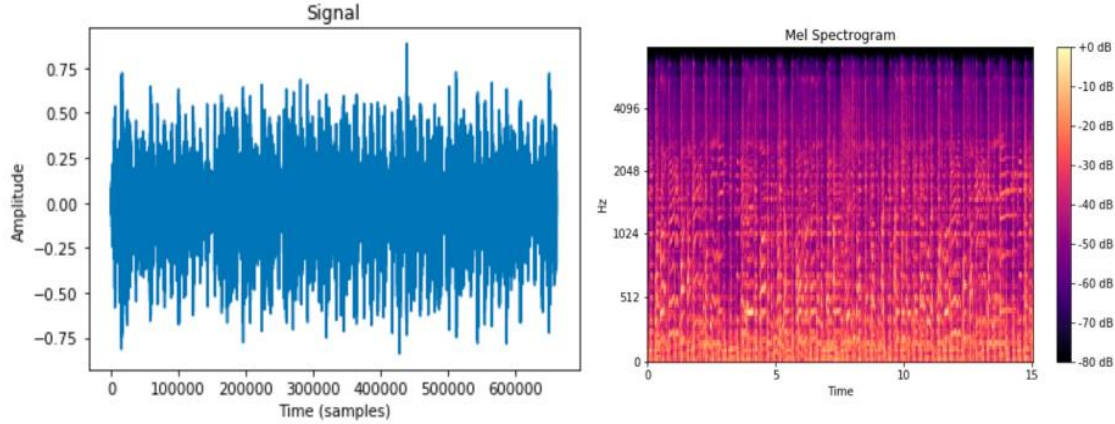
Ses verileri için özellikle CNN'ler, frekans alanında dağıtılan enerjinin zamanla nasıl değiştiği hakkında bilgi sağlayan sinyallerin zaman-frekans temsilleri (spektrogramlar) ile konvolüsyon katmanlarını besleyerek işlenmektedir. Konvolüsyon işleminden sonra ortaya çıkan özellik haritaları, sinyallerin akustik bilgisini temsil eden düşük ve yüksek seviyeli özellikleri içermektedir (Arias-Vergara ve ark., 2020).

Birçok çalışma, diğer yöntemlerden ziyade CNN ve spektrogramları, konuşma bozukluklarının otomatik tespiti (Vásquez-Correa, Orozco-Arroyave & Nöth, 2017; Wu ve ark., 2018; Alhussein & Muhammad, 2018), otomatik konuşma tanıma sistemleri için akustik modeller (Abdel-Hamid ve ark., 2014; Han ve ark., 2015) ve duygu algılama (Weißkirchen, Bock & Wendemuth, 2017) gibi farklı konuşma işleme uygulamalarında kullanmanın avantajlarını göstermiştir.

i. Mel Spektrogram

Mel-spektrogram, Mel-ölçeğindeki ses sinyalinin bir temsilidir. Sesin zaman-frekans gösterimi anlamına gelmektedir. Ses sinyallerine mel spektrogram uygulanmadan önce x ekseni zamanı, y ekseni sinyal genliğini gösterir. Mel spektrogramı uygulandıktan sonra oluşan grafikte x ekseni zamanı, y ekseni

frekansı gösterir. Mel-spektrogramın logaritmik formu, duyguları daha iyi anlamaya yardımcı olmaktadır (Venkataramanan & Rajamohan, 2019).



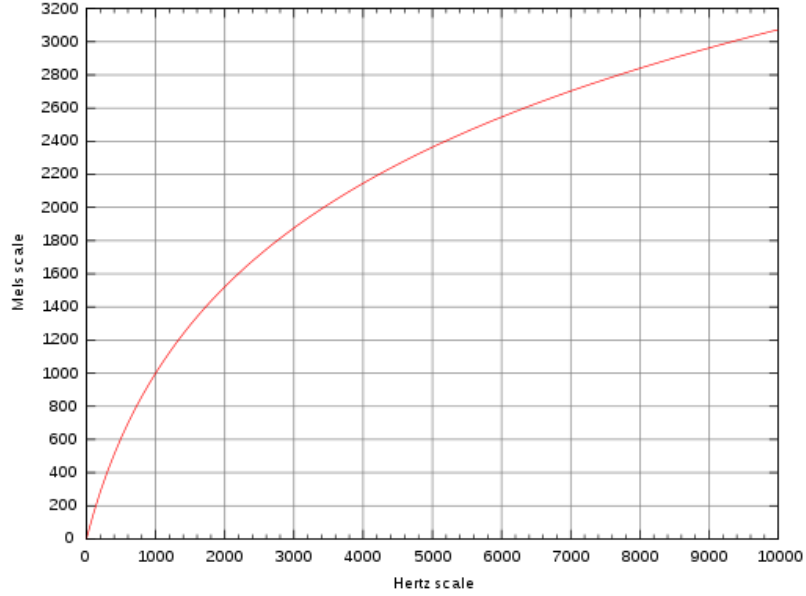
Şekil 17. Ses Sinyali ve Mel Spektrograma Dönüştürülmüş Hali

Kaynak: <https://medium.com/analytics-vidhya/understanding-the-mel-spectrogram-fca2afa2ce53> (Erişim Tarihi: 08.08.2020)

ii. MFCC (Mel-Frequency Cepstral Coefficients)

Öznelik çıkartma yöntemlerinden biri olan MFCC, sesteki spektral özelliklerin elde edilmesi için kullanılmaktadır. Mel ölçeği deneysel olarak hesaplanmıştır. 1 kHz 'in üstündeki frekanslar için logaritmik ve 1 kHz 'in altındaki frekanslar için doğrusala yakın olarak elde edilmiştir. MFCC, insan kokleasını taklit etmek için ses sinyalini bir dizi adımda dönüştürerek, sesin kısa vadeli güç spektrumunu temsil eder. Doğrusal ölçeklerin aksine insan kaynaklı ses algısına daha iyi yaklaşır (Alhlffee, 2020). MFCC, konuşma-duygu tanıma için güvenilir bir yöntemdir (Palo, Chandra & Mohanty, 2018).

MFCC' de, konuşmalar bölümlere ayrılır, ardından her bölüm kısa süreli ayrık Fourier dönüşümü kullanılarak frekans alanına dönüştürülür. Mel filtre bankası kullanılarak bir dizi alt bant enerjisi hesaplanır. Daha sonra bu alt bantların logaritması hesaplanır. Son olarak, MFCC'yi elde etmek için ters Fourier dönüşümü uygulanır. En yaygın kullanılan spektral özelliktir (Kuchibhotla ve ark., 2014).



Şekil 18. Mel Ölçeği

iii. Aktivasyon Fonksiyonları

Derin öğrenme algoritmalarında en iyi performansı elde edebilmek adına yapay sinir ağları açısından önemli olan aktivasyon fonksiyonların seçimi önemlidir. Yapay sinir ağlarının, doğrusal olmayan durumlarda dahi öğrenme işleminin gerçekleşmesi için aktivasyon fonksiyonları kullanılarak yapay sinir ağlarının doğrusal olmayan özellikleri de öğrenmesi sağlanmaktadır. Çıkış sinyali, aktivasyon fonksiyonunun kullanılmadığı durumlarda tek dereceli polinom olan doğrusal fonksiyonlardır. Doğrusal fonksiyonlar görüntü, ses, yazı ve video gibi çok parametreye sahip veriler için yetersizdir (Tan, 2019).

Aktivasyon fonksiyonları, yapay sinir ağlarında, giriş sinyalini bir sonraki katmana giriş olacak çıkış sinyaline dönüştürmek için özel olarak kullanılmaktadır. Yapay sinir ağında, girdilere karşılık gelen ağırlıklar hesaplanır ve son olarak o katmanın çıktısını almak ve onu bir sonraki katmana girdi olarak sağlamak için ona aktivasyon işlevi uygulanır. Girdiler 'x' ve ağırlıklar 'w' olarak tanımlanmaktadır. $F(x)$ fonksiyonu, başka bir katmanın giriş verisi olabileceği gibi çıkış verisi de olabilir (Sharma, 2017).

Yapay sinir ağlarının tahmin doğruluğu, kullanılan katman sayısına ve daha önemli olarak kullanılan aktivasyon fonksiyonu türüne bağlıdır. Yapay sinir ağlarının daha iyi sonuç ve doğruluk vermesi için kullanılacak minimum veya maksimum katman sayısını belirten bir kılavuz yoktur. Ancak temel olarak en az

2 katmanın kullanılması gerektiği bilinmektedir. Literatürde, kullanılması gereken aktivasyon fonksiyonu hakkında genel geçer bir kural yoktur (Gomes, Ludermir & Lima, 2011; Sibi, Jones & Siddarth, 2013).

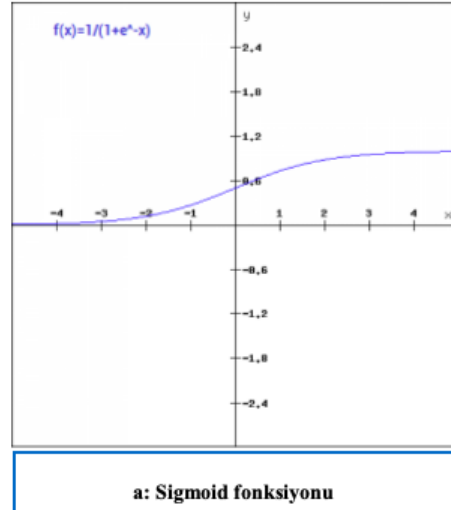
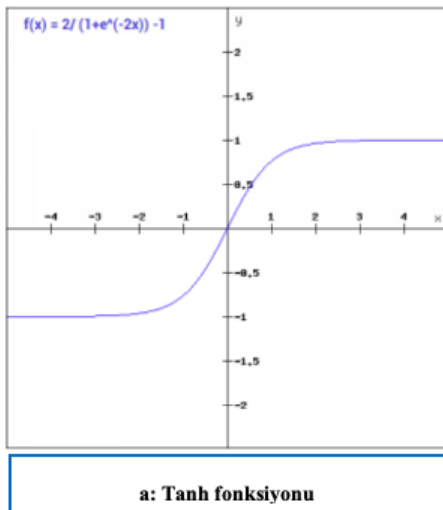
Aktivasyon fonksiyonlarına Binary Step Function, Linear, Sigmoid, Tanh (Hiperbolik Tanjant), ReLu (Rectified Linear Unit), Leakly ReLu, Parametrized ReLU, Exponential Linear Unit, Swish ve Softmax fonksiyonu örnek verilebilir (Sharma, 2017).

iv. Tanh Fonksiyonu

Tanh işlevi sigmoid aktivasyon fonksiyonuna benzer, ancak orijine göre simetriktir. Tanh fonksiyonu sürekli ve türevi alınabilir. Değerleri -1 ile 1 arasındadır. Sigmoid fonksiyonuna kıyasla tanh fonksiyonunun gradyanı daha diktir. Tanh, belirli bir yönde değişmekle sınırlı olmayan gradyanlara sahip olduğu ve sıfır merkezli olduğu için sigmoid fonksiyonundan daha fazla tercih edilmektedir (Sharma, 2017). Tanh ve sigmoid fonksiyonlarının her ikisi de, tanh çıktısının [-1,1] arasında değiştiği ve sigmoid çıktısının [0,1] arasında değiştiği "S" şeklinde eğri üretir (Namin ve ark., 2009). Şekil 19'da Tanh fonksiyonu (a) ve sigmoid fonksiyonu (b) grafikleri gösterilmiştir.

Tanh fonksiyonu aşağıdaki gibi tanımlanmaktadır:

$$f(x) = 2\text{sigmoid}(2x) - 1$$



Şekil 19. Tanh Fonksiyonu (-1,1) ve Sigmoid Fonksiyonu (0,1)

v. Softmax Fonksiyonu

Softmax fonksiyonu, birden çok sigmoid fonksiyonunun kombinasyonudur. İkili sınıflandırma için kullanılan sigmoid fonksiyonundan farklı olarak Softmax fonksiyonu, fazla sayıda sınıflandırma grubunun olduğu durumlarda tercih edilmektedir. Softmax fonksiyonu 0 ile 1 arasındaki girdi verilerini, toplamı 1 olacak olasılıklar şeklinde haritalamaktadır. Çoklu sınıflandırma için softmax fonksiyonu kullanılarak bir yapay sinir ağı oluşturulduğunda, çıktı katmanı hedefteki sınıf sayısı kadar nörona sahip olmaktadır (Sharma, 2017).

Softmax aktivasyon fonksiyonu, derin öğrenme mimarilerinde özellikle çıkış katmanında kullanılmaktadır (Çiğdem & Çırak, 2019).

Softmax fonksiyonu aşağıdaki gibi tanımlanmaktadır:

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{k=1}^K e^{z_k}} \quad \text{for } j = 1, \dots, K.$$

III.GEREÇ VE YÖNTEM

A. Çalışmanın Yürütüldüğü Birim

Bu çalışma İstanbul Aydın Üniversitesi Lisansüstü Eğitim Enstitüsü Odyoloji Anabilim Dalı, Odyoloji Programı Yüksek Lisans tezi olarak yapılmış olup, Üsküdar Üniversitesi Girişimsel Olmayan Klinik Araştırmalar Etik Kurulu tarafından 61351342/ 2020-449 kayıt numarası ile onay verilmiştir (Ek- 2). Katılımcılar, gönüllülük esasına dayalı olarak bilgilendirilmiş ve gönüllü onam formu imzalamışlardır (Ek- 1).

B. Çalışmanın Örneklemi

Normal işitmeye sahip ve koklear implant kullanan işitme kayıplı bireylerin nötr, mutluluk, hüznün, öfke, korku ve şaşkınlık ses duygu değerlendirmelerini karşılaştırmak için; normal işiten bireyler için İstanbul Aydın Üniversitesi'nde çalışan bireylerden (20 birey); koklear implant kullanan bireyler için İstanbul ilindeki Özel Eğitim ve Rehabilitasyon Merkez'lerinde eğitim görmekte olan işitme kayıplı bireylerden (20 birey) olmak üzere toplamda 40 birey çalışmaya dahil edilmiştir. Altı farklı ses duygu kaydı için normal işitmeye sahip bireylerden 600, koklear implant kullanan bireylerden 600 olmak üzere toplamda 1200 ses kaydı derin öğrenme ile değerlendirilmiştir.

Çalışmaya dahil edilme kriterleri:

- 18-40 yaş aralığında olma,
- Normal işitmeye sahip bireyler için kriter; otoskopik bakının normal olması, 125 Hz – 8.000 Hz frekansları arasında saf ses işitme eşiklerinin 25 dB HL veya daha iyi olması, Konuşmayı Ayırt Etme Skorunun %88 ve üzeri olması, immitansmetrik bulgularının normal olması,

- Koklear implant kullanan bireyler için kriter; ileri veya çok ileri derecede işitme kaybı olmak ve tek veya çift taraflı koklear implant operasyonu geçirmesi.

Çalışmaya dahil edilmeme kriterleri:

- Ana dilinin Türkçe olmaması,
- Psikiyatrik ve mental problemi olması,
- Fiziksel ek engel olması.

C. Katılımcıların Özellikleri

Çalışmaya 18- 40 yaş aralığında, normal işiten ve koklear implant kullanan olmak üzere toplamda 40 birey dahil edilmiştir. Normal işitmeye sahip bireylerin 10'u kadın, 10'u erkek ve koklear implant kullanan bireylerin 14'ü kadın, 6'sı erkektir (Çizelge 4).

Çizelge 4. Çalışmaya Dahil Edilen Bireylerin Sayısı

Gruplar	Kadın		Erkek		Toplam n
	n	%	n	%	
Normal İşitmeye Sahip Bireyler	10	50%	10	50%	20
Koklear İmplant Kullanan Bireyler	14	70%	6	30%	20
Toplam	24	60%	16	40%	40

D. Veri Toplama Araçları ve Çalışma Planı

Çalışma üç aşamada gerçekleştirilmiştir.

- Veri Setinin Oluşturulması
- Derin Öğrenme Algoritması Oluşturulması ve Eğitilmesi
- Prototipin Canlıya Geçirilmesi (Klinik Uygulama)

1. Veri Setinin Oluşturulması

a. Derin Öğrenme Eğitim Seti İçin Cümlelerin Hazırlanması

İlk olarak ses duygu analizini gerçekleştirecek derin öğrenme algoritmasında kullanılacak olan cümle veri setleri oluşturulmuştur. Cümleleri oluştururken günlük yaşamda sık kullanılan kelimeler olmasına dikkat edilmiştir. Türkiye'nin 7 coğrafik bölgesindeki 70 bireye, oluşturulan cümleler ses kaydı ile İstanbul Aydın Üniversitesi öğrencileri tarafından dinletilerek cümlelerin hedef yaş grubu için bilinebilirlikleri değerlendirilmiştir. Bilinebilirliği ve anlaşılabilirliği yüksek olan on cümle seçilmiş ve bu cümlelerin mutluluk, hüznün, öfke, korku, şaşkınlık duyguları ile ifade edilmesi için Drama ve Oyunculuk Bölümünde eğitim veren tiyatro sanatçılarından destek alınarak kısa hikayeler kurgulanmıştır. Cümleler Çizelge 5'de gösterilmiştir. Cümleler, iOS işletim sistemine sahip cep telefonlar kullanılarak 20 kadın ve 20 erkek olmak üzere toplamda 40 öğrenci tarafından seslendirilmiştir.

Seslendirme için ağız ile telefon mikrofonu arasındaki mesafenin 25 cm olması ve seslendirme yapılan ortamın gürültüsüz olması gerektiği ses kaydı öncesinde belirtilmiştir. Öğrenciler belirtilen noktalara dikkat ederek altı farklı duyguda cümleleri seslendirmiştir.

Çizelge 5. Oluşturulan Cümleler

CÜMLEL ER		DUYGU İFADELERİ				
	Nötr	Mutluluk	Hüzün	Öfke	Korku	Şaşkınlık
Hızlıca eve döndüm.	Hızlıca eve döndüm.	Annemin bana aldığı hediyeyi görmek için hızlıca eve döndüm.	Sınav sonucum açıklandığında moralim bozularak hızlıca eve döndüm.	İşyerinde patronumla tartıştıktan sonra hızlıca eve döndüm.	Ütüyü prizde unuttuğum aklıma gelince hızlıca eve döndüm.	Kötü geçen sınavdan çok yüksek puan aldığımı öğrenince ailemle paylaşmak için hızlıca eve döndüm.
Birden ağlamaya başladı.	Birden ağlamaya başladı.	Tekerlekli sandalye kullanan kardeşinin yürümeye başladığını görünce birden ağlamaya başladı.	Çok beğenerek aldığı elbisesinin yırtıldığını görünce birden ağlamaya başladı.	Kopya çekmediği halde öğretmeni tarafından sınıfın ortasında azarlanınca birden ağlamaya başladı.	Gece yarısı çalılarının arkasında kocaman bir yılan olduğunu görünce birden ağlamaya başladı.	En sevdiği arkadaşından hiç beklemediği sözleri duyunca birden ağlamaya başladı.

Çizelge 5. (devamı) Oluşturulan Cümleler

CÜMLEL ER	DUYGU İFADELERİ					
	Nötr	Mutluluk	Hüzün	Öfke	Korku	Şaşkınlık
Akşam erkenden uyudum.	Akşam erkenden uyudum.	Sabah tatile gidecek olmanın sevinciyle akşam erkenden uyudum.	Evde yaşanan tatsızlığı düşünmemek için akşam erkenden uyudum.	Erken saatlerde başlayan bitmek bilmeyen tadilat yüzünden uykusuz kalmamak için akşam erkenden uyudum.	Elektrik kesildikten sonra duyduğum tıkırtı sesleri karşısında ürpermek için akşam erkenden uyudum.	Bir sonraki gün akşam altıya aldığımı sandığım uçak biletinin sabah altıda olduğunu fark edince akşam erkenden uyudum.
Çatal elinden düştü.	Çatal elinden düştü.	Taraftarı olduğu takımın son dakika gol attığını görünce çatal elinden düştü.	Yaralı muhabbet kuşunun öldüğü haberini alınca çatal elinden düştü.	Yan masada kavgalı olduğu arkadaşının yemek yediğini görünce yaşadığı haksızlık aklına geldi ve çatal elinden düştü.	Bahçede dolaşan kedisi köpek tarafından karşısında ısırılınca çatal elinden düştü.	Hangi taraftan geldiğini anlamadığı basketbol topu yüzüne çarpınca çatal elinden düştü.
Hastaneye gitmek istemedi.	Hastaneye gitmek istemedi.	Halı saha maçında burkulan ayak bileğinin hızlıca iyileşmesinin ardından hastaneye gitmek istemedi.	Son günlerini evinde ailesiyle geçirmek isteyen büyükannem hastaneye gitmek istemedi.	Dikkatsizliği sonucu arabasının aldığı büyük hasar karşısında kendisini cezalandırarak hastaneye gitmek istemedi.	Kazaya sebep olup, olay yerinden kaçan sürücü yakalanmamak için hastaneye gitmek istemedi.	Karşılaşacağın 1 asla düşünmediği ilkokul arkadaşını yolda görünce hastaneye gitmek istemedi.
Dikkatleri üzerine çekti.	Dikkatleri üzerine çekti.	O kadar güzel giyinmişti ki dikkatleri üzerine çekti.	Oğlunu askere gönderecek olan teyzem saatlerce ağlayarak dikkatleri üzerine çekti.	Yolda siren çalan ambulansa geçit vermeyen araç dikkatleri üzerine çekti.	Otobanda makas atarak giden motor yürekleri ağza getiren araç dikkatleri üzerine çekti.	7 yaşında ayda 250 kitap bitiren çocuk dikkatleri üzerine çekti.
Yağmurda yürümeye başladı.	Yağmurda yürümeye başladı.	Eşiyle eski günlerdeki gibi el ele tutuşarak yağmurda yürümeye başladı.	Memlekete giden ailesinden ayrılmanın verdiği üzüntüyle yağmurda yürümeye başladı.	Saatlerce beklediği otobüs gelmeyince sinirden yağmurda yürümeye başladı.	Tanımadığı bir adam tarafından takip edildiğini fark edince hızlıca yağmurda yürümeye başladı.	Milyon dolarlık bir arabaya sahip olmasına rağmen yağmurda yürümeye başladı.
Beş çocuk gördüm.	Beş çocuk gördüm.	Birbirinden tatlı mı tatlı, güzel mi güzel beş çocuk gördüm.	Geçim kaynağı için çöp toplayan beş çocuk gördüm.	Düşüncesizce arabaya taş atan beş çocuk gördüm.	Hortumun hızlıca yaklaştığı mahallede beş çocuk gördüm.	Buz gibi havada şortla umursamazca gezen beş çocuk gördüm.

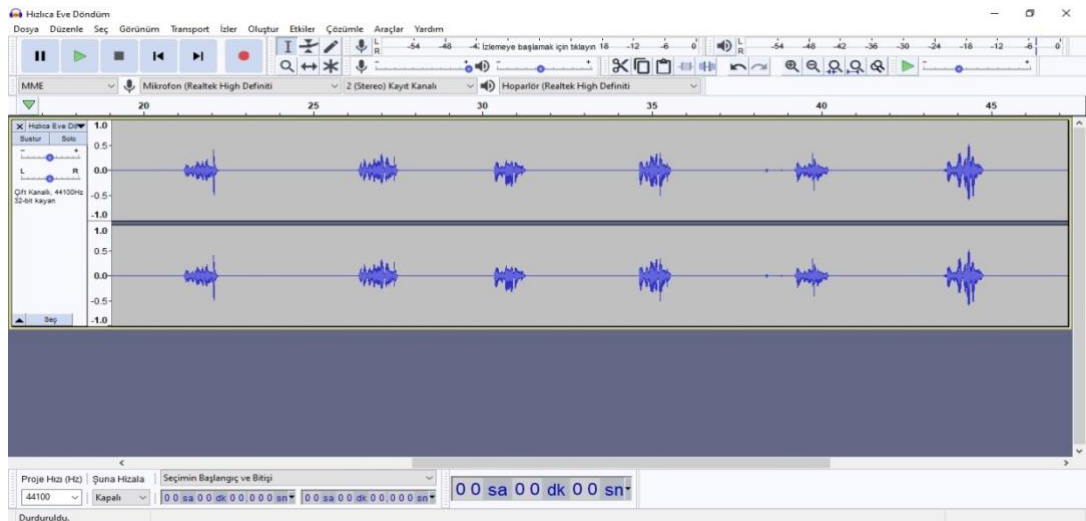
Çizelge 5. (devamı) Oluşturulan Cümleler

CÜMLEL ER	DUYGU İFADELERİ					
	Nötr	Mutluluk	Hüzün	Öfke	Korku	Şaşkınlık
Televizyo nun sesini açtı.	Televizyo nun sesini açtı.	Okulların tatil olduğu haberini alınca televizyonun sesini açtı.	Çok sevdiği yazarın ölüm haberini duyunca televizyonun sesini açtı.	Yan komşusunun bitmek bilmeyen gürültüsüne dayanamayıp televizyonun sesini açtı.	Savaş haberlerini öğrendiği anda televizyonun sesini açtı.	Oğlunu ansızın ekranda görünce televizyonun sesini açtı.
Uçak bileti aldı.	Uçak bileti aldı.	Yeni doğan kardeşini görecek olmanın heyecanıyla uçak bileti aldı.	Yaşamak istediği şehirde yaptığı iş görüşmesi olumsuz sonuçlanınca uçak bileti aldı.	Sınav tarihlerinin erkene alındığını öğrenince sabah gittiği tatilden dönmek için kızgınlıkla uçak bileti aldı.	Bulaşıcı hastalığın oturduğu şehirde hızla yayıldığını öğrenince aceleyle uçak bileti aldı.	Beklenmedik gelişmeler karşısında hiç aklında yokken uçak bileti aldı.

b. Veri Seti Ses Kayıtlarının Etiketlenmesi

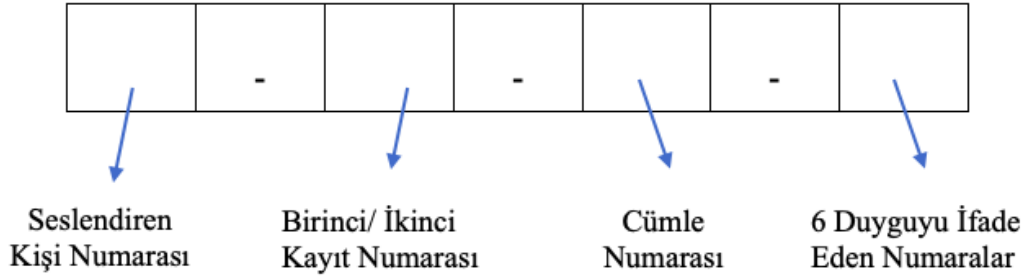
Tüm öğrenciler (40 birey), on cümleye ait altı farklı duyguyu ikişer kez seslendirmiştir. Her öğrenciye ait 120 ses kaydı, toplamda 4800 ses kaydı elde edilmiştir. Öğrencilerin her cümle arasında 5 saniye bırakarak oluşturduğu kayıt, Audacity programında her cümle farklı bir wav dosyası olacak şekilde kırıldıktan sonra yeniden etiketlenmiştir. Audacity birçok işletim sistemi ile uyumlu çalışan ve açık kaynak kodlu bir ses editörüdür (Şekil 20).

Drama ve Oyunculuk Bölümü'nde öğrenim gören öğrencilerden alınan kayıtların spektrogramları incelenerek, daha iyi spektral analize sahip olan 2400 ses kaydı derin öğrenme mimarisi girdi katmanında kullanılmıştır.



Şekil 20. Audacity Ekran Görüntüsü

Kayıtların etiketlenmesi Şekil 21’de gösterilmiştir.



i. Seslendiren Kişi Numarası

1’ den 40’ a kadar olan sayılar, her bir birey için farklı olacak şekilde verilmiştir.

ii. Birinci/ İkinci Kayıt Numarası

Her bir cümle için iki kez kayıt alınmıştır. Birinci kayıt 1; ikinci kayıt 2 olarak tanımlanmıştır.

iii. Cümle Numaraları

1’ den 10’a kadar olan sayılar, her bir cümle için farklı olacak şekilde verilmiştir. Çizelge 6’da gösterilmiştir.

Çizelge 6. Eğitim Aşaması / Cümle Numaraları

Cümleler	Numara
Hızlıca eve döndüm.	1
Birden ağlamaya başladı.	2
Akşam erkenden uyudum.	3
Çatal elinden düştü.	4
Hastaneye gitmek istemedi.	5
Dikkatleri üzerine çekti.	6
Yağmurda yürümeye başladı.	7
Beş çocuk gördüm.	8
Televizyonun sesini açtı.	9
Uçak bileti aldı.	10

iv. Duyguları İfade Eden Numaralar

1’den 6’ ya kadar olan sayılar, her bir sayı farklı duyguyu temsil edecek şekilde verilmiştir. Çizelge 7’de gösterilmiştir.

Çizelge 7. Eğitim Aşaması / Duyguları İfade Eden Numaralar

Duygu İfadeleri	Numara
Nötr	1
Mutluluk	2
Hüzün	3
Öfke	4
Korku	5
Şaşkınlık	6

2. Derin Öğrenme Algoritması Oluşturulması ve Eğitilmesi

Ses duygu analizi derin öğrenme algoritması, İstanbul Teknik Üniversitesi Elektrik- Elektronik Mühendisliği Bölümü Prof. Dr. Tahir Çetin Akıncı tarafından oluşturulmuştur. Derin öğrenme mimarisi, Anaconda Navigator’da bulunan Spyder bütünleşik geliştirme ortamında Python programlama dili kullanılarak kodlanmıştır.

Derin öğrenme mimarisinde akustik girdiler kullanıldığı durumlarda yapay sinir ağı öznelik çıkartıcı ve sınıflandırıcı görevlerini yerine getirir. Öznelik çıkarma, konuşma sinyaline ait tanımlayıcı ve ayırtedici özellikleri elde etmek için kullanılır (Ateş, 2019). Çalışmamızda akustik girdilerin kullanılması ve konvolüsyonel katmanların yalnızca görüntüye dönüştürülmüş verilere işlem yapabilmesi sebebiyle sırasıyla öznelik çıkartma ve sınıflandırma görevleri tamamlanmıştır.

Derin öğrenme, çok sayıda gizli katmana sahip gelişmiş sinir ağlarından oluşmaktadır. Çalışmamızda derin öğrenme algoritmalarından Convolutional Neural Network (CNN) kullanılarak ses duygu analizi gerçekleştirilmiştir. CNN’ler görüntü işleme üzerinden öğrenmeyi sağladığı için ses kayıtları MFCC ve mel spektrogram kullanılarak görüntüye çevrilmiştir. Görüntü üzerinde yer alan kare filtre kombinasyonlarına konvolüsyon işlemi uygulanmıştır. Her konvolüsyon katmanından sonra tanjant hiperbolik (Tanh) fonksiyonu kullanılmıştır. Maksimum havuzlama işlemi uygulanarak kendinden sonra gelecek olan konvolüsyon katmanına girdi olarak daha düşük boyutta veri sağlanmıştır. Dropout aşamasında ağıdaki düğümlerden bazıları kaldırılmıştır. Tam bağlantılı katman ile önceki katmanlardan gelen tüm özellikler birleştirilmiştir. CNN’ in son aşaması olan sınıflandırma aşamasında veriler Softmax fonksiyonu kullanılarak sınıflandırılmıştır.

Drama ve Oyunculuk Bölümü'nde eğitim gören 40 (20 kadın, 20 erkek) öğrenciye ait 6 farklı duyguda seslendirilmiş 10 cümleden oluşan toplamda 4800 cümle kaydı derin öğrenmenin veri setini oluşturmuştur. Cümleler, spektral analiz yapıp gürültüden arınmış ve daha kaliteli ses kayıtları elde etmek amacıyla 2 kez seslendirilmiştir. Spektral analiz sonrasında 4800 olan ses kaydı sayısı 2400'e düşmüştür. 2400 ses kaydının, derin öğrenme eğitimi (training) aşamasında %75'i (1800 ses kaydı); derin öğrenme test aşamasında ise %25'i (600 ses kaydı) kullanılmıştır.

a. Derin Öğrenme Algoritmasının Aşamaları

i. Mel Spektrogram

Konuşma sinyali, yarı durağan şekilde zaman düzleminde yavaş değişen bir sinyaldir. 5 ile 100 ms arasında incelendiğinde, özellikleri oldukça sabittir. Bununla birlikte, 0.2sn veya daha fazla süre boyunca sinyal özellikleri, konuşulan farklı konuşma seslerini yansıtacak şekilde değişir. Bu nedenle, kısa süreli spektral analiz konuşma sinyalini karakterize etmenin en yaygın yoludur (Hasan, Jamil & Rahman, 2004). İnsan kulağı 1000 Hz'e kadar olan frekansları doğrusal olarak algılamakta, 1000 Hz'in üzerinde olan frekansları logaritmik olarak algılamaktadır.

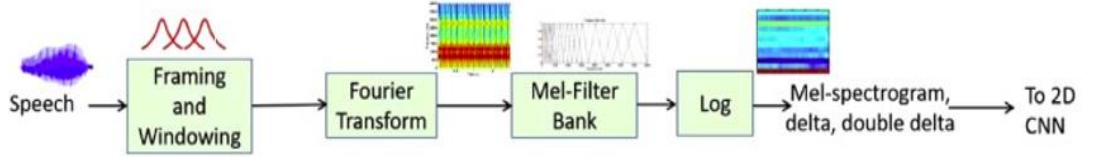
Ses kayıtlarından oluşan veri setine MFCC ve Mel Spektrogram uygulanmıştır. CNN modelleri girdi olarak görüntü kullanımını gerektirir. Bu sayede veriler zaman- frekans düzleminde haritalanarak CNN uygulanması için temel sağlanmıştır.

Mel spektrogram görüntü yoğunluk değerleri MFCC'ye benzer şekilde hesaplanır, ancak Discrete Cosine Transform (DCT) uygulanmadan, mel filtre bankası enerjileri kullanılarak hesaplanır. Mel filtre banka çıktısı filtre olarak belirlenebilir (Sharan & Moir, 2019).

Akustik verileri MFCC ve Mel spektrogram aracılığıyla zaman- frekans düzleminde göstermek amacıyla verilere sırasıyla uygulanan adımlar aşağıda sıralanmıştır.

- Ön vurgulama katsayısı
- Çerçeveleme ve Pencereleme

- Fast Fourier Transform (FFT)
- Mel Filtre Bankası
- Logaritma
- Mel spektrogram



Şekil 22. Mel Spektrogram Şeması

Kaynak: Hossain & Muhammad, 2019

ii. Önvurgulama katsayısı

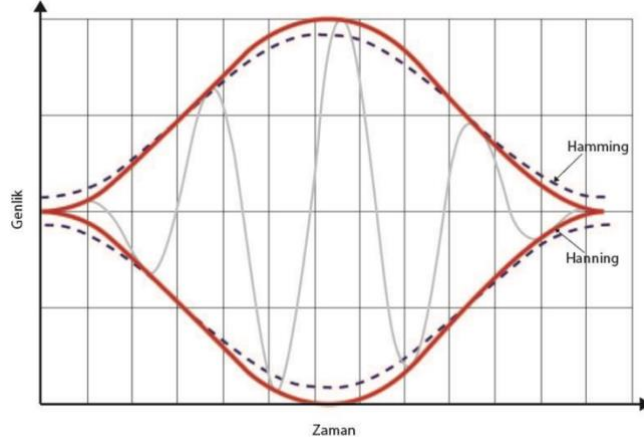
Yüksek frekanslardaki enerjinin belirgin hale getirilmesi için, önvurgulama katsayısı kullanılarak konuşma sinyali 0,9 ile 1,0 aralığında değişen bir değer ile çarpılmıştır.

iii. Çerçeveleme

Sesin spektral analizinin yapılabilmesi için konuşma sinyali belirli kısa zaman aralıklarına bölünüp çerçevenmiştir. Çerçeveler arası geçişin daha yumuşak olması için çerçevelerden her biri bir önceki çerçevenin bir kısmını örtmektedir.

iv. Pencereleme

Çerçeveleme işleminden sonra spektral bozulmaların minimize edilmesi amacıyla çerçevelenen konuşma sinyalinin başlangıç ve bitişine pencereleme işlemi uygulanmıştır. Bu işlem sıklıkla Hamming ve Hanning pencereleme fonksiyonları kullanılarak yapılmaktadır. Çalışmamızda kullanılan Hamming fonksiyonudur. Hamming fonksiyonu verileri sinyalin başında ve sonunda sıfıra yaklaştırır; Hanning fonksiyonu verileri ise sinyal başında ve sonunda sıfıra dokunur (Şekil 23). Hamming fonksiyonunda gürültü bastırma sonuçlarının daha iyi olması sebebiyle otomatik ses tanıma sistemlerinde daha fazla tercih edilmektedir (Ateş, 2019).



Şekil 23. Hamming ve Hanning Fonksiyonu Farkı

Hamming ve Hanning fonksiyonları aşağıdaki gibidir (Eskidere & Ertaş, 2009).

Hamming fonksiyonu:

$$w[k + 1] = 0.54 - 0.46 \cos\left(2\pi \frac{k}{N-1}\right) \quad k = 0, \dots, N-1$$

Hanning fonksiyonu:

$$w[k + 1] = 0.5 \left(1 - \cos\left(2\pi \frac{k}{N-1}\right)\right), \quad k = 0, \dots, N-1$$

v. Fast Fourier Transform

Konuşma sinyali FFT ile zaman alanından frekans alanına dönüştürülmüştür. FFT, güç spektrumu analizi ve filtre simülasyonu gibi sinyal analizini kolaylaştıran bir basamaktır (Cochran, 1967).

vi. Mel Filtre Bankası

FFT adımından sonra frekans spektrumuna insan kulağının algılama modeline benzer birbiriyle örtecek üçgensel pencere şeklindeki Mel ölçekli süzgeç bankası uygulanmıştır. 1 kHz'e kadar olan kısmı lineer, 1 kHz'in üstündeki kısımları logaritmik olarak dağılmıştır.

vii. Giriş Katmanı

CNN' in ilk basamağı olan giriş katmanı, ham verileri girdi olarak almaktadır. Giriş verilerini doğru etiketlemek performans ve eğitim süresi açısından önemlidir (Tan, 2019).

Bu çalışmada kullanılan ses dosyaları Drama ve Oyunculuk Bölümü'nde öğrenim gören öğrencilerin ses kayıtlarından oluşmaktadır. Ses veri tabanı 6 sınıfa (nötr, mutluluk, hüzün, korku, öfke, şaşkınlık) ait, sınıf başına 400 dosya olmak üzere toplam 2400 manuel olarak bölümlenmiş wav dosyalarından oluşmaktadır. Veri tabanındaki tüm sinyaller 24 bit çözünürlüğe ve 44100 Hz örnekleme frekansına sahiptir.

viii. Konvolüsyon Katmanı

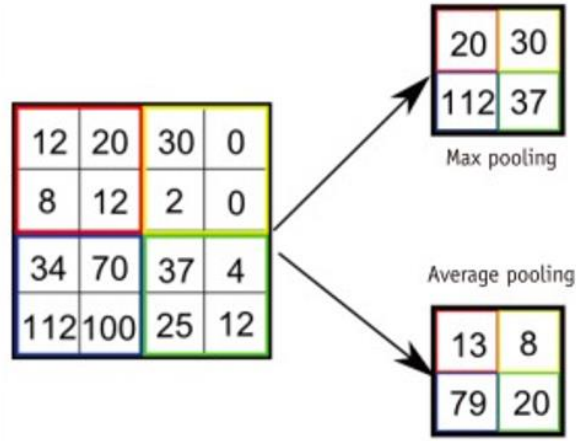
Görüntü üzerinde yer alan kare filtre kombinasyonlarından birinin sol üst köşesinden başlanarak, filtrenin merkezi görüntü üzerindeki her bir pikselden geçecek şekilde görüntünün taranmasına konvolüsyon işlemi denilmektedir (Güler, 2018). Konvolüsyon işlemi uygulanarak hem çıkış verisi oluşturulur hem de öznitelik haritası çıkarılır.

Konuşma sinyali için 2D CNN, video sinyali için ise 3D CNN kullanılmaktadır. Çalışmamızda konuşma sinyalleri kullanıldığı için 2D CNN kullanılmıştır. Tam bağlı katmanın çıktısına bir Tanh aktivasyon fonksiyonu uygulanmıştır. Tanh fonksiyonunun çıktısı daha sonra sınıflandırıcıya girdi olmuş ve Softmax fonksiyonu ile sınıflandırma yapılmıştır.

ix. Havuzlama Katmanı

Özellik azaltma havuzlama işleminin temel amacıdır. Havuzlama işlemi ile kendinden sonra gelecek olan konvolüsyon katmanına girdi olarak genişlik ve yükseklik bakımından daha düşük boyutta veri sağlanmıştır. Havuzlama sayesinde yapay sinir ağlarının ezberlemesinin önüne geçilir.

Maksimum havuzlama ve ortalama havuzlama olmak üzere iki çeşidi vardır (Şekil 24). Maksimum havuzlama katmanları olmayan CNN, tüm giriş görüntüsü üzerinde ilk katmandaki tüm konvolüsyonların hesaplanması ve ardından ortaya çıkan genişletilmiş haritalar üzerinde sonraki katmanlardaki tüm konvolüsyonların hesaplanmasıyla önemsiz bir şekilde uygulanır (Nagi ve ark., 2011). Çalışmamızda maksimum havuzlama yöntemi kullanılmıştır.



Şekil 24. Maksimum ve Ortalama Havuzlama Yöntemleri

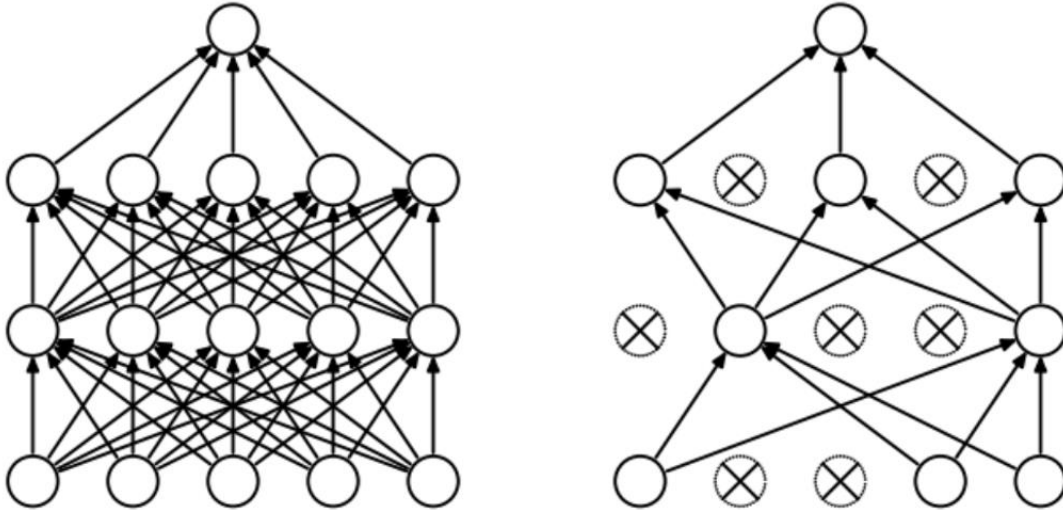
Kaynak: Lee ve ark., 2017

x. Tam Bağlantılı Katman

CNN' in son aşamalarından biridir. Yapay sinir ağının öğrenmesi bu aşamada gerçekleşmiştir. Önceki katmanlardan gelen tüm özellikler birleştirilerek sınıflandırma için zemin hazırlanmıştır.

xi. Dropout Katmanı

Yapay sinir ağının ezber yapmasının önüne geçmek amacıyla dropout aşamasında ağdaki düğümlerden bazıları kaldırılmıştır.



Şekil 25. Dropout Uygulanmadan Önce (A) ve Dropout Uygulandıktan Sonra CNN

xii. Sınıflandırma

CNN' in son aşaması olan sınıflandırma aşamasında veriler Softmax fonksiyonu kullanılarak sınıflandırılmıştır. Çalışmamızda kullanılan CNN mimarisi aşamalarına ait bilgiler Çizelge 8'de verilmiştir.

Çizelge 8. CNN Mimarisi

Layers (Type)	Kernel Size	Output Shape	No. Parameters
2D convolution	3*3	3158*64	640
Max-pooling	2*2	1579*64	0
2D convolution	3*3	1578*128	73856
Max-pooling	2*2	789*128	0
Droupout (0.5)	-	-	-
Dense	-	1*1024	1049600
Dense	-	1*7	7175

3. Prototipin Canlıya Geçirilmesi (Klinik Uygulama)

Çalışmamızın örneklem grubunu oluşturan normal işiten bireyler ve işitme kaybı olan bireylerin ses duygu kayıtları prototipin canlıya geçirilme aşamasında kullanılmıştır.

a. Normal İşiten Bireylerin Test Aşamaları

Çalışmaya dahil edilecek olan normal işiten bireyler, İstanbul Aydın Üniversitesi'nde çalışan akademisyenlerden seçilmiştir. İstanbul Aydın Üniversitesi Odyoloji Laboratuvarında otoskopik bakı, saf ses odyometri testi ve immitansmetrik değerlendirmeler yapıldıktan sonra belirli kriterleri sağlayan bireyler çalışmaya dahil edilmiştir.

Welch Allyn otoskop ünitesi ile dış kulak kanalı ve kulak zarı muayenesi yapılmıştır. 226 Hz Prob ton kullanılarak MADSEN Otoflex 100 marka immitansmetre ile immitansmetrik değerlendirme yapılmıştır. 0,3 ml'den daha büyük statik kompliyans değeri ve -100 ile +500 daPa aralığındaki timpanik tepe basınç değeri normal olarak kabul edilmiştir. Jerger sınıflaması kullanılarak Tip A timpanogram elde edilen bireyler çalışmaya dahil edilmiştir.

Saf ses odyometri testi, Otometrics Madsen Astera2 marka klinik odyometre cihazı ile yapılmıştır. Saf ses odyometri testi; hava yolu işitme eşikleri, 125 Hz-

8kHz aralığında olan tüm oktav frekanslarında, kemik yolu işitme eşikleri, 500 Hz- 4kHz aralığında olan tüm oktav frekanslarında ascending yöntemi kullanılarak yapılmıştır. Hava yolu işitme eşikleri belirlenirken Telephonics TDH-39 supraaural kulaklık; kemik yolu işitme eşikleri belirlenirken Radioear B-71 kemik vibratör kullanılmıştır. Saf ses odyometri test sonuçlarını değerlendirme de Goodman İşitme Kaybı Sınıflandırması kullanılmıştır. (Çizelge 9)

Çizelge 9. Goodman İşitme Kaybı Sınıflandırması

Goodman Sınıflandırması	Saf Ses Ortalamaları (dBHL) (0.5, 1 ve 2 kHz)
<26 dB	Normal
26- 40 dB	Çok Hafif
41- 55 dB	Hafif
56- 70 dB	Orta
71- 90 dB	İleri
91 dB ve üstü	Çok İleri

Konuşma odyometrisi testi kapsamında Konuşmayı Alma Eşiği ve Konuşmayı Ayırt Etme Skoru elde edilmiştir. Konuşma odyometrisi testleri TDH-39 supraaural kulaklık kullanılarak yapılmıştır.

Konuşmayı Alma Eşiği (KAE), Türkçe fonetik dengeli üç heceli kelime listesi kullanılarak ascending yöntemi ile değerlendirilmiştir. KAE, Konuşmayı Ayırt Etme Testinin uygulanacağı şiddet düzeyini belirlemede yol göstericidir.

Konuşmayı Ayırt Etme Skoru, bireylerin en rahat ettiği ses seviyesinde Türkçe fonetik tek heceli kelime listesi kullanılarak elde edilmiştir. Test, taşıyıcı cümle kullanılarak canlı ses ile yapılmıştır.

İstanbul Aydın Üniversitesi'nde çalışan 18- 40 yaş aralığındaki 20 akademisyen otoskopik muayene, immitansmetrik değerlendirme, saf ses odyometri testi ve konuşma odyometrisi ile değerlendirildikten sonra yukarıda belirtilen kriterleri sağlayan bireyler normal işiten grubunu oluşturmak üzere çalışmaya dahil edilmiştir.

Çalışmaya dahil edilen normal işiten bireyler, iOS işletim sistemine sahip cep telefonlarını kullanarak 5 cümleyi, 6 farklı duyguda seslendirmiştir. Seslendirme yapan normal işiten bireylere, ağızları ile telefon mikrofonu arasındaki mesafenin 25 cm olması ve seslendirme yapılan ortamın gürültüsüz olması gerektiği ses kaydı öncesinde belirtilmiştir.

b. İşitme Kayıplı Bireylerin Test Aşamaları

İleri- çok ileri derecede sensörinöral işitme kaybına sahip olup koklear implant kullanan 20 birey çalışmaya dahil edilmiştir. Koklear implant kullanan bireyler, İstanbul ilindeki Özel Eğitim ve Rehabilitasyon Merkezleri'nde rehabilitasyon eğitimi alan 18- 40 yaş aralığındaki bireylerden seçilmiştir.

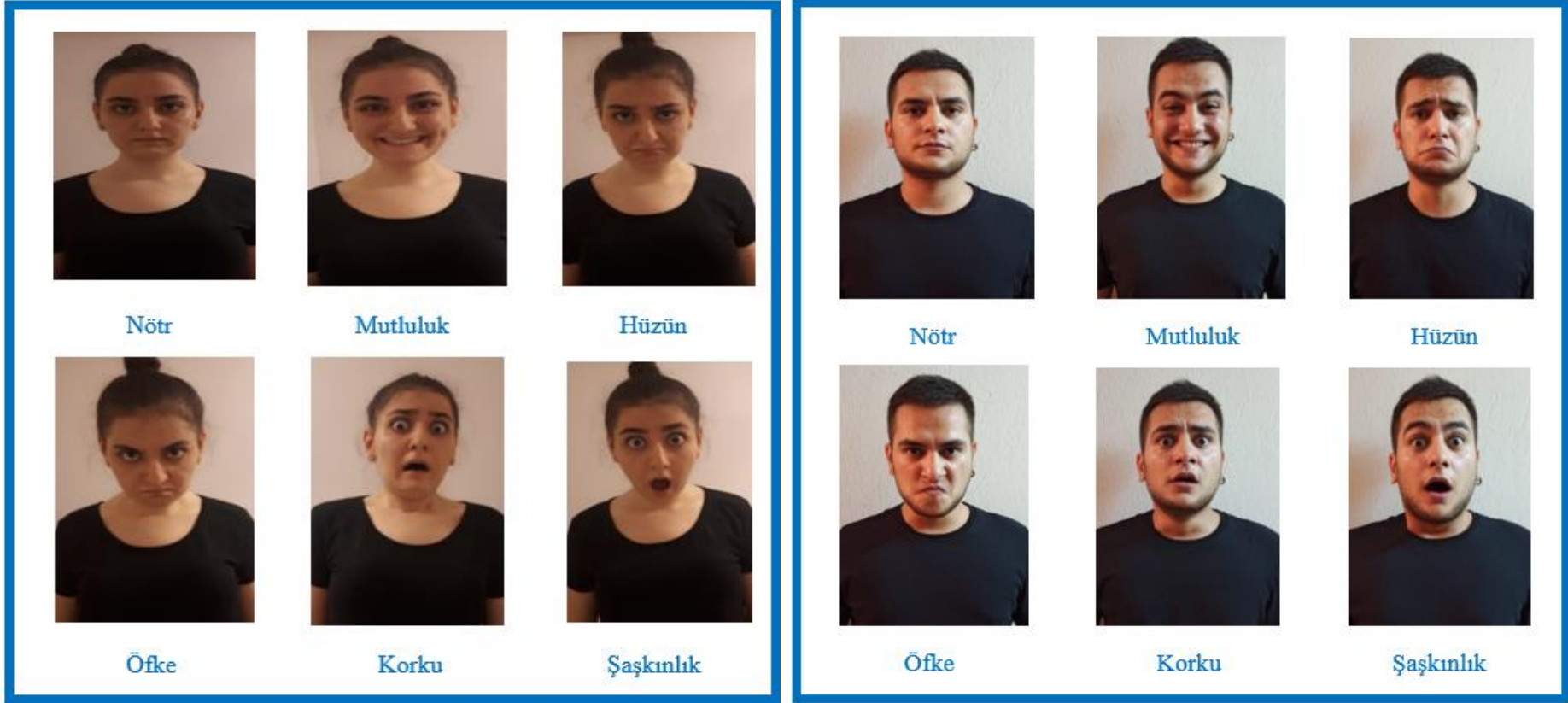
İşitme kayıplı bireylerin ses kayıtları eğitim öncesi ve eğitim sonrası olmak üzere toplamda iki aşamada alınmıştır. Koklear implant kullanan bireylerden eğitim öncesi, iOS işletim sistemine sahip cep telefonları ile 5 cümlenin, 6 farklı duyguda ses kaydı alınmıştır.

Eğitim için resim ve animasyon olmak üzere iki ayrı eğitim materyali kullanılmıştır. İAÜ Drama ve Oyunculuk Bölümü'nde eğitim gören bir kadın ve bir erkek öğrenci tarafından nötr, mutluluk hüznü, öfke, korku ve şaşkınlık olmak üzere altı farklı duyguda yüz ifadelerini gösteren resimler oluşturulmuştur. Kadın ve erkek yüz ifadeleri Şekil 26'da gösterilmiştir. Eğitim aşamasında önce altı farklı duyguyu ifade eden ses kayıtlarına uygun bu yüz ifadeleri gösterilmiştir.

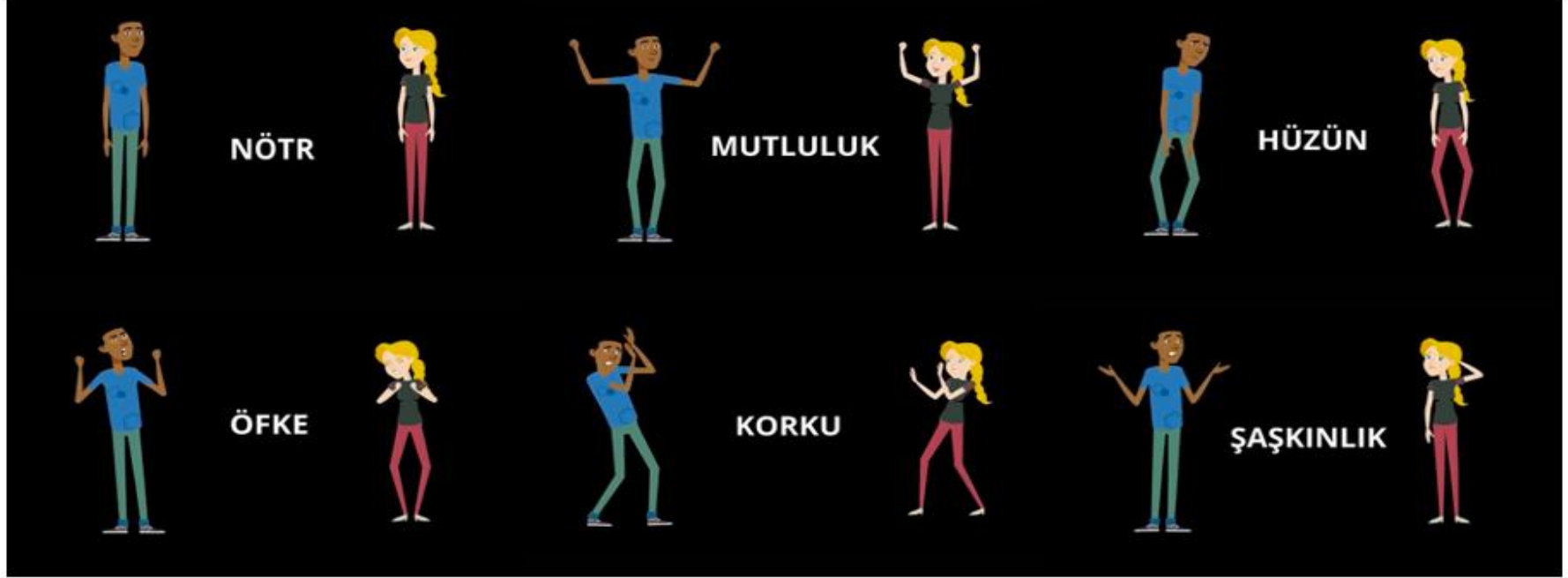
Daha sonra eğitimi pekiştirmek için bu bireylere, veri setindeki öğrencilerin ses duygu kayıtlarına göre görsel ve işitsel ipuçlarını birleştiren animasyon izletilmiştir (Şekil 27). Yukarıdaki beş cümlenin altı farklı duyguda ses kayıtlarına uygun, kadın ve erkek animasyon figürlerinin yüz ifadesi (mimik) ve diğer vücut hareketleri ile oluşturulan beş animasyon modülü eğitim materyali olarak kullanılmıştır.

Bu iki eğitim modülü Zoom aracılığıyla 3'lü, 4'lü gruplar halinde bir kez 2-3 saat toplanılarak araştırmacı tarafından yukarıdaki sırasıyla verilmiştir.

Eğitim aşaması tamamlandıktan sonra işitme kayıplı bireyler 5 cümleyi, 6 farklı duyguda tekrar seslendirmiştir. Seslendirme yapan işitme kayıplı bireylere, ağızları ile telefon mikrofonu arasındaki mesafenin 25 cm olması ve seslendirme yapılan ortamın gürültüsüz olması gerektiği ses kaydı öncesinde belirtilmiştir.



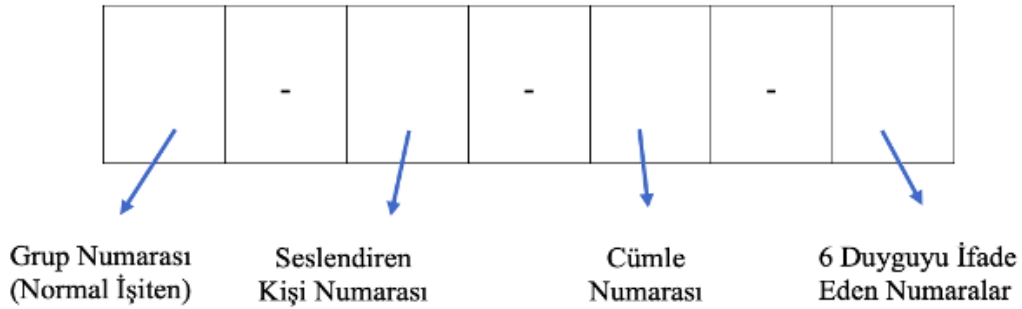
Şekil 26. Altı Farklı Duygu İçin Kadın ve Erkek Yüz İfadeleri



Şekil 27. Animasyon Modülü Eğitim Materyalinden Bir Kesit

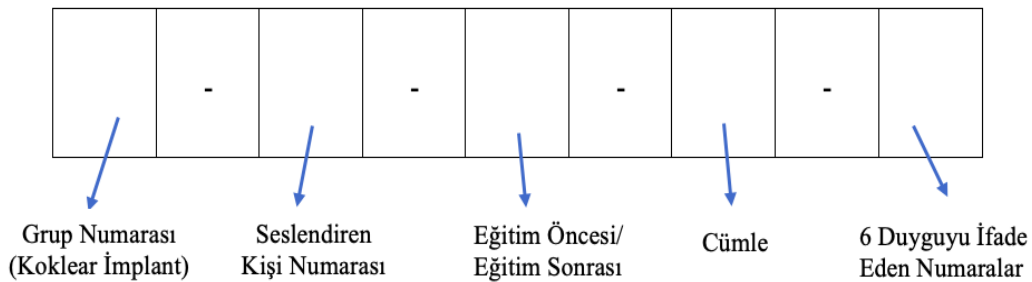
c. Normal İşiten ve İşitme Kayıplı Bireylerin Ses Kayıtlarının Etiketlenmesi

Normal işiten bireyler (20), beş cümleye ait altı farklı duyguyu seslendirmiştir. Her bireye ait 30 ses kaydı, toplamda 600 ses kaydı elde edilmiştir. Normal işiten bireylerin her cümle arasında 5 saniye bırakarak oluşturduğu kayıt, Audacity programında her cümle farklı bir wav dosyası olacak şekilde kırpıldıktan sonra yeniden etiketlenmiştir. Kayıtların etiketlenmesi Şekil 28'de gösterilmiştir.



Şekil 28. Normal İşiten Bireylerin Ses Kayıtlarının Etiketlenmesi

İşitme kayıplı bireyler (20), beş cümleye ait altı farklı duyguyu eğitim öncesi ve eğitim sonrası olmak üzere toplamda iki kez seslendirmiştir. Her bireye ait 60 ses kaydı, toplamda 1200 ses kaydı elde edilmiştir. İşitme kayıplı bireylerin her cümle arasında 5 saniye bırakarak oluşturduğu kayıt, Audacity programında her cümle farklı bir wav dosyası olacak şekilde kırpıldıktan sonra yeniden etiketlenmiştir. Kayıtların etiketlenmesi Şekil 29'da gösterilmiştir.



Şekil 29. Koklear İmplant Kullanan Bireylerin Ses Kayıtlarının Etiketlenmesi

i. Grup Numarası

Normal işiten bireyler 1, koklear implant kullanan bireyler 2 olarak temsil edilmiştir.

ii. Seslendiren Kişi Numarası

Normal işiten ve koklear implant kullanan bireylere 1’den 20’ye kadar olan sayılar verilmiştir.

iii. Eğitim Durumu

Sadece işitme kayıplı bireylerin ses kayıtlarının adlandırılmasında kullanılmıştır. İlk kayıtlar 0, eğitim sonrası seslendirilen ikinci kayıtlar 1 ile ifade edilmiştir.

iv. Cümle Numaraları

1’ den 5’e kadar olan sayılar, her bir cümle için farklı olacak şekilde verilmiştir.

Çizelge 10. Test Aşaması/ Cümle Numaraları

Cümleler	Numara
Hızlıca eve döndüm.	1
Birden ağlamaya başladı.	2
Akşam erkenden uyudum.	3
Çatal elinden düştü.	4
Hastaneye gitmek istemedi.	5

v. Duyguları İfade Eden Numaralar

1’den 6’ ya kadar olan sayılar, her bir sayı farklı duyguyu temsil edecek şekilde verilmiştir.

Çizelge 11. Test Aşaması/ Duyguları İfade Eden Numaralar

Duygu İfadeleri	Numara
Nötr	1
Mutluluk	2
Hüzün	3
Öfke	4
Korku	5
Şaşkınlık	6

Normal işiten (20 birey) ve işitme kayıplı bireylerden (20 birey) toplamda 1200 (5 cümle x 6 duygu x 40 birey) ses kaydı toplanmıştır. Derin öğrenme ile duyguları sözel ifade etme becerileri analiz edilmiştir. Normal işiten bireyler ile

koklear implant kullanan bireylerin eğitim öncesi ve eğitim sonrası duyguları sözel ifade etme becerileri grup içi ve gruplar arasında karşılaştırılmıştır.

IV. BULGULAR

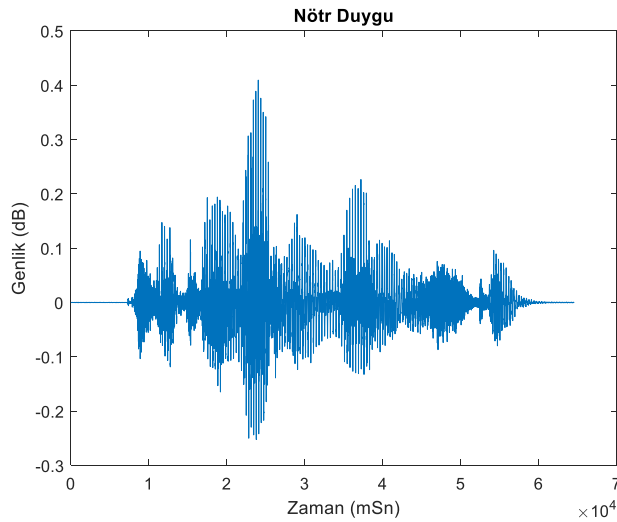
Çalışmamızda derin öğrenme yöntemi kullanılarak, normal işiten bireyler ile koklear implant kullanan bireylerin ve koklear implant kullanan bireylerin eğitim öncesi ve eğitim sonrası ses duygu analizleri karşılaştırılmıştır.

A. Derin Öğrenme Algoritması Sınıflandırma Bulguları

Bu bölümde verilerin çok olması sebebiyle yalnızca ‘Çatal Elinden Düştü’ cümlesi nötr, nötr, mutluluk, hüznün, öfke, korku ve şaşkınlık duygularına ait ses işareti, histogram, Rose, Ks yoğunluk analizi ve normal dağılım için olasılık analizleri değerlendirilmiştir. Yapılan ses kayıtlarının derin öğrenme analizleri aşağıdaki grafiklerde verilmiştir. Cümlelerin farklı altı duygu ile seslendirilmesi sonucunda her duygunun farklı bir karakteristik çizdiği analizlerden görülebilmektedir.

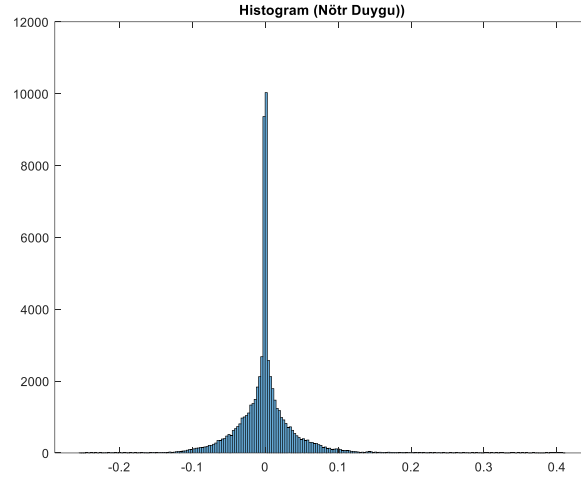
1. Nötr Duygu Analizi

Şekil 30’da ‘Çatal Elinden Düştü’ cümlesinin nötr duygusu ile seslendirilmiş ham ses işareti verilmiştir.



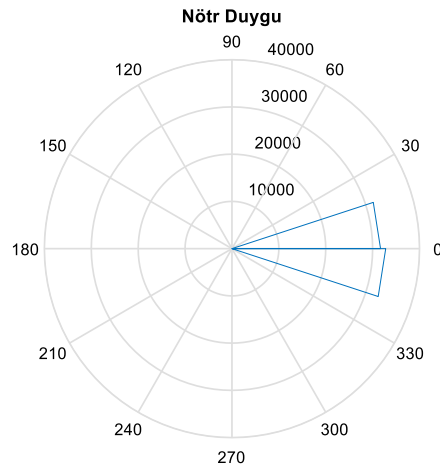
Şekil 30. ‘Çatal Elinden Düştü’ Cümlesine Ait Nötr Duygu Ham Ses İşareti

Görüntü histogramı, sayısal bir görüntü içerisinde yer alan renk değerlerinden her birinin kaç adet olduğunu gösteren bir grafikdir. Şekil 31’de ‘Çatal Elinden Düştü’ cümlesinin nötr duygu ile seslendirilmiş histogramı verilmiştir.



Şekil 31. ‘Çatal Elinden Düştü’ Cümlesine Ait Nötr Duygu Histogramı

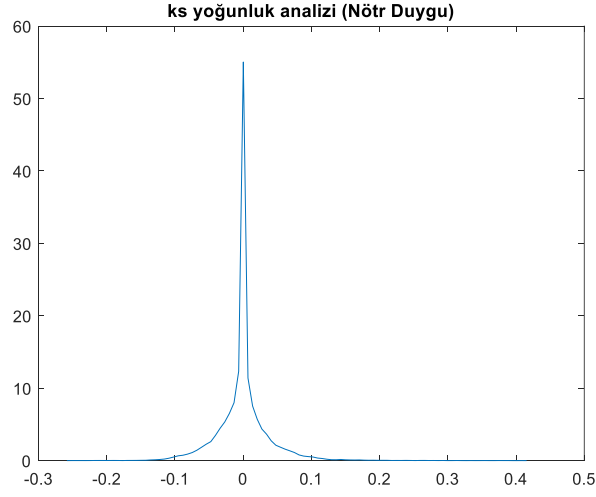
Şekil 32’de ‘Çatal Elinden Düştü’ cümlesinin nötr duygu ile seslendirilmiş kayıtlarının rose analizi verilmiştir. Rose analizi özellikle büyük farklılıkların belirlenmesinde oldukça ayırt edici bir yöntemdir. Rose analizinde küçük açısıl farklılıklar olmasına rağmen duygu seslendirmelerinde gözle ayırt edilebilecek büyük farklılıklar gözlemlenmez.



Şekil 32. ‘Çatal Elinden Düştü’ Cümlesine Ait Nötr Duygu Rose Analizi

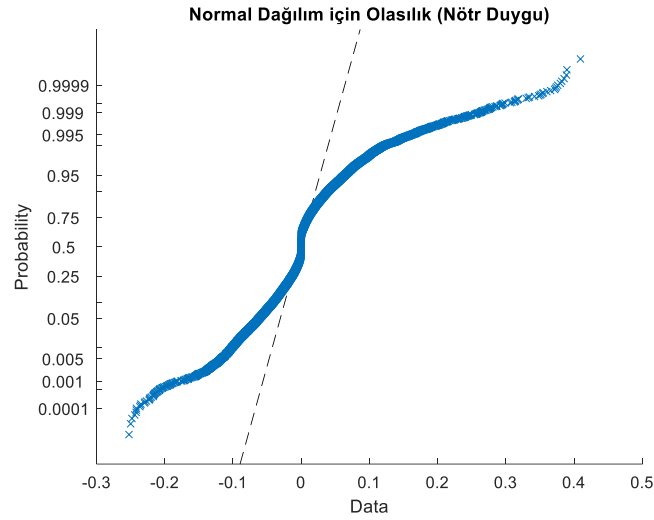
Şekil 33’ de ‘Çatal Elinden Düştü’ cümlesinin Ks yoğunluk analizi nötr duygusu için gösterilmiştir. Burada yatay eksendeki dağılım -0.1 ile +0.1 arasında

değişim gösterirken genlik 55 değerini almaktadır. Bu değerler diğer duygularda farklılıklar göstermekte olup bu analiz ayırt edici özelliklerden birisi olarak kullanılabilir.



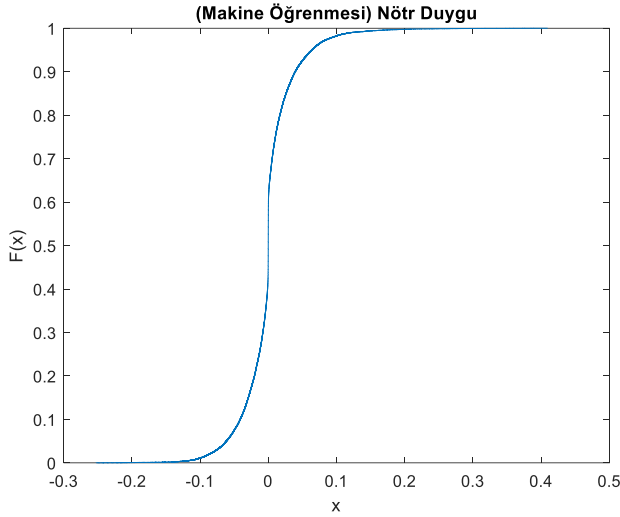
Şekil 33. 'Çatal Elinden Düştü' Cümlesine Ait Nötr Duygusu Ks Yoğunluk Analizi

Şekil 34' de 'Çatal Elinden Düştü' cümlesinin normal dağılım için olasılık grafiği nötr duygusu için analiz edilmiştir.



Şekil 34. 'Çatal Elinden Düştü' Cümlesine Ait Nötr Duygusu Normal Dağılım İçin Olasılık Grafiği

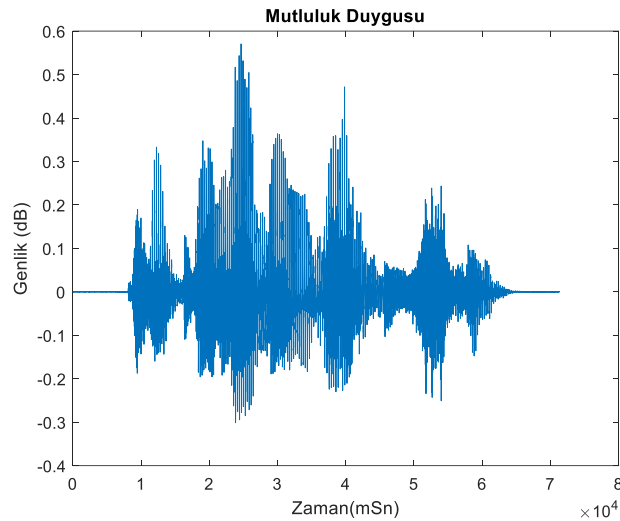
Şekil 35’de nötr duygusuna ait makine öğrenmesi sonucu verilmiştir. Diğer duyguların da makine öğrenmesi analizleri yapılmış, ancak fark görülemediğinden sadece nötr duygusuna ait analiz verilmiştir.



Şekil 35. ‘Çatal Elinden Düştü’ Cümlesine Ait Nötr Duygusu İçin Makine Öğrenmesi Analizi

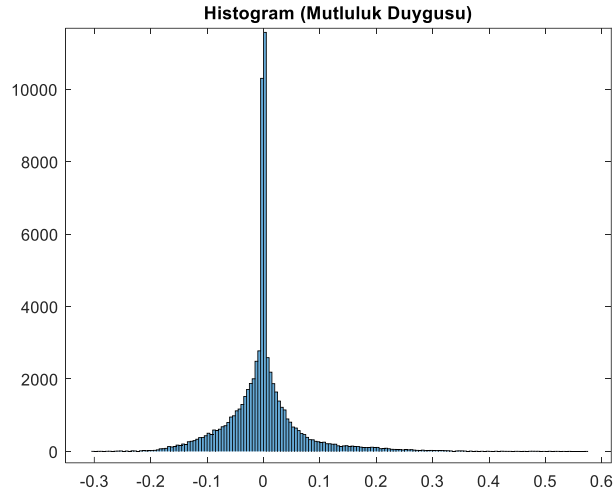
2. Mutluluk Duygu Analizi

Şekil 36’da ‘Çatal Elinden Düştü’ cümlesinin mutluluk duygusu ile seslendirilmiş ham ses işareti verilmiştir.



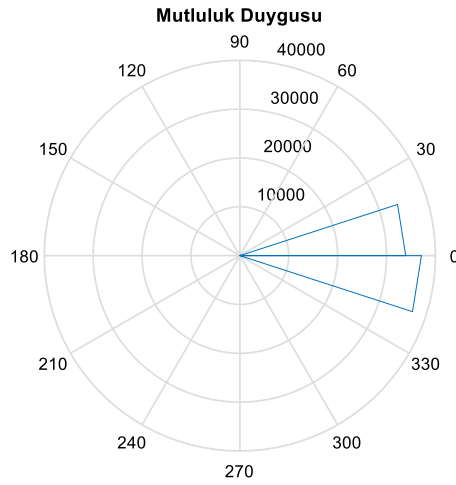
Şekil 36. . ‘Çatal Elinden Düştü’ Cümlesine Ait Mutluluk Duygusu Ham Ses İşareti

Şekil 37’ de ‘Çatal Elinden Düştü’ cümlesinin mutluluk duygusuna ait histogram analizi yapılmıştır. Burada genlik 1200 değerini, yatay eksendeki dağılım ise -0.3 ile +0.6 arasında değişim göstermektedir. Bu değişimler diğer duygularda farklı değerler alarak ayırt etmede kullanılabilir.



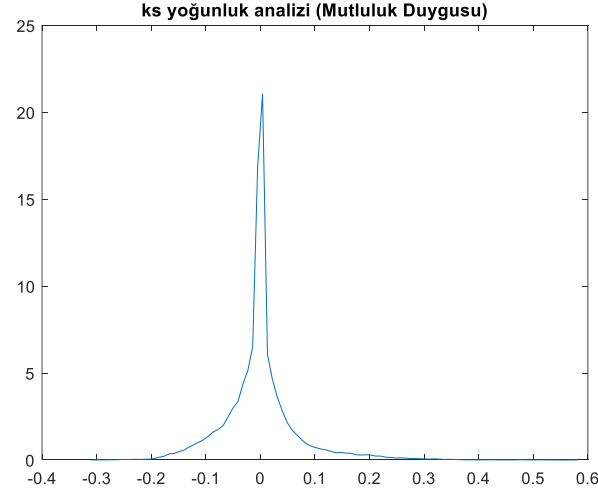
Şekil 37. . ‘Çatal Elinden Düştü’ Cümlesine Ait Mutluluk Duygusu Histogramı

Şekil 38’ de ‘Çatal Elinden Düştü’ cümlesinin mutluluk duygusu ile seslendirilmiş kayıtlarının Rose analizi verilmiştir.



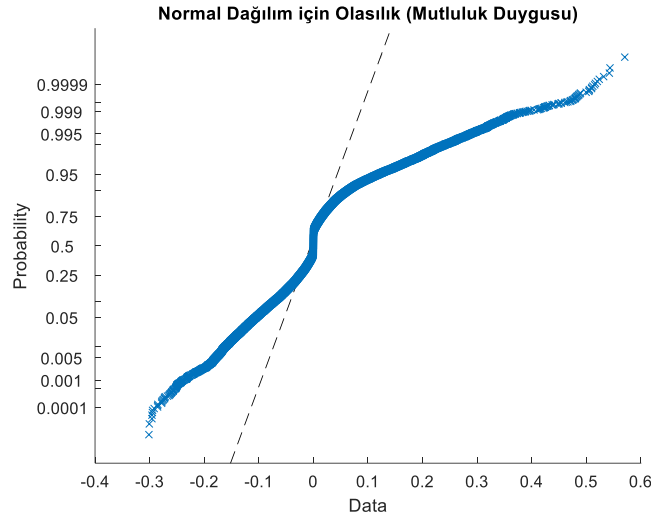
Şekil 38. ‘Çatal Elinden Düştü’ Cümlesine Ait Mutluluk Duygusu Rose Analizi

Şekil 39’da ‘Çatal Elinden Düştü’ cümlesinin Ks yoğunluk analizi mutluluk duygusu için gösterilmiştir.



Şekil 39. 'Çatal Elinden Düştü' Cümlesine Ait Mutluluk Duygusu Ks Yoğunluk Analizi

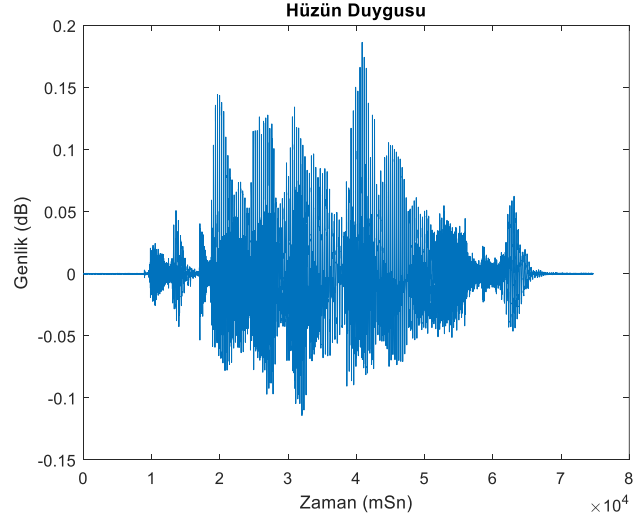
Şekil 40' da 'Çatal Elinden Düştü' cümlesinin normal dağılım için olasılık grafiği mutluluk duygusu için analiz edilmiştir.



Şekil 40. 'Çatal Elinden Düştü' Cümlesi Mutluluk Duygusuna Ait Normal Dağılım İçin Olasılık Grafiği

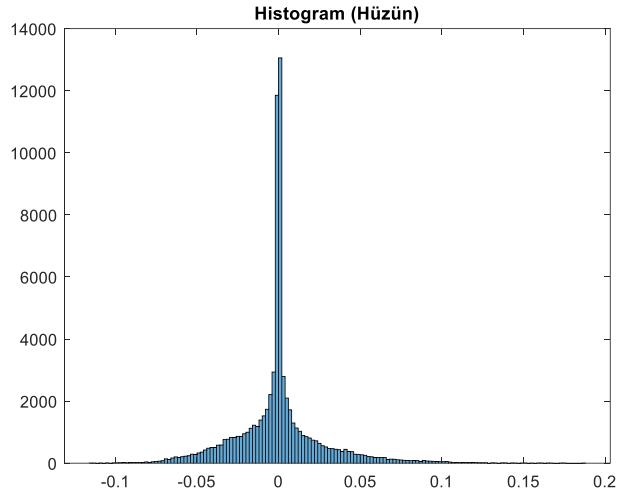
3. Hüzün Duygu Analizi

Şekil 41'de 'Çatal Elinden Düştü' cümlesinin hüzün duygusu ile seslendirilmiş kayıtlarının ham ses işareti verilmiştir.



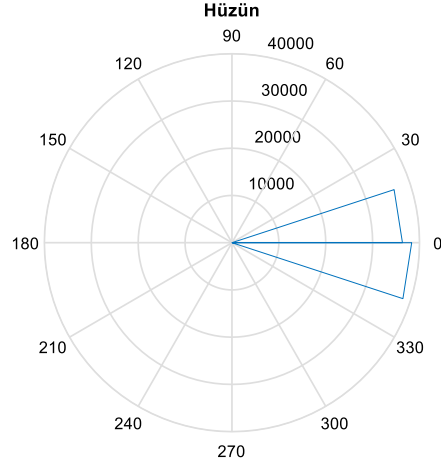
Şekil 41. 'Çatal Elinden Düştü' Cümlesine Ait Hüzün Duygusu Ham Ses İşareti

Şekil 42'de 'Çatal Elinden Düştü' cümlesinin hüzün duygu ile seslendirilmiş histogramı verilmiştir.



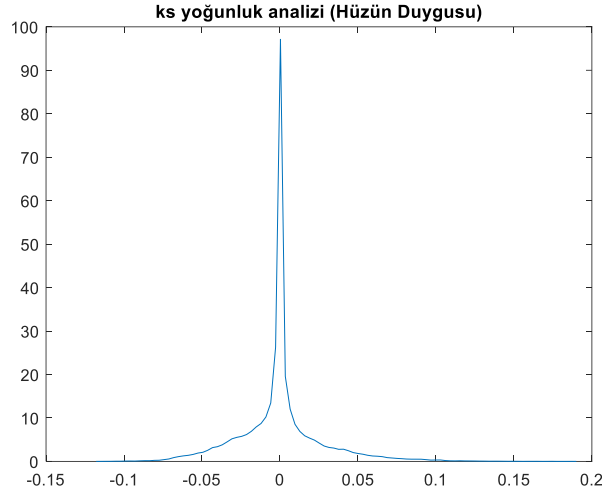
Şekil 42. 'Çatal Elinden Düştü' Cümlesine Ait Hüzün Duygusu Histogramı

Şekil 43' de 'Çatal Elinden Düştü' cümlesinin hüzün duygu ile seslendirilmiş kayıtlarının Rose analizi verilmiştir.



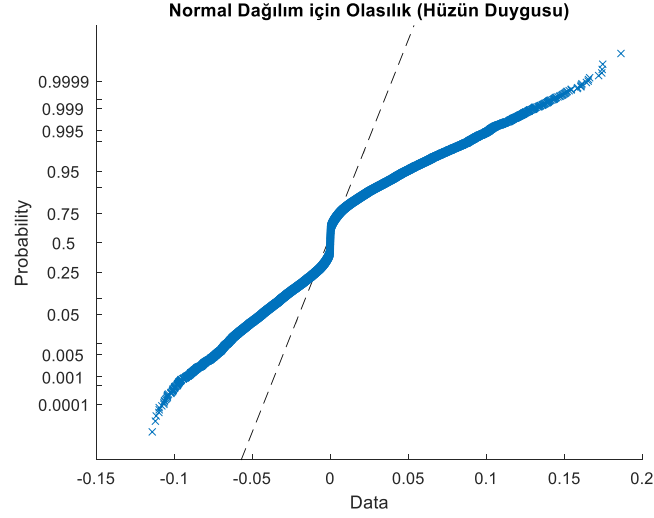
Şekil 43. ‘Çatal Elinden Düştü’ Cümlesine Ait Hüzün Duygusu Rose Analizi

Şekil 44’ de ‘Çatal Elinden Düştü’ cümlesinin Ks yoğunluk analizi hüzün duygusu için gösterilmiştir.



Şekil 44. ‘Çatal Elinden Düştü’ Cümlesine Ait Hüzün Duygusu Ks Yoğunluk Analizi

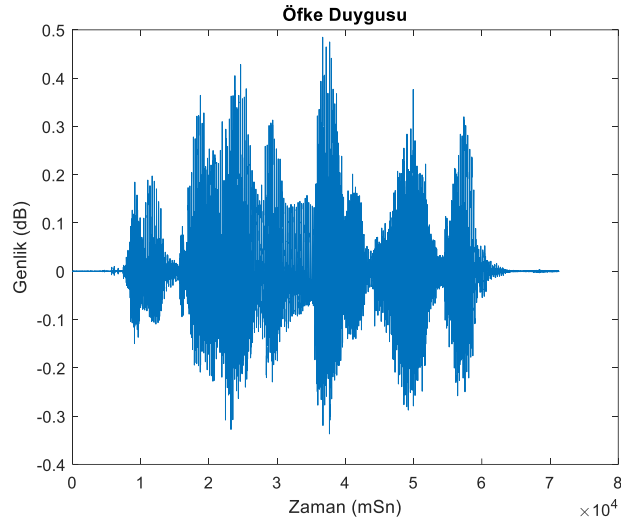
Şekil 45’ de ‘Çatal Elinden Düştü’ cümlesi hüzün duygusuna ait normal dağılım için olasılık grafiği gösterilmiştir. Dağılım x-y eksenine 45 derecelik açı şeklinde yer almakta olup, verilerin dağılımı karakteristik olarak 0 ile 1 olasılık dağılımı ile değişim göstermektedir. Bu dağılım diğer duygularla söylenmiş olasılık dağılımları ile farklılık gösterir. Dağılımların karşılaştırılması toplu halde sonuç bölümünde verilmiştir.



Şekil 45. 'Çatal Elinden Düştü' Cümlesi Hüzün Duygusuna Ait Normal Dağılım İçin Olasılık Grafiği

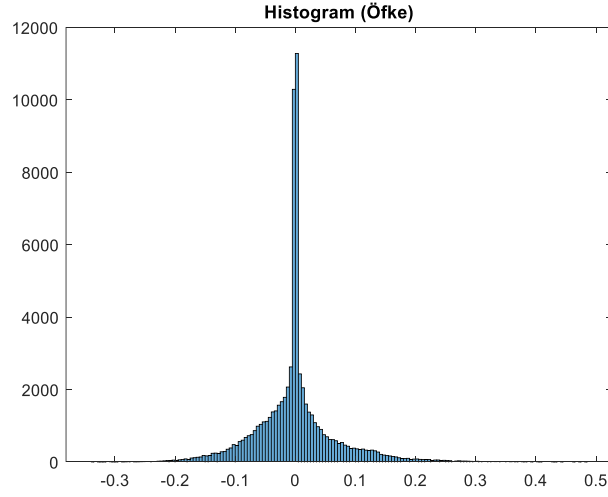
4. Öfke Duygu Analizi

Şekil 46' da 'Çatal Elinden Düştü' cümlesinin öfke duygusu ile seslendirilmiş kayıtlarının ham ses işareti verilmiştir.



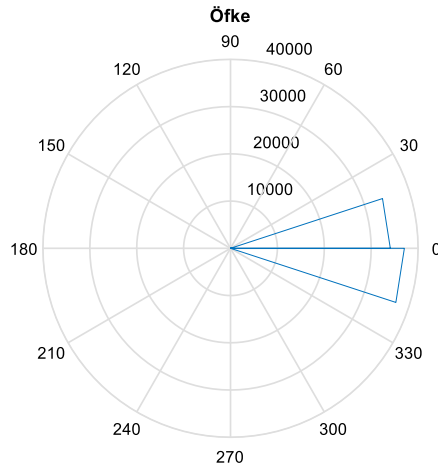
Şekil 46. 'Çatal Elinden Düştü' Cümlesine Ait Öfke Duygusu Ham Ses İşareti

Şekil 47' de 'Çatal Elinden Düştü' cümlesinin öfke duygusu ile seslendirilmiş kayıtlarının histogramı verilmiştir.



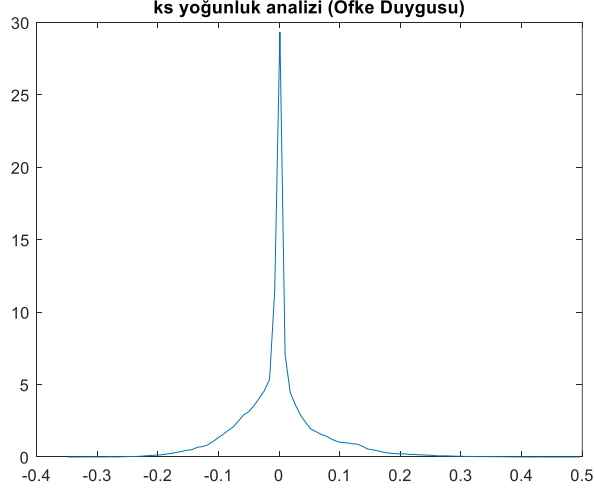
Şekil 47. 'Çatal Elinden Düştü' Cümlesine Ait Öfke Duygusu Histogramı

Şekil 48' de 'Çatal Elinden Düştü' cümlesinin öfke duygusu ile seslendirilmiş kayıtlarının Rose analizi verilmiştir.



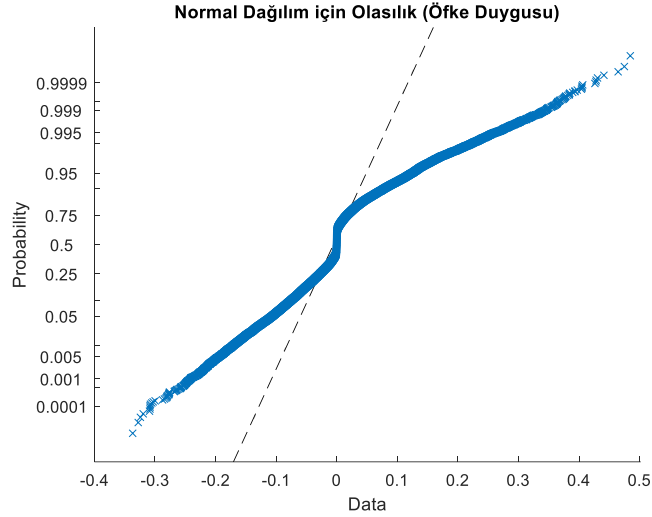
Şekil 48. 'Çatal Elinden Düştü' cümlesine ait öfke duygusu Rose analizi

Şekil 49' da 'Çatal Elinden Düştü' cümlesinin Ks yoğunluk analizi öfke duygusu için gösterilmiştir.



Şekil 49. 'Çatal Elinden Düştü' Cümlesine Ait Öfke Duygusu Ks Yoğunluk Analizi

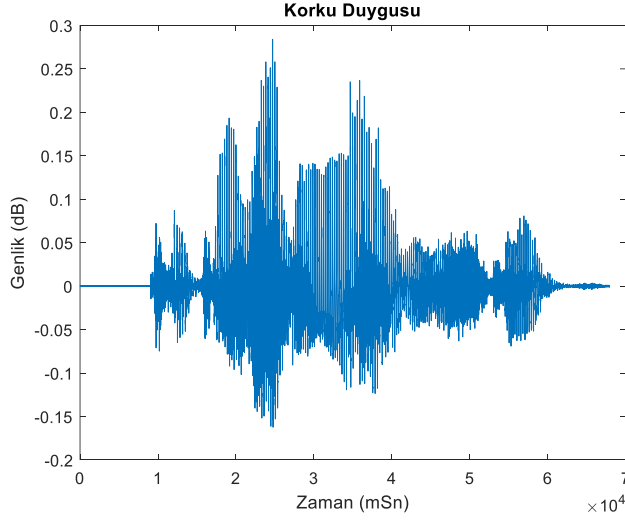
Şekil 50' de 'Çatal Elinden Düştü' cümlesi öfke duygusuna ait normal dağılım için olasılık grafiği gösterilmiştir. Dağılım x-y eksenine 45 derecelik açı şeklinde yer almakta olup, verilerin dağılımı karakteristik olarak 0 ile 1 olasılık dağılımı ile değişim göstermektedir. Bu dağılım diğer duygularla söylenmiş olasılık dağılımları ile farklılık gösterir.



Şekil 50. 'Çatal Elinden Düştü' Cümlesi Öfke Duygusuna Ait Normal Dağılım İçin Olasılık Grafiği

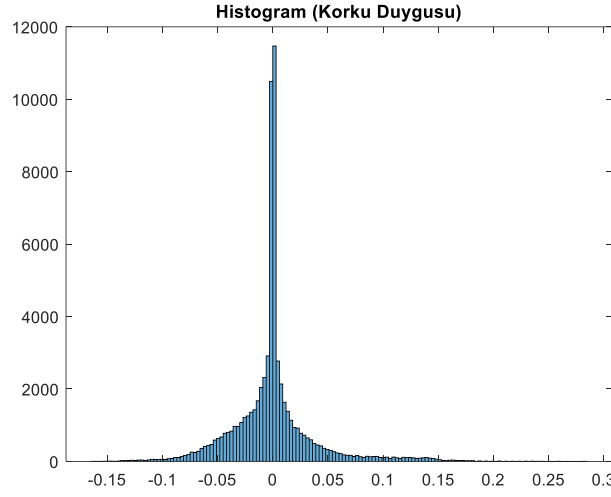
5. Korku Duygu Analizi

Şekil 51' de 'Çatal Elinden Düştü' cümlesinin korku duygusu ile seslendirilmiş kayıtlarının ham ses işareti verilmiştir.



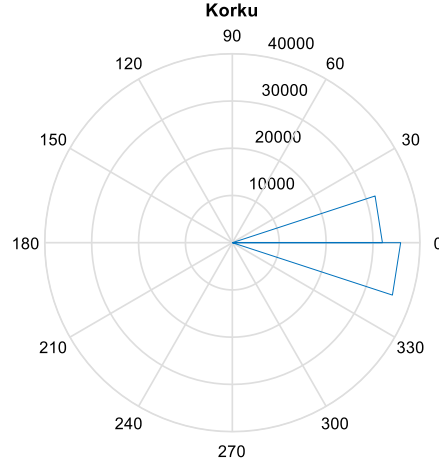
Şekil 51. 'Çatal Elinden Düştü' Cümlesine Ait Korku Duygusu Ham Ses İşareti

Şekil 52' de 'Çatal Elinden Düştü' cümlesinin korku duygusu ile seslendirilmiş kayıtlarının histogramı verilmiştir.



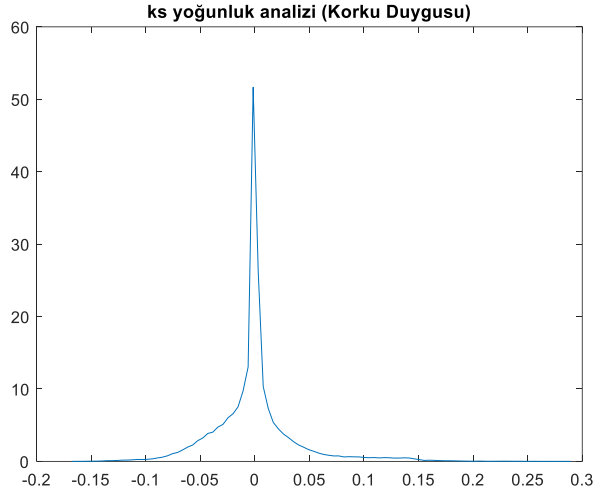
Şekil 52. 'Çatal Elinden Düştü' Cümlesine Ait Korku Duygusu Histogramı

Şekil 53' de 'Çatal Elinden Düştü' cümlesinin korku duygusu ile seslendirilmiş kayıtlarının Rose analizi verilmiştir.



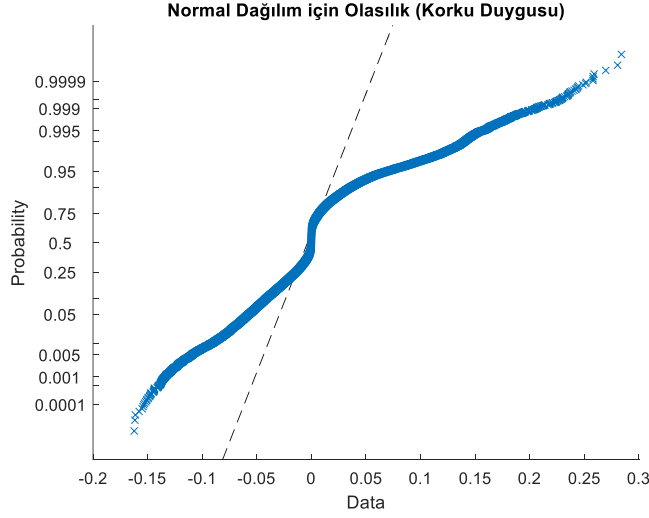
Şekil 53. 'Çatal Elinden Düştü' Cümlesine Ait Korku Duygusu Rose Analizi

Şekil 54' de 'Çatal Elinden Düştü' cümlesinin Ks yoğunluk analizi korku duygusu için gösterilmiştir.



Şekil 54. 'Çatal Elinden Düştü' Cümlesine Ait Korku Duygusu Ks Yoğunluk Analizi

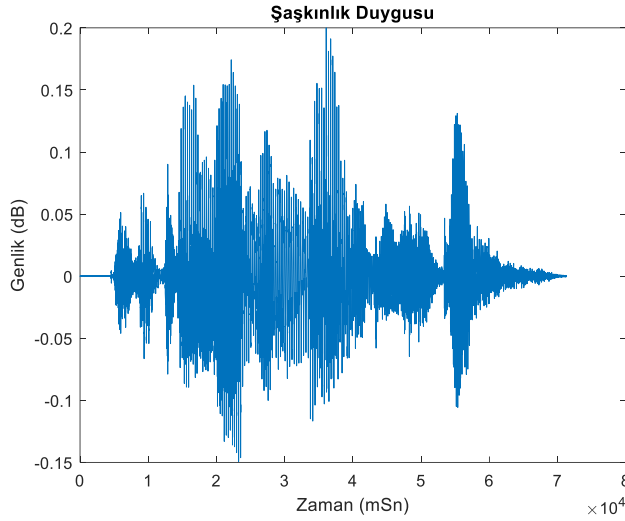
Şekil 55' de 'Çatal Elinden Düştü' cümlesi korku duygusuna ait normal dağılım için olasılık grafiği gösterilmiştir. Dağılım x-y eksenine 45 derecelik açı şeklinde yer almakta olup, verilerin dağılımı karakteristik olarak 0 ile 1 olasılık dağılımı ile değişim göstermektedir.



Şekil 55. 'Çatal Elinden Düştü' Cümlesi Korku Duygusuna Ait Normal Dağılım İçin Olasılık Grafiği

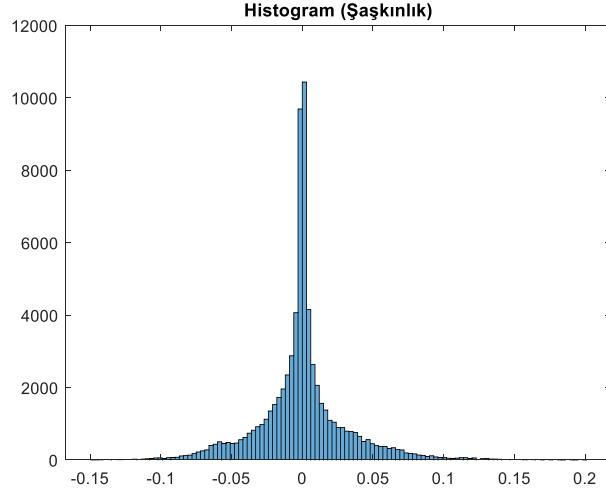
6. Şaşkınlık Duygu Analizi

Şekil 56' da 'Çatal Elinden Düştü' cümlesinin şaşkınlık duygusu ile seslendirilmiş kayıtlarının ham ses işareti verilmiştir.



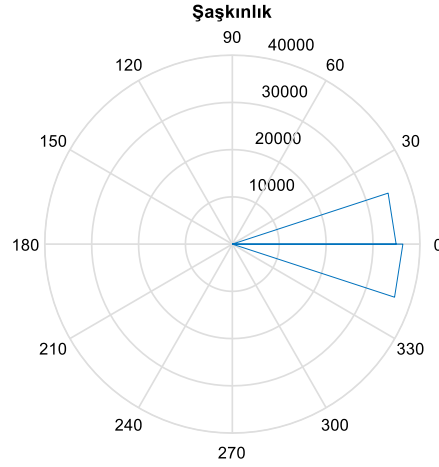
Şekil 56. 'Çatal Elinden Düştü' Cümlesine Ait Şaşkınlık Duygusu Ham Ses İşareti.

Şekil 57' de 'Çatal Elinden Düştü' cümlesinin şaşkınlık duygusu ile seslendirilmiş kayıtlarının histogramı verilmiştir.



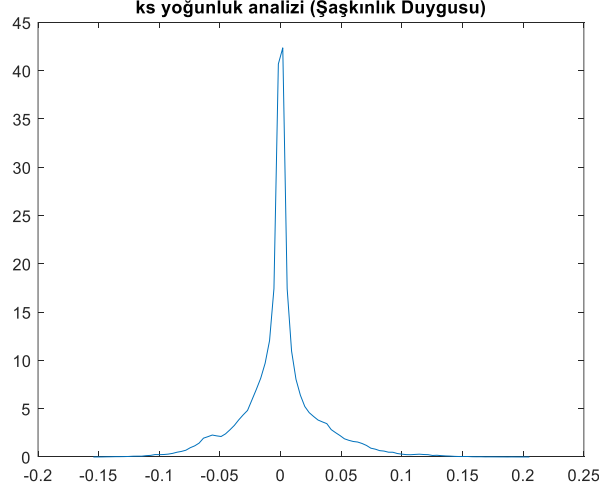
Şekil 57. ‘Çatal Elinden Düştü’ Cümlesine Ait Şaşkınlık Duygusu Histogramı

Şekil 58’ de ‘Çatal Elinden Düştü’ cümlesinin şaşkınlık duygusu ile seslendirilmiş kayıtlarının Rose analizi verilmiştir.



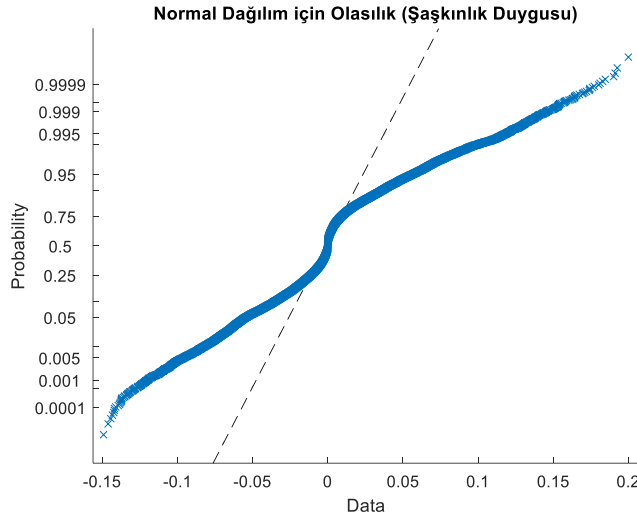
Şekil 58. ‘Çatal Elinden Düştü’ Cümlesine Ait Şaşkınlık Duygusu Rose Analizi

Şekil 59’ da ‘Çatal Elinden Düştü’ cümlesinin Ks yoğunluk analizi şaşkınlık duygusu için gösterilmiştir.



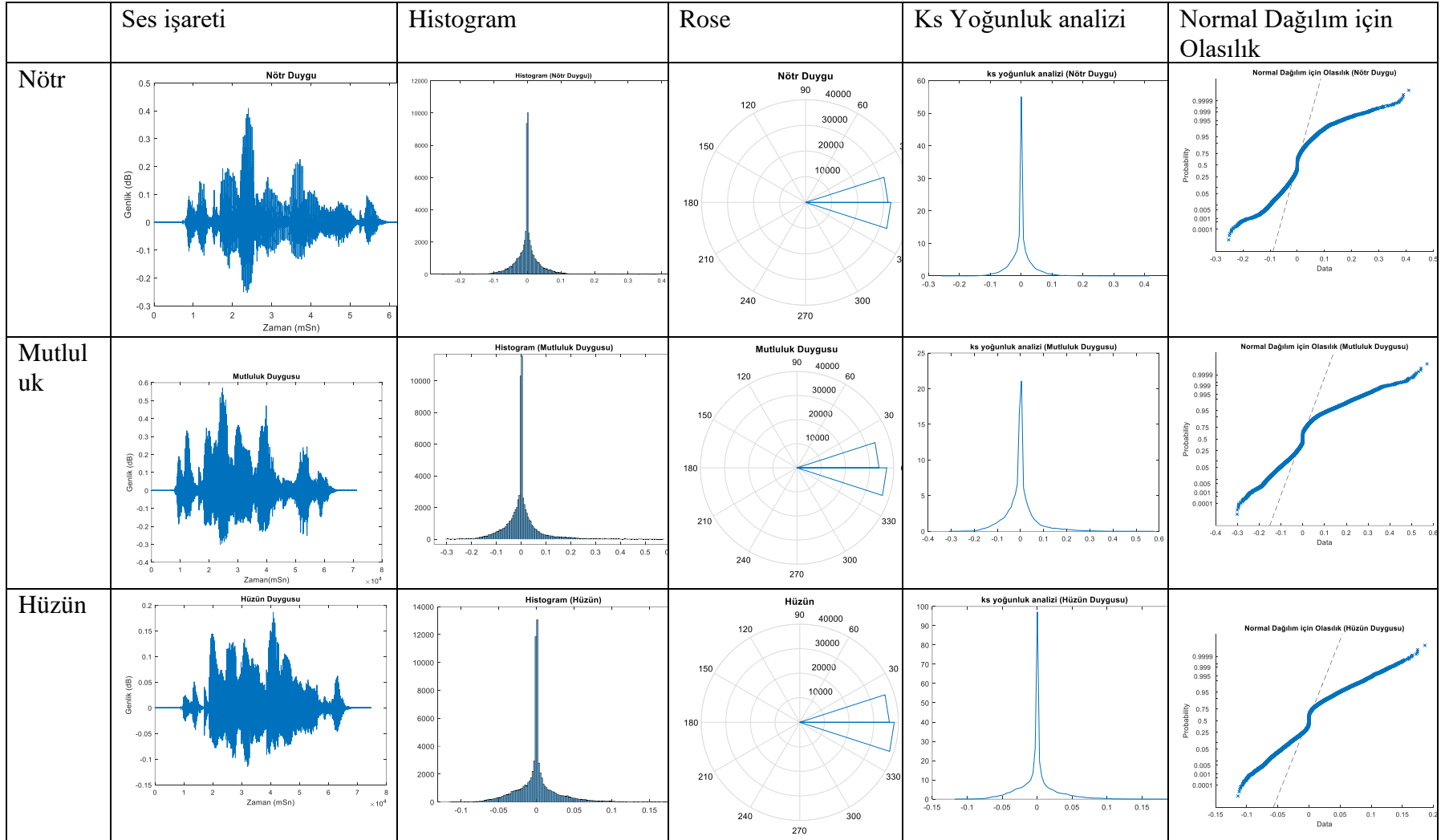
Şekil 59. Çatal Elinden Düştü Cümlesi Şaşkınlık Duygusuna Ait Ks Yoğunluk Analizi

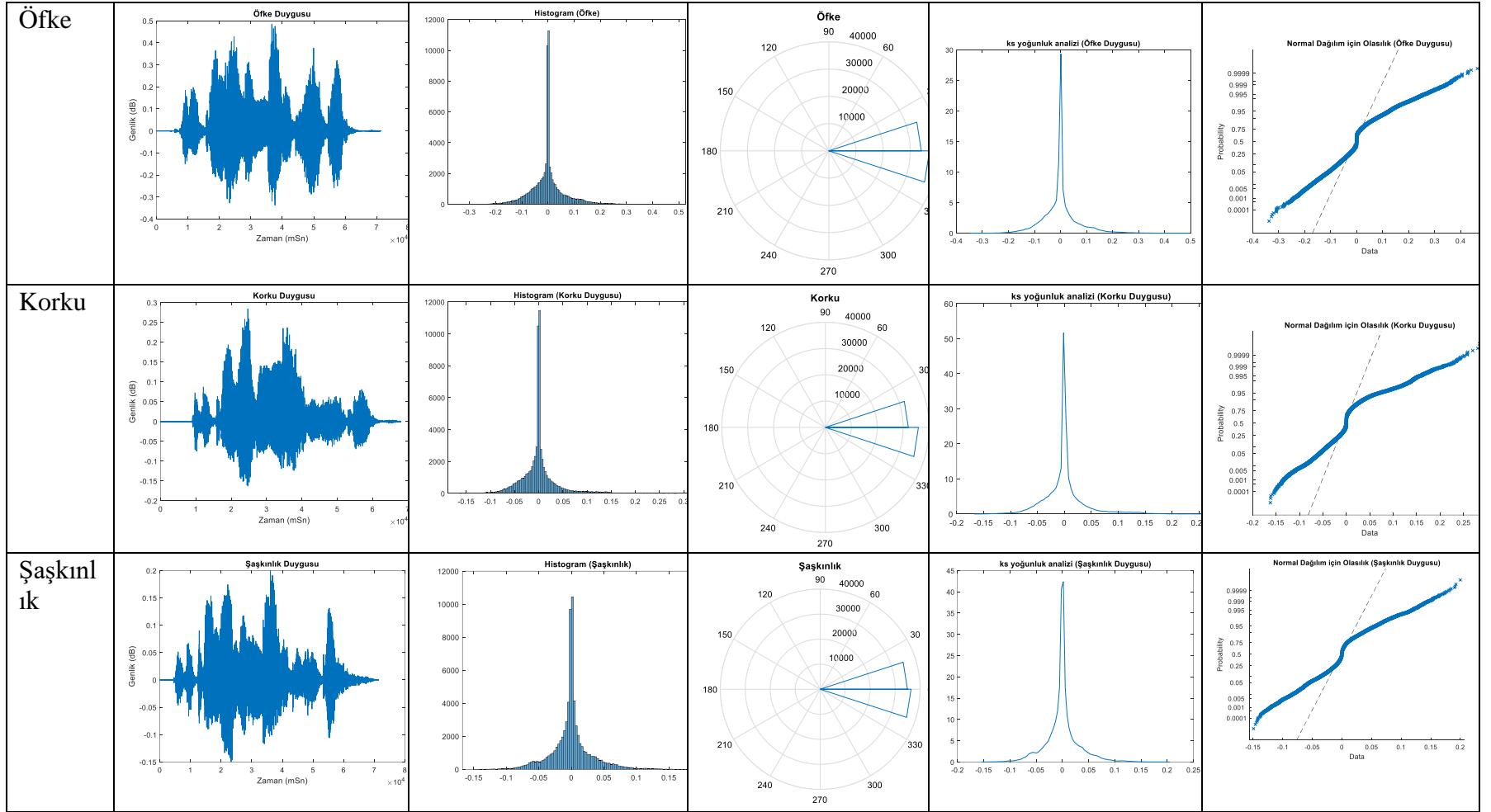
Şekil 60’ da ‘Çatal Elinden Düştü’ cümlesinin normal dağılım için olasılık grafiği şaşkınlık duygusu için analiz edilmiştir.



Şekil 60. ‘Çatal Elinden Düştü’ Cümlesi Şaşkınlık Duygusuna Ait Normal Dağılım İçin Olasılık Grafiği

Altı farklı duygu için ham ses işaretlerinin istatistiksel ve zaman-genlik özellikleri ile de birbirinden farklı oldukları göz ile ayırt edilebilmektedir. Ayrıca tüm analizler toplu olarak kıyaslama açısından Şekil 61’de bir arada verilmiştir. Bu grafikte farklı duyguların benzer ya da farklı özellikleri derin öğrenme uygulamaları ile gözle dahi ayırt edilebilecek şekilde görselleştirilmiştir. Duyguların sınıflandırılması gerçekte grafiksel görsel ayırım ile değil; yazılım tarafından grafiklerin değerlendirilmesi ile yapılmaktadır.

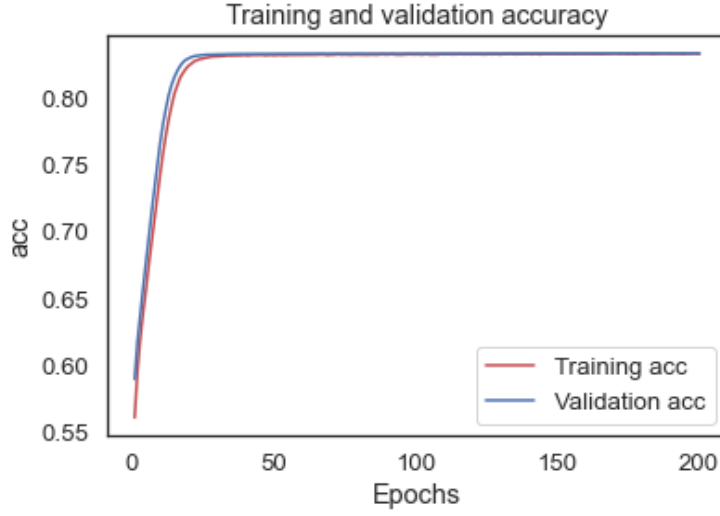




Şekil 61. Altı Farklı Duyguya Ait Derin Öğrenme Analizleri

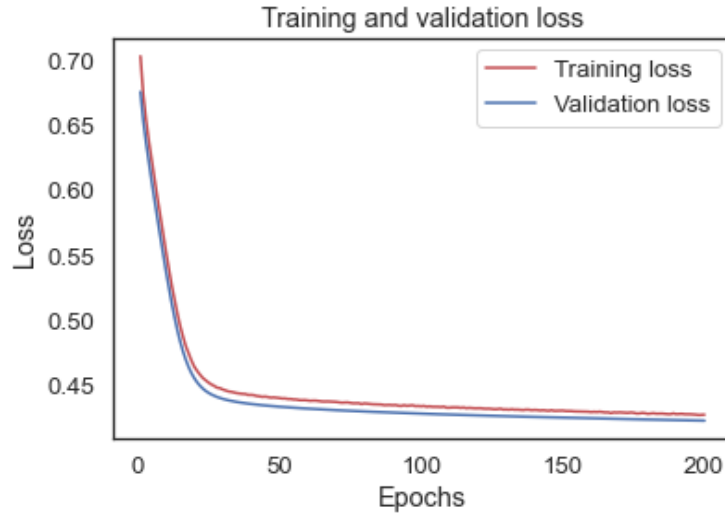
B. Derin Öğrenme Doğruluğu

Test aşamasında derin öğrenme doğruluğu %93 olarak elde edilmiştir. Şekil 62’ de derin öğrenme modelinin doğruluk grafiği verilmiştir.



Şekil 62. Derin Öğrenme Modeli Doğruluk Grafiği

Şekil 63’ de derin öğrenme modelinin kayıp grafiği verilmiştir.



Şekil 63. Derin Öğrenme Modeli Kayıp Grafiği

C. Normal İşiten ve Koklear İmplant Kullanan Bireylerin Sosyodemografik Özellikleri

Normal işitmeye sahip bireylerin 10'u kadın, 10'u erkek ve koklear implant kullanan bireylerin 14'ü kadın, 6'sı erkektir. Normal işitmeye sahip katılımcıların yaş ortalaması 24,85 (\pm 4,13); koklear implant kullanan katılımcıların yaş ortalaması 29,85 (\pm 6,53)' dir.

Koklear implant kullanan bireylerin 17' si tek taraflı, 3'ü çift taraflıdır. Koklear implant kullanım süresi ortalama 13,78 (\pm 5,97) 'dir (yıl). Koklear implant kullanım süresi hesaplanırken, çift taraflı koklear implant kullanıcıları için birinci koklear implant operasyonunun ardından geçen süre dikkate alınmıştır (Çizelge 12).

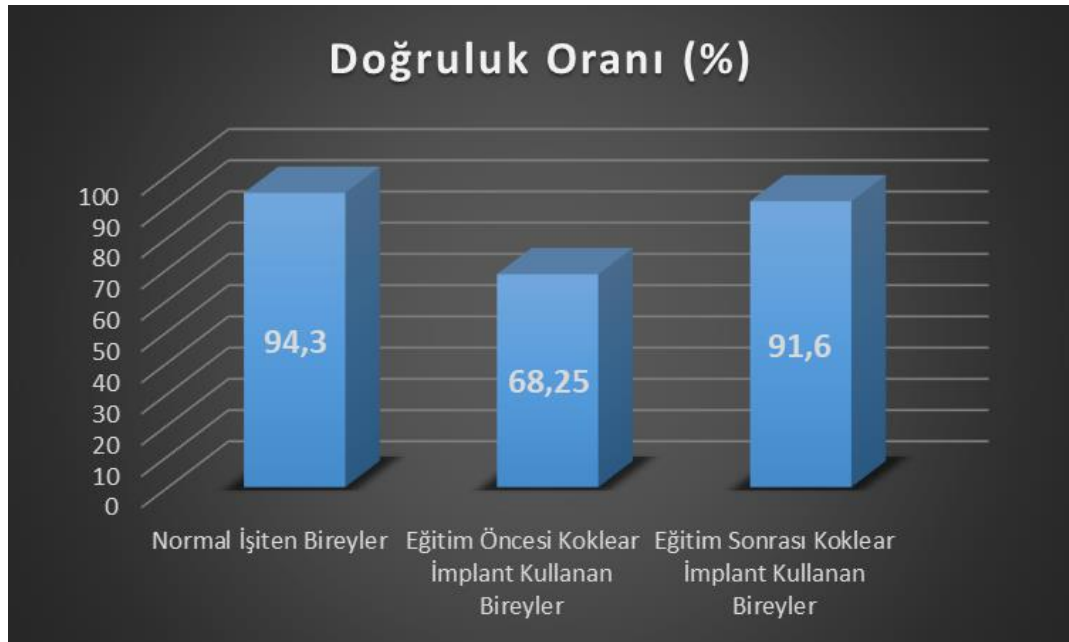
Çizelge 12. Koklear İmplant Kullanan Bireylere Ait Özellikler

Numara	Cinsiyet	Yaş	Tek Taraflı/ Çift Taraflı	Koklear İmplant Kullanım Süresi (Yıl)
1	Erkek	28	Tek Taraflı (Sağ)	10.4
2	Erkek	35	Tek Taraflı (Sağ)	12.4
3	Kadın	36	Tek Taraflı (Sol)	3.5
4	Kadın	38	Tek Taraflı (Sağ)	3.8
5	Kadın	26	Tek Taraflı (Sağ)	21.6
6	Kadın	40	Tek Taraflı (Sağ)	18.5
7	Kadın	39	Tek Taraflı (Sol)	18.2
8	Kadın	33	Tek Taraflı (Sağ)	7.7
9	Erkek	24	Çift Taraflı	Sağ 19.8 / Sol 9.4
10	Kadın	19	Tek Taraflı (Sağ)	15.4
11	Erkek	33	Tek Taraflı (Sağ)	6
12	Erkek	35	Tek Taraflı (Sağ)	12.3
13	Erkek	18	Tek Taraflı (Sağ)	10
14	Kadın	29	Tek Taraflı (Sağ)	9
15	Kadın	22	Çift Taraflı	Sağ 12.5 / Sol 15.8
16	Kadın	22	Çift Taraflı	Sağ 18 / Sol 13.3
17	Kadın	29	Tek Taraflı (Sol)	10.8
18	Kadın	33	Tek Taraflı (Sağ)	21.6
19	Kadın	29	Tek Taraflı (Sol)	21
20	Kadın	29	Tek Taraflı (Sağ)	19.8

D. Normal İşiten ve Koklear İmplant Kullanan Bireylerin Derin Öğrenme ile Ses Duygu Analizi Bulguları

Normal işiten (20 birey) ve koklear implant (20 birey) kullanan bireyler oluşturulan cümleleri altı farklı duyguda seslendirmiştir. Koklear implant kullanan bireylerden eğitim öncesi ve eğitim sonrası olmak üzere iki farklı zamanda ses kayıtları alınmıştır. Koklear implant kullanan bireylerden eğitim öncesi ses kayıtları toplandıktan sonra duygu yüz ifadeleri gösterilmiş ve görsel ipuçları barındıran animasyon seyrettirilmiştir. Animasyon, koklear implant kullanan bireylerin duyguları sözel ifade etme becerilerine yönelik eğitim kapsamında kullanılmıştır. Eğitim sonrası kayıtlar da animasyon seyrettirildikten sonra alınmıştır.

Normal işiten bireylerin ses duygu analiz doğruluk oranı %94.3, koklear implant kullanan bireylerin eğitim öncesi ses duygu analiz doğruluk oranı %68.25, koklear implant kullanan bireylerin eğitim sonrası ses duygu analiz doğruluk oranı %91.3 olarak elde edilmiştir (Şekil 64).



Şekil 64. Normal İşiten ve Koklear İmplant Kullanan Bireylerin Derin Öğrenme ile Ses Duygu Analizi Doğruluğu

V.TARTIŞMA

Derin öğrenme yöntemi kullanılarak normal işiten bireyler ile işitme kayıplı bireyler arasındaki duyguları sözel ifade etme becerisi farkını karşılaştırmak amacıyla yapılan bu çalışmadaki hipotezlerimiz; koklear implant kullanan bireylerin normal işiten bireylere göre duyguları sözel ifade etme becerisinin daha düşük olması ve koklear implant kullanan bireylerin eğitim sonrası sonuçlarının, eğitim öncesi sonuçlarına göre duyguları sözel ifade etmede daha yüksek doğruluk oranı elde edileceğidir. Çalışmamızın sonucu olarak, koklear implant kullanan bireyler, normal işiten bireylere göre daha iyi doğruluk vermiştir. Koklear implant kullanan bireylerin eğitim sonrası doğruluk oranı, eğitim öncesi doğruluk oranına göre anlamlı derecede artış göstermiştir.

İşitme kayıplı bireylerde derin öğrenme yöntemi kullanılarak ses duygu analizi ile ilgili literatürde benzer çalışmaya rastlanılmamıştır. Bu nedenle bu bölümde literatür çalışmalarındaki ses duygu analizi için kullanılan yapay zekâ yöntemleri karşılaştırılmış ve işitme kayıplı bireylerde MATLAB, Praat ve subjektif değerlendirmeler ile yapılan ses duygu analiz çalışmaları tartışılmıştır. Ayrıca tez çalışmasının işitme kayıplı bireyler için ilk olması ile literatüre yenilikçi bir katkı sağlayacağı beklenmektedir.

Çalışmamızda kullanılan derin öğrenme modeli CNN' dir. CNN modelleri kendi aralarında karşılaştırılırken kullanılan veri tabanlarının aynı veya ayrı olması doğruluk oranlarını etkilemektedir (Çizelge 1). Kullanılan veri tabanlarının kalitesi ve elde edilen performans duygu tanıma değerlendirmesinde en önemli faktörlerdir (Chen ve ark., 2018). The Interactive Emotional Dyadic Motion Capture (IEMOCAP), profesyonel 10 konuşmacı (5 kadın, 5 erkek) tarafından kaydedilen yaklaşık 12 saatlik İngilizce video, konuşma içeren görsel-ışitsel bir veri tabanıdır. Profesyonellerin, özellikle duygusal ifadeleri ortaya çıkarmak için seçilen doğaçlama veya senaryoları gerçekleştirdiği iki kişi (1K, 1E) olarak beş oturumdan oluşmaktadır. IEMOCAP veri tabanı, öfke, üzgün, mutlu ve nötr duygularını içermektedir. Berlin Emotional Speech Database

(EMO- DB), beş kadın ve beş erkek oyuncunun Almanca öfke, üzüntü, mutlu, nötr, korku, can sıkıntısı, tikslenme duygularından her biri için on cümleyi seslendirdiği bir veri tabanıdır (Anvarjon & Kwon, 2020). IEMOCAP veri tabanı kullanılarak test edilen CNN modellerinde Han ve ark. (2014) %51.24 doğruluk, Zhang ve ark. (2015) %40.02 doğruluk, Fayek ve ark. (2017) %64.78, Luo ve ark. (2018) %63.98 doğruluk, Chen ve ark. (2018) %64.74 doğruluk, Zhao ve ark. (2019a) %52.14 doğruluk, Guo ve ark. (2019) %57.10 doğruluk, Meng ve ark. (2019) %69.32 doğruluk, Zhao ve ark. (2019b) %66.50 doğruluk, Issa ve ark. (2020) %64.03 doğruluk, Mustaqeem ve ark. (2020) %72.25, Anvarjon & Kwon (2020) %77.01 doğruluk elde etmişlerdir. EMO- DB veri tabanı kullanılarak test edilen CNN modellerinde Chen ve ark. (2018) %82.82 doğruluk, Guo ve ark. (2019) %84.49 doğruluk, Meng ve ark. (2019) %88.99 doğruluk, Badshah ve ark. (2019) %80.79 doğruluk, Jiang ve ark. (2019) %84.53 doğruluk, Issa ve ark. (2020) %86.10 doğruluk, Mustaqeem ve ark. (2020) %85.57 doğruluk, Anvarjon & Kwon (2020) %92.02 doğruluk elde etmişlerdir. Çalışmamızda derin öğrenme CNN modeli için profesyoneller tarafından farklı duygularda seslendirilen cümlelerle ilk Türkçe eğitim veri tabanı oluşturulmuştur. Bu bağlamda CNN modeli ile %93 doğruluk elde edilmiş ve diğer çalışmalar ile karşılaştırıldığında çalışmamızın derin öğrenme doğruluk oranının EMO- DB veri tabanı gibi yüksek olduğu görülmüştür. Diğer yandan çalışmalardaki CNN modellerinin farklı katman sayıları, farklı filtre çeşitleri, farklı havuzlama çeşitleri (maksimum, ortalama) gibi farklı parametreler ile geliştirilmesi doğruluk oranlarının düşük veya yüksek olmasını etkileyeceği de göz önünde bulundurulmalıdır.

Gerçek duyguların anlaşılması, senaryolaştırılan duyguların anlaşılmasından daha zordur. Profesyonel tiyatro eğitimi almamış bireylerin ses kayıtları ile eğitilen derin öğrenme algoritması, abartılı olmayan, gerçeğe daha yakın sonuçların elde edilmesini sağlayabilir. Vera am Mittag (VAM) veri tabanı, Almanya'daki Karlsruhe Üniversitesi'nin INT enstitüsü tarafından toplanmıştır ve Alman televizyonu "Vera am Mittag" talk show programından alınan görsel-işitsel kayıtlardan oluşmaktadır. VAM programında, moderatör çoğunlukla konuklar arasındaki tartışmaları, ara sıra sorular sorarak yönetmektedir. Gerçek hayattan alınan konuşmalar tartışılan konulara bağlı olarak çok çeşitli duygusal durumlar barındırmaktadır. VAM veri tabanında duygu değerlik, baskınlık ve

uyarılma olmak üzere üç temelde ele alınmaktadır. Değerlik, bir durumun içsel hoşluğunu veya tatsızlığını (pozitiflikten negatifliğe); uyarılma, bir uyarının bir kişiyi artan veya azalan aktivite durumuna getirip getirmediğini (düşük uyarımdan yüksek uyarıma); baskınlık ise dinleyici üzerinde konuşmacının etkisinin zayıf ya da güçlü olması ile açıklamaktadır (Grimm ve ark., 2008). VAM veri tabanı kullanılarak eğitilen derin öğrenme algoritmalarında; Stuhlsatz ve ark. (2011) %68 doğruluk oranı (derin öğrenme modeli RBM + GerDa), Amer ve ark. (2014) %66.3 doğruluk oranı (derin öğrenme modeli CRF + CRBM) elde etmişlerdir. Çalışmamızda, Drama ve Oyunculuk Bölümü öğrencileri tarafından seslendirilen cümleler derin öğrenme eğitiminde (derin öğrenme modeli CNN) kullanılmış ve %93 doğruluk oranı elde edilmiştir. Literatürden farklı olarak tiyatro eğitimi olan bireylerin seçilmesi ve kullanılan derin öğrenme modelinin farklı olması sonuçlar arasındaki uyumsuzluğu açıklamaktadır. Ayrıca konuşmacıların spontane cümleleri ile eğitilen derin öğrenme algoritmasının, senaryo cümleleri ile eğitilen derin öğrenme algoritmasından daha düşük doğruluk oranı vermesi beklenen bir sonuçtur.

Çocuklarda duygu tanıma gelişimi, özellikle otistik çocuklar gibi bazı klinik gruplarda, yüz ifadeleri veya görsel-işitsel materyaller kullanılarak kapsamlı bir şekilde çalışılmıştır (Harms, Martin & Wallace, 2010; Rodger ve ark., 2015; Nagels ve ark., 2020). Literatür çalışmalarında ses duygularının tanınması için kullanılan yöntemler, ağırlıklı olarak ortalama temel frekans (F0) ve yoğunluk gibi ilgili vokal akustik ipuçlarının yanı sıra bu ipuçlarındaki dalgalanmalar ve konuşma hızına dayanmaktadır. Bununla birlikte, normal gelişen çocuklarda, vokal duygu tanımanın gelişimi hakkında çok şey bilinmemektedir. Çocuklar görsel ve işitsel uyarılardaki duyguları 5 yaşından itibaren duygu ortamları ile ilişkilendirebilirken, duygu tanıma gelişiminin yetişkin benzeri seviyelere gelmesinin geç çocukluk dönemine kadar sürdüğü görülmektedir. Ses duygu ipucu algısının gelişimi üzerine yapılan önceki araştırmalar, çocukların performansının işitsel kategorilerin sağlam temsillerinin bulunmaması nedeniyle yetişkinlere kıyasla daha düşük olduğunu göstermektedir (Nagels ve ark., 2020). Morton & Trehub (2001)' un 4- 10 yaş arası çocukları ve yetişkinleri dahil ettiği çalışmada, cümle içeriğine uygun olmayan bir şekilde cümleler seslendirilmiştir (örneğin, mutlu bir durumun üzgün bir şekilde ifade edilmesi) ve

çocuklar duyguları tanımlamak için çoğunlukla cümle içeriğine, yetişkinler ise çoğunlukla duygusal prozodiye yönelmiştir. Çocuklar çelişkili ipuçlarına sahip işitsel duygu ifadelerini ayırt etmede ve anlamlandırmada, çelişkili olmayan ipuçlarına göre daha fazla zorlanmıştır. Ek olarak hem çocuklar hem de yetişkinler, sesli duygu tanımaya göre yüzdeki duygu tanımada daha iyidir (Nelson & Russell, 2011; Chronaki ve ark., 2015). Nelson ve Russel (2011) 3-5 yaş aralığındaki 144 çocuğa yüz, postür, ses ve çoklu işaret (yüz ifadesi/ postür/ ses) ifadeleriyle aktarılan duygular sunmuştur. Mutluluk, üzüntü, öfke ve korku için dört işaret koşulunda dinamik ipuçlarıyla iletilen duyguyu etiketlemeleri istenmiştir. Çocuklar yüz, postür ve çok işaretli ifadeleri çoğunlukla doğru bir şekilde etiketlemiştir (doğruluk > %70), ancak ortalama olarak sadece ses ifadelerini tanıma değeri düşüktür (doğruluk %34). Ses, duygu durumlarından yalnızca üzüntüyü tanımada yüksek doğruluk vermiştir (doğruluk %72). Chronaki ve ark (2014), yüz ve ses ifadeleriyle aktarılan duyguları 4-5, 6- 9, 10- 11 yaş aralığındaki çocuklar ve yetişkinler ile karşılaştırmıştır. Sonuçlar, yüz ifadesinden duygu tanımının 11 yaşında yetişkin düzeylerine ulaştığını, sestem duygu tanımının geç çocukluk döneminde gelişmeye devam ettiğini göstermiştir. Tüm bu gözlemler bir araya geldiğinde, subjektif testlerle yapılan işitsel- görsel duygu tanıma çalışmalarının doğruluk oranları farklı elde edilmiştir. Derin öğrenme ile ses duygu analizi, subjektif testlerle yapılan çalışmalardaki farklı değerlendiricilerin farklı doğruluk oranı oluşturma dezavantajını ortadan kaldırarak objektif sonuçlar vermektedir.

Yüz ifadeleri, zorlu dinleme koşullarında faydalı bilgiler sağlamaktadır. Dinleyiciyle doğrudan yüz yüze gelinmeyen durumlarda ise konuşmacıların ses duygu içeriği sosyal iletişimde önemli rol oynamaktadır. Bu nedenle hem rehabilitasyon sürecinde hem de konuşma kodlama stratejileri geliştirmede duygusal prozodinin aktarımını ve algısını iyileştirmeye daha fazla dikkat edilmesi gerektiği açıktır. Şahin (2017), normal işiten (40 birey) ve koklear implant kullanan (20 birey) 7- 12 yaş arası okul çağı çocuklarda resimli kartlar ile görsel duygu tanıma becerileri ve içerik bağımlı/ içerik bağımsız cümlelerin kullanılması ile ses tonundan duygu çıkarımı becerilerini değerlendirmiştir. Çalışmaya dahil edilen tüm çocukların görsel ipuçlarını kullanarak duyguları anlama ve ayırt etme görevlerinde sorun yaşamadığı ancak linguistik ve/veya

paralinguistik ipuçlarını kullanarak, ses tonundan duygu durum çıkarımında koklear implant kullanan çocukların, normal işiten çocuklara göre başarısız olduğu sonucu elde edilmiştir. Duyguları doğru vurgu ve tonlama ile sözel ifade edebilmenin ön koşulu, duyguları tanımadır. Koklear implant kullanan bireyler, sese geç erişim gibi kısıtlamalar sebebiyle normal işiten akranlarından farklı olarak duygu tanıma gelişimini tamamlamamıştır. Bu bağlamda, çalışma örneklem grubumuz olan 18- 40 yaş aralığındaki koklear implant kullanıcıları, gelişim döneminde bu farkındalığa sahip olamadıkları için ses duygu ifade etme becerileri de zayıf kalmıştır.

İşitme kaybı, sese verilen duygusal tepkileri bozabilmektedir. Bununla birlikte, uyarıcı modalitesinin, bu bozulma üzerindeki etkisi ve bununla ilgili altta yatan mekanizmalar belirsizdir. Konuşma dışı uyaran kullanılarak yapılan çalışmalar, cümle duygu içeriği ipucunu ortadan kaldırdığı için elde edilen bulguların daha objektif olduğu söylenebilir. Buona ve ark. (2020), alçak ve yüksek geçişli filtreleme ile sunulan azaltılmış işitilebilirliğin duygu algısı üzerindeki etkilerini araştırmak amacıyla 24 normal işitmeye sahip yetişkin birey üzerinde çalışma yapmıştır. İşitsel uyaran olarak, International Affective Digitized Sounds (IADS; Bradley & Lang, 2007) veri tabanından alınan konuşma dışı sinyalleri kullanmışlardır. Mevcut çalışma için 26 hoş, 26 hoş olmayan ve 13 nötr olmak üzere toplamda 65 ses veri tabanından seçilmiştir. IADS uyaranlarının alçak geçişli filtrelenmiş versiyonları, 800 Hz'lik bir üst kesme frekansı kullanılarak oluşturulmuştur. IADS uyaranlarının yüksek geçişli filtrelenmiş sürümleri, 2000 Hz'lik düşük bir kesme frekansı kullanılarak oluşturulmuştur. Görsel uyaranlar International Affective Picture System (IAPS; Bradleyy, Cuthbert & Lang, 2005) veri tabanından seçilmiştir. IADS belirteçleriyle aynı üç kategoriye (hoş, hoş olmayan ve nötr) oluşturmak için 65 resim kullanılmıştır. Her katılımcı beş uyarıcı koşulunda duyguları derecelendirilmiştir. Bu beş koşul yalnızca işitsel (AO), yalnızca alçak veya yüksek geçişli filtreli işitsel (fAO), yalnızca görsel (VO), işitsel-görsel (AV) ve alçak veya yüksek filtrelenmiş işitsel-görsel (fAV- işitsel bileşeni filtrelenmiş) olacak şekilde sağlanmıştır. Katılımcıların yarısına alçak geçişli, diğer yarısına da yüksek geçişli filtre sunulmuştur. Duygu algısının etkileri, filtre tipine (alçak geçişli 800 Hz veya yüksek geçişli 2000 Hz), uyarıcı kategorisine (hoş, hoş olmayan, nötr) ve uyarıcı

modalitesine (AO, fAO, VO, AV ve fAV) bağılı olarak deęişmiştir. Katılımcıların, filtrelenmiş uyanları, filtrelenmemiş uyanlara göre hoř, hoř olmayan ve nötr olarak derecelendirilme olasılıęı daha düşük elde edilmiştir. Bu sonuçlar, işitilebilir spektral bilgilerdeki deęişikliklerin, işitme kaybına eşlik eden duygusal tepkilerdeki deęişikliklerden kısmen sorumlu olabileceğini göstermektedir (Buono ve ark., 2020). Örneklem grubunu normal işiten bireylerin oluşturduğu bu çalışmada alçak ve yüksek filtre uygulanarak işitilebilirliğin azaltılması ve konuşma dışı uyanlar kullanılması ile ses duygu tanıma görevlerinde başarısız sonuçlar elde edilmiştir. Çalışmamızın temel bulgularından biri olan işitme kayıplı bireylerde, normal işiten bireylere göre ses duygu ifade etme becerisinin daha düşük elde edilmesinin, literatürde örneęi bulunan bozulmuş akustik uyan varlığında (işitme kaybı simülasyonu) duyguları tanıma boyutundaki eksikliklerden kaynaklandığını söyleyebiliriz.

İşitme kayıplı bireylerin alıcı ve ifade edici dil becerilerini geliştirmeye yönelik verilen işitsel rehabilitasyon; mekân, zaman ve ekonomik sorunlar sebebiyle çoğunlukla yetersiz kalmaktadır. Bu kısıtlılıklar sebebiyle işitme kayıplı bireylerde verilmesi gereken duyguları ifade etme modülleri ise fonem üretme modüllerinin gerisinde kalmaktadır. Normal işiten bireyler, mutlu oldukları durumlarda daha yüksek frekans modülasyonları ile daha yüksek sesle ve daha hızlı; üzgün olduklarında ise daha düşük frekans modülasyonları ile daha monoton ve daha yavaş konuşmaktadırlar (Damm ve ark., 2019). Chatterjee ve ark. (2019), koklear implant kullanan bireylerin implantasyon öncesinde akustik girdiye erişim kısıtlılığının, ses duygu ifade etme becerisini nasıl etkilediğiyle ilgili çalışma yapmıştır. Bu kapsamda, konjenital işitme kaybı olan 2 yaşında implante edilen çocukların (13 birey) ürettiği vokal duyguların akustik özelliklerini, normal işitmeye sahip çocuklar (9 birey) / yetişkinler (9 birey) ve postlingual işitme kaybı olan koklear implant kullanan yetişkinlerle (10 birey) karşılaştırılmıştır. Bu çalışma için materyal olarak duygu hakkında açık bir semantik ipucu içermeyen 20 basit cümle (Örneğin: She is back / Time for bed / I see a dog) kullanılmıştır. Her katılımcı 20 cümleyi mutlu ve üzgün olmak üzere iki farklı şekilde (toplamda 40 kayıt) seslendirmiştir. Praat yazılım paketi kullanılarak ortalama frekans (F0, Hz), F0 standart sapması, ortalama yoğunluk, spektral ağırlık merkezi ve her bir ifadenin süresi hesaplanmıştır. Duygu bazında

bakıldığında ortalama frekans, F0 standart sapması, ortalama yoğunluk ve spektral ağırlık merkezi mutluluk için üzgünden daha fazladır; cümle süresi ise üzgün için mutluluktan daha fazladır. Sonuçlar koklear implant kullanan çocuk ve yetişkinlerin ses duygu ifade etme becerilerinin değişken olduğunu göstermektedir. Postlingual işitme kaybı olup koklear implant kullanan yetişkinlerin yüksek derecede bozulmuş anormal konuşma girdisine yıllarca maruz kalmalarına rağmen erken gelişim yıllarında akustik bilgiye erişimlerinin olmasının ses duygu ifade etme becerisinin korunmuş olduğu sonucuna varılmıştır. Diğer yandan konjenital işitme kaybı olup koklear implant kullanan çocukların ses duygu ifade etme becerisinin yetersiz olduğu sonucu elde edilmiştir. Aksine, koklear implant kullanan çocukların cümlelerdeki kelimeleri doğru telaffuz ettiği gözlenmiştir. Bu bağlamda verilen rehabilitasyon eğitiminin konuşma prozodisinden çok fonem üretimine odaklandığı sonucuna varılmıştır. Çalışmamızda, koklear implant kullanan bireylerde duygu eğitimi sonrası doğruluk oranı, duygu eğitimi öncesi doğruluk oranına göre daha yüksek elde edilmiştir. Bu sonuç, işitme kayıplı bireylerde duygu eğitiminin önemini göstermektedir. İşitme kayıplı bireylerde duygu eğitimi farkındalığının artırılması için daha fazla çalışma yapılması gerekmektedir.

Müziğin, konuşma uyaranlarından daha belirgin akustik ipucu barındırması sebebiyle duygular müzikte daha kolay kategorize edilebilmektedir (Gabrielsson & Juslin, 2003; Trainor & Corrigall, 2010; Mohn, Argstatter & Wilker, 2011). Siu ve Cheung (2017), 15 aylık bebekleri, çeşitli duyguların yüz ifadeleri ve vücut hareketleri eşliğinde üç ay boyunca eğitmişlerdir. Duygu eğitimi için bebekler rastgele olarak iki gruba ayrılmış ve müzik varlığında ya da müzik olmadan eğitimleri tamamlanmıştır. Sonuçlar, müzikal deneyimin duygusal durumlar ve eylem arasındaki ilişkiyi anlamalarını desteklediğini ve bebekleri duygusal temsiller oluşturmaya teşvik ettiğini göstermektedir (Siu & Cheung, 2017). Çalışmamızda koklear implant kullanan bireylere verilen duygu eğitiminde yüz ifadeleri ve animasyon (ses, yüz ifadesi ve duruş) kullanılmış ve ses duygu ifade etme becerisine olumlu etkileri olmuştur. Literatür verileri ışığında duygu eğitimine müziğin dahil edildiği çalışmalarda daha başarılı sonuçlar elde edilebileceği görülmektedir.

VI.SONUÇ VE ÖNERİLER

Çalışmamızda, işitme kayıplı bireylerin ses duygu ifade etme becerileri normal işiten bireylere göre daha düşük elde edilmiştir. Ayrıca koklear implant kullanan bireylerin eğitim sonrası doğruluk oranları, eğitim öncesi doğruluk oranlarına göre anlamlı derecede artış göstermiştir. Bu çalışma, işitme kayıplı bireyler için ilk olması ile literatüre yenilikçi bir katkı sağlayacağı öngörülmektedir.

1. Koklear implant kullanan bireylerin derin öğrenme ile ses duygu ifade etme becerisi, normal işiten bireylerin ses duygu ifade etme becerisine göre daha düşük elde edilmiştir.
2. Koklear implant kullanan bireylerin eğitim öncesi ses duygu ifade etme becerisi, koklear implant kullanan bireylerin eğitim sonrası ses duygu ifade etme becerisine göre daha düşük elde edilmiştir.
3. Derin öğrenme eğitimi için tiyatro öğrencilerinin seslendirmelerinden oluşan ilk Türkçe ses duygu veri tabanı literatüre kazandırılmıştır.
4. İşitme kayıplı bireylerde test edilen, ilk derin öğrenme ses duygu analizi yazılımı oluşturulmuştur.
5. İşitme kayıplı bireylerde ses duygu eğitimi için ilk Türkçe ön prototip eğitim modülü oluşturulmuş ve düşük doğruluk oranları eğitim sonrasında olumlu sonuçlar göstermiştir.
6. Derin öğrenme algoritması eğitimi, gürültülü ortam ve birden çok konuşmacı varlığı gibi zorlu dinleme koşullarında duygu ifadesi içeren sesler ile yapıldığında, gerçek hayatla daha uyumlu verilerin ortaya çıkması sağlanabilir.
7. Koklear implant kullanan ve normal işiten çocukların duyguları sözel ifade etme becerisi farklı duyguları içeren cümle materyalleri ile derin öğrenme yöntemi kullanılarak değerlendirilebilir.

8. Duyguları sözel ifade etmeyi geliştirecek rehabilitasyon programları geliştirilebilir ve erişimin kolay olduğu mobil uygulamalara entegre edilerek herkes tarafından ulaşılabilir olması sağlanabilir.
9. Gerçek duyguların anlaşılmasının, senaryolaştırılan duyguların anlaşılmasından daha zor olduğunu göz önünde bulundurarak, drama ve oyunculuk eğitimi almamış bireylerin ses kayıtları ile derin öğrenme mimarisi eğitilerek çalışmamızda elde ettiğimiz bulgular karşılaştırılabilir.
10. Koklear implant kullanan bireylerin, spektral, zamansal ve tını açısından konuşmadan daha kompleks olan müzik içerisindeki duyguları tanıma ve şarkı söyleme görevi esnasında duyguları ifade etme becerileri derin öğrenme yöntemi kullanılarak analiz edebilir.
11. Ses duygu analizi derin öğrenme yöntemi kullanılarak, işitme kayıplı birey sayısının daha fazla olduğu örneklem grubu ile çalışma yapılabilir.
12. Aynı çalışma içerisinde yapay zekâ (objektif) ve klinisyen (subjektif) ses duygu analizi karşılaştırmasını içeren çalışma tasarlanabilir.

VII. KAYNAKÇA

KİTAPLAR

- ADIT, V. V., RUBESH, C. V., BHARATHI, S. S., SANTHIYA, G., & ANURADHA, R. (2020) A Comparison of Deep Learning Algorithms for Plant Disease Classification. In **Advances in Cybernetics, Cognition, and Machine Learning for Communication Technologies**. Springer. Singapore.
- AFOLABI, O. A., ALABİ, B. S., SEGUN-BUSARI, S., & AREMU, S. K. (2012). Conductive Hearing Loss Due to Trauma. **Hearing Loss**, 157.
- ALSHUAIB, W. B., AL-KANDARI, J. M., & HASAN, S. M. (2015). Classification of Hearing Loss. Bahmed, F (Ed) **Update On Hearing Loss** içinde.
- BARRETT, L. F., & SALOVEY, P. (Eds.). (2002). **The wisdom in feeling: Psychological processes in emotional intelligence**. New York: Guilford Press
- BELLIS, T. J. (2003). **Assessment and management of central auditory processing disorders in the educational setting: From science to practice** (2nd ed.). Clifton Park, NY: Delmar Learning.
- BELGİN, E., & ÇALIŞKAN, M. (2004). **Çalışma yaşamında gürültü ve işitmenin korunması**. Ankara: Türk Tabipleri Birliği Yayınları.
- BRADHAM, T. S., & HOUSTON, K. T. (2015). **Assessing listening and spoken language in children with hearing loss**. San Diego, CA: Plural Publishing.
- BUCK, R. (1983). Emotional development and emotional education. In R. Plutchik & H. Kellerman (Eds.), **Emotion in early development** (pp. 259–293). New York: Academic Press.

- BUDUMA, N., & LOCASCIO, N. (2017). **Fundamentals of deep learning: Designing next-generation machine intelligence algorithms.** " O'Reilly Media, Inc."
- CHERMAK, G. D., & MUSIEK, F. E. (1997). **Central auditory processing disorders: New perspectives.** San Diego, CA: Singular.
- CHERMAK, G. D., & MUSIEK, F. E. (Eds.). (2014). **Handbook of central auditory processing disorder: Auditory neuroscience and diagnosis** (Vol. I). San Diego, CA: Plural Publishing.
- CHOLLET, F. (2018). **Deep learning with Python** (Vol. 361). New York: Manning.
- COLE, E. B., & FLEXER, C. (2019). **Children with hearing loss: Developing listening and talking, birth to six.** Plural Publishing.
- DAHL, G., RANZATO, M. A., MOHAMED, A. R., & HINTON, G. E. (2010). Phone recognition with the mean-covariance restricted Boltzmann machine. **Advances in neural information processing systems**, 23, 469-477.
- DALE, P. (1976). **Language development. Structure and function.** 2nd ed. New York: Holt, Rinehart and Winston.
- DAMASIO, A. (1999). **The feeling of what happens: Body and emotion in the making of consciousness.** New York: Harcourt, Inc.
- DARWIN, C., (1872). **The Expression of Emotions in Man and Animals.** John Murray, London (third ed., P. Ekman (Ed.), 1998, Harper Collins, London).
- DARWIN, C., (1871). **The Descent of Man and Selection in Relation to Sex.** John Murray, London.
- DAVIDSON, R. J., SCHERER, K. R. VE GOLDSMITH, H. (2002). Introduction. **Handbook of affective sciences** içinde (xiii-xv). Cary, NC: Oxford University Press.
- DILLON, H. (2012). **Hearing aids** (2nd ed.). New York, NY: Thieme Medical.
- DUMAN, B. (2007), **Neden Beyin Temelli Öğrenme**, Pegem Yayıncılık. Ankara.

- EKMAN, P. (1994). All emotions are basic. In P. Ekman & R. Davidson (Eds.), **The nature of emotion: Fundamental questions** (pp. 15–19). New York: Oxford University Press.
- EKMAN, P. (1999). Facial expressions. **Handbook of Cognition and Emotion**,
- EKMAN, P., FRIESEN, W. V., & ELLSWORTH, P. (1972). **Emotion in the Human Face: Guidelines for Research and an Integration of Findings: Guidelines for Research and an Integration of Findings**. Pergamon.
- FELDMAN, R.S., RIME, B. (Eds.), 1991. **Fundamentals of Nonverbal Behavior**. Cambridge University Press, New York.
- FITZPATRICK, E. M., & DOUCET, S. P. (2013). **Pediatric audiologic rehabilitation: From infancy to adolescence**. New York, NY: Thieme Medical.
- GABRIELSSON, A., & JUSLIN, P. N. (2003). **Emotional expression in music**. Oxford University Press.
- GETTYS, W. E., KELLER, F. J., & SKOVE, M. J. (1995). **Fizik 1**. Cilt. İstanbul: Literatür yayıncılık.
- HALL, J. W., III. (2014). **Introduction to audiology today**. New York, NY: Pearson.
- HARPER, R. G., WIENS, A. N., & MATARAZZO, J. D. (1978). **Nonverbal communication: The state of the art**. John Wiley & Sons.
- HUAWEILI, H VE CHAI R. (Ed.). (2019). **Hearing Loss: Mechanisms, Prevention and Cure** (1st ed.). Springer
- IZARD, C. E. (1977). **Human emotions**. New York: Plenum Press.
- JAMES, W. (1890). **The Principles of Psychology**, Vol. 1. New York: Henry Holt and Company.
- JOHNSTON, I. (2002). **Measured Tones: The Interplay Between Physics and Music**. Institute of Physics Publishing Bristol and Philadelphia, IOP Publishing Ltd, USA, 40- 86-144-151.
- KAYA U., YILMAZ A. (2019). **Derin Öğrenme**, 1-2, ISBN:978-605-2118-399.

- KAYAALP, K., & SÜZEN, A. A. (2018). **Derin Öğrenme ve Türkiye'deki Uygulamaları**. Yayın Yeri: IKSAD International Publishing House, Basım sayısı, 1.
- KELTNER, D., & HAİDT, J. (2001). Social functions of emotions. In T. J. Mayne & G. A. Bonanno (Eds.), **Emotions: Current issues and future directions** New York: Guilford Press.
- KENNEDY, G. A. (1972). **The Art of Rhetoric in the Roman World: 300 BC-AD 300**. Princeton University Press, Princeton, NJ.
- KNAPP, M.L. (1972). **Nonverbal Communication in Human Interaction**. Holt Rinehart and Winston, New York.
- KRAMER, S., & BROWN, D. K. (2019). **Audiology: Science to practice** (3rd ed). San Diego, CA: Plural Publishing.
- LAPP, D. R. (2003). **The physics of music and musical instruments**. Wright Center for Innovative Science Education, Tufts University.
- LECUN, Y., BOSER, B. E., DENKER, J. S., HENDERSON, D., HOWARD, R. E., HUBBARD, W. E., & JACKEL, L. D. (1990). Handwritten digit recognition with a back-propagation network. In **Advances in neural information processing systems**.
- LEE, N. (2007). **Music, Waves, Physics**. Rice University, Houston, Texas.
- LEWIS, M. (2007). Self-conscious emotional development. In J. L. Tracy, R. W. Robins & J. P. Tangney (Eds.), **The self-conscious emotions: Theory and research**. New York: Guilford Press.
- MACLEAN, P. D. (1993). Cerebral evolution of emotion. In M. Lewis & J. M. Haviland (Eds.), **Handbook of emotions**. New York: Guilford Press.
- MAHL, G.F., SCHULZE, G., 1964. Psychological research in the extralinguistic area. In: Sebeok, T.A., Hayes, A.S., Bateson, M.C. (Eds.), **Approaches to Semiotics**. Mouton, The Hague.
- MOORE, B. C., & CARLYON, R. P. (2005). Perception of pitch by people with cochlear hearing loss and by cochlear implant users. In **Pitch**. Springer, New York, NY.

- MOSES, P. (1954). **The Voice of Neurosis**. Grune and Stratton, New York.
- NAYEBI, A., BEAR, D., KUBILIUS, J., KAR, K., GANGULI, S., SUSSILLO, D., ... & YAMINS, D. L. (2018). Task-driven convolutional recurrent models of the visual system. In **Advances in Neural Information Processing Systems**.
- NIEDENTHAL, P. M., & RIC, F. (2017). **Psychology of emotion**. Psychology Press.
- NIPARKO, J. K. (2004). Cochlear implants: Clinical application. In F. G. Zeng, A. Popper, & R. Fay (Eds.), **Cochlear implants: Auditory prosthesis and electric hearing**. New York, NY: Springer-Verlag.
- PALO, H. K., CHANDRA, M., & MOHANTY, M. N. (2018). Recognition of human speech emotion using variants of mel-frequency cepstral coefficients. In **Advances in Systems, Control and Automation**. Springer, Singapore.
- PANKSEPP, J. (1998). **Affective neuroscience: The foundations of human and animal emotions**. Oxford: Oxford University Press.
- PEAR, T. H. (1931). **Voice and personality**. Chapman and Hall, London.
- PITTENGER, R. E., HOCKETT, C. F., & DANEHY, J. J. (1960). **The first five minutes: A sample of microscopic interview analysis**. Martineau, Ithaca, NY.
- POLLAK, S. D. (2013). Emotion and Learning: New approaches to the old nature-nurture debate. In S. Gelman & M. Banaji (Eds.), **Navigating the social world: What infants, children, and other species can teach us**. New York: Oxford University Press.
- RUSSELL, S VE NORVIG. P, (2009) **Artificial intelligence: A modern approach**, Edition 3, Prentice Hall.
- SAARNI, C. (2008). The interface of emotional development with social context. In M. Lewis, J. M. Haviland-Jones, L. F. Barrett, M. Lewis, J. M. Haviland-Jones & L. F. Barrett (Eds.), **Handbook of emotions** (3rd ed.) New York: Guilford Press

- SAARNI, C., CAMPOS, J. J., CAMRAS, L. A., & WITHERINGTON, D. (2008). Principles of emotion and emotional competence. In W. Damon, R. M. Lerner, D. Kuhn, R. Siegler & N. Eisenberg (Eds.), **Child and adolescent development: An advanced course**. Hoboken, NJ: John Wiley & Sons.
- SCHERER, K. R. (1982). **Vokale Kommunikation: Nonverbale Aspekte des Sprachverhaltens**. Beltz.
- SETHARES, W. A. (2005). **Tuning, timbre, spectrum, scale**. (Second Edition). United States of America: Springer-Verlag London Limited.
- STORY, B. H. (2015). Mechanisms of voice production. **The handbook of speech production**, 1st ed. Oxford: Wiley Blackwell Publishing.
- ŞEN, Z. (2004). **Yapay Sinir Ağları İlkeleri**. İstanbul: Su Vakfı Yayınları
- TOMKINS, S. S. (1962). **Affect, imagery, consciousness: Vol. I: The positive affects**. New York: Springer-Verlag.
- TOMKINS, S. S. (1963). **Affect, imagery, consciousness: Vol. 2: The negative affects**. New York: Springer-Verlag.
- TRAINOR, L. J., & CORRIGALL, K. A. (2010). Music acquisition and effects of musical experience. **Music perception**. Springer, New York, NY.
- WOLFE, J. (2018). **Cochlear implants: audiologic management and considerations for implantable hearing devices**. Plural Publishing.
- WOLFE, J. (2019). **Cochlear implants: Audiologic management and considerations for implantable hearing devices**. San Diego, CA: Plural Publishing.

MAKALELER

- ABDEL-HAMID, O., MOHAMED, A. R., JIANG, H., DENG, L., PENN, G., & YU, D. (2014). Convolutional neural networks for speech recognition. **IEEE/ACM Transactions on audio, speech, and language processing**, 22(10), 1533-1545.

- ADAM, B., & SMITH, I. F. (2008). Reinforcement learning for structural control. **Journal of Computing in Civil Engineering**, 22(2), 133-139.
- AKÇAY, M. B., & OĞUZ, K. (2020). Speech emotion recognition: Emotional models, databases, features, preprocessing methods, supporting modalities, and classifiers. **Speech Communication**, 116, 56-76.
- ALHLFFEE, M. (2020). MFCC-Based Feature Extraction Model for Long Time Period Emotion Speech Using CNN. **Revue d'Intelligence Artificielle**. Argentina.
- ALHUSSEİN, M., & MUHAMMAD, G. (2018). Voice pathology detection using deep learning on mobile healthcare framework. **IEEE Access**, 6, 41034-41041.
- ALLPORT, G. W., & CANTRIL, H. (1934). Judging personality from voice. **The Journal of Social Psychology**, 5(1), 37-55.
- ALPERT, M., KURTZBERG, R. L., & FRIEDHOFF, A. J. (1963). Transient voice changes associated with emotional stimuli. **Archives of General Psychiatry**, 8(4), 362-365.
- ANVARJON, T., & KWON, S. (2020). Deep-net: A lightweight CNN-based speech emotion recognition system using deep frequency features. **Sensors**, 20(18), 5212.
- AOUANI, H., & AYED, Y. B. (2020). Speech Emotion Recognition with deep learning. **Procedia Computer Science**, 176, 251-260.
- ARIAS-VERGARA, T., KLUMPP, P., VASQUEZ-CORREA, J. C., NOTH, E., OROZCO-ARROYAVE, J. R., & SCHUSTER, M. (2020). Multi-channel spectrograms for speech processing applications using deep learning methods. **Pattern Analysis and Applications**, 1-9.
- AVCI, D. E., & Yağbasan, R. (2008). Beyin Yarı Kürelerinin Baskın Olarak Kullanılmasına Yönelik Öğretim Stratejileri. **Gazi University Journal of Gazi Educational Faculty (GUJGEF)**, 28(2).
- BACHOROWSKI, J. A. (1999). Vocal expression and perception of emotion. **Current directions in psychological science**, 8(2), 53-57.

- BADSHAH, A. M., RAHIM, N., ULLAH, N., AHMAD, J., MUHAMMAD, K., LEE, M. Y., ... & BAIK, S. W. (2019). Deep features-based speech emotion recognition for smart affective services. **Multimedia Tools and Applications**, 78(5), 5571-5589.
- BÄNZIGER, T., MORTILLARO, M., AND SCHERER, K. R. (2012). Introducing the Geneva multimodal expression corpus for experimental research on emotion perception. **Emotion** 12, 1161–1179.
- BARRETT, K. C., CHATTERJEE, M., CALDWELL, M. T., DEROCHE, M. L., JIRADEJVONG, P., KULKARNI, A. M., & LIMB, C. J. (2020). Perception of Child-Directed Versus Adult-Directed Emotional Speech in Pediatric Cochlear Implant Users. **Ear and hearing**, 41(5), 1372-1382.
- BEAM, A. L., & KOHANE, I. S. (2018). Big data and machine learning in health care. **Jama**, 319(13), 1317-1318.
- BEKKERMAN, R., EL-YANIV, R., TISHBY, N., & WINTER, Y. (2002). Distributional word clusters vs. words for text categorization. **Journal of Machine Learning Research**, 1(1-48).
- BILGER RC (1977) Psychoacoustic evaluation of present prostheses. **Ann Otol Rhinol Laryngol Suppl** 86:92–104
- BINI, S. A. (2018). Artificial intelligence, machine learning, deep learning, and cognitive computing: what do these terms mean and how will they impact health care?. **The Journal of arthroplasty**, 33(8), 2358-2361.
- BLALOCK, D., ORTIZ, J. J. G., FRANKLE, J., & GUTTAG, J. (2020). What is the state of neural network pruning?. **arXiv preprint** arXiv:2003.03033.
- BUCK, R. (1999). The biological affects: A typology. **Psychological Review**, 106(2), 301.
- BUCK, R., LOSOW, J., MURPHY, M., & COSTANZO, P. (1992). Social facilitation and inhibition of emotional expression and communication. **Journal of Personality and Social Psychology**, 63(6), 962–968.
- BUONO, G. H., CRUKLEY, J., HORNSBY, B. W., & PICOU, E. M. (2020). Loss of high-or low-frequency audibility can partially explain effects of

- hearing loss on emotional responses to non-speech sounds. **Hearing Research**, 108153.
- CAFFI, C., & JANNEY, R. W. (1994). Toward a pragmatics of emotive communication. **Journal of pragmatics**, 22(3-4), 325-373.
- CANNON, S. A., & CHATTERJEE, M. (2019). Voice Emotion Recognition by Children with Mild to Moderate Hearing Loss. **Ear and hearing**, 40(3), 477.
- CENGIL, E., & ÇINAR, A. (2016). A New Approach For Image Classification: Convolutional Neural Network. **European Journal of Technic**, 6(2).
- CHATTERJEE, M., & PENG, S. C. (2008). Processing F0 with cochlear implants: Modulation frequency discrimination and speech intonation recognition. **Hearing research**, 235(1-2), 143-156.
- CHATTERJEE, M., KULKARNI, A. M., SIDDIQUI, R. M., CHRISTENSEN, J. A., HOZAN, M., SIS, J. L., & DAMM, S. A. (2019). Acoustics of emotional prosody produced by prelingually deaf children with cochlear implants. **Frontiers in psychology**, 10, 2190.
- CHATTERJEE, M., ZION, D. J., DEROCHE, M. L., BURIANEK, B. A., LIMB, C. J., GOREN, A. P., ... & CHRISTENSEN, J. A. (2015). Voice emotion recognition by cochlear-implanted children and their normally-hearing peers. **Hearing research**, 322, 151-162.
- CHAVHAN, Y., DHORE, M. L., & YESAWARE, P. (2010). Speech emotion recognition using support vector machine. **International Journal of Computer Applications**, 1(20), 6-9.
- CHEN, M., HE, X., YANG, J., & ZHANG, H. (2018). 3-D convolutional recurrent neural networks with attention model for speech emotion recognition. **IEEE Signal Processing Letters**, 25(10), 1440-1444.
- CHEN, X. W., & LIN, X. (2014). Big data deep learning: challenges and perspectives. **IEEE access**, 2, 514-525.
- CHING, T. Y. C., DILLON, H., LEIGH, G., & CUPPLES, L. (2018). Learning from the Longitudinal Outcomes of Children with Hearing Impairment

- (LOCHI) study: Summary of 5-year findings and implications. **International Journal of Audiology**, 57(S-2), S-105–S-111.
- CHRISTENSEN, J. A., SIS, J., KULKARNI, A. M., & CHATTERJEE, M. (2019). Effects of age and hearing loss on the recognition of emotions in speech. **Ear and hearing**, 40(5), 1069.
- CHRONAKI, G., HADWIN, J. A., GARNER, M., MAURAGE, P., & SONUGA-BARKE, E. J. (2015). The development of emotion recognition from facial expressions and non-linguistic vocalizations during childhood. **British Journal of Developmental Psychology**, 33(2), 218-236.
- CHUENWATTANAPRANITHI, S., XU, Y., THIPAKORN, B., & MANEEWONGVATANA, S. (2008). Encoding emotions in speech with the size code. **Phonetica**, 65(4), 210-230.
- COCHRAN, W. T., COOLEY, J. W., FAVIN, D. L., HELMS, H. D., KAENEL, R. A., LANG, W. W., ... & WELCH, P. D. (1967). What is the fast Fourier transform?. **Proceedings of the IEEE**, 55(10), 1664-1674.
- COHEN, O., ZHU, B., & ROSEN, M. S. (2018). MR fingerprinting deep reconstruction network (DRONE). **Magnetic resonance in medicine**, 80(3), 885-894.
- COLE, P. M., MARTIN, S. E., & DENNIS, T. A. (2004). Emotion regulation as a scientific construct: Methodological challenges and directions for child development research. **Child development**, 75(2), 317-333.
- CUNNINGHAM, M., COX, E. O., & Committee on Practice and Ambulatory Medicine. (2003). Hearing assessment in infants and children: recommendations beyond neonatal screening. **Pediatrics**, 111(2), 436-440.
- CÜCELOĞLU, D. (1998), **İnsan ve Davranışı**. Remzi Kitabevi, İstanbul.
- ÇİĞDEM, A. C. I., & ÇIRAK, A. (2019). Türkçe Haber Metinlerinin Konvolüsyonel Sinir Ağları ve Word2Vec Kullanılarak Sınıflandırılması. **Bilişim Teknolojileri Dergisi**, 12(3), 219-228.
- ÇUHADAR, M., & KAYACAN, C. (2005). Yapay Sinir Ağları Kullanılarak Konaklama İşletmelerinde Doluluk Oranı Tahmini: Türkiye'deki

Konaklama İşletmeleri Üzerine Bir Deneme. *Anatolia: Turizm Araştırmaları Dergisi*, 16(1).

- DALLA BELLA, S., PERETZ, I., ROUSSEAU, L., & GOSSELIN, N. (2001). A developmental study of the affective value of tempo and mode in music. *Cognition*, 80(3), B1-B10.
- DAMM, S. A., SIS, J. L., KULKARNI, A. M., & CHATTERJEE, M. (2019). How vocal emotions produced by children with cochlear implants are perceived by their hearing peers. *Journal of Speech, Language, and Hearing Research*, 62(10), 3728-3740.
- DAREKAR, R. V., & DHANDE, A. P. (2018). Emotion recognition from Marathi speech database using adaptive artificial neural network. *Biologically Inspired Cognitive Architectures*, 23, 35-42.
- DEROCHE, M. L., FELEZEU, M., PAQUETTE, S., ZEITOUNI, A., & LEHMANN, A. (2019). Neurophysiological Differences in Emotional Processing by Cochlear Implant Users, Extending Beyond the Realm of Speech. *Ear and hearing*, 40(5), 1197-1209.
- DEROCHE, M. L., ZION, D. J., SCHURMAN, J. R., & CHATTERJEE, M. (2012). Sensitivity of school-aged children to pitch-related cues. *The Journal of the Acoustical Society of America*, 131(4), 2938-2947.
- DESJARDIN, J. L., & EISENBERG, L. S. (2007). Maternal contributions: Supporting language development in young children with cochlear implants. *Ear and hearing*, 28(4), 456-469.
- DETTMAN, S. J., DOWELL, R. C., CHOO, D., ARNOTT, W., ABRAHAMS, Y., DAVIS, A., . . . BRIGGS, R. S. (2016). Long term communication outcomes for children receiving cochlear implants younger than 12 months: A multi-centre study. *Otology and Neurotology*, 37(2), e82–e95.
- DICARLO, J. J., ZOCCOLAN, D., & RUST, N. C. (2012). How does the brain solve visual object recognition?. *Neuron*, 73(3), 415-434.

- DJOURNO A, EYRIES C (1957) Auditory prosthesis by means of a distant electrical stimulation of the sensory nerve with the use of an indwelt coiling. **Presse Med** 65(63):1417–1417
- DORMAN, M. F., & GIFFORD, R. H. (2017). Speech understanding in complex listening environments by listeners fit with cochlear implants. **Journal of Speech, Language, and Hearing Research**, 60, 3019–3026.
- DUNN, C. C., TYLER, R. S., WITT, S., JI, H., & GANTZ, B. J. (2012). Sequential bilateral cochlear implantation: speech perception and localization pre-and post-second cochlear implantation. **American journal of audiology**.
- DYCK, M. J., & DENVER, E. (2003). Can the emotion recognition ability of deaf children be enhanced? A pilot study. **Journal of Deaf Studies and Deaf Education**, 8(3), 348-356.
- DYCK, M. J., FARRUGIA, C., SHOCHET, I. M., & HOLMES-BROWN, M. (2004). Emotion recognition/understanding ability in hearing or vision-impaired children: do sounds, sights, or words make the difference?. **Journal of Child Psychology and Psychiatry**, 45(4), 789-800.
- EGILMEZ, O. K., & KALCIOGLU, M. T. (2015). Cochlear implant: indications, contraindications and complications. **Scripta Scientifica Medica**, 47(4), 9-16.
- EKMAN, P. (1992). An argument for basic emotions. **Cognition & Emotion**, 6(3–4), 169–200.
- EKMAN, P. (1992). Are there basic emotions? **Psychol Rey**. Pp. 550- 553
- EKMAN, P., & CORDARO, D. (2011). What is meant by calling emotions basic. **Emotion Review**, 3(4), 364–370.
- EKMAN, P., & FRIESEN, W. V. (1971). Constants across cultures in the face and emotion. **Journal of Personality and Social Psychology**, 17(2), 124.
- ESKİDERE, Ö., & ERTAŞ, F. (2009). Mel frekansı kepstrem katsayılarındaki değişimlerin konuşmacı tanımaya etkisi. **Uludağ University Journal of The Faculty of Engineering**, 14(2).

- ESMAEILPOUR, M., CARDINAL, P., & KOERICH, A. L. (2020). From Sound Representation to Model Robustness. **arXiv preprint** arXiv:2007.13703.
- ESTABROOKS, W. (Ed). (2012). **101 Frequently asked questions about auditory-verbal practice**. Washington DC: The Alexander Graham Bell Association for the Deaf and Hard of Hearing.
- EVERHARDT, M. K., SARAMPALIS, A., COLER, M., BASKENT, D., & LOWIE, W. (2020). Meta-Analysis on the Identification of Linguistic and Emotional Prosody in Cochlear Implant Users and Vocoder Simulations. **Ear and Hearing**: Volume 41 - Issue 5 - p 1092-1102
- FAULKNER, A., ROSEN, S., & SMITH, C. (2000). Effects of the salience of pitch and periodicity information on the intelligibility of four-channel vocoded speech: implications for cochlear implants. **J Acoust Soc Am**, 108, 1877–1887.
- FAYEK, H. M., LECH, M., & CAVEDON, L. (2017). Evaluating deep learning architectures for Speech Emotion Recognition. **Neural Networks**, 92, 60-68.
- FLOCK, Å., & ORMAN, S. (1983). Micromechanical properties of sensory hairs on receptor cells of the inner ear. **Hearing research**, 11(3), 249-260.
- FREDRICKSON, B. L. (2001). The role of positive emotions in positive psychology: The broaden-and-build theory of positive emotions. **American Psychologist**, 56(3), 218.
- FRETZ, R. J., & FRAVEL, R. P. (1985). Design and function: a physical and electrical description of the 3M House cochlear implant system. **Ear and Hearing**, 6(3 Suppl), 14S-19S.
- FRICK, R. W. (1985). Communicating emotion: The role of prosodic features. **Psychological bulletin**, 97(3), 412.
- FU, Q. J., & NOGAKI, G. (2005). Noise susceptibility of cochlear implant users: the role of spectral resolution and smearing. **Journal of the Association for Research in Otolaryngology**, 6(1), 19-27.
- GEERS, A. E., MITCHELL, C. M., WARNER-CZYZ, A., WANG, N. Y., EISENBERG, L. S., & THE CDACI INVESTIGATIVE TEAM. (2017).

- Early sign language exposure and cochlear implantation benefits. **Pediatrics**, 140(1), e20163489.
- GERÇEKER, M., YORULMAZ, İ., & URAL, A. (2000). Ses ve konuşma. **KBB ve Baş Boyun Cerrahisi Dergisi**, 8(1), 71-78.
- GEURTS, L., & WOUTERS, J. (2001). Coding of the fundamental frequency in continuous interleaved sampling processors for cochlear implants. **J Acoust Soc Am**, 109, 713– 726.
- GFELLER, K., TURNER, C., MEHR, M., WOODWORTH, G., FEARN, R., KNUTSON, J. F., ... & STORDAHL, J. (2002). Recognition of familiar melodies by adult cochlear implant recipients and normal-hearing adults. **Cochlear implants international**, 3(1), 29-53.
- GILBERS, S., FULLER, C., GILBERS, D., BROERSMA, M., GOUDBEEK, M., FREE, R., & BAŞKENT, D. (2015). Normal-hearing listeners' and cochlear implant users' perception of pitch cues in emotional speech. **i-Perception**, 6(5), 0301006615599139.
- GOMES, G. S. D. S., LUDERMIR, T. B., & LIMA, L. M. (2011). Comparison of new activation functions in neural network for forecasting financial time series. **Neural Computing and Applications**, 20(3), 417-439.
- GOSSELIN, N., PERETZ, I., JOHNSEN, E., & ADOLPHS, R. (2007). Amygdala damage impairs emotion recognition from music. **Neuropsychologia**, 45(2), 236-244.
- GREEN, T., FAULKNER, A., & ROSEN, S. (2002). Spectral and temporal cues to pitch in noise-excited vocoder simulations of continuous-interleaved-sampling cochlear implants. **J Acoust Soc Am**, 112, 2155–2164.
- GREEN, T., FAULKNER, A., & ROSEN, S. (2004). Enhancing temporal cues to voice pitch in continuous interleaved sampling cochlear implants. **J Acoust Soc Am**, 116, 2298 –2310.
- GRIMES, A. (2017). Kids need two ears. **Audiology Today**, 29(1), 54–58.
- GUO, L., WANG, L., DANG, J., LIU, Z., & GUAN, H. (2019). Exploration of complementary features for speech emotion recognition based on kernel extreme learning machine. **IEEE Access**, 7, 75798-75809.

- GÜÇLÜ, U., & VAN GERVEN, M. A. (2015). Deep neural networks reveal a gradient in the complexity of neural representations across the ventral stream. **Journal of Neuroscience**, 35(27), 10005-10014.
- HARGREAVES, W. A., STARKWEATHER, J. A., & BLACKER, K. H. (1965). Voice quality in depression. **Journal of Abnormal Psychology**, 70(3), 218.
- HARMS, M. B., MARTIN, A., & WALLACE, G. L. (2010). Facial emotion recognition in autism spectrum disorders: a review of behavioral and neuroimaging studies. **Neuropsychology review**, 20(3), 290-322.
- HAUSER, M. D. (1997). Artifactual kinds and functional design features: what a primate understands without language. **Cognition**, 64(3), 285-308.
- HAUSER, M. D. (1993). The evolution of nonhuman primate vocalizations: effects of phylogeny, body weight, and social context. **The American Naturalist**, 142(3), 528- 542.
- HENGLIN, M., STEIN, G., HUSHCHA, P. V., SNOEK, J., WILTSCHKO, A. B., & CHENG, S. (2017). Machine learning approaches in cardiovascular imaging. **Circulation: Cardiovascular Imaging**, 10(10), e005614.
- HERZOG, H. (1933). Stimme und Persönlichkeit. **Zeitschrift für Psychologie**, 190, 300-369.
- HOPYAN-MISAKYAN, T. M., GORDON, K. A., DENNIS, M., & PAPSIN, B. C. (2009). Recognition of affective speech prosody and facial affect in deaf children with unilateral right cochlear implants. **Child Neuropsychology**, 15(2), 136-146.
- HORNÍK, K., STINCHCOMBE, M., & WHITE, H. (1989). Multilayer feedforward networks are universal approximators. **Neural networks**, 2(5), 359-366.
- HORSTMANN, G. (2003). What do facial expressions convey: Feeling states, behavioral intentions, or actions requests? **Emotion**, 3(2), 150.
- HOSSAIN, M. S., & MUHAMMAD, G. (2019). Emotion recognition using deep learning approach from audio–visual emotional big data. **Information Fusion**, 49, 69-78.

- HOUSE WF, URBAN J (1973) Long term results of electrode implantation and electronic stimulation of the cochlea in man. **Ann Otol Rhinol Laryngol** 82(4):504–517
- HUNG, C. P., KREIMAN, G., POGGIO, T., & DICARLO, J. J. (2005). Fast readout of object identity from macaque inferior temporal cortex. **Science**, 310(5749), 863-866.
- IAKOVIDES, S. A., ILIADOU, V. T., BIZELI, V. T., KAPRINIS, S. G., FOUNTOULAKIS, K. N., & KAPRINIS, G. S. (2004). Psychophysiology and psychoacoustics of music: Perception of complex sound in normal subjects and psychiatric patients. **Annals Of General Hospital Psychiatry**, 3(1), 1-4.
- IMAIZUMI, S., MORI, K., KIRITANI, S., KAWASHIMA, R., SUGIURA, M., FUKUDA, H., ... & KOJIMA, S. (1997). Vocal identification of speaker and emotion activates different brain regions. **Neuroreport**, 8(12), 2809-2812.
- ISSA, D., DEMİRCİ, M. F., & YAZICI, A. (2020). Speech emotion recognition with deep convolutional neural networks. **Biomedical Signal Processing and Control**, 59, 101894.
- ISSERLIN, M. (1925). Psychologisch- phonetische untersuchungen. **Zeitschrift für die gesamte Neurologie und Psychiatrie**, 94(1), 437-448.
- IZARD, C. E. (2007). Basic emotions, natural kinds, emotion schemas, and a new paradigm. **Perspectives on Psychological Science**, 2(3), 260–280.
- JIAM, N. T., CALDWELL, M., DEROCHE, M. L., CHATTERJEE, M., & LIMB, C. J. (2017). Voice emotion perception and production in cochlear implant users. **Hearing research**, 352, 30-39.
- JIANG, P., FU, H., TAO, H., LEI, P., & ZHAO, L. (2019). Parallelized convolutional recurrent neural network with spectral features for speech emotion recognition. **IEEE Access**, 7, 90368-90377.
- JIN, K. H., MCCANN, M. T., FROUSTEY, E., & UNSER, M. (2017). Deep convolutional neural network for inverse problems in imaging. **IEEE Transactions on Image Processing**, 26(9), 4509-4522.

- JOHNSON-LAIRD, P. N., & OATLEY, K. (1992). Basic emotions, rationality, and folk theory. **Cognition & Emotion**, 6(3–4), 201–223.
- KAYA, U., YILMAZ, A., & DİKMEN, Y. (2019). Sağlık Alanında Kullanılan Derin Öğrenme Yöntemleri. **Avrupa Bilim ve Teknoloji Dergisi**, (16), 792-808.
- KELTNER, D., & GROSS, J. J. (1999). Functional accounts of emotions. **Cognition & Emotion**, 13(5), 467-480.
- KELTNER, D., & KRING, A. M. (1998). Emotion, social function, and psychopathology. **Review of General Psychology**, 2(3), 320.
- KETELAAR, L., RIEFFE, C., WIEFFERINK, C. H., & FRIJNS, J. H. (2013). Social competence and empathy in young children with cochlear implants and with normal hearing. **The laryngoscope**, 123(2), 518-523.
- KHALIGH-RAZAVI, S. M., & KRIEGESKORTE, N. (2014). Deep supervised, but not unsupervised, models may explain IT cortical representation. **PLoS computational biology**, 10(11), e1003915.
- KILIÇ, M. A. (2002). Larenksin fonksiyonel anatomisi ve ses fizyolojisi. **Türkiye Klinikleri KBB Dergisi**, 2(3), 1-8.
- KILGARD, M. P., VAZQUEZ, J. L., ENGINEER, N. D., & PANDYA, P. K. (2007). Experience dependent plasticity alters cortical synchronization. **Hearing Research**, 229, 171–179.
- KONG, Y. Y., CRUZ, R., JONES, J. A., & ZENG, F. G. (2004). Music perception with temporal cues in acoustic and electric hearing. **Ear and hearing**, 25(2), 173-185.
- KOROTCOV, A., TKACHENKO, V., RUSSO, D. P., & EKINS, S. (2017). Comparison of deep learning with multiple machine learning methods and metrics using diverse drug discovery data sets. **Molecular pharmaceuticals**, 14(12), 4462-4475.
- KRAL, A., & LENARZ, T. (2015). How the brain learns to listen: Deafness and the bionic ear. **E-Neuroforum**, 6(1), 21–28.

- KRAL, A., KRONENBERGER, W. G., PISONI, D. B., & O'DONOGHUE, G. M. (2016). Neurocognitive factors in sensory restoration of early deafness: A connectome model. **The Lancet Neurology**, 15(6), 610–621.
- KUCHIBHOTLA, S., VANKAYALAPATI, H. D., VADDI, R. S., & ANNE, K. R. (2014). A comparative analysis of classifiers in emotion recognition through acoustic features. **International Journal of Speech Technology**, 17(4), 401-408.
- KUTLU, M. O. (2010). Beyin Temelli Öğrenmenin Sosyal Bilgiler Dersi Öğretiminde Uygulanması. **Cukurova University Faculty of Education Journal**, 39.
- LAMMERS, M. J., VENEKAMP, R. P., GROLMAN, W., & VAN DER HEIJDEN, G. J. (2014). Bilateral cochlear implantation in children and the impact of the inter-implant interval. **The Laryngoscope**, 124(4), 993-999.
- LATIF, S., RANA, R., QADIR, J., & EPS, J. (2017). Variational autoencoders for learning latent representations of speech emotion: A preliminary study. **arXiv preprint** arXiv:1712.08708.
- LECUN, Y., BENGIO, Y., & HINTON, G. (2015). Deep learning. **Nature**, 521(7553), 436- 444.
- LEE, J. G., JUN, S., CHO, Y. W., LEE, H., KIM, G. B., SEO, J. B., & KIM, N. (2017). Deep learning in medical imaging: general overview. **Korean journal of radiology**, 18(4), 570.
- LEE, J., LEE, D., CHOI, J. Y., SHIN, D., SHIN, H. G., & LEE, J. (2020). Artificial neural network for myelin water imaging. **Magnetic Resonance in Medicine**, 83(5), 1875-1883.
- LIEBERMAN, P., & MICHAELS, S. B. (1962). Some aspects of fundamental frequency and envelope amplitude as related to the emotional content of speech. **The Journal of the Acoustical Society of America**, 34(7), 922-927.
- LIMB, C. J., & ROY, A. T. (2014). Technological, biological, and acoustical constraints to music perception in cochlear implant users. **Hearing research**, 308, 13-26.

- LIN, J. C., WU, C. H., & WEI, W. L. (2011). Error weighted semi-coupled hidden Markov model for audio-visual emotion recognition. **IEEE Transactions on Multimedia**, 14(1), 142-156.
- LITOVSKY, R. (2010). Bilateral cochlear implants: Are two ears better than one? **ASHA Leader**, 12(2), 14–17.
- LOIZOU PC (1999) Introduction to cochlear implants. **IEEE Engineering in Medicine and Biology Magazine**.18(1):32–42
- LOVETT, R., ELIZABETH, S., VICKERS, D. A., & SUMMERFIELD, A. Q. (2015). Bilateral cochlear implantation for hearing-impaired children: Criterion of candidacy derived from an observational study. **Ear and Hearing**, 36(1), 14–23.
- LU, H., LI, Y., CHEN, M., KIM, H., & SERIKAWA, S. (2018). Brain intelligence: go beyond artificial intelligence. **Mobile Networks and Applications**, 23(2), 368-375.
- LUO, X., & FU, Q. J. (2004). Enhancing Chinese tone recognition by manipulating amplitude envelope: Implications for cochlear implants. **The Journal of the Acoustical Society of America**, 116(6), 3659-3667.
- LUO, X., FU, Q. J., & GALVIN III, J. J. (2007). Cochlear implants special issue article: Vocal emotion recognition by normal-hearing listeners and cochlear implant users. **Trends in amplification**, 11(4), 301-315.
- MAJAJ, N. J., HONG, H., SOLOMON, E. A., & DICARLO, J. J. (2015). Simple learned weighted sums of inferior temporal neuronal firing rates accurately predict human core object recognition performance. **Journal of Neuroscience**, 35(39), 13402-13418.
- MCDERMOTT, H. J. (2004). Music perception with cochlear implants: a review. **Trends in amplification**, 8(2), 49-82.
- MEHU, M., & DUNBAR, R. I. (2008). Naturalistic observations of smiling and laughter in human group interactions. **Behaviour**, 145(12), 1747–1780.
- MENG, H., YAN, T., YUAN, F., & WEI, H. (2019). Speech emotion recognition from 3D log-mel spectrograms with deep learning network. **IEEE access**, 7, 125868-125881.

- MOHN, C., ARGSTATTER, H., & WILKER, F. W. (2011). Perception of six basic emotions in music. **Psychology of music**, 39(4), 503-517.
- MOHRI, M., ROSTAMIZADEH, A., & TALWALKAR, A. (2018). Foundations of machine learning. **MIT press**.
- MORTON, E. S. (1977). On the occurrence and significance of motivation-structural rules in some bird and mammal sounds. **The American Naturalist**, 111(981), 855-869.
- MORTON, J. B., & TREHUB, S. E. (2001). Children's understanding of emotion in speech. **Child development**, 72(3), 834-843.
- MOST, T., & AVINER, C. (2009). Auditory, visual, and auditory–visual perception of emotions by individuals with cochlear implants, hearing aids, and normal hearing. **Journal of Deaf Studies and Deaf Education**, 14(4), 449-464.
- MOST, T., & MICHAEĹIS, H. (2012). Auditory, visual, and auditory–visual perceptions of emotions by young children with hearing loss versus children with normal hearing. **Journal of Speech, Language, and Hearing Research**
- MURRAY, I. R., & ARNOTT, J. L. (1993). 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.
- MURUGASU, E., YONG, T. T., & YOON, C. P. (2004). Invasive middle ear cholesterol granuloma involving the basal turn of the cochlea with profound sensorineural hearing loss. **Otology & Neurotology**, 25(3), 231-235.
- MUSIEK, F. E. (2009). The human auditory cortex: Interesting anatomical and clinical perspectives. **Audiology Today**, 21(4), 26–37.
- NAGELS, L., GAUDRAIN, E., VICKERS, D., LOPES, M. M., HENDRIKS, P., & BASKENT, D. (2020). Development of vocal emotion recognition in school-age children: The EmoHI test for hearing-impaired populations. **PeerJ**, 8, e8773.

- NAJAFABADI, M. M., VILLANUSTRE, F., KHOSHGOFTAAR, T. M., SELIYA, N., WALD, R., & MUHAREMAGIC, E. (2015). Deep learning applications and challenges in big data analytics. **Journal of big data**, 2(1), 1-21.
- NAKATA, T., TREHUB, S. E., & KANDA, Y. (2012). Effect of cochlear implants on children's perception and production of speech prosody. **The Journal of the Acoustical Society of America**, 131(2), 1307-1314.
- NELSON, N. L., & RUSSELL, J. A. (2011). Preschoolers' use of dynamic facial, bodily, and vocal cues to emotion. **Journal of experimental child psychology**, 110(1), 52-61.
- NIEDENTHAL, P. M., & BRAUER, M. (2012). Social functionality of human emotion. **Annual Review of Psychology**, 63(1), 259–285.
- OHALA, J. J. (1984). An ethological perspective on common cross-language utilization of F₀ of voice. **Phonetica**, 41(1), 1-16.
- ORTONY, A., & TURNER, T. J. (1990). What's basic about basic emotions? **Psychological Review**, 97(3), 315–331.
- OSBERGER, M. J., ROBBINS, A. M., MIYAMOTO, R. T., et al. (1991). Speech perception abilities of children with cochlear implants, tactile aids, or hearing aids. **Am J Otol**, 12, 105S–115S.
- OSTWALD, P. F. (1964). Measurement of verbal and vocal behavior. Acoustic manifestations of emotional disturbance. **Research Publications-Association for Research in Nervous and Mental Disease**, 42, 450-465.
- OHMAN, A., & MINEKA, S. (2001). Fears, phobias, and preparedness: Toward an evolved module of fear and fear learning. **Psychological Review**, 108(3), 483.
- PANCH, T., SZOLOVITS, P., & ATUN, R. (2018). Artificial intelligence, machine learning and health systems. **Journal of global health**, 8(2).
- PANKSEPP, J. (2007). Criteria for basic emotions: Is DISGUST a primary 'emotion'? **Cognition and Emotion**, 21(8), 1819–1828.

- PENG, S. C., TOMBLIN, J. B., & TURNER, C. W. (2008). Production and perception of speech intonation in pediatric cochlear implant recipients and individuals with normal hearing. **Ear and hearing**, 29(3), 336-351.
- PENG, S. C., TOMBLIN, J. B., CHEUNG, H., LIN, Y. S., & WANG, L. S. (2004). Perception and production of Mandarin tones in prelingually deaf children with cochlear implants. **Ear and hearing**, 25(3), 251-264.
- PENG, S., TOMBLIN, J. B., SPENCER, L. J., et al. (2007). Imitative production of rising speech intonation in pediatric cochlear implant recipients. **J Speech Lang Hear Res**, 50, 1210–1227.
- PERCY-SMITH, L., TONNING, T. L., JOSVASSEN, J. L., MIKKELSEN, J. H., NISSEN, L., DIELEMAN, E., CAYE-THOMASEN, P. (2018). Auditory-verbal habilitation is associated with improved outcome for children with cochlear implant. **Cochlear Implants International**, 19(1), 38– 45.
- PINTO, N., COX, D. D., & DICARLO, J. J. (2008). Why is real-world visual object recognition hard?. **PLoS Comput Biol**, 4(1), e27.
- POLLAK, S. D. (2012). The role of parenting in the emergence of human emotion: New approaches to the old nature-nurture debate. **Parenting**, 12(2–3), 232–242.
- PRATT, R. L., & DOAK, P. E. (1976). A subjective rating scale for timbre. **Journal of Sound and Vibration**, 45(3), 317-328.
- PURWINS, H., LI, B., VIRTANEN, T., SCHLUTER, J., CHANG, S. Y., & SAINATH, T. (2019). Deep learning for audio signal processing. **IEEE Journal of Selected Topics in Signal Processing**, 13(2), 206-219.
- QIU, J., WU, Q., DING, G., XU, Y., & FENG, S. (2016). A survey of machine learning for big data processing. **EURASIP Journal on Advances in Signal Processing**, 2016(1), 67.
- RAO, K. S., KOOLAGUDI, S. G., & VEMPADA, R. R. (2013). Emotion recognition from speech using global and local prosodic features. **International journal of speech technology**, 16(2), 143-160.
- RAO, K. S., KUMAR, T. P., ANUSHA, K., LEELA, B., BHAVANA, I., & GOWTHAM, S. V. S. K. (2012). Emotion recognition from

- speech. **International Journal of Computer Science and Information Technologies**, 3(2), 3603-3607.
- REZAEI, N., & SALEHI, A. (2006). An introduction to speech sciences (acoustic analysis of speech). **Iranian Rehabilitation Journal**, 4(1), 5-14.
- RODGER, H., VIZIOLI, L., OUYANG, X., & CALDARA, R. (2015). Mapping the development of facial expression recognition. **Developmental science**, 18(6), 926-939.
- RUBEN, R. J. (1997). A time frame of critical/sensitive periods of language development. **Acta oto-laryngologica**, 117(2), 202-205.
- RUSSELL, J. A. (1980). A circumplex model of affect. **Journal of personality and social psychology**, 39(6), 1161.
- SAJJAD, M., & KWON, S. (2020). Clustering-based speech emotion recognition by incorporating learned features and deep BiLSTM. **IEEE Access**, 8, 79861-79875.
- SAMMLER, D., GRIGUTSCH, M., FRITZ, T., & KOELSCH, S. (2007). Music and emotion: electrophysiological correlates of the processing of pleasant and unpleasant music. **Psychophysiology**, 44(2), 293-304.
- SARANT, J., HARRIS, D., BENNET, L., & BANT, S. (2014). Bilateral versus unilateral cochlear implants in children: A study of spoken language outcomes. **Ear and Hearing**, 35(4), 396-409.
- SATALOFF, R. T., HEMAN-ACKAH, Y. D., & HAWKSHAW, M. J. (2007). Clinical anatomy and physiology of the voice. **Otolaryngologic clinics of north America**, 40(5), 909-929.
- SAUTER, D. A., PANATTONI, C., & HAPPÉ, F. (2013). Children's recognition of emotions from vocal cues. **British Journal of Developmental Psychology**, 31(1), 97-113.
- SCHACHTER, S., & SINGER, J. (1962). Cognitive, social, and physiological determinants of emotional state. **Psychological Review**, 69(5), 379.
- SCHERER, K. R. (2003). Vocal communication of emotion: A review of research paradigms. **Speech communication**, 40(1-2), 227-256.

- SCHMIDHUBER, J. (2015). Deep learning in neural networks: An overview. **Neural networks**, 61, 85-117.
- SCHUBERT, E. (1999). Measuring emotion continuously: Validity and reliability of the two-dimensional emotion-space. **Australian Journal of Psychology**, 51(3), 154-165.
- SHAMI, M., & VERHELST, W. (2007). An evaluation of the robustness of existing supervised machine learning approaches to the classification of emotions in speech. **Speech communication**, 49(3), 201-212.
- SHANNON, R. V. (1983). Multichannel electrical stimulation of the auditory nerve in man. I. Basic psychophysics. **Hear Res**, 11, 157–189
- SHARAN, R. V., & MOIR, T. J. (2019). Acoustic event recognition using cochleagram image and convolutional neural networks. **Applied Acoustics**, 148, 62-66.
- SHARMA, A., DORMAN, M. F., & KRAL, A. (2005). The influence of a sensitive period on central auditory development in children with unilateral and bilateral cochlear implants. **Hearing Research**, 203, 134–143.
- SHARMA, A., TOBEY, E., DORMAN, M., BHARADWAJ, S., MARTIN, K., & GILLEY, P. (2004). Central auditory maturation and babbling development in infants with cochlear implants. **Archives of Otolaryngology–Head and Neck Surgery**, 130(5), 511–516.
- SHARMA, S. (2017). Activation functions in neural networks. **Towards Data Science**, 6.
- SHAW, A., VARDHAN, R. K., & SAXENA, S. (2016). Emotion recognition and classification in speech using Artificial neural networks. **International Journal of Computer Applications**, 145(8), 5-9.
- SIBI, P., JONES, S. A., & SIDDARTH, P. (2013). Analysis of different activation functions using back propagation neural networks. **Journal of theoretical and applied information technology**, 47(3), 1264-1268.
- SIMMONS FB, EPLEY JM, LUMMIS RC, GUTTMAN N, FRISHKOPF LS, HARMON LD, ZWICKER E (1965) Auditory nerve: electrical stimulation in man. **Science** 148:104–106

- SIU, T. S. C., & CHEUNG, H. (2017). Infants' sensitivity to emotion in music and emotion-action understanding. **PloS one**, 12(2), e0171023.
- SMITH, C. A., & ELLSWORTH, P. C. (1985). Patterns of cognitive appraisal in emotion. **Journal of personality and social psychology**, 48(4), 813.
- SMITH, J. T., WOLFE, J., & DETTMAN, S. (2018). Compelling evidence supports early implantation. (2018). **The Hearing Journal**, 71(6), 36–38.
- SRIVASTAVA, N., HINTON, G., KRIZHEVSKY, A., SUTSKEVER, I., & SALAKHUTDINOV, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. **The journal of machine learning research**, 15(1), 1929-1958.
- SUSSKIND, J. M., LEE, D. H., CUSI, A., FEIMAN, R., GRABSKI, W., & ANDERSON, A. K. (2008). Expressing fear enhances sensory acquisition. **Nature Neuroscience**, 11(7), 843–850.
- ŞEKER, A., DİRİ, B., & BALIK, H. H. (2017). Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. **Gazi Mühendislik Bilimleri Dergisi (GMBD)**, 3(3), 47-64.
- TING, D. S., LIU, Y., BURLINA, P., XU, X., BRESSLER, N. M., & WONG, T. Y. (2018). AI for medical imaging goes deep. **Nature medicine**, 24(5), 539-540.
- TOBEY, E. (2010). The changing landscape of pediatric cochlear implantation: Outcomes influence eligibility criteria. **ASHA Leader**, 12(2), 10–13.
- TONKS, J., WILLIAMS, W. H., FRAMPTON, I., YATES, P., & SLATER, A. (2007). Assessing emotion recognition in 9–15-years olds: Preliminary analysis of abilities in reading emotion from faces, voices and eyes. **Brain injury**, 21(6), 623-629.
- TOS, M., STANGERUP, S. E., CAYÉ-THOMASEN, P., TOS, T., & THOMSEN, J. (2004). What is the real incidence of vestibular schwannoma?. **Archives of Otolaryngology– Head & Neck Surgery**, 130(2), 216-220.
- TRACY, J. L., & ROBİNS, R. W. (2004). Show your pride: Evidence for a discrete emotion expression. **Psychological Science**, 15(3), 194–197.

- UMUT, K., YILMAZ, A., & DİKMEN, Y. (2019). Sağlık alanında kullanılan derin öğrenme yöntemleri. **Avrupa Bilim ve Teknoloji Dergisi**, (16), 792-808.
- VENAIL, F., SICARD, M., PIRON, J. P., LEVI, A., ARTIERES, F., UZIEL, A., & MONDAIN, M. (2008). Reliability and complications of 500 consecutive cochlear implantations. **Archives of Otolaryngology–Head & Neck Surgery**, 134(12), 1276-1281.
- VENKATARAMANAN, K., & RAJAMOHAN, H. R. (2019). Emotion recognition from speech. **arXiv preprint** arXiv:1912.10458.
- VIDAS, D., DİNGLE, G. A., & NELSON, N. L. (2018). Children's recognition of emotion in music and speech. **Music & Science**, 1, 2059204318762650.
- VOLKOVA, A., TREHUB, S. E., SCHELLENBERG, E. G., PAPSIN, B. C., & Gordon, K. A. (2013). Children with bilateral cochlear implants identify emotion in speech and music. **Cochlear Implants International**, 14(2), 80-91.
- VOLTA (1800) On the electricity excited by mere contact of conducting substances of different kinds. **Royal Soc Philos Trans** 90:403–431
- WALTZMAN, S. (2005). Expanding patient criteria for cochlear implantation. **Audiology Today**, 17(5), 20–21.
- WANG, D. J., TREHUB, S. E., VOLKOVA, A., & VAN LİESHOUT, P. (2013). Child implant users' imitation of happy-and sad-sounding speech. **Frontiers in psychology**, 4, 351.
- WEI, C. G., CAO, K., & ZENG, F. G. (2004). Mandarin tone recognition in cochlear-implant subjects. **Hearing research**, 197(1-2), 87-95.
- WENINGER, F., EYBEN, F., SCHULLER, B. W., MORTILLARO, M., & SCHERER, K. R. (2013). On the acoustics of emotion in audio: what speech, music, and sound have in common. **Frontiers in psychology**, 4, 292.
- WIEFFERINK, C. H., RIEFFE, C., KETELAAR, L., DE RAEVE, L., & FRIJNS, J. H. (2013). Emotion understanding in deaf children with a cochlear implant. **Journal of deaf studies and deaf education**, 18(2), 175-186.

- STEVENS, K. N., & WILLIAMS, C. E. (1969). On determining the emotional state of pilots during flight- An exploratory study(Pilot emotional state during stressful situations from tape recorded vocal utterances of air to ground radio communications using spectrographic analysis). **Aerospace Medicine**, 40, 1369-1372.
- STEVENS, K. N. & WILLIAMS, C. E., (1972). Emotions and speech: Some acoustical correlates. **The Journal of the Acoustical Society of America**, 52(4B), 1238-1250.
- WOLLMER, M., KAISER, M., EYBEN, F., SCHULLER, B., & RIGOLL, G. (2013). LSTM- Modeling of continuous emotions in an audiovisual affect recognition framework. **Image and Vision Computing**, 31(2), 153-163.
- XIN, Y., KONG, L., LIU, Z., CHEN, Y., LI, Y., ZHU, H., ... & WANG, C. (2018). Machine learning and deep learning methods for cybersecurity. **IEEE Access**, 6, 35365-35381.
- XU, L., ZHOU, N., CHEN, X., LI, Y., SCHULTZ, H. M., ZHAO, X., & HAN, D. (2009). Vocal singing by prelingually-deafened children with cochlear implants. **Hearing research**, 255(1-2), 129-134.
- YAMINS, D. L., HONG, H., CADIEU, C. F., SOLOMON, E. A., SEIBERT, D., & DICARLO, J. J. (2014). Performance-optimized hierarchical models predict neural responses in higher visual cortex. **Proceedings of the National Academy of Sciences**, 111(23), 8619-8624.
- YAWN, R., HUNTER, J. B., SWEENEY, A. D., & BENNETT, M. L. (2015). Cochlear implantation: a biomechanical prosthesis for hearing loss. **F1000prime reports**,
- YOON, J., GONG, E., CHATNUNTAWECH, I., BILGIC, B., LEE, J., JUNG, W., ... & KIM, E. Y. (2018). Quantitative susceptibility mapping using deep neural network: QSMnet. **Neuroimage**, 179, 199-206.
- YU, K. H., BEAM, A. L., & KOHANE, I. S. (2018). Artificial intelligence in healthcare. **Nature biomedical engineering**, 2(10), 719-731.

- ZAHNERT, T. (2011). The differential diagnosis of hearing loss. **Deutsches ärzteblatt international**, 108(25), 433.
- ZENG, F. G. (2002). Temporal pitch in electric hearing. **Hearing research**, 174(1-2), 101-106.
- ZENG, Z., PANTIC, M., ROISMAN, G. I., & HUANG, T. S. (2008). A survey of affect recognition methods: Audio, visual, and spontaneous expressions. **IEEE transactions on pattern analysis and machine intelligence**, 31(1), 39-58.
- ZENNER, H. P. (1997). The role of outer hair cell damage in the loss of hearing. **Ear, nose & throat journal**, 76(3), 140-144.
- ZHANG, S., ZHANG, S., HUANG, T., & GAO, W. (2017). Speech emotion recognition using deep convolutional neural network and discriminant temporal pyramid matching. **IEEE Transactions on Multimedia**, 20(6), 1576-1590.
- ZHAO, J., MAO, X., & CHEN, L. (2019a). Speech emotion recognition using deep 1D & 2D CNN LSTM networks. **Biomedical Signal Processing and Control**, 47, 312-323.
- ZHAO, Z., BAO, Z., ZHAO, Y., ZHANG, Z., CUMMINS, N., REN, Z., & SCHULLER, B. (2019b). Exploring deep spectrum representations via attention-based recurrent and convolutional neural networks for speech emotion recognition. **IEEE Access**, 7, 97515-97525.
- ZWOLAN, T. A., ASHBAUGH, C. M., ALARFAJ, A., KILENY, P. R., ARTS, H. A., EL-KASHLAN, H. K., & TELIAN, S. A. (2004). Pediatric cochlear implant patient performance as a function of age at implantation. **Otology and Neurotology**, 25(2), 112-120

ANSİKLOPEDİLER

CRYSTAL D. (1987). *The Cambridge Encyclopedia of Language*. New York: Cambridge University Press.

TÜRKÇE SÖZLÜK (2011), TDK Yayınları, Ankara (11. Baskı).

ELEKTRONİK KAYNAKLAR

<https://www.ninds.nih.gov/Disorders/Patient-Caregiver-Education/Life-and-Death-Neuron> Erişim Tarihi: 09.10.2020

<https://www.who.int/news-room/fact-sheets/detail/deafness-and-hearing-loss> Erişim Tarihi:24.08.2020

<http://http://yann.lecun.com/exdb/lenet/> Erişim Tarihi:24.07.2020

TEZLER

ASLAN F. (2010). Normal işiten çocuklar ile koklear implant kullanan çocukların duygu ifadelerini tanıma becerilerinin karşılaştırılması. Yüksek Lisans Tezi. Hacettepe Üniversitesi, Sağlık Bilimleri Enstitüsü.

ATEŞ, E. (2019). Derin öğrenme ile sesli komut tanıma. Yüksek Lisans Tezi. ESOĞÜ, Fen Bilimleri Enstitüsü.

BAŞBUĞ, A. M. (2019). Ses olayı tanıma ve akustik sahne geri getirme. Yüksek Lisans Tezi. Başkent Üniversitesi Fen Bilimleri Enstitüsü..

GATLIFFE, E. M. (2010). Influence of Ethnicity on English Vocabulary Measurement: Using the Montgomery Assessment of Vocabulary Acquisition in One Rural Southeastern New Mexico School. University of Denver.

GÜLER, S. (2018). Derin öğrenme ile RFID tabanlı konu belirleme uygulaması. KTO Karatay Üniversitesi. Yüksek Lisans Tezi.

GÜRER YÜCEL, F. (2013). Ses Bilgisi Ve Akustik Konusunda Geliştirilen Etkinliklerin Müzik Ve Fizik Öğretmen Adaylarının Tutum ve Başarı Düzeylerine Olan Etkisinin Araştırılması. Hacettepe Üniversitesi. Doktora Tezi.

ÖZCAN, H., (2014). Çok Düşük Çözünürlüklü Yüz İmgelerinde Derin Öğrenme Uygulamaları. Yüksek Lisans Tezi, Deniz Harp Okulu, Bilgisayar Mühendisliği Bölümü.

- PEREIRA, C. M. O. (2000). Perception and expression of emotion in speech. Macquarie University. Doktora tezi.
- ŞAHİN, Ö. (2017). Linguistik ve paralinguistik ipuçlarını kullanarak sestem duygu çıkarımında koklear implant kullanan okul çağı çocuklarının normal işiten akranları ile karşılaştırılması. Yüksek Lisans Tezi. Başkent Üniversitesi Sağlık Bilimleri Enstitüsü.
- TAN, Z. (2019). Derin Öğrenme Yardımıyla Araç Sınıflandırma. Yüksek Lisans Tezi. Fırat Üniversitesi, Fen Bilimleri Enstitüsü.
- Yavuz, F. K. (2009). Duygusal şemalar ve Leahy duygusal şema ölçeği'nin Türkçe uyarlaması, geçerlik ve güvenilirliği. Uzmanlık Tezi, Ankara Dışkapı Yıldırım Beyazıt Eğitim ve Araştırma Hastanesi.

DİĞER KAYNAKLAR

- AGUS, T. R., THORPE, S. J., SUIED, C., & PRESSNITZER, D. (2010). Characteristics of human voice processing. In Proceedings of 2010 IEEE International Symposium on Circuits and Systems (pp. 509-512). IEEE.
- AĞAOĞLU, Ş. N., ÖZER, M., CALİM, A., & UZUNTARLA, M. (2017). Effects of synaptic time delay on vibrational resonance in neuronal networks. In 2017 25th Signal Processing and Communications Applications Conference (SIU) (pp. 1-4). IEEE.
- AKIN, F., DUMAN, E. (2019, Nisan). Yüz Tanıma Sürecinde Farklı CNN Modellerinin Performans Karşılaştırması. Uluslararası Mardin Artuklu Multidisipliner Çalışmalar Kongresi'nde sunulan bildiri, Fırat Üniversitesi, Elazığ.
- AMER, M. R., SIDDIQUIE, B., RICHEY, C., & DIVAKARAN, A. (2014). Emotion detection in speech using deep networks. In 2014 IEEE international conference on acoustics, speech and signal processing (ICASSP) (pp. 3724-3728). IEEE.
- AMERICAN SPEECH-LANGUAGE-HEARING ASSOCIATION. (1996). Central auditory processing: Current status of research and implications for clinical practice. American Journal of Audiology, 5, 41–54.

- ANDERSON, D., & MCNEILL, G. (1992). Artificial neural networks technology. *Kaman Sciences Corporation*, 258(6), 1-83.
- BERTERO, D., & FUNG, P. (2017). A first look into a convolutional neural network for speech emotion detection. In *2017 IEEE international conference on acoustics, speech and signal processing (ICASSP)* (pp. 5115-5119). IEEE.
- BRADLEY, M. M., & LANG, P. J. (2007). *The International Affective Digitized Sounds (;IADS-2): Affective ratings of sounds and instruction manual*. University of Florida, Gainesville, FL, Tech. Rep. B-3.
- BRADLEY, M. M., CUTHBERT, B. N., & LANG, P. J. (2005). *International Affective Picture System: Technical manual and affective ratings*. NIMH Center for the Study of Emotion and Attention.
- DENG, J., BERG, A. C., LI, K., & FEI-FEI, L. (2010). What does classifying more than 10,000 image categories tell us?. In *European conference on computer vision* (pp. 71-84). Springer, Berlin, Heidelberg.
- EGHBAL-ZADEH, H., LEHNER, B., DORFER, M., & WIDMER, G. (2016). CP-JKU submissions for DCASE-2016: A hybrid approach using binaural i-vectors and deep convolutional neural networks. *IEEE AASP Challenge on Detection and Classification of Acoustic Scenes and Events (DCASE)*, 6, 5024-5028.
- ESKİMEZ, S. E., DUAN, Z., & HEINZELMAN, W. (2018). Unsupervised learning approach to feature analysis for automatic speech emotion recognition. In *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 5099-5103). IEEE.
- GRIMM, M., KROSCHEL, K., AND NARAYANAN, S. (2008). "The Vera am Mittag German audio-visual emotional speech database," in *Proceeding of the IEEE International Conference on Multimedia and Expo (ICME)* (Hannover: IEEE), 865– 868.
- GU, X., ZHANG, H., ZHANG, D., & KIM, S. (2016). Deep API learning. In *Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering* (pp. 631-642).

- HAN, K., HE, Y., BAGCHI, D., FOSLER-LUSSIER, E., & WANG, D. (2015). Deep neural network based spectral feature mapping for robust speech recognition. In Sixteenth Annual Conference of the International Speech Communication Association.
- HAN, K., YU, D., & TASHEV, I. (2014). Speech emotion recognition using deep neural network and extreme learning machine. In Fifteenth annual conference of the international speech communication association.
- HARÁR, P., BURGET, R., & DUTTA, M. K. (2017). Speech emotion recognition with deep learning. In 2017 4th International Conference on Signal Processing and Integrated Networks (SPIN) (pp. 137-140). IEEE.
- HASAN, M. R., JAMIL, M., & RAHMAN, M. G. R. M. S. (2004). Speaker identification using mel frequency cepstral coefficients. variations, 1(4). 3rd International Conference on Electrical & Computer Engineering ICECE.
- ISLAM, M. R., MASSICOTTE, D., NOUGAROU, F., MASSICOTTE, P., & ZHU, W. P. (2020). S-ConvNet: A shallow convolutional neural network architecture for neuromuscular activity recognition using instantaneous high-density surface EMG images. In 2020 42nd Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC) (pp. 744-749). IEEE.
- JONES, N. (2014). Computer science: The learning machines. Nature News, 505(7482), 146.
- LEE, D., JUNG, W., LEE, J., KO, J., SHIN, H. G., EUN, H., & NAM, Y. (2018). SafeNet: Artificial neural network for real-time T2 mapping with quality assurance. In Joint Annual Meeting ISMRM-ESMRMB. ISMRM (Vol. 2277).
- LÌ, Y., ZHAO, T., & KAWAHARA, T. (2019). Improved End-to-End Speech Emotion Recognition Using Self Attention Mechanism and Multitask Learning. In Interspeech (pp. 2803-2807).
- LIN, Y. L., & WEI, G. (2005). Speech emotion recognition based on HMM and SVM. In 2005 International conference on machine learning and cybernetics (Vol. 8, pp. 4898-4901). IEEE.

- LUO, D., ZOU, Y., & HUANG, D. (2018). Investigation on Joint Representation Learning for Robust Feature Extraction in Speech Emotion Recognition. In *Interspeech* (pp. 152-156).
- NAGI, J., DUCATELLE, F., DI CARO, G. A., CIRESAN, D., MEIER, U., GIUSTI, A., ... & GAMBARDELLA, L. M. (2011). Max-pooling convolutional neural networks for vision-based hand gesture recognition. In *2011 IEEE International Conference on Signal and Image Processing Applications (ICSIPA)* (pp. 342-347). IEEE.
- NAMIN, A. H., LEBOEUF, K., MUSCEDERE, R., WU, H., & AHMADI, M. (2009). Efficient hardware implementation of the hyperbolic tangent sigmoid function. In *2009 IEEE International Symposium on Circuits and Systems* (pp. 2117-2120). IEEE.
- NIU, Y., ZOU, D., NIU, Y., HE, Z., & TAN, H. (2018). Improvement on speech emotion recognition based on deep convolutional neural networks. In *Proceedings of the 2018 International Conference on Computing and Artificial Intelligence* (pp. 13-18).
- SEIDE, F., LI, G., & YU, D. (2011). Conversational speech transcription using context-dependent deep neural networks. In *Twelfth annual conference of the international speech communication association*.
- STUHLSTADT, A., MEYER, C., EYBEN, F., ZIELKE, T., MEIER, G., & SCHULLER, B. (2011). Deep neural networks for acoustic emotion recognition: Raising the benchmarks. In *2011 IEEE International conference on acoustics, speech and signal processing (ICASSP)* (pp. 5688-5691). IEEE.
- TAN, K. R., VILLARINO, M. L., & MADERAZO, C. (2019). Automatic music mood recognition using Russell's twodimensional valence-arousal space from audio and lyrical data as classified using SVM and Naïve Bayes. In *IOP Conference Series: Materials Science and Engineering* (Vol. 482, No. 1, p. 012019). IOP Publishing.
- TIPARÍ M. (2018) Digit Classification in Images Using Convolutional Neural Network. Technological University. Dublin.

- TIWARI, U., SONI, M., CHAKRABORTY, R., PANDA, A., & KOPPARAPU, S. K. (2020). Multi-Conditioning and Data Augmentation Using Generative Noise Model for Speech Emotion Recognition in Noisy Conditions. In ICASSP 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 7194-7198). IEEE.
- TRIGEORGIS, G., RINGEVAL, F., BRUECKNER, R., MARCHI, E., NICOLAOU, M. A., SCHULLER, B., & ZAFEIRIOU, S. (2016). Adieu features? end-to-end speech emotion recognition using a deep convolutional recurrent network. In 2016 IEEE International conference on acoustics, speech and signal processing (ICASSP) (pp. 5200-5204). IEEE.
- US DEPARTMENT OF HEALTH AND HUMAN SERVICES. (2013). National Institute on Deafness and Other Communication Disorders (NIDCD): Quick statistics.
- VALENTI, M., SQUARTINI, S., DIMENT, A., PARASCANDOLO, G., & VIRTANEN, T. (2017). A convolutional neural network approach for acoustic scene classification. In 2017 International Joint Conference on Neural Networks (IJCNN) (pp. 1547-1554). IEEE.
- VÁSQUEZ-CORREA, J. C., OROZCO-ARROYAVE, J. R., & NOTH, E. (2017). Convolutional Neural Network to Model Articulation Impairments in Patients with Parkinson's Disease. In Interspeech (pp. 314-318).
- WEIßKIRCHEN, N., BOCK, R., & WENDEMUTH, A. (2017). Recognition of emotional speech with convolutional neural networks by means of spectral estimates. In 2017 Seventh International Conference on Affective Computing and Intelligent Interaction Workshops and Demos (ACIIW) (pp. 50-55). IEEE.
- WU, H., SORAGHAN, J., LOWIT, A., & DI CATERINA, G. (2018). A deep learning method for pathological voice detection using convolutional deep belief networks. Interspeech 2018.
- ZHANG, B., QUAN, C., & REN, F. (2016). Study on CNN in the recognition of emotion in audio and images. In 2016 IEEE/ACIS 15th International Conference on Computer and Information Science (ICIS) (pp. 1-5). IEEE.

ZHENG, W. Q., YU, J. S., & ZOU, Y. X. (2015). An experimental study of speech emotion recognition based on deep convolutional neural networks. In 2015 international conference on affective computing and intelligent interaction (ACII) (pp. 827-831). IEEE.

EKLER

Ek-1: Bilgilendirilmiş Gönüllü Olur Formu

Ek.2: Etik Kurul Raporu

EK-1: Bilgilendirilmiş Gönüllü Olur Formu

BİLGİLENDİRİLMİŞ GÖNÜLLÜ OLUR FORMU

Sayın Gönüllü,

Bilimsel araştırmanın adı "İşitme Kayıplı Bireylerin Derin Öğrenme Tabanlı Ses Duygu Analizi" dir. Bu araştırmanın amacı; işitme kayıplı bireylerin, normal işiten bireylere göre konuşma esnasında duyguları sözel ifade etmede farklılık olup olmadığını derin öğrenme ile analiz etmektir.

Çalışmaya 18- 40 yaş aralığında, 20 normal işiten birey ve 20 koklear implant kullanan birey olmak üzere toplamda 40 birey dahil edilecektir. Çalışmanın normal işiten grubunu belirlemek amacıyla İstanbul Aydın Üniversitesi Odyoloji Laboratuvarı'nda odyolojik testler uygulanacaktır. Çalışmaya dahil edilecek olan normal işiten bireyler ve koklear implant kullanan bireyler çalışma kapsamında oluşturduğumuz cümleleri altı farklı duyguda seslendirecektir. Koklear implant kullanan bireylere kısa bir eğitim uygulanacaktır. Çalışma yaklaşık 25 dakika sürecektir.

Bu çalışma sırasında uygulanacak testlerin ve araştırma ile ilgili gerçekleştirilecek diğer işlemlerin masrafları size veya güvencesi altında bulunduğunuz resmi ya da özel hiçbir kurum veya kuruluşa ödetilmeyecektir. Size ait tüm tıbbi ve kimlik bilgileriniz gizli tutulacaktır ve araştırma yayınlansa bile kimlik bilgileriniz verilmeyecektir, ancak araştırmanın izleyicileri, etik kurullar ve resmi makamlar gerektiğinde tıbbi bilgilerinize ulaşabilir. Siz de istediğinizde kendinize ait tıbbi bilgilere ulaşabilirsiniz. Çalışmamızda yer alan değerlendirme yöntemlerinin hiçbirinde herhangi bir risk bulunmamaktadır.

Yukarıda yer alan ve araştırmadan önce gönüllüye verilmesi gereken bilgileri okudum ve katılmam istenen çalışmanın kapsamını ve amacını, gönüllü olarak üzerime düşen sorumluluktan anladım. Çalışma hakkında yazılı ve sözlü açıklama aşağıda adı belirtilen araştırmacı tarafından yapıldı. Kişisel bilgilerimin özenle korunacağı konusunda yeterli güven verildi.

Bu koşullarda söz konusu araştırmaya kendi isteğimle, hiçbir baskı ve telkin olmaksızın katılmayı kabul ediyorum.

Gönüllünün,

Adı Soyadı:
Telefon:
İmza:

Açıklama yapan araştırmacıım,

Adı Soyadı: Büşra Uludağ
Telefon:
İmza:

Bu araştırmada yer almak tamamen sizin isteğinize bağlıdır. Araştırmada yer almayı reddedebilirsiniz ya da herhangi bir aşamada araştırmadan ayrılabilirsiniz. İmzalı bu form kağıdının bir kopyası size verilecektir.

EK-2: Etik Kurul Kararı



T.C.
İSTANBUL AYDIN ÜNİVERSİTESİ REKTÖRLÜĞÜ
Yazı İşleri Müdürlüğü

Sayı : 17123671-044
Konu : Büşra ULUDAĞ Hk.

LİSANSÜSTÜ EĞİTİM ENSTİTÜSÜ MÜDÜRLÜĞÜNE

İlgi : 09/10/2020 tarihli ve 12848 sayılı yazı.

Enstitünüz *Y1816.070029* numaralı Odyoloji Anabilim Dalı Odyoloji Tezli Yüksek Lisans programı öğrencilerinden **Büşra ULUDAĞ**'ın "İşitme Kayıplı Bireylerin Derin Öğrenme Tabanlı Ses Duygu Analizi" adlı tez çalışması gereği, Üniversitemiz Odyoloji laboratuvarını kullanması talebiniz uygun görülmüştür.

Bilgilerinizi ve gereğini rica ederim.

e-imzadır
Prof. Dr. Yadiğar İZMİRLİ
Rektör





T.C.
ÜSKÜDAR ÜNİVERSİTESİ
GİRİŞİMSEL OLMAYAN ARAŞTIRMALAR
ETİK KURULU BAŞKANLIĞI

SAYI: 61351342/ 2020-449

29/09/2020

Sayın Prof.Dr.Bahriye Özlem KONUKSEVEN
(Büşra ULUDAĞ)

Üsküdar Üniversitesi Girişimsel Olmayan Araştırmalar Etik Kurulunun 28/09/2020 Tarihinde Yapılan 11 No.lu Toplantısında "İşitme Kayıplı Bireylerin Derin Öğrenme Tabanlı Ses Duygu Analizi" adlı araştırma projenizin kurum izni getirme koşulu ile şerhli olarak etik açıdan uygun olduğuna karar verilmiştir.

Bilgilerinize rica ederim.

Doç. Dr. Cumhuri TAŞ
Girişimsel Olmayan Araştırmalar Etik
Kurulu Başkanı

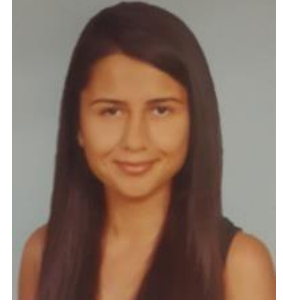
ÖZGEÇMİŞ

Ad- Soyad: Büşra ULUDAĞ

Doğum Yeri ve Tarihi: 23.11.1995 / Kızılcahamam

E-Posta Adresi: buludag@aydin.edu.tr

busraauludag@gmail.com



EĞİTİM BİLGİSİ

Yüksek Lisans: İstanbul Aydın Üniversitesi, Lisansüstü Eğitim Enstitüsü, Odyoloji Yüksek Lisans Programı (2018- Devam Ediyor)

Lisans: İstanbul Aydın Üniversitesi, Sağlık Bilimleri Fakültesi, Odyoloji Lisans Programı (2014- 2018)

Lise: Aldemir Atilla Konuk Anadolu Lisesi

İŞ TECRÜBESİNE AİT GENEL BİLGİLER

- İstanbul Aydın Üniversitesi, Lisansüstü Eğitim Enstitüsü, Araştırma Görevlisi (2020- Devam Ediyor)
- İstanbul Aydın Üniversitesi, Sağlık Bilimleri Fakültesi, Odyoloji Bölümü, Araştırma Görevlisi (2019- Devam Ediyor)
- İstanbul Aydın Üniversitesi, Sağlık Hizmetleri Meslek Yüksekokulu, Odyometri Bölümü, Öğretim Görevlisi (2018- 2019)

STAJ DENEYİMLERİ

- Yeni Yüzyıl Üniversitesi Özel Gaziosmanpaşa Hastanesi- İstanbul, Türkiye (2018)
- Maxtone İşitme Merkezleri- İstanbul, Türkiye (2017)
- Akdeniz Üniversitesi Hastanesi- Antalya, Türkiye (2017)
- Şişli Hamidiye Etfal Eğitim ve Araştırma Hastanesi- İstanbul, Türkiye (2016)
- Antalya Eğitim ve Araştırma Hastanesi- Antalya, Türkiye (2016)

BAŞARILAR

- Sağlık Bilimleri Fakültesi Birincisi

- Odyoloji Programı Bölüm Birincisi

YABANCI DİL BİLGİSİ

- İngilizce

YAYINLAR

1. Uluslararası bilimsel toplantılarda sunulan ve bildiri kitabında (Proceedings) basılan bildiriler

1.1 Kartal, A., Meral, M., Konukseven, Ö., Cetinkaya, Ü.C., **Uludağ, B.**, Adalı, İ., C-VEMP

Test Parameters in Patients with Motion Sickness: Latency, Amplitude and Asymmetric Ratio,

14th Congress of the European Federation of Audiology Societies. 22-25 May, 2019. Lisbon, Portugal

1.2 Cetinkaya, Ü.C., Konukseven, Ö., Kartal, A., Meral, M., **Uludağ,**

B., Evaluation Of Hearing And Balance System In Childhood With Chronic Renal Failure, 14th Congress of the European Federation of Audiology Societies. 22-25 May, 2019. Lisbon, Portugal

1.3 Meral, M., Konukseven, Ö., Cetinkaya, Ü.C., Kartal, A., **Uludağ, B.**, Adalı, İ., Age-Related Treatment Effect On Sudden Hearing Loss,

14th Congress of the European Federation of Audiology Societies. 22-25 May, 2019. Lisbon, Portugal

1.4 Adalı, İ., Konukseven, Ö., Cetinkaya, Ü.C., Kartal, A., Meral, M., **Uludağ,**

B., What are the Differences Between Canalolithiasis and Cupulolithiasis in patients with BPPV: Latency, Direction, Duration and Age Distribution,

14th Congress of the European Federation of Audiology Societies. 22-25 May, 2019. Lisbon, Portugal

2. Yazılan uluslararası kitaplar veya kitaplarda bölümler

2.1 Uludağ B., Meral M., Vestibüler Hipofonksiyonlu Hastalarda

Rehabilitasyon, Temel Vestibüler Rehabilitasyon El Kitabı,

Ed. Özlem Konukseven, US Akademi, İstanbul, 2019, ISBN978-605-9358-71-2

3. Ulusal hakemli dergilerde yayınlanan makaleler

3.1 Adalı, İ., **Uludağ, B.**, Meral, M., Kartal, A., Erbaşaran, E., İskender, S.,

& Konukseven, Ö. (2019). Odyoloji Bölümü Öğrencilerinin Mesleğe Yönelik

Kaygı ve Umutsuzluk Düzeylerinin

Belirlenmesi. *Turkish Journal of Audiology and Hearing Research*.

3.2 Adalı, İ., Meral, M., **Uludağ, B.**, Kartal, A., Bülbül, A., Çakıcı, F.,
& Konukseven, Ö. 18–24 Yaş Arası Gençlerde Telefon Kullanımının İşitme
Fonksiyonu Üzerine

Etkisi. *Turkish Journal of Audiology and Hearing Research*.

4. Ulusal bilimsel toplantılarda sunulan ve bildiri kitabında basılan bildiriler

4.1. Külekçi S, **Uludağ B**, Erten C, Pekdemir C, Kayıkcı H, Lülecioğlu A,
Kartal A, Meral M, Adalı İ, Konukseven Ö, “İşitsel

Uyarılmış Beyinsapı Yanıtlarında Uyarın Tekrar Oranındaki (Rate)
Değişikliğinin Latans Değerleri Üzerindeki Etkisi”, 9. Ulusal Odyoloji ve
Konuşma Bozuklukları Kongresi, 11-13 Ekim, İstanbul, Türkiye.

5. Bilimsel ve Mesleki Kuruluşlara Üyelikler

İstanbul Odyologlar Derneği

