

ELM368 – DÖNEM PROJESİ

KEDİ İLE KÖPEK SESLERİ ARASINDA FREKANS ANALİZİ İLE TESPİT YAPILMASI

Berat KIZILARMUT, Ömer Emre POLAT
171024086, 1801022037
beratkizilarmut@gtu.edu.tr, o.polat2018@gtu.edu.tr

ABSTRACT (ÖZET)

Kedilerin sesleri köpeklere kıyasla daha tiz tona sahiptir. Projede bu özellikten faydalanarak verilen örnek ses dosyalarının kedilere mi köpeklere mi ait olduğu filtreleme ve analiz operasyonları ile bulundukları frekans bantlarına bakılarak tespit edilecektir.

ANAHTAR KELİMELEER

FIR Filtre, Spektrum Analizi, Frekans Analizi, Alçak Geçiren Filtre, Bant Geçiren Filtre, Ses Ayırımı, Spectrogram, Fourier Dönüşümü, Genlik Spektrumu

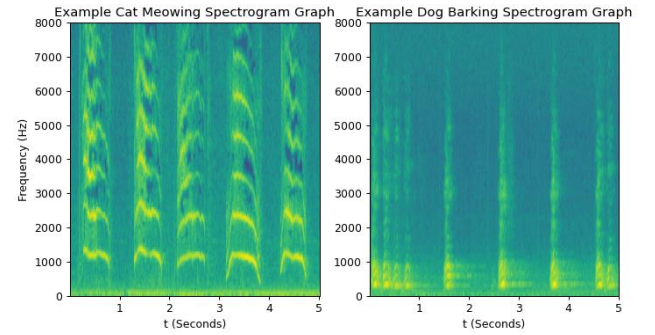
1. Giriş

Kedi ve köpek sesleri arasında tespit yapılması için bu iki hayvanın seslerinde tiz ve pes ses farklarından faydalanılmıştır. Örnek ses dosyalarının fourier dönüşümleri alınıp frekans analizleri yapılmıştır. Genlik enerji spektrumları [1] ve spektrogramları [2] incelenmiştir. Örnekler farklı filtrelerden geçirilerek aynı analizler tekrarlanmıştır. Filtrelenmiş ve filtrelenmemiş veriler, grafikler ve ortalama değerler karşılaştırılmıştır. Analizlerden elde edilen veriler doğrultusunda, verilen bir ses dosyasının kedilere mi köpeklere mi ait olduğunu tespit edebilen bir algoritma oluşturulmuştur.

2. Deneyler ve Analiz

Problemin çözümünde öncelikle kedi ve köpek seslerindeki farklar tespit edildi. Farkların tespit edilmesinin ardından bu karakteristik farkların, verilen ses örneklerinden nasıl tespit edilebileceği hakkında araştırma yapıldı. Analizlerde kullanmak üzere veri setlerinden [3] uygun ses örnekleri bulundu. Teorik olarak sahip olunan fikirler pratik olarak uygulandı ve sonuçlar gözlemlendi. Gözlemlere ve elde edilen verilere dayanarak kedi ve köpek sesleri arasında ayrım yapabilen algoritma oluşturuldu. Veriler manuel olarak tekrar analiz edildi. Manuel elde edilen sonuçlar ile algoritmanın sonuçları karşılaştırıldı.

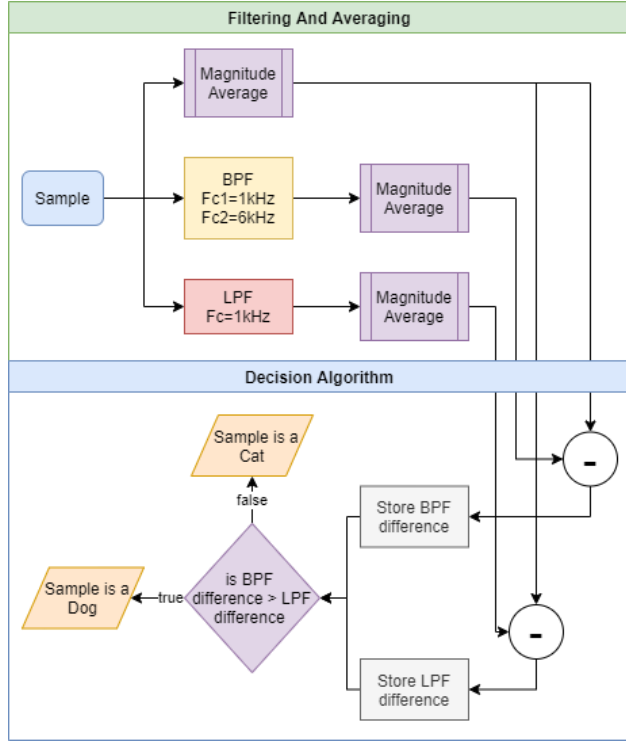
2.1. Kedi ve Köpek Sesleri Arasındaki Farklar



Şekil 1

Köpeklerin kedilere kıyasla çok daha az vokal ses çeşitliliği bulunmaktadır. Köpekler sadece havlama ve hırlama gibi sesleri çıkartırken, kediler miyavlama, mırlama, kuş seslerini taklit etme, hırlama ve daha birçok farklı sesi çıkartabilmektelerdir. Köpek ve kedilerin bu çeşitli vokal iletişim yollarının frekans analizlerine bakıldığında örnek olarak; köpeklerin havlamaları ile kedilerin hırlama veya mırlama seslerinin frekanslarının benzer frekans aralığında bulunması gibi problemler ortaya çıkmaktadır. Bu durumdan dolayı projede frekans aralıkları arasında işlenebilir bir fark bulunan ve bu iki hayvan türünün de en çok çıkardıkları sesler olan havlama ve miyavlama seslerine odaklanılmıştır. Şekil 1’de görüldüğü üzere, köpek havlama sesleri 400 Hz ile 1000 Hz aralığında ağırlıklı ortalama frekans değerlerine sahip olmaktadır. Kedi miyavlama sesleri ise 1000 Hz ile 3000 Hz aralığında ağırlıklı ortalama frekans değerlerine sahip olmaktadır.

2.2. Tespit Metodu

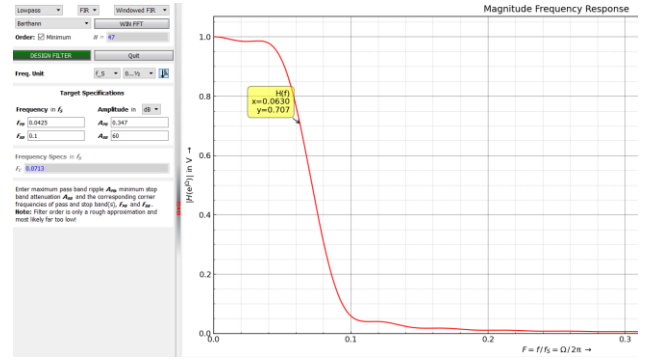


Şekil 2

Yapılan araştırmalara ve elde edilen örneklerin frekans spektrumlarının incelenmesi ile iki farklı filtrenin oluşturulmasına karar verilmiştir. Bu filtrelerden biri kesim frekansı 1000 Hz olan Alçak Geçiren Filtre, ikincisi ise kesim frekansları 1000 Hz ile 6000 Hz olan bir Bant Geçiren Filtredir. Şekil 2’de de gözüktüğü gibi örneklerin kopyaları bu iki filtreden de geçirilecektir. Örnek işaretin kendisinin ve iki filtrenin çıkışındaki işaretlerin ortalama genlik değerleri alınmıştır. Bu elde edilen ortalama genlik değerleri karşılaştırılarak işaretin enerjisinin çoğunluğunun 1000 Hz frekansı değerinin üzerinde mi altında mı olduğuna karar verilecektir. Karar verme işlemi için işaretin kendisinin ortalama genlik değerinden filtrelenmiş genlik değerleri çıkarılacaktır. Alçak Geçiren Filtre uygulanmış olan işaretin fark değeri, Bant Geçiren Filtre uygulanmış olan işaretin fark değerinden küçük ise işaretin köpek havlaması olduğu, tersi durumda ise kedi miyavlaması olduğuna karar verilmiştir.

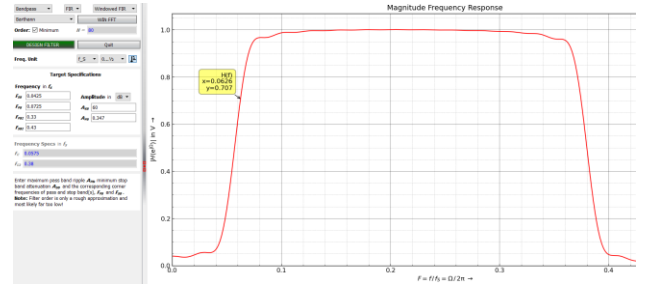
2.3. Kullanılacak Filtrelerin Oluşturulması

Sistemde kullanılmak istenilen Alçak Geçiren ve Bant Geçiren filtreler pyFDA programı aracılığı ile oluşturuldu.



Şekil 3

İlk filtre olarak, Şekil 3’te gösterilen, FIR Alçak Geçiren Filtre oluşturulmuştur. Bu filtre oluşturulurken Windowed FIR Barthann Metodu kullanılmıştır. Kesim frekansı 1000 Hz olması istenilen filtreyi oluşturmak için filtre ve işaret frekans indislerinin ilişkisi bulundu ve orantıları yapıldı. İşaretin kendisinin frekans cevabı 0-8000 Hz aralığında iken pyFDA aralığı 0-0.5 rad/sample olmaktadır. Bu ilişkiye dayanarak 1000 Hz noktası 0.0625 rad/sample noktasına gelmektedir. Belirlenen 0.0625 rad/sample noktasında filtrenin gerilim değerinin en az 0.707 olması için filtre özellikleri ayarlanmıştır. Geçirme bandı köşe frekansı $F_{PB} = 0.0425$ rad/sample, durdurma bandı köşe frekansı $F_{SB} = 0.1$ rad/sample olarak alınmıştır. Filtrenin derecesi $N = 47$ olmuştur.



Şekil 4

İkinci filtre olarak, Şekil 4’te gösterilen, FIR Bant Geçiren Filtre oluşturulmuştur. Bu filtre oluşturulurken Windowed FIR Barthann Metodu kullanılmıştır. Alçak kesim frekansı 1000 Hz, yüksek kesim frekansı 6000 Hz olması istenilen bu filtreyi oluşturmak için bir önceki adımda da yapılan işlemlere dayanarak, 1000 Hz kesim noktası 0.0625 rad/sample noktasına, 6000 Hz kesim noktası ise 0.375 rad/sample noktasına gelmesi gerektiği hesaplanmıştır. Belirlenen 0.0625 rad/sample ve 0.375 rad/sample noktalarında filtrenin gerilim değerinin en az 0.707 olması için filtre özellikleri ayarlanmıştır. Alçak durdurma bandı köşe frekansı $F_{SB} = 0.0425$ rad/sample, alçak geçirme bandı köşe frekansı $F_{PB} = 0.0725$ rad/sample, yüksek geçirme bandı köşe frekansı $F_{PB2} = 0.33$ rad/sample, yüksek durdurma bandı köşe frekansı $F_{SB2} = 0.43$ rad/sample olarak alınmıştır. Filtrenin derecesi $N = 80$ olmuştur.

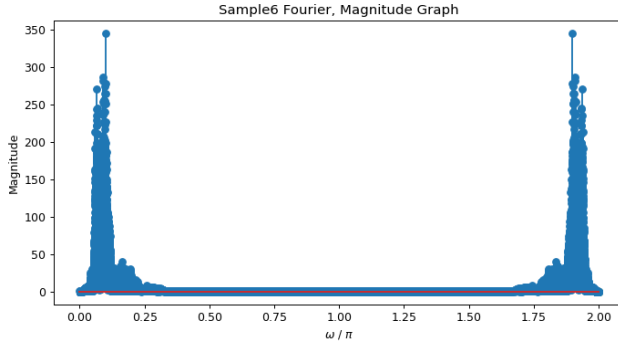
2.4. Uygulama

Hesaplamalar, analizler ve uygulamalar Python dilinde, Jupyter Notebook ortamında oluşturulmuş ve kullanılmıştır. Uygulama dosyasında öncelikle giriş aşaması olarak numpy, matplotlib, pyplot ve scipy kütüphaneleri çağırılmış, işaret işleme ve analiz işlemlerinde kullanılacak olan xn, Fs, N vb. birçok değişken tanımlanmıştır. Uygulamanın kullanımının rahatlığını arttırmak için index isimli bir örnek değişkeni oluşturulmuştur ve fonksiyonlarda index değerleri kullanılmıştır. Bütün değişkenler dizi olarak oluşturulmuştur ve aktif olan örneğin index değerine göre işlevlerini yerine getirmektedirler.

Uygulamanın ikinci adımında pyFDA’de oluşturulan filtrelerin b ve a değerleri oluşturulan değişkenlerine atanmıştır. İki filtre için de, genlik grafikleri, faz grafikleri ve grup gecikmesi grafikleri pyFDA’de gözlemlendiği ve beklenildiği gibi elde edilmiştir.

Uygulamanın üçüncü adımında her örnek için tekrar edecek şekilde; örneklerin oynatılabildiği bir medya oynatıcı oluşturulmuş ve zaman domeninde grafikleri çizdirilmiştir. Bu işlemlerin ardından örneklerin fourier dönüşümleri alınmış, genlik spektrumları oluşturulmuş ve spectrogram grafikleri oluşturulmuştur.

2.4.1. Fourier Dönüşümü



Şekil 5

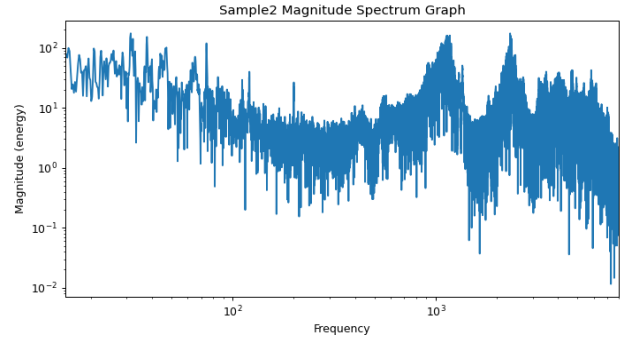
Fourier dönüşümü alınması için, “scipy.fftpack” kütüphanesinden “fft” (fast fourier transform) fonksiyonu çağırılmıştır. FFT fonksiyonunda kullanılması amacıyla ayrık zamanlı işaretin örnek sayısı ve örnekleme frekansı kullanılarak açısal frekans ve frekans değişkenleri tanımlanmıştır.

Tablo 1

$xw[index] = (np.fft.fft(xn[index], N[index]) / N[index])$
--

Tablo 1’de gözlemlendiği gibi, FFT fonksiyonu sinyalin kendisini ve örnek sayısını giriş olarak, çıkışta örnek sayısı ile çarpılmış bir şekilde sinyalin frekans domenindeki halini vermektedir. Çıkış örnek sayısı ile çarpılmış durumda olduğu için ekstra işlem olarak örnek sayısı ile bölünmüştür. Şekil 5’te fourier dönüşümünün genlik grafiği verilmiştir.

2.4.2. Genlik Spektrumu



Şekil 6

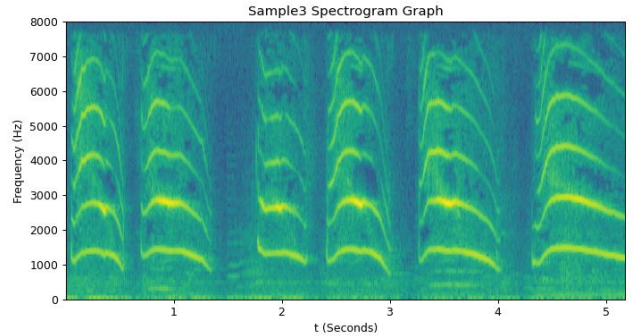
Genlik spektrumu elde etmek için “pyplot” kütüphanesinde bulunan “magnitude_spectrum” fonksiyonu kullanılmıştır.

Tablo 2

$plt.magnitude_spectrum(xn[index], Fs=Fs[index])$
--

Tablo 2’de gösterilen genlik spektrumu fonksiyonu, işaretin zaman domenindeki ayrık zamanlı halini ve örnekleme frekansını giriş olarak alır. Çıkış olarak genlik spektrumunu lineer şekilde verir. Orantılama fonksiyonları ile genlik değeri olan dikey, frekans değeri olan yatay indisler logaritmik hale getirilmiştir. Logaritmik halde çıkış Şekil 6’da verilmiştir.

2.4.3. Spectrogram



Şekil 7

Spectrogram grafiğinin elde edilmesi için “pyplot” kütüphanesinde bulunan “specgram” fonksiyonu kullanılmıştır.

Tablo 3

$plt.specgram(xn[index], Fs=int(Fs[index]))$
--

Tablo 3’te verilen “specgram” fonksiyonu, giriş olarak sinyalin kendisi ve işaretin örnekleme frekansını alır. Çıkış olarak Şekil 7’de ki gibi, dikey indisi frekans değerleri, yatay indisi zaman olan spectrogram grafiğini verir.

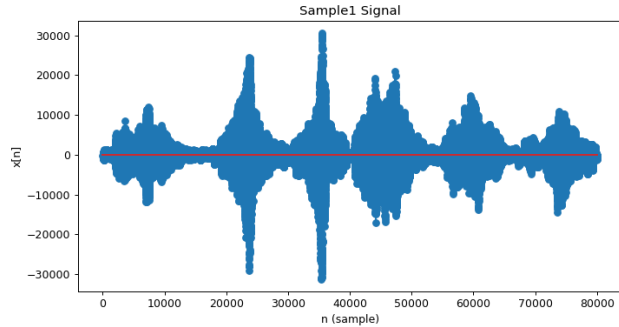
2.4.4. Karar Algoritması

Karar Algoritması, işaretinin frekans domenindeki genlik ortalamasından, iki filtrenin çıkışındaki işaretlerin frekans domenindeki genlik ortalaması arasındaki farkı kullanarak, işaretin enerjisinin çoğunluğunun hangi bölgede olduğunu tespit etmektedir. İşaretin enerjisinin büyük çoğunluğu alçak geçiren filtreden geçmiş olan işarete ise bu durum işaretin enerjisinin çoğunluğunun 1000 Hz bandının altında olduğunu, yani işaretin bir köpek sesi olduğuna karar verilir. Bu durumun tam tersinde ise, bant geçiren filtrenin enerjisinin daha yüksek olduğu, enerjinin çoğunluğunun 1000 Hz – 6000 Hz bandında olduğu tespit edilir ise işaretin bir kedi sesi olduğuna karar verilir.

2.5. Örnek Analizi

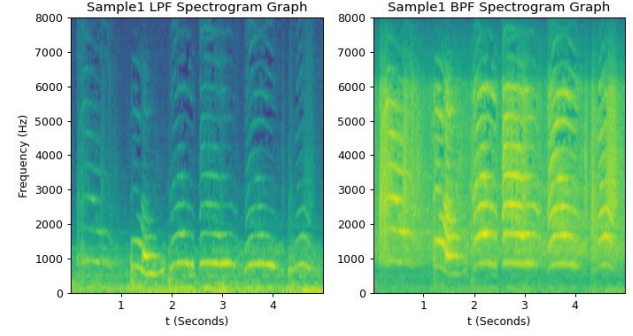
Bütün örnekler için, işaretin zaman domenindeki grafiği, frekans domenindeki genlik grafiği, genlik spektrumu, ortalama genlik değeri ve spektrogram grafiği incelenecektir. Ardından aynı incelemeler alçak geçiren filtre ve bant geçiren filtre çıktıları için tekrarlanacaktır. Elde edilen üç farklı kaynaklı veriler karşılaştırılacaktır.

2.5.1. Örnek 1



Şekil 8

Toplam olarak 5 saniye olan bu ses dosyasının, Şekil 8’de, 6 parçadan oluştuğu gözlenmiştir. Frekans domenindeki genlik grafiğine bakıldığında hem yüksek frekanslarda hem düşük frekanslarda tepe değerler gözlenmektedir. Genlik spektrum grafiği incelendiğinde 120 Hz, 840 Hz, 1670 Hz ve 2560 Hz değerlerinde yüksek genlikler gözlemlenmiştir. Spektrogram grafiği incelendiğinde enerjinin çoğunluğunun yüksek frekanslarda olduğu gözlemlenir.

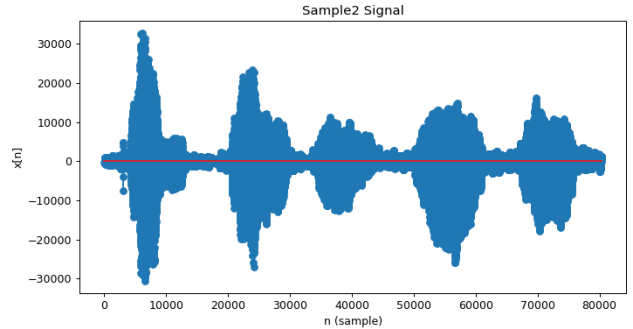


Şekil 9

Filtrelenmiş sonuçlar incelendiğinde ise, frekans domeninde genlik grafiklerine bakıldığında BPF çıkışında enerji yoğunluğunun daha fazla olduğu gözlemlenmiştir. Genlik spektrumları incelendiğinde de BPF çıkışında daha yüksek tepe noktalar bulunduğu gözlemlenir. Şekil 9’da verilen spektrogram grafiğine ve ortalama genlik değeri çıktılarına bakıldığında (LPF = 1.51, BPF = 4.71) bant geçiren filtre çıktısında daha yüksek enerji bulunduğunu, yani örneğin kedi miyavlaması olduğu tespit edilir. Algoritma da bu örneği kedi olarak değerlendirmiştir.

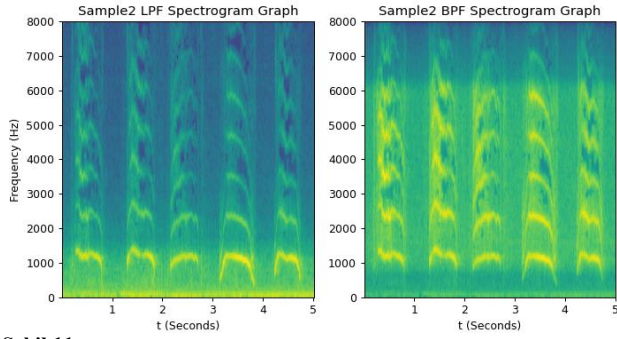
Örnek kedi olmasına rağmen düşük frekanslarda beklenilenden daha yüksek enerji tespit edilmiştir, bu durumun sebebi bu ses kaydında oldukça fazla ortam gürültüsü bulunmasıdır.

2.5.2. Örnek 2



Şekil 10

Bu ses dosyasının, Şekil 10’da, 5 parçadan oluştuğu gözlenmiştir. Frekans domenindeki genlik grafiğine bakıldığında ağırlıklı olarak yüksek frekanslarda tepe değerler gözlenmektedir, fakat frekans genişliği az olmasına rağmen belirli bir düşük frekansta da oldukça yüksek genlik değeri almıştır. Genlik spektrum grafiği incelendiğinde 1150 Hz ve 2300 Hz noktalarında yüksek genlik değerleri almıştır. Spektrogram grafiği incelendiğinde enerjinin çoğunluğunun yüksek frekanslarda olduğu gözlemlenir.

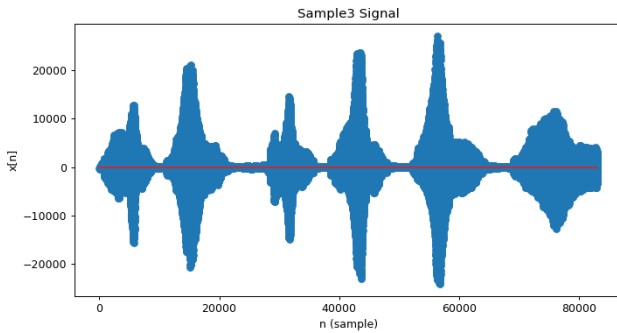


Şekil 11

Filtrelenmiş sonuçlar incelendiğinde, frekans domeninde genlik grafiklerine bakıldığında BPF çıkışında enerji yoğunluğu daha fazla olmasına rağmen dar bir düşük frekans bandında da tepe değerler bulunmaktadır. Genlik spektrumları incelendiğinde de BPF çıkışında daha yüksek enerji bulunduğu gözlemlenir. Şekil 11’de verilen spektrogram grafiğine ve ortalama genlik değeri çıktılarına bakıldığında (LPF = 1.95, BPF = 8.11) bant geçiren filtre çıktısında oldukça daha yüksek enerji bulunduğunu, yani örneğin kedi miyavlaması olduğu tespit edilir. Algoritma da bu örneği kedi olarak değerlendirmiştir.

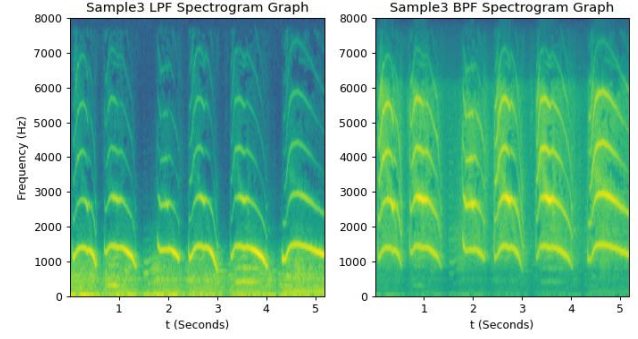
Alçak geçiren filtre çıkışı ile bant geçiren filtre çıkışı ortalama genlik değerleri arasında ciddi miktarda fark bulunmasına rağmen bütün grafiklerde gözlemlenebilen, spesifik bir düşük frekans noktasında (30 Hz) bulunan bir gürültü vardır. Bu gürültü yüksek değerli fakat dar bir bantta bulunmaktadır. Bu gürültünün sebebi arka planda çalışan düşük frekansa sahip bir havalandırma aleti vb. olabilir.

2.5.3. Örnek 3



Şekil 12

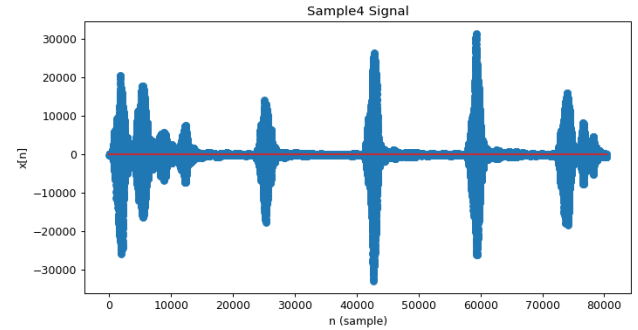
Bu ses dosyasının, Şekil 12’de, 6 parçadan oluştuğu gözlemlenmiştir. Frekans domenindeki genlik grafiğine bakıldığında sadece yüksek frekans değerlerinde tepe noktaları gözlemlenmektedir. Genlik spektrum grafiği incelendiğinde 1380 Hz ve 2700 Hz noktalarında yüksek değerler almışlardır. Spektrogram grafiği incelendiğinde enerjinin hepsinin yüksek frekanslarda olduğu gözlemlenir.



Şekil 13

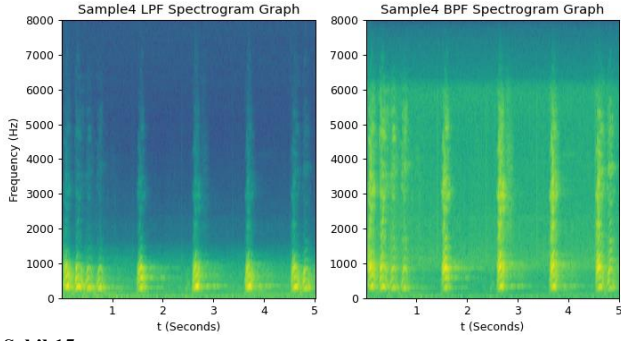
Filtrelenmiş sonuçlar incelendiğinde, frekans domeninde genlik grafiklerine bakıldığında BPF çıkışında tepe noktaları LPF’ye kıyasla yaklaşık 5 kat olarak gözlemlenmiştir. Genlik spektrumları incelendiğinde de BPF çıkışında tepe değerlerde daha yüksek genlik gözlemlenir. Şekil 13’te verilen spektrogram grafiğine ve ortalama genlik değeri çıktılarına bakıldığında (LPF = 0.52, BPF = 4.79) bant geçiren filtre çıktısında daha yüksek enerji bulunduğu, örneğin kedi miyavlaması olduğu tespit edilir. Algoritma da bu örneği kedi olarak değerlendirmiştir. Bu örnekte gürültü miktarı oldukça az olduğu için düşük frekans bandında değerler düşük gelmiştir.

2.5.4. Örnek 4



Şekil 14

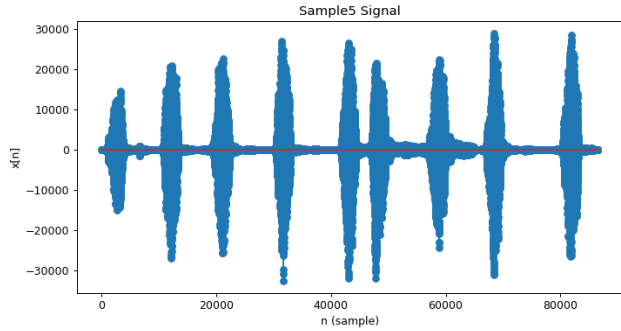
Bu ses dosyasının, Şekil 14’te, 5 parçadan oluştuğu gözlemlenmiştir. Frekans domenindeki genlik grafiğine bakıldığında sadece düşük frekans noktalarında genlik değeri gözlemlenmektedir. Genlik spektrum grafiği incelendiğinde sadece 320 Hz ve 570 Hz noktalarında yüksek değer almışlardır. Spektrogram grafiği incelendiğinde enerjinin hepsinin düşük frekanslar bandında olduğu gözlemlenir.



Şekil 15

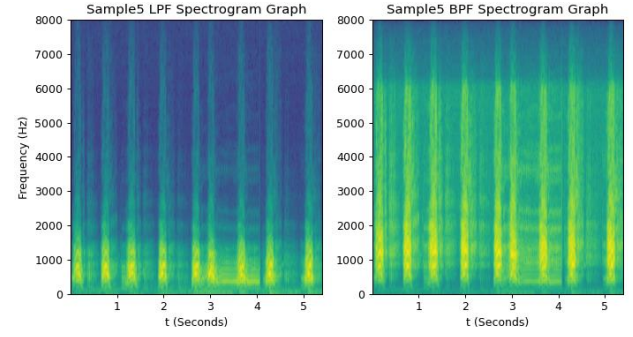
Filtrelenmiş sonuçlar incelendiğinde, frekans domeninde genlik grafiklerine bakıldığında LPF çıkışının tepe noktalarının BPF'ye kıyasla yaklaşık 15 kat olduğu gözlemlenmiştir. Genlik spektrumları incelendiğinde de LPF çıkışının tepe değerlerinde daha yüksek enerji miktarı gözlemlenir. Şekil 15'te verilen spektrogram grafiğine ve ortalama genlik değeri çıktılarına bakıldığında (LPF = 2.14, BPF = 0.42) alçak geçiren filtre sonucunda daha yüksek enerji bulunduğu gözlemlenir. Bu durumdan dolayı bu işaretin köpek havlaması sesi olduğu tespit edilir. Algoritma da bu örneği köpek olarak değerlendirmiştir. Gürültü miktarı bu örnekte problem olmamıştır.

2.5.5. Örnek 5



Şekil 16

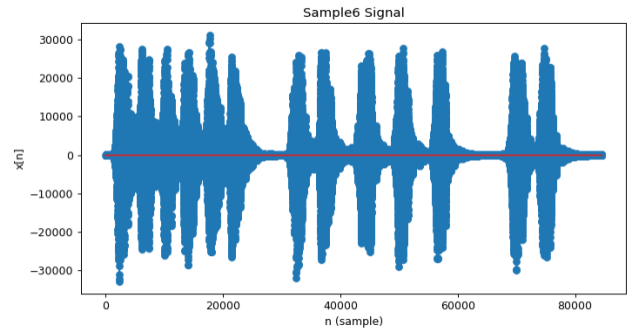
Bu ses dosyasının, Şekil 16'da, 9 parçadan oluştuğu gözlenmiştir. Frekans domenindeki genlik grafiğine bakıldığında sadece düşük frekans noktalarında genlik değerleri gözlemlenmektedir. Genlik spektrum grafiği incelendiğinde sadece 560 Hz ve 820 Hz noktalarında yüksek değerler almışlardır. Spektrogram grafiği incelendiğinde enerjinin hepsinin orta-düşük frekanslarda olduğu gözlemlenir.



Şekil 17

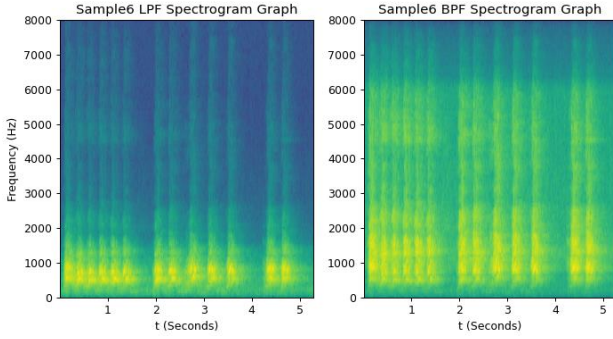
Filtrelenmiş sonuçlar incelendiğinde, frekans domeninde genlik grafiklerine bakıldığında LPF çıkışının tepe değerleri, BPF çıkışına kıyasla yaklaşık 4 kat olarak gözlemlenmiştir. Genlik spektrumları incelendiğinde de LPF çıkışında tepe değerlerde daha yüksek enerji miktarı gözlemlenir. Şekil 17'de verilen spektrogram grafiğine ve ortalama genlik değeri çıktılarına bakıldığında (LPF = 3.16, BPF = 1.83) alçak geçiren filtre sonucunda daha yüksek enerji bulunduğu gözlemlenir. Bu durumdan dolayı bu işaretin köpek havlaması sesi olduğu tespit edilir. Algoritma da bu örneği köpek olarak değerlendirmiştir. Bu örnekte bant geçiren filtrenin genlik değeri beklenilenden daha yüksek çıkmıştır, bu durumun sebebi bu sesin ortalama frekansının daha yüksek olması, örnekteki köpeğin diğerlerine kıyasla daha küçük olması olabilir.

2.5.6. Örnek 6



Şekil 18

Bu ses dosyasının, Şekil 18'de, 13 parçadan oluştuğu gözlenmiştir. Frekans domenindeki genlik grafiğine bakıldığında sadece orta-düşük frekans noktalarında genlik değeri gözlemlenmektedir. Genlik spektrum grafiği incelendiğinde sadece 460 Hz, 520 Hz ve 780 Hz noktalarında yüksek değer almışlardır. Spektrogram grafiği incelendiğinde enerjinin hepsinin orta-düşük frekanslarda olduğu gözlemlenir.



Şekil 19

Filtrelenmiş sonuçlar incelendiğinde, frekans domeninde genlik grafiklerine bakıldığında LPF çıkışının tepe noktalarında, BPF çıkışına kıyasla yaklaşık 3 kat değer aldığı gözlemlenmiştir. Genlik spektrumları incelendiğinde de LPF çıkışının tepe değerlerinde daha yüksek enerji miktarı gözlemlenir. Şekil 19’da verilen spektrogram grafiğine ve ortalama genlik değeri çıktılarına bakıldığında (LPF = 4.22, BPF = 1.73) alçak geçiren filtre sonucunda daha yüksek enerji bulunduğu gözlemlenir. Bu durumdan dolayı bu işaretin köpek havlaması sesi olduğu tespit edilir. Algoritma da bu örneği köpek olarak değerlendirmiştir.

2.6. Analiz Sonuçları

Bütün örnekler incelendiğinde ve dinlendiğinde, gözlemler ve analizlerle varılan sonuçların, işitilen ile örtüştüğü tespit edilmiştir.

Analiz grafiklerinin çıktılarında da gözlemlenebilir olduğu gibi, örnek 1 ve 2 de oldukça fazla arka plan gürültüsü bulunmaktadır. Ayrıyeten örnek 5’te ki köpeğin sesi diğer iki köpeğe kıyasla daha tizdir.

Analizlerle varılan ve grafiklerde elde edilen sonuçlar, işitilen ve beklenen sonuçlar ile örtüşmüştür.

Analiz sonuçlarını iyileştirmek için, örnek 1 ve 2’nin aksine daha az gürültüye sahip ses örnekleri kullanılabilir. Filtreleme işlemlerinde kullanılan FIR Barthann filtreleri yerine daha ideal çalışan filtreler kullanılabilir.

3. Sonuç ve Yorum

Kazanım olarak, genlik spektrumu, spectrogram ve ortalama genlik değerlerine göre yorum yapılması konusunda deneyim elde edildi. Duruma ve ihtiyaca göre filtre tasarlanması konusunda deneyim elde edildi.

Çözümün kısıtları ve problemleri, arka planında yüksek frekanslarda gürültü bulunan bir köpek sesi veya düşük frekanslarda gürültü bulunan bir kedi sesi yanlış analiz edilebilir. Çözümün geliştirilmesi için gürültü azaltıcı algoritmalar kullanılabilir. Çözüm kedilerin veya köpeklerin çıkardığı miyavlama ve havlama seslerinden farklı sesleri ayırt edememektedir.

Bu çalışmada bir probleme çözüm ararken elde bulunan veri setlerinin önemi fark edildi. Örnek verilerin kalitesinin veya kalitesizliğinin çözüm üretme sürecinde

birçok farklı probleme sebep olabileceği ve bu problemlerin nasıl çözülebileceği öğrenildi.

Kaynaklar

1. matplotlib.pyplot.magnitude_spectrum, matplotlib.org, https://matplotlib.org/stable/api/as_gen/matplotlib.pyplot.magnitude_spectrum.html
2. matplotlib.pyplot.specgram, matplotlib.org, https://matplotlib.org/stable/api/as_gen/matplotlib.pyplot.specgram.html
3. Audio Cats and Dogs, kaggle, <https://www.kaggle.com/mmoreaux/audio-cats-and-dogs>