

ELM368 – DÖNEM PROJESİ

FFT ve STFT Kullanarak Gerçek Zamanlı Ses Analizi ve Kuş Sesi Sınıflandırması

Hüseyin Kaan Öner, Muhittin Hamza Hanedar, Kağan Alper

210102002001, 200102002039, 210102002073

h.oner2021@gtu.edu.tr, m.hanedar2020@gtu.edu.tr, k.alper2021@gtu.edu.tr

ABSTRACT (ÖZET)

Bu çalışmada, Python kullanılarak Kakapo ve Finch kuş seslerini analiz etmek ve sınıflandırmak için iki farklı yöntem karşılaştırılmıştır. İlk yöntem, ses sinyalini bloklara ayırarak her blok üzerinde hızlı Fourier dönüşümü (FFT) uygulanmasına ve bu blokların zirve (peak) frekanslarına odaklanarak tür belirlenmesine dayanmaktadır. İkinci yöntem ise, tüm sinyal üzerinde uygulanan kısa-zamanlı Fourier dönüşümü (STFT) kullanarak belirli frekans bantlarında enerji yoğunluklarını hesaplamakta ve bu yoğunluklar üzerinden karar vermektedir. Her iki yaklaşım da tür baskınlığını gerçek zamanlı olarak belirlemeyi, elde edilen frekans spektrumlarını görselleştirerek kullanıcılara anlık bilgi sunmayı ve kullanılan tekniklerin güçlü ve zayıf yönlerini ortaya koymayı hedeflemiştir. Elde edilen sonuçlar, frekans-temelli analiz yaklaşımlarının sınıflandırma başarısını ve sahada kullanılabilirliğini değerlendirmek açısından önem taşımaktadır.

ANAHTAR KELİMELELER

Ses Analizi, FFT, STFT, Frekans Spektrumu, Kuş Sınıflandırma, Enerji Analizi

1. Giriş

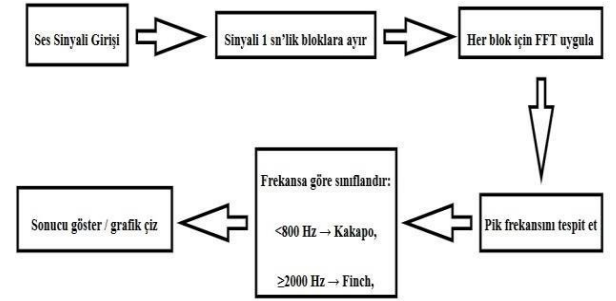
Bu projede amaç, farklı türdeki kuş seslerini frekans tabanlı analiz yöntemleriyle sınıflandırmak ve gerçek zamanlı görselleştirme sağlamaktır. Özellikle Kakapo (düşük frekans) ve Finch (yüksek frekans) türleri ele alınmış, sinyallerin hızlı analiz edilmesi ve anında sınıflandırılması hedeflenmiştir. İki yöntem incelenmiştir: basit ve hızlı FFT (Fast Fourier Transform) analizi ile detaylı STFT (Short-Time Fourier Transform) tabanlı enerji analiz yöntemi.

2. Deneyler ve Analiz

2.1 Yöntem 1: Blok Bazlı FFT Analizi

Bu yöntemde ses sinyali, 1 saniyelik bloklara ayrılmış ve her blokta FFT uygulanmıştır. Spektrumda en yüksek genlikli frekans bulunmuş, bu pik frekansın değerine göre blok Kakapo, Finch veya sessizlik olarak

sınıflandırılmıştır. Blok başına canlı grafik çizilmiş ve ses çalma gerçekleştirilmiştir.

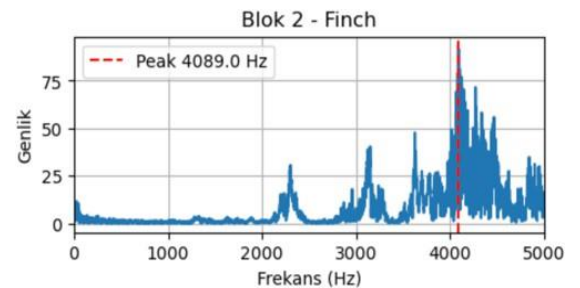


Şekil 1: FFT Yöntemi Akış Diagramı

Şekil 1’de, FFT tabanlı kuş sesi sınıflandırma yönteminin adım adım işleyişi akış diyagramı şeklinde sunulmuştur.

Şekil 2’de görüldüğü üzere, analiz edilen ikinci blokta frekans spektrumunun zirvesi 4089 Hz olarak belirlenmiş ve bu nedenle blok Finch türü olarak sınıflandırılmıştır. Benzer şekilde, Şekil 4’te verilen zaman-frekans dağılımında Kakapo ve Şekil 5’te verilen zaman-frekans dağılımında Finch bantları görselleştirilmiş, tür tespiti için kullanılan enerji yoğunlukları net biçimde ayrıştırılmıştır. Elde edilen bulgular, kullanılan iki yöntemin (blok bazlı FFT ve STFT enerji analizi) analiz performansını doğrudan karşılaştırmalı biçimde incelemektedir.

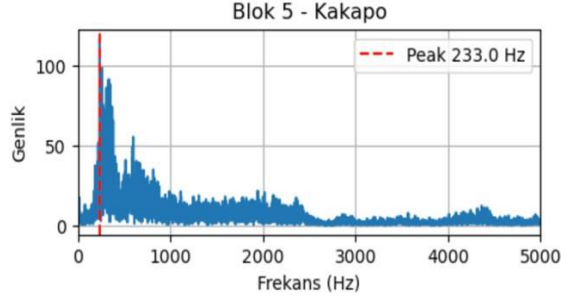
Blok 2: Finch (Peak: 4089.0 Hz)



Şekil 2: Karışık ses dosyasında blok 2 anında algılanan Finch sesi spektrumu.

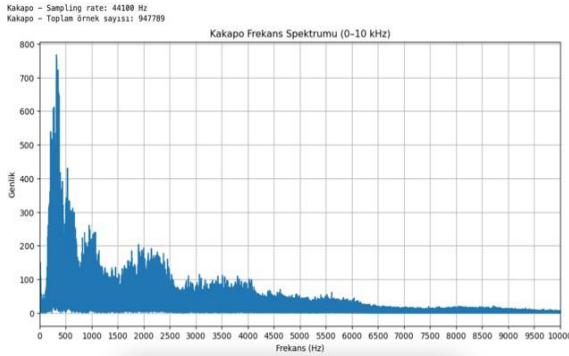
Şekil 3’de görüldüğü üzere, karışık ses dosyasının 5. bloğunda 233 Hz’de bir pik tespit edilerek bu blok Kakapo türü olarak sınıflandırılmıştır; spektrumun düşük frekans yoğunluğu bu kararı desteklemektedir.

Blok 5: Kakapo (Peak: 233.0 Hz)

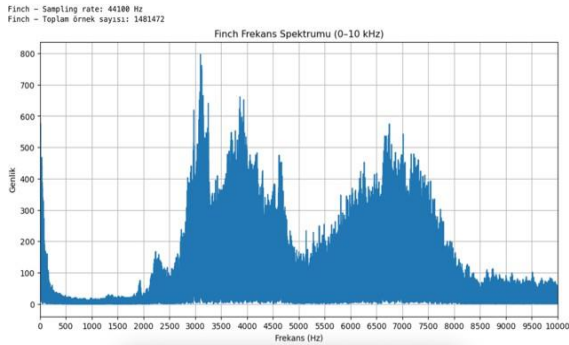


Şekil 3: Karışık ses dosyasında blok 5 anında algılanan Finch sesi spektrumu.

Şekil 4 ve Şekil 5'te görüldüğü üzere, Kakapo ve Finch türlerinin ses frekans spektrumları belirgin şekilde farklılık göstermektedir. Kakapo sesi ağırlıklı olarak 500 Hz altındaki düşük frekanslarda yoğunlaşırken, Finch sesi 3000–7000 Hz aralığında güçlü zirveler oluşturarak yüksek frekanslarda baskın hale gelmiştir. Bu spektrum farkları, analiz ve sınıflandırma algoritmalarının doğru tür tespiti yapabilmesi için temel dayanak oluşturmaktadır.



Şekil 4: Kakapo sesi frekans spektrumu.



Şekil 5: Finch sesi frekans spektrumu.

Parametreler:

- Blok süresi: 1 saniye
- Pik frekans eşikleri: <100 Hz sessizlik, <800 Hz Kakapo, ≥2000 Hz Finch

Avantajlar: Bu yöntem hızlı, basit ve gerçek zamanlı uygulamalar için oldukça uygundur; özellikle düşük hesaplama gücüne sahip sistemlerde bile anlık analiz yapılmasına olanak tanır. Blok bazlı çalışması sayesinde

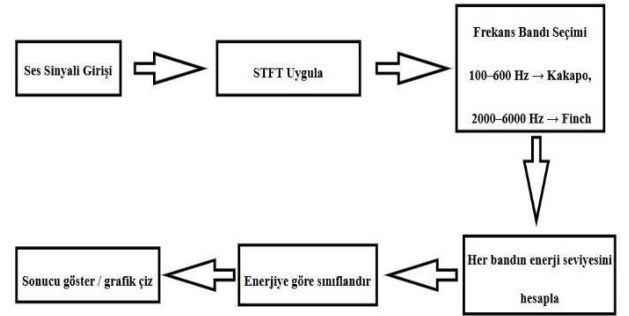
ses akışını anlık olarak değerlendirip sınıflandırma sunabilir, bu da özellikle sahada veya canlı gözlemlerde pratik bir çözüm oluşturur.

Kısıtlar: Ancak yalnızca en yüksek (peak) frekans bileşenine odaklanması, tüm enerji dağılımını ve spektrumun genel yapısını göz ardı etmesine neden olur. Bu durum, özellikle gürültü veya birden fazla frekans bileşeni içeren karmaşık sinyallerde hatalı sınıflandırmalara yol açabilir. Enerji yoğunluğu düşük ancak biyolojik açıdan önemli olan ses parçaları bu analizde gözden kaçabilir.

2.2 Yöntem 2: STFT Tabanlı Enerji Analizi

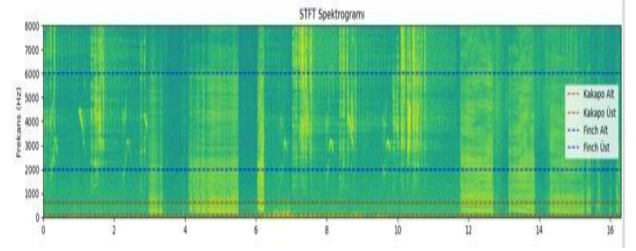
İkinci yöntemde, tüm ses sinyali üzerinde STFT uygulanmakta ve elde edilen zaman-frekans matrisinden Kakapo (100–600 Hz) ve Finch (2000–6000 Hz) frekans bantlarına ait enerji seviyeleri hesaplanmaktadır. Enerji değerleri normalize edilmekte ve dinamik eşik değerleri kullanılarak her bir zamanda hangi türün baskın olduğu belirlenmektedir.

Bu yöntem yalnızca analiz amaçlıdır; ses bileşenlerini yeniden sentezleme veya ayırma işlemi gerçekleştirilmemektedir.

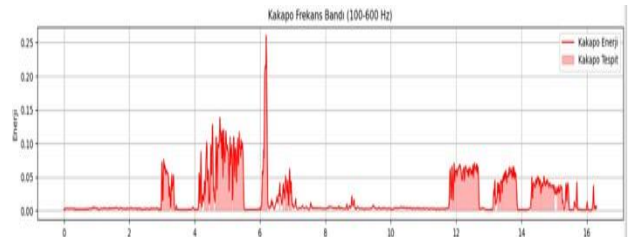


Şekil 6: STFT Akış Diagramı

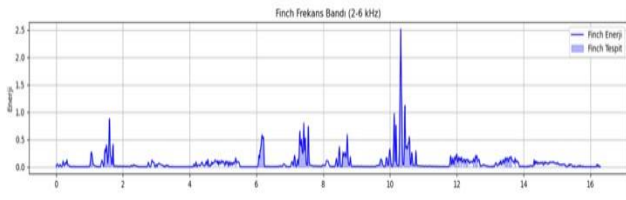
Şekil 6, STFT yöntemiyle enerji bazlı kuş sesi sınıflandırma sürecinin akış diyagramını göstermektedir.



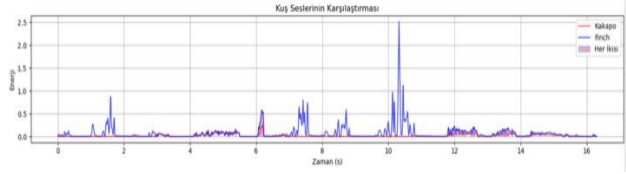
Şekil 7: STFT Spektrogramı.



Şekil 8: Kakapo sesi frekans bandı.



Şekil 9: Finch sesi frekans bandı.



Şekil 10: Kuş sesleri enerji karşılaştırması.

Şekil 7’te yer alan STFT spektrogramı, sinyalin zaman-frekans düzlemindeki dağılımını göstermektedir. Bu görselde Kakapo ve Finch’e ait frekans bantları işaretlenmiş, sinyalin belirli zaman dilimlerinde hangi frekanslarda yoğunlaştığı gözlemlenmiştir. Şekil 8 ve Şekil 9’da, sırasıyla Kakapo (100–600 Hz) ve Finch (2000–6000 Hz) frekans bantlarındaki enerji değişimleri ayrıntılı olarak sunulmuştur. Her iki tür için tespit edilen enerji yoğunlukları grafiklerde açıkça seçilmektedir. Şekil 10 ise, her iki türün enerji düzeylerini karşılaştırmalı olarak vermekte ve aynı anda tespit edildikleri zaman aralıklarını vurgulayarak analizde eşzamanlılık durumlarını ortaya koymaktadır. Bu görseller, STFT tabanlı yöntemin enerji bazlı tespitlerde sağladığı avantajları açıkça ortaya koymaktadır.

SANİYE BAZINDA DETAYLI ANALİZ

Saniye	Baskın Tür	Kakapo %	Finch %	Durum
0	Sessizlik	0.0%	1.1%	● Sessiz
1	Finch	0.0%	30.2%	● Net
2	Sessizlik	2.3%	0.0%	● Sessiz
3	Kakapo	29.1%	0.0%	● Net
4	Kakapo	67.4%	3.5%	● Net
5	Kakapo	47.7%	2.3%	● Net
6	Kakapo	37.2%	16.3%	● Net
7	Finch	0.0%	31.0%	● Net
8	Finch	0.0%	25.6%	● Net
9	Finch	0.0%	12.8%	● Net
10	Finch	0.0%	47.7%	● Net
11	Kakapo	17.4%	11.6%	● Net
12	Kakapo	69.8%	31.4%	● Net
13	Kakapo	58.1%	20.9%	● Net
14	Kakapo	67.4%	1.2%	● Net
15	Kakapo	23.0%	0.0%	● Net
16	Kakapo	7.7%	0.0%	● Net

Tablo 1: Kuş sesleri enerji karşılaştırması.

Tablo 1’de görüldüğü üzere, analiz edilen ses verisi saniyelik bloklara ayrılarak her blokta hangi kuş türünün baskın olduğu belirlenmiştir. Başlangıçta birkaç saniyelik sessizlik tespit edilse de, özellikle 3. saniyeden itibaren Kakapo ve Finch türleri net bir şekilde ayırt edilebilmiştir. Kakapo sesi 4–6 ve 11–14. saniyeler arasında belirgin şekilde baskınken, Finch sesi ise 1, 7, 8 ve 10. saniyelerde öne çıkmıştır. Bu tablo, STFT tabanlı enerji analizinin zaman çözünürlüğü açısından oldukça başarılı olduğunu ve kuş türlerini saniye düzeyinde ayırtılabildiğini göstermektedir. Ancak yöntemin blok başına enerji hesaplaması ve grafik üretimi gibi işlemleri nedeniyle diğer

yöntemlere kıyasla daha yavaş çalıştığı gözlemlenmiştir. Bu durum, özellikle gerçek zamanlı uygulamalarda performans sınırlayıcı bir faktör olabilir.

Avantajlar:

- Daha kapsamlı ve doğru tespit
- Tüm frekans bandındaki enerji dağılımını dikkate alır
- Zaman ve frekans bazında detaylı içgörüler sunar

Kısıtlar:

- Daha yüksek hesaplama maliyeti
- Daha karmaşık kodlama ve işlem yapısı
- Sadece analiz yapar, sesleri ayırmaz veya yeniden üretmez

Her iki yöntem aynı ses kaydı üzerinde uygulanmıştır. FFT yaklaşımı hızlı ve blok bazlı sınıflandırma sağlamış, konsola anında tür tespitleri yazdırmış ve eş zamanlı spektrum grafiği çizmiştir. Ancak, bu yöntem yalnızca zirve frekansa odaklandığı için karmaşık veya örtüşen sinyallerde hata yapma eğilimindedir.

STFT tabanlı yöntem ise, frekans bantlarındaki enerji dağılımını detaylı inceleyerek daha güvenilir tespitler sağlamıştır. Enerji trendleri grafiklerle sunulmuş, ancak yöntem yalnızca analiz amaçlı kullanılmış, ses bileşenlerinin ayrıştırılması veya yeniden üretilmesi yapılmamıştır.

Genel Karşılaştırma:

- FFT yöntemi, hızlı ve gerçek zamanlı izleme için uygundur.
- STFT yöntemi, detaylı enerji analizi ve güçlü tespit yetenekleri sunar.

3. Sonuç ve Yorum

Bu çalışmada, Kakapo ve Finch kuş türlerine ait ses verileri üzerinde iki farklı frekans analizi yöntemi uygulanarak sınıflandırma işlemi gerçekleştirilmiştir. İlk yöntem olan FFT (Hızlı Fourier Dönüşümü), ses sinyalinin her bir blokunda baskın frekans bileşenini tespit ederek tür tahmini yapmaya odaklanırken, ikinci yöntem olan STFT (Kısa Süreli Fourier Dönüşümü) belirli frekans bantlarındaki enerji yoğunluğunu zamana bağlı olarak analiz etmektedir. FFT yöntemi, düşük işlem maliyeti ve hızlı çalışabilmesi sayesinde özellikle gerçek zamanlı uygulamalarda kullanılmaya uygundur. Ancak yalnızca tepe frekansa (peak frequency) odaklanması nedeniyle enerji spektrumunun genel dağılımını göz ardı edebilmekte ve bu da sınıflandırma doğruluğunu düşürebilmektedir.

STFT yöntemi ise sinyalin zaman-frekans düzlemindeki dağılımını ayrıntılı olarak analiz ederek daha güvenilir sonuçlar vermektedir. Bu yöntem sayesinde kuş seslerinin belirli frekans aralıklarındaki davranışları izlenebilmekte ve zamana bağlı olarak sesin baskın türü tespit edilebilmektedir. Ancak daha karmaşık hesaplamalar

gerektirdiđi iin zellikle uzun ses kayıtlarında yavaş alıřabilmekte ve anlık kararlar gerektiren durumlarda sınırlayıcı olabilmektedir.

Elde edilen sonular gstermektedir ki her iki yntemin avantajları bir araya getirilerek daha verimli bir hibrit sistem tasarlanabilir. rneđin, FFT yntemiyle hızlı bir n tarama yapılarak belirli zaman aralıklarında dikkat eken blmler belirlenebilir ve ardından bu blmler zerinde STFT yntemi ile daha derinlemesine analiz gerekleřtirilerek sınıflandırma dođruluđu artırılabilir. Bu tr bir yaklařım, hem iřlem sresi aısından tasarruf sađlayacak hem de analiz gvenilirliđini koruyacaktır.

Sonu olarak bu alıřma, iki farklı analiz yntemiyle kuř seslerini bařarıyla sınıflandırmıř ve yntemlerin gl ve zayıf yönlerini ortaya koymuřtur. Elde edilen sonular, evresel izleme ve ses tanıma gibi alanlarda pratik kullanım potansiyeli tařımaktadır.

Kaynaklar

- [1] A. V. Oppenheim, R. W. Schaffer, “Discrete-Time Signal Processing,” Prentice Hall, 2010.
- [2] B. Gold, N. Morgan, D. Ellis, “Speech and Audio Signal Processing,” Wiley, 2011.
- [3] Librosa Ktphanesi Dokmantasyonu, <https://librosa.org>
- [4] Scipy Signal İřleme Dokmantasyonu, <https://docs.scipy.org/doc/scipy/reference/signal.html>