

ELM368 – DÖNEM PROJESİ

Genç ve Yaşlı Erkek Seslerinin Analizi

Şule Nur Demirdaş, Furkan Kanbut, Mehmet Batuhan Abay
210102002053, 220102002180, 171024070
s.demirdas2021@gtu.edu.tr, fkabut@gtu.edu.tr, mehmet.abay@gtu.edu.tr

ABSTRACT (ÖZET)

Bu proje, dijital sinyal işleme tekniklerini kullanarak erkek seslerini gençlik ve yaşlılık arasında sınıflandırmayı amaçlamaktadır. Bu amaç doğrultusunda, spektrum analizi, temel frekans hesabı, IIR filtre tasarımı, güç hesabı ve Fourier dönüşümü kullanıldı.

Proje kapsamında, öncelikle ses dosyaları yüklendi ve IIR filtreler kullanılarak gerekli ön işlemler yapıldı. Ardından, spektrum analizi için Fourier dönüşümü uygulandı ve her sesin temel frekansı hesaplandı. Bu temel frekanslar, gençlik ve yaşlılık arasındaki ses farklarını belirlemek için önemli bir özellik olarak kullanıldı. Ayrıca, güç hesabı yapılarak her sesin frekans bileşenlerindeki enerjisi hesaplandı.

ANAHTAR KELİMELELER

Spektrum analizi, Temel frekans hesabı, IIR filtre tasarımı, güç hesabı, Fourier dönüşümü

1. Giriş

Bu çalışma, dijital sinyal işleme teknikleri kullanılarak erkek seslerinin yaş gruplarına (genç ve yaşlı) göre sınıflandırılmasını hedeflemektedir. İnsan sesi, yaş ilerledikçe belirgin değişiklikler gösterir. Genç bireylerin sesleri genellikle daha yüksek temel frekanslara (pitch) sahipken, yaşlı bireylerin sesleri daha düşük temel frekans ve farklı spektral özellikler sergiler. Bu akustik farklılıklar, fiziksel yapısı, akciğer kapasitesi ve diğer biyolojik faktörler tarafından belirlenir. Bu projede, bu farkları dijital sinyal işleme yöntemleri ile tespit etmek ve sınıflandırmak amaçlanmaktadır.

Problemin Tanımı

Erkek seslerinin yaşa göre sınıflandırılması:

1. Ses Sinyallerinin Toplanması ve Hazırlanması:

- Farklı yaş gruplarındaki erkek bireylerden dijital ortamda ses verisi toplandı.
- Bu ses kayıtları, dijital işleme için uygun formata dönüştürülür ve normalize edildi.

2. Frekans Domaininde Analiz:

- Hızlı Fourier Dönüşümü (FFT) kullanılarak ses sinyallerinin frekans spektrumları elde edildi.
- Frekans spektrumları, ses sinyalinin bileşen frekanslarını ve bu bileşenlerin genliklerini gösterir.
- Dominant frekans, temel frekans ve formantlar gibi önemli özellikler belirlendi.

3. Filtreleme:

- Gürültüyü azaltmak ve belirli frekans bantlarını vurgulamak için dijital filtreler uygulanır.
- Farklı filtrelerin (örneğin, düşük geçiş, yüksek geçiş, bant geçiş) etkileri incelenir.

3. Güç Analizi:

- Ses sinyallerinin güç spektrumları hesaplandı.
- Güç analizi ile sesin enerji yoğunluğu ve genel şiddeti belirlendi.

5. Genç ve Yaşlı Seslerin Karşılaştırılması:

- Elde edilen özellikler (temel frekans, formantlar, güç) kullanılarak genç ve yaşlı sesler arasındaki farklar tespit edildi.
- Bu farklar, seslerin yaş kategorilerine göre sınıflandırılması için kullanılır.

Çalışmanın Amacı

Bu çalışmanın temel amacı, dijital sinyal işleme tekniklerini kullanarak erkek seslerini genç ve yaşlı olarak sınıflandırabilecek bir sistem geliştirmektir. Böyle bir sistem, konuşma tanıma, yaş tespiti ve sesle ilgili diğer biyometrik uygulamalarda kullanılabilir. Ayrıca, yaşlı ve genç seslerin akustik özelliklerini derinlemesine incelemek ve bu özelliklerin yaşa bağlı olarak nasıl değiştiğini anlamak açısından da önemlidir.

Literatürdeki Yeri ve Uygulama Alanları

Bu çalışmanın çıktıları, birçok uygulama alanında değerlidir:

- Konuşma Tanıma Sistemleri: Yaşa dayalı ses tanıma, konuşma tanıma sistemlerinin doğruluğunu artırabilir.

- Sağlık ve Terapi: Yaşa bağlı ses değişikliklerinin tespiti, konuşma terapisi ve yaşa bağlı ses bozukluklarının erken teşhisinde kullanılabilir.
- Biyometrik Doğrulama: Ses tabanlı biyometrik sistemlere yaş doğrulama özelliği eklenebilir.
- **Duygu Analizi**: Sesin yaşa bağlı değişimleri, duygu analizinde ek bilgi sağlayabilir.

Yöntem

Bu çalışmada, yukarıda belirtilen analiz adımları sistematik bir şekilde uygulandı. Ses sinyalleri öncelikle dijital formata dönüştürülüp normalize edildi. Daha sonra, zaman ve frekans domainlerinde analiz edildi. Frekans domainindeki analizlerde, FFT kullanılarak spektral bileşenler elde edildi. Belirli frekans bantlarını vurgulamak amacıyla çeşitli dijital filtreler uygulandı. Son olarak, ses sinyallerinin güç spektrumları hesaplanarak genç ve yaşlı sesler arasındaki farklar belirlendi.

2. Deneyler ve Analiz

Adım 1: Ses Dosyasının Yüklenmesi ve Normalizasyonu

Amaç: Ses dosyasını okuyup, analiz etmek için normalize etmek.

Yöntem: read(audio) fonksiyonu kullanılarak ses dosyası okunur. Peak normalizasyonu yapılarak ses sinyalinin maksimum genlik değeri 1'e çekilir.

Adım 2: Fourier Dönüşümü ve Frekans Spektrumunun Hesaplanması

Amaç: Ses sinyalinin frekans bileşenlerini belirlemek.

Yöntem: FFT (Hızlı Fourier Dönüşümü) uygulanarak frekans spektrumu hesaplanır ve normalize edilir. Spektrumun genlik ve faz bileşenleri elde edilir.

```
import numpy as np
import matplotlib.pyplot as plt
from scipy.io.wavfile import read
from scipy.fft import fft

def generateGraph(audio, title):
    # Ses dosyasını oku
    Fs, x = read(audio)

    # Stereo dosya için tek kanala indirgeme
    try:
        # Stereo dosya varsa, sadece ilk kanalı al
        xn = [x[i][0] for i in range(len(x))]
    except IndexError:
        # Mono dosya varsa, olduğu gibi kullan
        xn = [x[i] for i in range(len(x))]

    # Ses şiddetini normalize et (Peak Normalizasyonu)
    def peak_normalize(audio):
        max_val = np.max(np.abs(audio))
        return audio / max_val
```

```
normalized_audio = peak_normalize(xn)
```

```
# Zamansal Sinyal Grafiği
```

```
N = len(xn)
n = np.arange(0, N)
```

```
# Zamansal sinyal grafiğini çiz
```

```
plt.figure()
plt.title(title)
plt.stem(n, normalized_audio)
plt.xlabel("n (örnek)")
plt.ylabel("x[n]")
plt.show()
```

```
# Frekans Spektrumu için açısal frekanslar (0 ile  $\pi$  arasında)
```

```
w = np.linspace(0, np.pi, N // 2)
```

```
# Fourier Dönüşümü ile frekans spektrumunu hesapla
```

```
Xw = fft(normalized_audio)
```

```
# FFT Sonucunu Normalize Etme
```

```
Xw_mag = np.abs(Xw[:N // 2]) / np.max(np.abs(Xw[:N // 2])) # Genlik spektrumu
```

```
Xw_phs = np.unwrap(np.angle(Xw[:N // 2])) # Faz spektrumu
```

```
# Frekans spektrumu grafiğini çiz
```

```
fig, (ax1, ax2) = plt.subplots(2, 1)
```

```
# Genlik grafiği
```

```
ax1.plot(w / np.pi, Xw_mag)
ax1.set_ylabel("Magnitude")
ax1.set_title("Frequency Spectrum (Magnitude)")
```

```
# Faz grafiği
```

```
ax2.plot(w / np.pi, Xw_phs)
ax2.set_ylabel("Phase (radians)")
ax2.set_xlabel("Frequency ( $\pi$  rad/sample)")
ax2.set_title("Frequency Spectrum (Phase)")
```

```
plt.tight_layout()
plt.show()
```

Adım 3: Filtre Uygulama

Amaç: Belirli frekans aralıklarındaki değerleri bastırmak.

Yöntem: Bant geçiren filtre uygulanır. applyFilter fonksiyonu kullanılarak filtre uygulanmış ses sinyali elde edilir. Bu filtre cutoff frekansları 0.05-0.2 aralığında bant geçiren IIR Butterworth bir filtredir.

```
def applyFilter(filter, X):
    # Filtre katsayılarını içeren .mat dosyasını yükle
    filter_data = loadmat(filter)

    # Filtre katsayılarını alın ve float tipine dönüştürün
    Coeffs = filter_data["ba"].astype(float)
```

```

# FIR filtre olduğunda a katsayıları tanımlanmaz, bu
# yüzden a'yı 1 olarak ayarlıyoruz
if Coeffs.shape[1] == 1:
    b = Coeffs[:,0] # FIR filtresi için b katsayıları
    w, Hw = sgnl.freqz(b) # Frekans yanıtı
    a = np.array([1]) # FIR filtresi için a = 1
else:
    b = Coeffs[:,0] # IIR filtresi için b katsayıları
    a = Coeffs[:,1] # IIR filtresi için a katsayıları
    w, Hw = sgnl.freqz(b, a) # Frekans yanıtı

# Frekans yanıtının genlik ve faz spektrumunu hesapla
Hw_mag = abs(Hw) # Genlik spektrumu
Hw_phs = np.unwrap(np.angle(Hw)) # Faz spektrumu

# Sinyali filtrele
filtered_signal = sgnl.lfilter(b, a, X)

# Frekans eksenini oluştur
N = len(filtered_signal)
w = np.linspace(0, np.pi, N/2)

# Filtrelenmiş sinyalin FFT'sini hesapla
Xw = fft(filtered_signal)

# FFT Sonucunu Normalize Etme
Xw_mag = np.abs(Xw[:N // 2]) / np.max(np.abs(Xw[:N // 2])) # Genlik spektrumu

```

Adım 4: Pitch Hesaplama

Amaç: Sesin temel frekansını (pitch) belirlemek.

Yöntem: Frekans spektrumunda en büyük genliğe sahip frekans bileşeni bulunur ve bu bileşen Hz cinsine dönüştürülür.

Açısal frekans – Frekans Dönüşümü:

$$F = \frac{F_s}{2\pi} * w \quad [Hz]$$

```

dominant_frequency_index = np.argmax(Xw_mag)
dominant_frequency = w[dominant_frequency_index] *
(Fs / (2 * np.pi))

```

```

print("Temel frekans (pitch):", dominant_frequency,
"Hz")

```

Çocuklarda veya genç erkeklerde pitch (temel frekans) değeri 250-400 hatta daha da yüksek değerler alabilir. Yaşlı erkeklerdeyse bu değer 50 Hz ye kadar düşebilir. Bu değer bir çok farklı etmene bağlı olmasına rağmen yaşlı erkeklerde genç erkeklere ve çocuklara kıyasla bariz bir şekilde daha düşüktür.

Adım 5: Güç Hesaplama

Amaç: Sinyalin gücünü belirlemek.

Yöntem: Sinyalin RMS (Root Mean Square) değeri hesaplanarak güç değeri elde edilir.

Aşağıda verilen fonsiyon kullanılarak güç hesabı yapılmıştır. Öncelikle sinyal normalize edilmiştir. Bu işlem sinyalin genliğini belli bir skalaya getirmemize yarar.

```

def calculate_power(signal):
    # Sinyalin gücünü hesapla
    power = np.sum(np.abs(signal) ** 2) / len(signal)
    return power

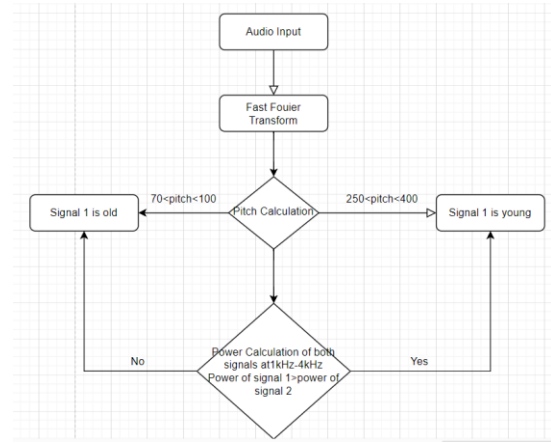
```

Sinyalin enerjisi filtrelenen aralıkta yüksek ise ses bir gence ait demektir.

Adım 6: Sonuçların Karşılaştırılması

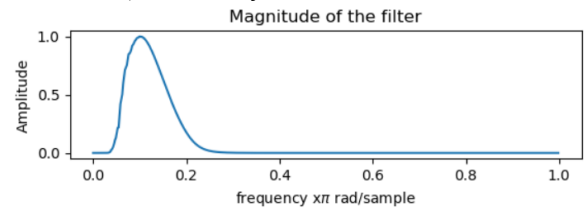
Amaç: Yaşlı ve genç sesler arasındaki farkları belirlemek.

Yöntem: Filtrelenmiş sinyallerin güç ve pitch değerleri karşılaştırılarak analiz yapılır. Yaşlı seslerde daha düşük pitch ve daha düşük güç beklenir.

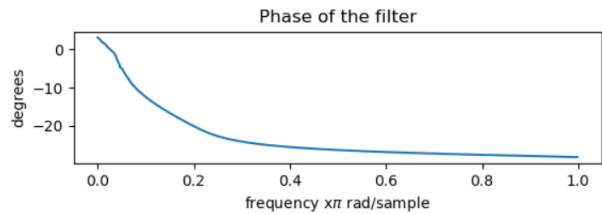


Şekil 1: Akış Diagramı

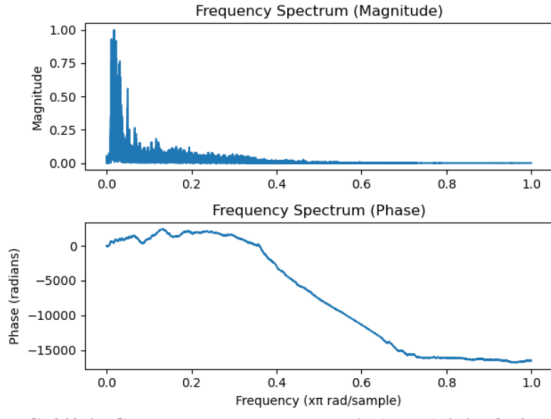
2.1. Grafik, Tablo ve Şekiller



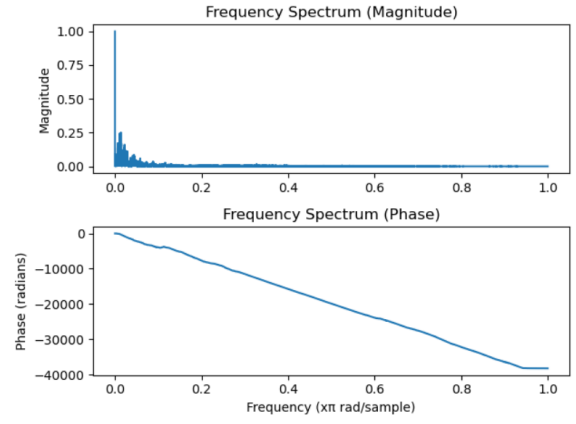
Şekil 2: Kullanılan IIR Butterworth filtresinin genlik grafiği



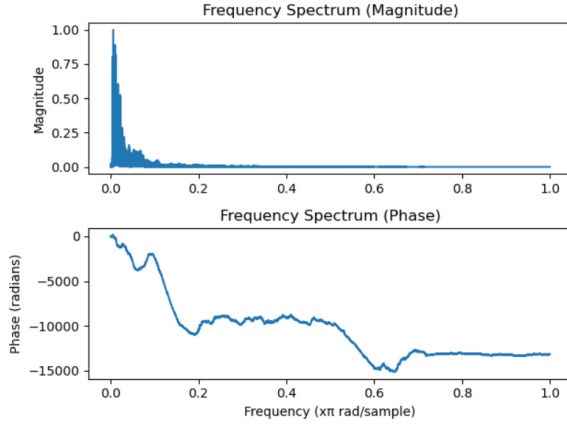
Şekil 3: Kullanılan IIR Butterworth filtrenin faz grafiği



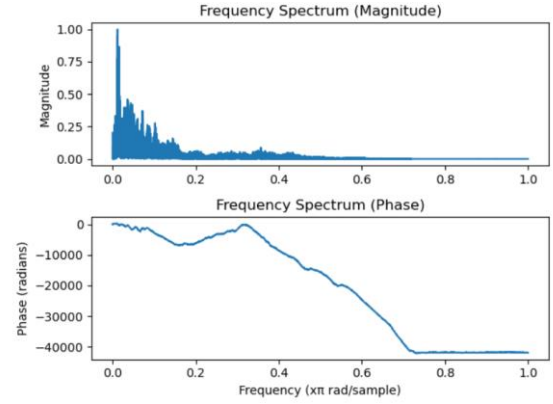
Şekil 4: Genç ses (young man train1.wav) için frekans spektrumu genlik ve faz grafikleri



Şekil 7: Yaşlı ses (old man train3.wav) için Genlik ve Faz grafikleri

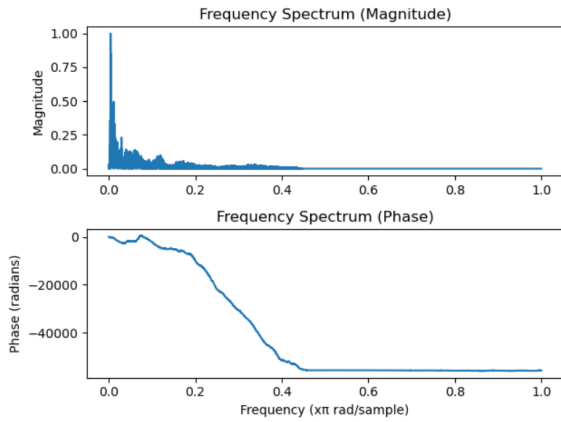


Şekil 5: Yaşlı ses (old man train1.wav) için Genlik ve Faz grafikleri

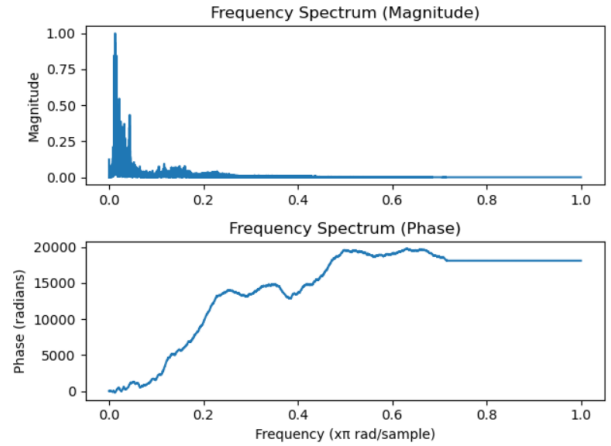


Şekil 8: Genç ses (young man train2.wav) için Genlik ve Faz grafikleri

Şekil 4 genlik grafiği ve Şekil 5 genlik grafiğinden yola çıkarak yaşlı ve genç sesleri frekans spektrumları arasındaki bariz farklar tespit edilde ve bu farklara göre (farklı frekans değerlerinde genlik farkları) butterworth filtresi Şekil 2 de verilen filtre tasarlandı. 4 diğer test verisi kullanılarak bu gözlem teyit edildi.



Şekil 6: Yaşlı ses (old man train2.wav) için Genlik ve Faz grafikleri



Şekil 9: Genç ses (young man train3.wav) için Genlik ve Faz grafikleri

Temel Frekans Değerleri (Pitch)		
	Genç	Yaşlı
Data 1	388 Hz	142 Hz
Data 2	262 Hz	96 Hz
Data 3	286 Hz	0 Hz

Tablo 1: Eğitim yapılan verinin temel frekans değerleri

Filtrelenen Aralıkta Güç Değerleri		
	Genç	Yaşlı
Data 1	0.0080 W	0.0024 W
Data 2	0.0079 W	0.0031 W
Data 3	0.0026 W	0.0038 W

Tablo 2: Eğitim yapılan verinin filtrelenen aralıkta güç hesabı

Tablo 1’ de elde edilen değerler teorik olarak elde edilen değerlerle uyum sağlamaktadır.

Tablo 2’ de elde edilen değerler yüzde 83 oranında beklenen değerlerle uyum sağlamakta yalnızca genç ses data 3’te elde edilen veri üzerindeki gürültü sebebiyle hatalı sonuç elde edildiği düşünülmektedir. Gürültü engelleyici filtreler ile bu hata giderilebilir.

3. Sonuç ve Yorum

Bu çalışmada, dijital sinyal işleme (DSP) teknikleri kullanılarak genç ve yaşlı erkek seslerinin ayrımını yapmak için bir yöntem geliştirildi. Fast Fourier Transform (FFT) ve IIR bandpass Butterworth filtrelerinin kullanımıyla, ses sinyallerinin frekans alanındaki özellikleri analiz edildi. Elde edilen sonuçlar, genç ve yaşlı sesler arasındaki temel frekans (pitch) ve formant frekanslarındaki farklılıkların başarılı bir şekilde tespit edilmesini sağladı.

Yapılan Çalışmada Elde Edilen Deneyim

Bu çalışma sırasında çeşitli önemli deneyimler kazanılmıştır:

1. FFT'nin Gücü ve Kullanımı: FFT'nin ses sinyallerinin frekans bileşenlerini hızlı ve verimli bir şekilde analiz etmek için ne kadar etkili olduğu görülmüştür. Zaman alanında karmaşık görünen sinyallerin, frekans alanında daha anlaşılır hale geldiği gözlemlenmiştir.

2. Filtre Tasarımı ve Uygulanması: Pyfda kütüphanesi kullanılarak IIR bandpass Butterworth filtrelerinin tasarımı ve uygulanması sırasında, filtrelerin belirli frekans aralıklarını izole etme yeteneği öğrenilmiştir. Bu filtrelerin kullanımıyla genç ve yaşlı seslerin karakteristik frekanslarının başarılı bir şekilde ayrıştırıldığı görülmüştür.

3. Pitch: Temel frekans (pitch) ve güç hesabının, ses sinyallerinin yaşına dair önemli bilgiler sunduğu ve bu frekansların doğru tespiti ile yaş ayrımının yapılabileceği anlaşılmıştır.

Önerilen Çözümün Avantajları ve Kısıtları

Avantajlar

- Yüksek Doğruluk FFT ve IIR bandpass filtreler kullanılarak elde edilen sonuçlar, genç ve yaşlı erkek seslerinin başarılı bir şekilde ayrılmasını sağladı.

- Hız ve Verimlilik FFT algoritmasının hızlı olması, büyük veri kümelerinin bile etkin bir şekilde işlenmesini sağladı.

- Kolay Uygulanabilirlik: Butterworth filtrelerinin tasarımı ve uygulanması, diğer karmaşık DSP yöntemlerine göre daha kolaydır.

Dezavantajlar

- Veri Seti Sınırlamaları Sadece belirli sayıda ses örneği kullanıldı. Daha geniş bir veri seti ile çalışmak, sonuçların genel geçerliliğini artırabilir.

- Gürültü ve Parazit Gerçek dünya uygulamalarında, ses sinyallerindeki gürültü ve parazit, frekans analizi ve filtreleme performansını olumsuz etkileyebilir. Bu durumlarda, ek gürültü azaltma tekniklerinin uygulanması gerekebilir.

Çözümü Geliştirmek İçin Neler Yapılabilir?

- Daha Geniş Veri Seti: Çeşitli yaş gruplarından daha geniş bir ses veri seti kullanılarak, modelin genel performansı artırılabilir.

- Gelişmiş Gürültü Azaltma Teknikleri: Ses sinyallerindeki gürültü ve paraziti azaltmak için gelişmiş DSP teknikleri kullanılabilir.

- Ek Özellikler: Temel frekans ve formant frekanslarının yanı sıra, diğer akustik özellikler (örneğin, jitter, shimmer) de analize dahil edilerek modelin doğruluğu artırılabilir.

- Makine Öğrenmesi: Ek özelliklerin çıkarılması ve sınıflandırılması için makine öğrenmesi algoritmaları entegre edilebilir, bu da manuel eşik belirleme ihtiyacını ortadan kaldırabilir.

Çalışmadan Öğrenilenler ve Katkıları

Bu çalışma, dijital sinyal işlemenin temel prensiplerini ve uygulamalarını öğrenmek için önemli bir fırsat sağladı. Ses sinyallerinin frekans analizinin nasıl yapıldığını, FFT ve IIR bandpass filtrelerinin nasıl tasarlandığını ve uygulandığını pratik olarak görme şansı sundu. Ayrıca, genç ve yaşlı seslerin akustik özelliklerinin nasıl farklılaştığını ve bu farkların nasıl tespit edilebileceğini öğrenmek, ses sinyali işlemenin pratik uygulamalarına dair değerli bilgiler kazandırdı.

Sonuç olarak, bu çalışma dijital sinyal işleme alanında önemli deneyimler kazandırıp ve ses sinyali analizinde kullanılabilecek etkili yöntemler sundu. Bu tür projeler, sinyal işleme ve analiz tekniklerinin daha geniş bir yelpazede uygulanmasına olanak tanıyarak, gelecekteki çalışmalara sağlam bir temel oluşturur.