

# Mısır Meyve Yararasası Çağrılarının Derin Öğrenme Yöntemleriyle Sınıflandırılması

## Classification of Egyptian Fruit Bat Calls with Deep Learning Methods

Doğukan MESCI<sup>1</sup>, Anıl KOLUAÇIK<sup>2</sup>, Batuhan YILMAZ<sup>3</sup>, Melih ŞEN<sup>4</sup>, Engin MAŞAZADE<sup>5</sup>, Vedat BEŞKARDEŞ<sup>6</sup>

<sup>1,2,3,5</sup>Marmara Üniversitesi, Mühendislik Fakültesi, Elektrik ve Elektronik Mühendisliği Bölümü, İstanbul, 34840, Türkiye.

<sup>4</sup>Marmara Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü İstanbul 34840, Türkiye.

<sup>6</sup>İstanbul Üniversitesi – Cerrahpaşa, Orman Fakültesi, Orman Mühendisliği Bölümü, İstanbul, 34473, Türkiye.

<sup>1</sup>dogukanmesci@marun.edu.tr, <sup>2</sup>anilkoluacik@marun.edu.tr, <sup>3</sup>batuhanyilmaz@marun.edu.tr, <sup>4</sup>melihsen@marun.edu.tr,

<sup>5</sup>engin.masazade@marmara.edu.tr, <sup>6</sup>vkardes@istanbul.edu.tr

**Özetçe—** Yarasalar, tüm canlı varlıkların yaşamlarının devamı ve biyoçeşitlilik için büyük önem taşımaktadır. Bu çalışma, yayılış alanının en kuzeyi Türkiye olan Mısır meyve yararasasının topluluk halinde çıkardıkları farklı içerikteki çağrıları derin öğrenme yöntemleri CNN ve LSTM kullanarak ve MFCC (Mel Frequency Cepstral Coefficients) özneliklerinden faydalanılarak sınıflandırmayı amaçlamaktadır. Türe özgü çağrıların sınıflandırılması sayesinde türün habitat tercihi, sosyal ilişkileri, yiyecek arama, üreme, hareketlilik ve göç gibi durumlarının gözlemlenmesi mümkün olmaktadır. Bu çalışmada elde edilen sınıflandırma sonuçları, özellikle belirli çağrıların ayırt edilmesinde önceki çalışmaya göre önemli artışlar sağlamaktadır.

**Anahtar Kelimeler —** Derin Öğrenme, MFCC, Mısır Meyve Yararasası, İşaret İşleme

**Abstract—** Bats are of great importance for the survival of all living beings and for biodiversity. This study aims to classify the collective calls of the Egyptian fruit bat, whose northernmost distribution is in Turkey, using deep learning methods CNN and LSTM and utilizing MFCC (Mel Frequency Cepstral Coefficients) features. Thanks to the classification of species-specific calls, it is possible to observe the habitat preference, social relations, foraging, reproduction, mobility and migration of the species. The classification results obtained in this study provide significant increases compared to the previous study, especially in distinguishing certain calls.

**Keywords —** Deep Learning, MFCC, Egyptian Fruit Bat, Signal Processing

### I. GİRİŞ

Yarasalar, ekosistemde yırtıcı, omurgalılar için av, parazitler için konak olmaları; tozlaşma, tohum dağılımı, toprak verimliliği ve besin dağılımına fayda sağlamaları; biyoindikatör olmaları dolayısıyla ekolojik ve ekonomik olarak önemli rollere sahiptir ve zararlılar için biyolojik kontrol ajanı olarak hizmet ederler [1]. Mısır meyve yararasaları, besin olarak meyve ve nektarlarını tüketir ve bu yolla Türkiye'deki diğer yarsa türünden ayrılır [2]. Dışkılarında bulunan yüksek miktarda fosfat, nitrojen ve potasyum içeren guano, tarımda gübre olarak kullanılan maddelerdir ve bu yolla yarasarlar, tarımdaki verimin artırılmasında büyük rol oynarlar. Mısır meyve yararasası, yayılış alanının en kuzeyine ulaşan Afro-tropik türler olması nedeniyle Türkiye'nin biyoçeşitliliği açısından önemlidir [3]. Polen taşıma

konusunda az etkin olmalarına karşın dışkılama yoluyla meyve çekirdeklerini uzak bölgelere taşımaları dolayısıyla bitki türlerinin yayılmasına olanak sağlamaktadır [4]. Mısır meyve yararasaları sosyal hayvanlardır. Mağaralarda 500 bireye kadar çıkabilen kolonilerle yaşarlar [4]. Böylece, habitat tercihi, sosyal ilişkiler, yiyecek arama, üreme, hareketlilik ve göç gibi durumlar ile kolonilerdeki hastalıklar, gıda kıtlığı, az nüfus ve aşırı nüfus gibi sorunların gözlemlenmesi mümkün olmaktadır. Bu değişim ve sorunların anlamsal bazda tespiti, ekolojik dengenin korunması ve tarım uygulamalarının doğru biçimde gerçekleştirilmesinde düzenleyici bir rol oynayabilir. Örneğin, çiftleşme çağrılarının ve yavru yarsa seslerinin tespiti, üreme mevsiminin erken veya geç başladığının göstergesi olarak ele alınabilir. Öte yandan türün beslenme sürecinde ürettiği seslerin azalması ve agresif çağrıların artması kolonide yiyecek kıtlığının başladığını, akabinde türün habitatı terk edebileceğinin işaretini verebilmektedir. Bu ve benzeri durumların erkenden tespiti, tarım uygulamalarında tohum ekim zamanı ve gübreleme gibi verimi doğrudan veya dolaylı olarak etkileyen karar süreçlerine önemli katkılarda bulunabilir. Edinilen bilgiler ışığında zamanında alınacak önlemlerle; olası mahsul kayıplarından kaçınılması, çevreye verilecek zararın önüne geçilmesi ve daha fazla verim elde edilmesi sağlanabilir.

Ses işleme alanında yapılan çalışmaların büyük çoğunluğu açık kaynaklı veri kümeleri kullanılarak günlük hayatta sıklıkla duyulabilen sesleri tespit etmek için yapılmıştır. Genellikle geleneksel makine öğrenmesi ile derin öğrenme yöntemleri kullanılarak alınan sonuçlar kıyaslanmıştır. Piczak [5], evrişimli sinir ağlarının 3 farklı ve az sayıda veri içeren veri kümesi ile performansını, 2 tane evrişim katmanını takip eden maksimum ortaklama katmanı ile kurduğu derin öğrenme ağı ile test etmiştir. Öznelik çıkarmak için MFCC kullanmıştır. Çalışmaları sonucunda ESC-50 veri kümesi ile %64,5'lik bir doğruluk oranı elde ederek [6]'da yapılan gözetimsiz öğrenmeden daha iyi bir performans gösterdiği sonucuna ulaşmıştır. Ek olarak, ses çok kısa süreli bile olsa evrişimli sinir ağlarının denetimsiz öğrenme tekniklerinden daha başarılı olduğunu iddia etmiştir. [7]'de yazarlar, çevresel sesleri Kırmızı-Yeşil-Mavi renk uzayına çevirmiştir. Kırmızı kanalda Log-Mel Spektrogram, yeşil kanalda Skalogram ve mavi kanalda ise MFCC değerleri saklanmıştır. ESC-50 veri kümesi ile kısıtlı miktarda veri kullandıkları için transfer öğrenimi kullanmışlardır. Önceden öğretilmiş model olan Xception ile beraber kendi kurdukları ve bir evrişim katmanını takip eden maksimum ortaklama katmanı, seyreltme katmanı ve son olarak yoğun katman içeren derin öğrenme ağı ile yaptıkları çalışma sonucu %90,5'lik bir doğruluk oranına ulaşmışlardır. Bu oran, Piczak'ın [5] modeli ile

karşılaştırıldığında çok daha doğru sonuçlar vermektedir. Graves, Fernandes ve Schmidhuber [8], tek yönlü ve iki yönlü LSTM (Long Short Time Memory) modelleri üzerinde çalışmışlardır. Çalışmalarında LSTM ve RNN (Recurrent Neural Network) metodlarını kıyaslamışlardır. Farklı LSTM ve RNN modellerinin kullanıldığı 7 farklı eğitimi kare gecikmesi ve HMM (Hidden Markov Model) ile gerçekleştiren yazarlar, LSTM/HMM'in RNN'den daha iyi performans gösterdiği sonucuna varmışlardır.

Çalışmaların bazıları, hayvan seslerini içeren açık kaynaklı veri kümeleri kullanılarak, bu sesler yardımıyla hayvanları sınıflayan evrişimli sinir ağları modeli oluşturmak ve modelden elde edilen doğruluk oranını yükseltmek için yapılmıştır. Şaşmaz ve Tek [9], hayvan seslerini tespit eden bir evrişimli sinir ağı modeli oluşturmaya çalışmış ve optimize edicilerin dengeli ve dengesiz veri kümeleri ile doğruluk değeri üzerindeki etkilerini araştırmışlardır. Yazarlar, 3 evrişim katmanı, bir maksimum ortaklama katmanı ve üç tane yoğun katman içeren bir ağı kurmuşlar ve çalışmalarının sonucunda NADAM (Nastarov-Accelerated Adaptive Moment Estimation) optimize edicinin %75'lik doğruluk oranı ile en iyisi olduğu sonucuna varmışlardır. Ek olarak yazarlar, büyük bir veri kümesi ile çalışılmasını önermektedir. [10]'da yazarlar, daha önceki makalelerde kurulmuş derin öğrenme ağlarını karşılaştırmış ve kuş seslerini uygun bir şekilde tespit edebilmek için bu ağları birleştirmeyi denemişlerdir.

TABLO I. CNN [4]- (CNN, LSTM) KARŞILAŞTIRMASI

No	Kombinasyon	[4]'teki CNN Doğruluk Değerleri	Bu Çalışmadaki CNN Doğruluk Değerleri	Bu Çalışmadaki LSTM Doğruluk Değerleri
1	2-3-4-6-9-12	0.620	0.674	0.577
2	12 – (2,3,4,6,9)	0.734	0.802	0.701
3	2-3-4-6-9	0.742	0.746	0.689
4	2-3-4-6	0.822	0.819	0.771
5	2-3-4	0.803	0.793	0.747
6	2-3	0.881	0.869	0.825
7	2-6	0.977	0.976	0.933
8	2-4	0.898	0.863	0.832
9	3-9	0.892	0.867	0.830
10	2-9	0.782	0.832	0.752
11	4-9	0.832	0.856	0.751

Buna ek olarak yazarlar, kuş sesleri özelliklerini açığa çıkarmak için 3 farklı frekans temsili yolunu araştırmışlardır. Yazarların elde ettiği sonuç, aynı veri kümesinin kullanıldığı makale [11]'dekinin aksine Mel-bin sayısını arttırmanın doğruluk değerini azalttığı yönündedir. Yazarlar, evrişimli sinir ağlarına girdi olarak mel-spectrogram verildiğinde modelin daha doğru sonuçlar verdiğini görmüşlerdir. Ek olarak, yapılan model birleştirmeler sonucu doğruluk değerlerinin arttırılabileceği sonucuna varmışlardır. [12]'deki çalışmada yazarlar, iki yönlü LSTM modeli ile kuş sesi tespit etmeyi amaçlamışlardır. Kullanılan veri kümesi, 20 sınıftan toplam 14311 kayıt içermektedir. Yazarlar, ilk olarak yığın boyutunun doğruluk ve kayıp üzerine etkisini incelemişler ve boyut 64 olduğu zaman modelin en doğru sonucu verdiğini gözlemlemişlerdir. Ek olarak yazarlar, farklı önceden

eğitilmiş modelleri kıyaslamışlar ve DenseNet'in en doğru sonuçları ürettiği sonucunu elde etmişlerdir.

Yarasaların canlılar için önemi, araştırmacıları yarasa türlerinin ve çıkardıkları seslerin sınıflandırılmasını yapacak çalışmalar yapmaya teşvik etmiştir. [13]'te yazarlar, yarasa çağrılarının sınıflandırılması için kullanılan ve Amerika Birleşik Devletleri Balık ve Yaban Hayatı Servisi tarafından önerilen yazılımların yetersizliği üzerine, yarasa seslerini resimlere çevirerek bir evrişimli sinir ağı modeli oluşturmuşlardır. Yazarlar, resim elde etmek için kendi kütüphanelerini oluşturmuşlar ve modellerini 13 farklı veri kümesi üzerinde denemişlerdir. Çalışma sonucunda, veri kümeleri içindeki yarasa sesleri yüksek test ve doğrulama doğruluk değerleri ile türlerine göre sınıflandırılmıştır. Diğer bir çalışmada [4] yazarlar, Mısır Meyve Yarasa türünün topluluk halinde çıkardıkları sesleri tespit eden bir derin öğrenme modeli oluşturmaya amaçlamışlardır. Çalışmalarında, açık kaynaklı bir veri kümesinin [14] 6 sınıfını kullanan yazarlar, bir derin öğrenme yöntemi olan CNN (Convolutional Neural Network) ile klasik makine öğrenmesi tekniklerini kıyaslamışlardır. Frekans temsili için Mel-spektrogram yöntemini kullanmışlardır. Sonuç olarak yazarlar, derin öğrenme modelinin klasik makine öğrenmesi tekniklerine göre en az %10 daha tutarlı sonuçlar verdiği sonucuna ulaşmışlardır. Ek olarak, sesler birbirine çok benzese bile, derin öğrenmenin klasik makine öğrenmesine göre daha iyi iş çıkardığını da söylemektedir. Öte yandan [4]'te yazarlar, çalışma sırasında eğitim süresini kısaltmak ve daha az veri kullanmak üzere, gerçek Mel-spektrogram verileri yerine Mel-spektrogram fotoğraflarını kullanmışlardır.

Bu çalışma, [4]'te yapılan çalışma sonuçlarına iyileştirme sağlamayı amaçlamaktadır. Yazarlar, aynı veri kümesini [14] kullanarak [4]'e göre daha doğru sonuçlar veren evrişimli sinir ağı modeli elde etmeyi amaçlamış, bunun yanında değişken uzunluktaki sesleri belli bir uzunluğa sabitlemek yerine sesleri olduğu gibi işleyebilecek bir LSTM modeli kurmayı hedeflemişlerdir. Bu çalışmada, [4]'ün aksine geliştirilen derin öğrenme modellerine fotoğraf yerine gerçek MFCC değerlerinin verilmiştir. Evrişimli sinir ağı modeli için elde edilen sonuçlara göre, Tablo I' de gösterilen kombinasyonların bazılarında daha yüksek doğruluk sonuçları elde edilirken, bazılarında yakın sonuçlar elde edilmiştir. LSTM modeli başarılı bir şekilde çalıştırılmış ve LSTM eğitimleri sonucunda [4]'te bulunan CNN eğitimleri sonuçlarına yakın doğruluk değerleri elde edilmiştir. Böylece, sabit uzunlukta ses işlemek yerine sesi olduğu gibi işleme hedefi başarılmış ve sesi kırpmadan (veya sıfır eklemeyen) kaynaklanabilecek özellik kayıplarının önüne geçilmiştir.

## II. YÖNTEM

Bu çalışmada derin öğrenme yöntemleri olan CNN ve LSTM kullanılmıştır. Kayıtların girdi olarak sunulabilmesi için öznetelik elde edilmesinde MFCC yöntemi tercih edilmiştir. Çalışmada kullanılan veri kümesindeki [4] kayıtlar toplamda 13 davranış sesi ve toplamda 293,238 kayıt içermektedir [14]. Yazarlar [14]'te sunulan veri kümesinden dengeli veri sayısı ile çalışabilmek ve diğer davranış türlerine nazaran daha çok kayıt sayısı içeren davranışlarla çalışmak üzere bu davranışlar içerisinden toplamda 6 davranış seçmiştir. Bu davranışlar 'ısıрма (2)', 'beslenme (3)', 'dövüşme (4)', 'izolasyon (yavru yarasa sesleri) (6)', 'çiftleşme isteğine tepki (9)' ve 'uyku (12)' olarak seçilmiştir. Aynı davranış çeşitlerinin [4]'te kullanılması sebebiyle ilgili doğruluk kıyaslamaları yapılabilmektedir. Davranışların yanlarında bulunan numaralar ilgili davranışın veri kümesindeki [14] numaralarını belirtmektedir.

Veri kümesindeki kayıtlar ilgili davranış klasörlerine ayrılmadan rastgele klasörlere dağıtılmış şekilde paylaşılmıştır. Veri kümesi içerisinde ilgili kaydın hangi davranışa ait olduğunu belirten dosyalar kayıtlarla beraber sunulmuştur. Yazarlar, rahat kullanım sağlamak için

ilgili bilgi dosyalarından faydalanarak ve MATLAB kullanarak seçilen davranış biçimlerine ait kayıtları oluşturulan davranış klasörlerine ayırmışlardır. Bahsedilen eğitimleri gerçekleştirmek üzere Google Colab platformu seçilmiştir. Derin öğrenme kütüphanelerinin platforma kolayca eklenebilmesi dolayısıyla ve Google'ın sağladığı yüksek işlem gücüne sahip cihazları kullanabilmek adına bu platform tercih edilmiştir. Kayıtların MFCC değerlerini alma aşamasında orijinal örnekleme oranı olan 250 kHz kullanılmıştır ve MFCC değerlerini elde etmek için Librosa kütüphanesinden yararlanılmıştır. MFCC değerleri alınacak zaman aralıkları, ardışık 512 örnek uzunluğunda aralıklar olacak şekilde ayarlanmıştır ve her bir aralıktan 16 adet MFCC değeri alınmıştır. FFT (Fast Fourier Transform) sayısı ise 2048 olarak tercih edilmiştir. Bu noktaya kadar gerçekleşen durumlar CNN ve LSTM yöntemleri için ortak durumlardır. Bu noktadan sonra her bir yöntem için izlenen yol farklıdır ve her birinden ayrıca bahsedilecektir.

#### A. CNN Yöntemi

Geliştirilen CNN modelinde 4 adet evrişim katmanı, 4 adet maksimum ortaklama katmanı, 1 adet düz katmanı takip eden tam bağlantılı yoğun katman ve 1 adet çıktı katmanı bulunmaktadır. CNN yönteminde tüm girdi işaretlerinin boyutlarının aynı olması gerekmektedir. Kayıtların sabitleneceği süreyi bulmak üzere tüm kayıt uzunluklarının sürelerini gösteren kod Python programlama dili kullanılarak yazılmıştır ve tüm kayıt sürelerinin 1 saniyeden uzun olduğu gözlemlenmiştir. Dolayısıyla sabitlenecek süre 1 saniye olarak belirlenmiştir.

Her 512 örneklik zaman aralığının MFCC değeri hesaplanacağı için her bir kayıttan gelebilecek toplam aralık sayısı denklem (1)'de gösterilmiştir.

$$\text{tavan} \left[ \frac{\text{kayıt uzunluğu}}{\text{her bir aralıktaki örnek miktarı}} \right] = \quad (1)$$

$$\text{tavan} \left[ \frac{250000}{512} \right] = 489$$

250 kHz örnekleme oranına sahip bir kayıta 1 saniyelik uzunluk 250000 örneklik yer kaplar. 250000 örnek içerisinde 512 örnek uzunluğuna sahip aralıklardan 488 adet bulunur. Kayıtların sonunda kalan 144 örneklik kısmın MFCC değerleri de ayrıca hesaplanmıştır ve (1)'de tavan fonksiyonunun kullanılma sebebi budur. Dolayısıyla geliştirilen CNN modeline verilen matrislerin boyutu, 489 adet zaman aralığı ve her bir aralık için 16 adet MFCC değeri elde edilmesinden dolayı (489x16) şeklindedir.

Süre belirlendikten sonra her bir kaydı belirlenen süreye sabitlemek üzere kayıtların 1 saniyeden sonraki kısımlarını silen ve sabitlenen kayıtların MFCC değerlerini elde edip sonrasında geliştirilen CNN modeline verilmek üzere bir JSON (JavaScript Object Notation) dosyasına yazan kod Python dili kullanılarak geliştirilmiştir.

#### B. LSTM Yöntemi

Geliştirilen LSTM modelinde 2 adet birbirini takip eden LSTM katmanı ve 1 adet yoğun katmanı takip eden çıktı katmanı bulunmaktadır. LSTM yönteminin en büyük avantajı girdi boyutlarının sabit olmasına gerek olmamasıdır. Sabit girdi boyutu gerektirmediği için herhangi uzunlukta verilen kaydın sınıflandırmasını yapabilmektedir. Böylece, yarasa çağrılarının sürdüğü uzunluğa bakılmaksızın geliştirilen LSTM modeliyle bu çağrılar sınıflandırılabilir. Geliştirilen LSTM modelinin girdi matrisi boyutları (None x 16) şeklindedir. Buna göre kayıtlardan gelen 512 örnek uzunluğundaki zaman aralığı sayısı herhangi bir sayı olabilir, ancak her bir zaman aralığından 16 adet MFCC değeri elde edilmesi durumu CNN de olduğu gibi LSTM modelinde de geçerlidir.

MFCC değerlerinin yazıldığı JSON dosyasının boyutunu küçük tutmak ve aynı zamanda yeterince fazla veri sağlayabilmek üzere kayıt uzunlukları bar grafikleri kullanılarak incelenmiştir. İncelemeler

sonucunda 2,5 saniyeden uzun tüm kayıtların silinmesi halinde her bir davranıştan CNN yönteminde kullanılan kayıt sayısına çok yakın miktarda veri kaldığı gözlemlenmiştir. Bu kısımda da elde kalan 2,5 saniyeden kısa değişken uzunluktaki kayıtların MFCC değerleri elde edilerek ilgili JSON dosyasına yazılmıştır ve bu bilgiler daha sonrasında geliştirilen LSTM modeline girdi olarak verilmiştir.

### III. ÇALIŞMA SONUÇLARI

Tablo I' de yer alan bu çalışmaya ait kombinasyonlar en az 3 kez veri kümesindeki [14] kayıtlar karıştırılarak eğitilmiştir ve gelen doğruluk değerlerinin ortalamaları tabloya kaydedilmiştir. Bu çalışmada elde edilen sonuçlar incelendiğinde Tablo I 'deki özelliklerle 1, 2 ve 10 ile numaralandırılan kombinasyonların doğruluk değerlerinde önemli artış gözlemlenmektedir. 1 numaralı kombinasyon için artış miktarı %5,4; 2 numaralı kombinasyon için artış miktarı %6,8 ve 10 numaralı kombinasyon için artış miktarı %5 kadardır. Yazarlar, MFCC değerlerini kullanarak ve [4]'te kullanılan 1.5 saniyelik kayıt süresi yerine daha kısa süren 1 saniyelik kayıt süresi kullanarak hemen hemen [4]'teki doğruluk değerlerine yakın değerler elde etmiş ve özellikle yukarıda bahsedilen kombinasyonların doğruluk değerlerinde önemli artışlar sağlamıştır. [4]'te yazarlar, çalışma sırasında eğitim süresini kısaltmak ve daha az veri kullanmak üzere, Mel-spektrogram fotoğraflarını kullandıkları için bu fotoğraflar tekrar girdi matrislerine çevrildiğinde bilgi kaybı ortaya çıkmaktadır. Halbuki bu çalışmada MFCC değerleri doğrudan derin öğrenme yöntemlerine verildiği için fotoğraflamadan doğan bilgi kaybı ortadan kalkmaktadır.

Tablo I' de en sağ sütun incelendiğinde, elde edilen LSTM doğruluk değerlerinin [4]'te verilen değerlere göre daha düşük geldiği söylenebilir. Ancak değişken uzunlukta girdi verilebildiği göz önüne alındığında ve doğruluk değerleri arasındaki farklar incelendiğinde, LSTM modelinde elde edilen doğruluk değerlerinin [4]'te elde edilen doğruluk değerlerine çok yakın geldiği ve LSTM modellerinin kullanılması durumunda [4]'teki modellere çok yakın performans sergileyeceği söylenebilir. Mısır meyve yarasaları gündüzleri uyuyup geceleri avlanmaktadır [15]. 2 numaralı kombinasyondaki iyileştirme sonucunda koloninin uykuda çıkardığı seslerin tespiti ile uyku düzeni hakkında önceki çalışmaya göre daha net bilgi edinilebilir, bu sayede sürüyü olumlu ya da olumsuz etkileyen etmenlerin zamana bağlı tespitinin önü açılabilir. Buna örnek olarak gündüz yapılan inşaat, tünel ve maden çalışmalarının koloninin uyku düzenini bozup bozmadığı anlaşılabilir. Isırma ve çiftleşme isteğine tepki seslerinin birbirinden ayırt edilmesi (10 numaralı kombinasyon), sürünün stres etmenlerinin bulunmasına ve stresin doğurabileceği sonuçların öngörülmesine olanak sağlamaktadır. Örneğin; yavru üretimine olan direncin fazla olması, sürünün barındığı yeri güvenli görmemesinden kaynaklanabilir. Bu sebeple sonraki dönemde sürü popülasyonunun artmayacağı veya azalacağı öngörülebilir.

Mevcut tüm sınıfların düzgün olarak belirlenmesi (1 numaralı kombinasyon), sürü içindeki bireylerin davranışları hakkında genellenmiş bir tespit üretmeye yardımcı olmaktadır. Tablo II' de tüm sınıfları içeren sınıflandırma modelinin sonuçları verilmiştir. Parantez içindeki değerler bir önceki [4] çalışmasında elde edilen değerlerdir. Önceki çalışma sonuçlarının da gösterdiği üzere [4], en iyi ayırt edilebilen çağrılar yavru yarasalara (izolasyon) aittir ve izolasyon seslerinin ayırt edilmesi kolonideki yavru sayısının arttığını, bu da yarasaların bulundukları yaşam şartlarının uygun olduğunu gösterebilir. Benzer şekilde beslenme seslerinin artması yarasaların beslenebildiğini ve yaşam koşullarının iyi olduğunu gösterebilir. Bu modelin gerçek bir sistemde kullanılması durumunda gelen tüm çağrılar birbirinden ayırt edebilmesi hayvanlarla ilgili alınacak kararlar açısından önemlidir.

Bu çalışmanın sonucunda ısırma, beslenme ve çiftleşme isteğine tepki seslerinin tespitinde model kesinliği ile F1 skorunun arttığı görülmüştür. Buna karşılık, uykudaki seslerin duyarlılığı ve F1 skorunda artış görülsede tespit için yeterli düzeye erişememektedir. Bu sebeple, uyuma ve uyku dışı ses tespitindeki başarımları ayrı olarak incelenmiştir. Sonuçlar Tablo III' te listelenmiştir. Çalışma sonucunda, önceki değerlendirme metrikleriyle karşılaştırıldığında önemli bir artış gözlenmektedir. Bu sonuçlar, araştırmacılara koloninin uyudukları bölgeyi yaşama alanı olarak benimseyip benimsemedikleri hakkında bilgiler verebilir, koloni üzerindeki stresi gösterebilir ve bu sorunların oluşmasını engelleyecek önlemler alınmasını sağlayabilir.

#### IV. SONUÇ VE GELECEK ÇALIŞMALAR

Bu çalışmada Mısır meyve yarasası çağrılarının sınıflandırılması için geliştirilen CNN ve LSTM modelleri, daha önce aynı amaçla geliştirilmiş olan [4] başka bir modelle kıyaslanmış ve gelen sonuçlar değerlendirilmiştir. Geliştirilen CNN modeli [4]'te elde edilen bazı sonuçlardan daha yüksek sonuçlar üretirken [4]'teki çalışmaya göre iyileştirmeler sağlamıştır. Geliştirilen LSTM modelinde [4]'e kıyasla daha düşük doğruluk değerleri gözlenmiş olsa bile bu değerler [4]'teki değerlere çok yakın geldiği için başarılı olabilmektedir. Ayrıca bu çalışmada elde edilen MFCC değerleri doğrudan modellere girdi olarak verildiği için [4]'te yapılan mel-spektrogram fotoğrafına çevirmeden kaynaklı veri kaybetme sorununu da ortadan kaldırmıştır. LSTM modellerinin eğitimi sırasında döngü değerinin artmasının bazı LSTM modellerinin doğruluk değerini arttırabileceği gözlemlenmiştir. Sonraki aşamada, daha güçlü cihazlarla çalışması durumunda, LSTM modellerine verilen girdi miktarı ve model eğitimindeki döngü miktarı arttırılarak LSTM modellerinin doğruluk değerlerinin artırılması hedeflenmektedir. Ayrıca gelecek planı olarak veri kümesindeki kayıtların bir mikrodnetleyici üzerinde çalıştırılması planlanmaktadır. Uygun derin öğrenme yapılarının mikrodnetleyici üzerinde çalıştırılarak çağrılarının sınıflandırılması amaçlanmaktadır. İstenilen sonuçların elde edilmesi halinde mikrodnetleyiciler birbirleriyle haberleştirilip sınıflandırma sonuçlarını birbirleriyle paylaşacaklardır. Bu şekilde bu sistem, Mısır meyve yarasasının uzaktan takip edilmesine olanak sağlayacaktır.

TABLO II. BU ÇALIŞMA VE [4]'TEKİ ISIRMA, BESLENME, DÖVÜŞME, İZOLASYON ÇİFTLEŞME İSTEĞİNE TEPKİ VE UYKU SESLERİNİN CNN SONUÇLARI

	Isır.	Bes.	Döv.	İzo.	Ç. İst Tepki	Uyu.
<b>Isırma</b>	273 (154)	5 (10)	11 (3)	1 (1)	9 (19)	41 (10)
<b>Beslenme</b>	82 (50)	245 (128)	4 (4)	2 (3)	4 (10)	3 (4)
<b>Dövüşme</b>	98 (37)	19 (13)	198 (120)	2 (1)	11 (15)	12 (8)
<b>İzolasyon</b>	11 (9)	11 (1)	3 (1)	310 (187)	4 (0)	1 (0)
<b>Ç. İst Tepki</b>	117 (69)	14 (7)	12 (10)	1 (0)	193 (107)	3 (5)
<b>Uyuma</b>	207 (103)	7 (13)	28 (16)	4 (1)	13 (26)	81 (36)
<b>Kesinlik</b>	0.346 (0.365)	0.814 (0.744)	0.773 (0.779)	0.969 (0.969)	0.825 (0.605)	0.575 (0.571)
<b>Duyarlılık</b>	0.803 (0.781)	0.721 (0.643)	0.582 (0.619)	0.912 (0.944)	0.568 (0.540)	0.238 (0.185)
<b>F1 Skoru</b>	0.592 (0.498)	0.765 (0.690)	0.664 (0.690)	0.940 (0.957)	0.673 (0.571)	0.337 (0.279)
<b>Yitim</b>	0.878 (1.00)					
<b>Doğruluk</b>	0.674 (0.620)					

TABLO III. BU ÇALIŞMA VE [4]'TEKİ UYUMA VE UYUMAMA SESLERİNİN CNN SONUÇLARI

	Uyumama	Uyuma
<b>Uyumama</b>	905 (688)	295 (298)
<b>Uyuma</b>	188 (224)	1012 (752)
<b>Kesinlik</b>	0.827 (0.754)	0.774 (0.716)
<b>Duyarlılık</b>	0.754 (0.698)	0.843 (0.770)
<b>F1 Skoru</b>	0.789 (0.725)	0.807 (0.742)
<b>Yitim</b>	0.459 (0.503)	
<b>Doğruluk</b>	0.802 (0.734)	

#### KAYNAKLAR

- [1] M. Geda, M. Balakrishnan, "Ecological and Economic Importance of Bats (Order Chiroptera)", ISRN Biodiversity, 1-9. 10.1155/2013/187415, 2013.
- [2] Ş. Bulut, Ş. Özkurt, "Mammals of Turkey Türkiye Memelileri", Panama Yayıncılık, 978-605-7739-57-5, 2021.
- [3] P. Benda, K. Faizolahi, M. Andreas, J. Obuch, A. Reiter, M. Sevcik, M. Uhrin, P. Vallo, S. Ashrafi, "Bats (Mammalia: Chiroptera) of the Eastern Mediterranean and Middle East. Part 10. Bat fauna of Iran." Acta Soc. Zool. Bohem. 76: 163–582, 2012.
- [4] B. Yılmaz, M. Sen, E. Masazade, and V. Beskardes. "Behavior Classification of Egyptian Fruit Bat (Rousettus aegyptiacus) From Calls With Deep Learning." In Handbook of Research on New Investigations in Artificial Life, AI, and Machine Learning. edited by Habib, Maki K. 60-98. Hershey, PA: IGI Global, 2022.
- [5] K. J. Piczak, "Environmental sound classification with convolutional neural networks," 2015 IEEE 25th International Workshop on Machine Learning for Signal Processing (MLSP), pp. 1-6, doi: 10.1109/MLSP.2015.7324337, 2015.
- [6] J. Salamon and J. P. Bello, "Unsupervised feature learning for urban sound classification," 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 2015, pp. 171-175, doi: 10.1109/ICASSP.2015.7177954.
- [7] J. Lu, R. Ma, G. Liu and Z. Qin, "Deep Convolutional Neural Network with Transfer Learning for Environmental Sound Classification," 2021 International Conference on Computer, Control and Robotics (ICCCR), pp. 242-245, doi: 10.1109/ICCCR49711.2021.9349393, 2021.
- [8] A. Graves, S. Fernández, en J. Schmidhuber, "Bidirectional LSTM Networks for Improved Phoneme Classification and Recognition", in Artificial Neural Networks: Formal Models and Their Applications, 2005.
- [9] E. Şaşmaz and F.B. Tek., "Animal sound classification using A convolutional neural network", Paper presented at the 2018 3rd International Conference on Computer Science and Engineering (UBMK), 625-629. doi:10.1109/UBMK.2018.8566449
- [10] J. Xie, K. Hu, M. Zhu, J. Yu and Q. Zhu, "Investigation of Different CNN-Based Models for Improved Bird Sound Classification", in IEEE Access, vol. 7, pp. 175353-175361, doi: 10.1109/ACCESS.2019.2957572, 2017.
- [11] K. J. Piczak. "The details that matter: Frequency resolution of spectrograms in acoustic scene classification", In Proceedings of the Detection and Classification of Acoustic Scenes and Events 2017 Workshop, Munich, Germany, 2017
- [12] H. Liu, C. Liu, T. Zhao and Y. Liu, "Bird Song Classification Based on Improved Bi-LSTM-DenseNet Network", 4th International Conference on Robotics, Control and Automation Engineering (RCAE), pp. 152-155, doi: 10.1109/RCAE53607.2021.9638962, 2021.
- [13] M. A. Tabak, K. L. Murray, A. M. Reed, J. A. Lombardi, K. J. Bay. Automated classification of bat echolocation call recordings with artificial intelligence. Ecological Informatics Volume 68, May 2022
- [14] Y. Prat, M. Taub, E. Prattand Y. Yovel, "An annotated dataset of Egyptian fruit bat vocalizations across varying contexts and during vocal ontogeny". Scientific data, (4),1, pp.1-7, Nature Publishing Group , 2017
- [15] G. Kwiecinski, T. Griffiths, "Rousettus aegyptiacus. Mammalian Species", 611: 1-9, 1999.