



## Dog Behavior Recognition and Tracking based on Faster R-CNN

Emre Dandil<sup>1\*</sup>, Rukiye Polattimur<sup>2</sup>

<sup>1</sup>Department of Computer Engineering, Bilecik Seyh Edebali University, Bilecik, 11210, Turkey

<sup>2</sup>Open and Distance Learning Center, Bilecik Seyh Edebali University, Bilecik, 11210, Turkey

### Highlights:

- A novel video dataset is created for the recognition of dog behavior
- Dog behaviors are recognized with high accuracy using faster R-CNN
- Recognized dog behaviors are tracked on video

### Keywords:

- Computer Vision
- Deep Learning
- Faster Regional-Convolutional Neural Networks Animal Behavior Analysis
- Dog Behavior Recognition

### Article Info:

Research Article

Received: 19.03.2019

Accepted: 28.07.2019

### DOI:

10.17341/gazimmfd.541677

### Acknowledgement:

This study was supported by Bilecik Seyh Edebali University BAPK with Project No: 2017-01.BŞEÜ.03-09.

### Correspondence:

Author: Emre Dandil

e-mail:

emre.dandil@bilecik.edu.tr

phone: +90 228 214 16 13

### Graphical/Tabular Abstract

In this study, a deep learning based system is proposed for the detection and classification of dog's behaviour. In the study, firstly, a dataset is created by collecting videos containing the behavior of dogs which don't avoid contact with people. After the necessary analysis on the obtained videos, a customized dataset consisting of more meaningful sections is developed by extracting determined behaviors in videos. It is recognized the behavior with the Faster R-CNN (Faster Regional-Convolutional Neural Networks) by selecting key frames from these customized video sections. In the last stage, the related behaviors in videos are followed by video tracker after the behavior of the dog is recognized.

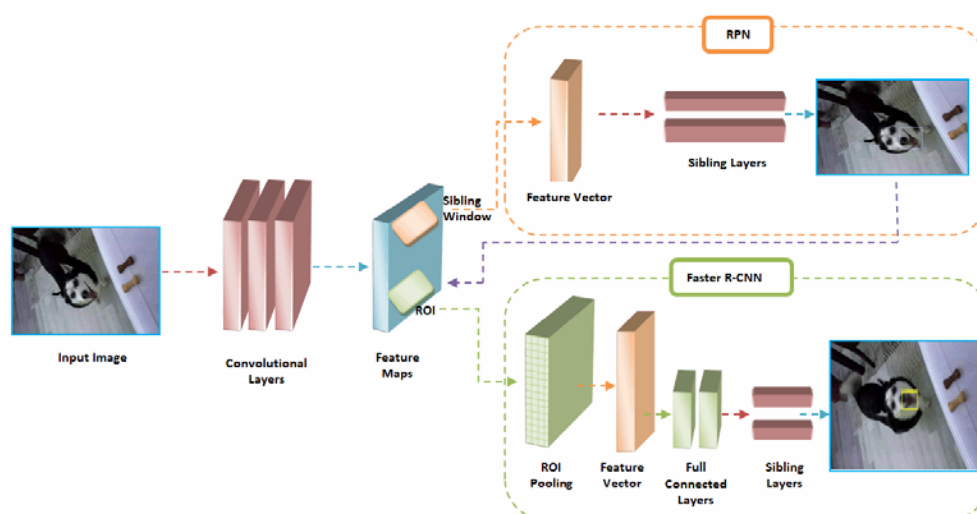


Figure A. The architecture of the proposed faster R-CNN model for the recognition of dog behavior

**Purpose:** In this study, it is aimed to develop a holistic system consisting of stages such as literature review, provision of tools and equipment, collection of data, obtaining key frames from video data, feature extraction from video frames, recognition of dog behaviors and dog tracking on video for the analysis and classification of dog behaviors.

### Theory and Methods:

The proposed method uses faster R-CNN, which is very successful in regional object recognition on video. The behavior of dogs is recognized using faster R-CNN after local region feature propositions are obtained in the system.

### Results:

In experimental studies, the behaviors of dog such as opening the mouth, sticking out the tongue, sniffing, rearing the ear, swinging the tail and playing are examined and accuracy rates 94.00%, 98.00%, 99.33%, 99.33%, 98.00% and 98.67% are obtained for these behaviors, respectively.

### Conclusion:

With the results obtained in the study, it is seen that our proposed method based on key frame selection and determination of regions of interest are successful in recognition the behavior of dogs.



## Daha hızlı bölgesel evrişimsel sinir ağları ile köpek davranışlarının tanınması ve takibi

Emre Dandil<sup>1\*</sup>, Rukiye Polattimur<sup>2</sup>

<sup>1</sup>Bilgisayar Mühendisliği Bölümü, Bilecik Şeyh Edebali Üniversitesi, Bilecik, 11210, Türkiye

<sup>2</sup>Açık ve Uzaktan Öğrenme Merkezi, Bilecik Şeyh Edebali Üniversitesi, Bilecik, 11210, Türkiye

### Ö N E Ç İ K A N L A R

- Köpek davranışlarının tanınması için yeni bir video veriseti üretilmiştir
- Daha Hızlı Bölgesel-Evrişimsel Sinir Ağları ile yüksek başarımda köpek davranışları tanınmıştır
- Tanınan köpek davranışları video üzerinde takip edilmiştir

#### Makale Bilgileri

Araştırma Makalesi

Geliş: 19.03.2019

Kabul: 28.07.2019

#### DOI:

10.17341/gazimmfd.541677

#### Anahtar Kelimeler:

Bilgisayarlı görü,  
derin öğrenme,  
daha hızlı bölgesel-  
evrişimsel sinir ağları,  
hayvan davranış analizi,  
köpek davranış tanıma

#### ÖZET

Hayvan yüzlerinin, vücut duruşlarının, davranışlarının ve fiziksel hareketlerinin tespiti ve tanınması son zamanlarda disiplinlerarası bir alan olarak ön plana çıkmıştır. Bilgisayarlı görü yöntemi ile hayvanların davranışlarının tespitine, sonraki davranışların öngörülmesine ve hayvanların evcilleştirilmesine katkı sunabilir. Bu çalışmada, köpeklerin davranışlarının tespit edilmesi ve sınıflandırılması için derin öğrenmeye dayalı bir sistem önerilmiştir. Çalışmada öncelikle, insanlar ile temastan kaçınmayan köpeklerin davranışlarını içeren videolar toplanarak bir veri seti oluşturulmuştur. Elde edilen videolar üzerinde gerekli analizler yapıldıktan sonra belirlenen davranışlar videolardan çıkarılarak, daha anlamlı bölümlerden oluşan özelleştirilmiş bir veri seti geliştirilmiştir. Bu anlamlı video bölümlerinden anahtar çerçeveler seçilerek Daha Hızlı Bölgesel-Evrişimsel Sinir Ağları (DH B-ESA) ile davranışlar tanınmıştır. Son aşamada ise, köpeğin davranışı tanındıktan sonra, video üzerinde ilgili davranışlar takipçi ile izlenmiştir. Yapılan deneysel çalışmalar sonucunda, köpeklerin ağız açma, dil çıkarma, koklama, kulak dikme, kuyruk sallama ve oyun oynama davranışları incelenmiş ve bu davranışlar için sırasıyla %94,00, %98,00, %99,33, %99,33, %98,00, %98,67 doğruluk oranı elde edilmiştir. Çalışmada elde edilen sonuçlar ile anahtar çerçeve seçimi ve ilgi bölgelerin belirlenmesine dayalı önerilen yöntemin, köpeklerin davranışlarını tanımda başarılı olduğu görülmüştür.

## Dog behavior recognition and tracking based on faster R-CNN

### H I G H L I G H T S

- A novel video dataset is produced for the recognition of dog behavior
- Dog behaviors are recognized high accuracy using faster R-CNN
- Recognized dog behaviors are tracked on video

#### Article Info

Research Article

Received: 19.03.2019

Accepted: 28.07.2019

#### DOI:

10.17341/gazimmfd.541677

#### Keywords:

Computer vision,  
deep learning,  
faster regional-convolutional  
neural networks animal  
behavior analysis,  
dog behavior recognition

#### ABSTRACT

Recently, detection and recognition of animal faces, body postures, behaviors, and physical movements is became an interdisciplinary field. Computer vision methods can contribute to determine behaviors of animals and predict the following behavior of animals. Moreover, these methods would contribute to domesticate animals. In this study, a deep learning based system is proposed for the detection and classification of dog's behaviour. In the study, firstly, a dataset is created by collecting videos containing the behavior of dogs which don't avoid contact with people. After the necessary analysis on the obtained videos, a customized data set consisting of more meaningful sections is developed by extracting determined behaviors in videos. It is recognized the behavior with the Faster R-CNN (Faster Regional-Convolutional Neural Networks) by selecting key frames from these customized video sections. In the last stage, the related behaviors in videos are followed by video tracker after the behavior of the dog is recognized. As a result of experimental studies, the behaviors of dog such as opening the mouth, sticking out the tongue, sniffing, rearing the ear, swinging the tail and playing are examined and accuracy rates 94.00%, 98.00%, 99.33%, 99.33%, 98.00% and 98.67% are obtained for these behaviors, respectively. With the results obtained in the study, it is seen that our proposed method based on key frame selection and determination of regions of interest are successful in recognition the behavior of dogs.

\*Sorumlu Yazar/Corresponding Author: emre.dandil@bilecik.edu.tr, rukiyeozubek@gmail.com / Tel: +90 228 214 16 13

## 1. GİRİŞ (INTRODUCTION)

İnsanların birçok alanda yardımcısı olan köpekler, sahip oldukları yetenekleri eğitimlerle daha da özelleştirerek, uzun yıllardan bu yana, birçok alanda işlerin kolaylaştırılmasına katkı sağlamaktadırlar. Köpekler kendilerine özgü karşılaşılan bazı problemlere hızlı ve kolay çözüm bulabilmektedirler. Örneğin köpekler günlük hayatta polis köpekleri, bekçi köpekleri, duyma-görme engelli kişiler için özel eğitilmiş köpekler, enkaz arama köpeği, kızak köpekleri, mineral arama köpekleri, mantar bulma köpekleri olarak birçok özel alanda karşımıza çıkmaktadır [1]. Hayatımızda bu kadar var olan köpeklerle ilgili olarak bazı durumlarda onların psikolojilerinin araştırılması ve davranışlarının analiz edilerek tanınması gerekebilir. Dahası, sahiplendirildiklerinde değişen davranış yapılarının incelendiği [2], gösterdikleri ruhsal veya davranış bozukluklarının tespit edilerek hangi eğitimlerin verilmesi gerektiği gibi daha çok köpeklerin davranışlarının incelenmesine [3] yönelik araştırmalar da yapmak gerekmektedir.

Genel olarak köpeklerin ne tip davranış sergilediğini eğitmenler, veterinerler veya sahipleri geleneksel yöntemlerle anlamaya çalışmaktadırlar [4]. Ancak böyle durumlarda bazen köpeklerin davranışları tam olarak kestirilemediğinden ortaya istenmeyen sonuçlar çıkabilmektedir. Bu gibi durumları ortadan kaldırmak için günümüzde video üzerinden alınan görüntüler ile köpekler için yakın mesafe davranış analizi ve hareket tanıma işlemleri yapılabilmektedir [5]. Bunun yanında eğitilecek köpeklerde görme, görme perspektifi, görme keskinliği, form algısı ve renk görüşü gibi çeşitli görme araştırmalarının yapılması, uzman köpeklerin görevlerdeki performansını nasıl etkileyebileceğinin belirlenmesi, belirtilen sorunların ortaya çıkmasını engelleyecektir [6]. Nitekim hayvanların davranışlarını daha iyi anlamak, onların sağlık sorunlarına hızlı çözümler bulmak, psikolojik olarak kendilerini daha güvende hissedecekleri ortamlar oluşturmak için bu tip çalışmaların geliştirilmesi, köpeklerin sağladığı hizmetlerin kalitesini arttıracak ve onlarla daha uyumlu yaşamımıza olanak verecektir [7].

Köpekler doğal ortamlarında yaşarken tıpkı insanlar gibi belli hareketleri içeren günlük aktivitelerini yapmaktadırlar. Köpeklerin yaptıkları bazı hareketlerin günlük hayatta insanlar tarafından anlaşılması oldukça zor ve masraf gerektiren işlemler olabilmektedir. Bazen insani duygular ve hızlı karar verme söz konusu olduğunda, insanların bu hareketleri anlaması ve tanımlaması oldukça zor olmaktadır. Bu problemi ortadan kaldırmak için otomatik karar verebilen bilgisayarlı görü destekli makine öğrenimi yaklaşımlarından yararlanılabilmektedir [8]. Bu yaklaşımlar sayesinde, günlük hayatımızda hareketlerinde farklılık gözlenen köpekler için mevcut sorunlarına en iyi yaklaşımı getirerek çözüm bulunması sağlanmakta ve onların eğitim süreçlerine yön verilebilmektedir. Bunlara ek olarak köpeklerde davranış örüntülerinin tanınması amacıyla yeni yaklaşımlar da

önerilmektedir. Ladha vd. [8] çalışmalarında, giydirilebilir sürekli kayıt yapan üç eksenli ivmeölçer yardımıyla köpeklerden veriler toplamışlardır. Çalışmalarında 18 köpeğin 17 hareketi (yürüme, koşu, atlama, kazma, uzanma, oturma, havlama, çiğneme, yemek yeme, içme, esneme, koklama, sarsılma, idrar yapma, tuvalet yapma, silkelenme) için Temel Bileşen Analizi (TBA) tabanlı özellik çıkarımı ve örnek tabanlı bir öğrenme sınıflandırması metodu önermişlerdir. Benzer olarak bir diğer çalışmada ise Leos-Barajas vd.[9] bir köpeğe giydirilmiş ivmeölçer ile hayvan davranışlarının analizi çalışmasını Saklı Markov Modeli (SMM) kullanarak gerçekleştirmiştir. Denetimli ve denetimsiz öğrenme metodlarını, köpek davranışlarının durum tahmini ve köpeklerin biyolojik olarak daha anlamlı hareketlerin sınıflandırılması için kullanmışlardır. Gerencser vd. [10] ise, köpeğe giydirilmiş ivmeölçer ve kamera ile alınan verilerin değerlendirildiği farklı bir çalışmada ise takip edilecek hareketleri belirleyerek (durma, yürüme, koşma, vb.), Destek Vektör Makineleri (DVM) öğrenme algoritması kullanarak sınıflandırma işlemi yapmışlardır. Brugarolas vd. [11] ise makine öğrenmesi yöntemlerinden karar ağaçları ve SMM kullanarak kablosuz makine arayüzü ile sensörlerden aldıkları titreşim verileri yardımıyla köpeklerin davranışlarının tanınmasını önermişlerdir. Bu hareketleri statik ve dinamik olarak ikiye ayırmışlardır. Statik hareketler için; oturmak, ayakta durmak, uzanmak, iki ayak üzerinde durmak ve beslenmek, dinamik hareketler için ise yürümek, merdiven çıkmak ve bir rampadan aşağı inmek gibi hareketleri tanımlamışlardır.

Hayvanların davranışlarının tanınmasında bilgisayarla görme uygulamaları, hayvanlarla doğrudan etkileşime girme ihtiyacı olmadan ilgili özelliklerinin ölçülmesini sağladığı için cazip hale gelmiştir. Video üzerinden gerçek zamanlı olarak köpeklerin davranışlarının incelenmesinde bilgisayar görüşü ve makine öğrenmesi alanlarındaki gelişmeler, bu alanda güçlü sistemlerin geliştirilmesine olanaklar sağlamaktadır. Ayrıca son yıllarda videolar üzerinden nesne tanıma amaçlı derin öğrenme yöntemlerinden yararlanma kayda değer bir şekilde artış göstermektedir [12]. Derin öğrenme sistem çalışmalarında önemli başarı oranları elde edilmesi ve bilgisayar donanım yapılarının (GPU) gelişmesiyle birlikte derin öğrenme temelli yaklaşımların davranış analizinde sıklıkla kullanıldığı görülmektedir [13-19]. Köpeklerin davranışlarının analizinde bilgisayarlı görü temelli yaklaşımlarda, kameralar ile alınan video görüntüleri bilgisayar algoritmaları ile incelenerek davranış analizi işlemleri gerçekleştirilmektedir. Son yıllarda birçok alanda kullanımı yaygın hale gelen düşük fiyatlı ve yüksek kaliteli giyilebilir kameralar davranış tanıma uygulamalarında videoların elde edilmesinde sıklıkla kullanılmaktadır [20-23]. Iwashita vd. [24] çalışmalarında egosantrik (kafa bölgesine giydirilmiş) kameralar yardımıyla farklı köpeklerin video görüntüleri üzerinde 10 farklı davranışın (topla oynama, arabanın geçmesini bekleme, su içme, beslenme, kafasını sola doğru çevirme, başını sağa çevirme, sevgi gösterisinde bulunma, silkelenme, koklamak ve yürümek) analizini gerçekleştirmişlerdir. Dodge ve Karam

[25] yine farklı tür köpeklerden elde edilmiş veriseti üzerinde yaptıkları çalışmada, çözünürlüğü iyi olan resimlerde makine ve insan neredeyse aynı başarı oranına sahipken bozuk görüntüler tanımda insanların daha başarılı olduğu göstermiştir. Ek olarak başarıyı artırmak için Evrişimsel Sinir Ağları (ESA) tabanlı bir model önerilmiştir. Diğer bir çalışma da, Ehsani vd. [26], ESA ve Uzun Kısa Hafıza Ağları (*Long Short Term Memory*) ile karma bir model oluşturarak köpeklerin kafa bölgesine giydirilen kamera ile alınan görüntüleri alarak köpeğin mevcut duruma göre nasıl hareket edeceğinin öngörüsünü yapmışlardır. Son zamanlarda, nesne tanıma problemlerinde manuel özellik çıkarımına dayalı geleneksel yöntemlerin düşük doğruluk oranına sahip olmasından dolayı derin öğrenme destekli uygulamalar tercih edilmektedir [27-29]. Derin öğrenme yöntemlerinden birisi olan ESA, özellik seçim işlemini otomatik yaparak görüntüden nesne tanıma uygulamalarında yaygın olarak kullanılmaktadır. Durağan görüntüler üzerinde nesne tanınmasında ESA tabanlı algoritmalar etkili olsa da, Bölgesel-Evrişimsel Sinir Ağları (B-ESA) gibi altyapılar bölgesel nesne tanıma öncelikli olduğundan videolar için daha fazla tercih edilmektedir[30]. Bölge öneri yöntemi olarak bilinen B-ESA nesne tanıma problemlerinde sıklıkla kullanılmaktadır. Bu yöntemde binlerce yerel aday bölge çıkarıldıktan sonra bu bölgeler ESA ile birleştirilir. Bölgesel karakteristikler çıkarıldıktan sonra, farklı bölgeler hedef bölge veya arkaplan olarak sınıflandırılır[31]. Ancak bazı durumlarda B-ESA'da birbirine benzer çok sayıda hesaplama oluşabilmektedir. Buna benzer sorunları ortadan kaldırmak için Hızlı B-ESA kullanılır. DH B-ESA ise Hızlı B-ESA ve Bölge Öneri Ağının (BÖA) birleşiminden oluşmaktadır. DH B-ESA metodunda, bölge önerileri durağan görüntülerden BÖA ile elde edilerek, nesnelere bu önerilere Hızlı B-ESA tarafından konumlandırılır[15, 31, 32]. Son zamanlarda, DH B-ESA'nın hayvanların davranışlarının tespitinde yaygın bir şekilde kullanıldığı görülmektedir[14, 15, 33]. Buradan yola çıkarak, DH B-ESA hayvanların davranışlarının tanınmasında etkili bir şekilde kullanılabilir.

Bu çalışmanın katkısı ise, köpeklerin sıklıkla sergiledikleri davranışlarının yüksek doğruluk ve etkinlikle tanınması için video dizileri kullanarak bütüncül bir bilgisayarlı görüş yönteminin önerilmesidir. Çalışmada öncelikle köpeklerin günlük hayatlarında sıklıkla sergiledikleri hareketlerden olan ağız açma, dil çıkarma, koklama, kulak dikme, kuyruk sallama ve oyun oynama olmak üzere altı farklı hareket temel alınarak köpeklerin doğal ortamlarında videoları toplanmıştır. Daha sonra, önerilen yöntem ile akan videodan anahtar çerçeveler çıkarılmıştır. Önerilen yöntemde video üzerinde bölgesel nesne tanıma işlemlerinde oldukça başarılı olan DH B-ESA kullanılmıştır. Burada dedektör ile yerel bölge çıkarımları elde edildikten sonra köpeklerin davranışları DH B-ESA ile tanınmıştır. Böylece hem köpeklerin belirtilen davranışları yüksek başarımla tanınmış hem de akan video üzerinde köpeklerin davranışlarının izlenmesinde iyi bir performans elde edilmiştir. Çalışmanın geri kalan bölümleri şu şekilde organize edilmiştir. İkinci bölümde Materyal ve Metot kısmı

ile veriseti ve ESA'nın mimari yapısı detaylandırılmıştır. Üçüncü bölümde DH B-ESA ile köpek davranışlarının tanınmasının süreçleri sunulmuştur. Dördüncü bölümde deneysel çalışmalar ve son bölümde ise elde edilen sonuçlardan bahsedilmiştir.



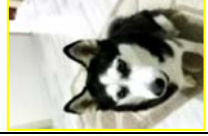

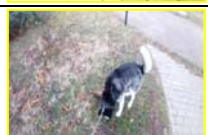

## 2. MATERYAL VE METOT (MATERIAL AND METHOD)

### 2.1. Köpek Video Veri Seti Oluşturma (Construction of Dog Video Dataset)

En bilinen haliyle aktivite tanıma işlemi; veri toplama, görüntü ön işleme ve bölütleme, özellik çıkarma ve belirleme, eğitim ve sınıflandırma aşamalarını içermektedir [34]. Köpeklerin duygusal durumlarını ve niyetlerini çevrelerindeki diğer kişilere iletmelerini sağlayan kendilerine has bir vücut dilleri vardır. Köpekler çoğunlukla kendilerini sesli olarak (havlayarak) anlatmaya çalışsa da yüz ifadelerini, kendilerine has vücut dillerini ve duruşlarını, anlık ruhsal ve fiziksel durumlarını ifade etmek için kullanırlar. Bu hareketler köpeğin davranışını anlamak, davranışı hakkında fikir sahibi olmak, endişeli veya sinirli olduğunun anlaşılması ya da bir şeyler yemeğe ihtiyacı olduğu zamanları anlayabilmek gibi birçok anlamlı bilgi içerirler. Bu bilgileri anlamak için köpeğin yüz mimiklerine ve vücut diline bakmak gereklidir [35]. Bu çalışmada köpeklerin günlük hayatlarında sıklıkla sergiledikleri hareketlerden olan ağız açma, dil çıkarma, koklama, kulak dikme, kuyruk sallama ve oyun oynama olmak üzere 6 farklı hareket temel alınarak köpeklerin doğal ortamlarında videoları toplanmıştır. Önerilen çalışmada kullanılacak davranış biçimlerinin köpeğin hangi duygusuna karşılık geldiği Tablo 1'de sunulmuştur.

Çalışma süresince farklı zaman dilimlerinde, kamera yardımı ile insanlar ile temastan kaçınmayan iki farklı köpek üzerinde görüntü kayıt işlemleri yapılmıştır. Görüntü alınan köpekler sağlıklı kilolarında ve yaklaşık 2 yaşlarındadır. Bu görüntüler iki farklı kamera kullanılarak video verileri toplanmıştır. Bu kameranın birincisi insan tarafından kullanılarak direkt elde edilen veriler iken diğeri köpeğe giydirilmiş bir aktivite kamerası yardımıyla elde edilmiş görüntülerdir. Teknik ekipman temin edildikten sonra köpekler belirli süre izlenmiştir ve daha sonra altı farklı davranışın (ağız açma, dil çıkarma, koklama, kulak dikme, kuyruk sallama, oyun oynama) üzerinde tanıma ve sınıflandırma yapılmasına karar verilmiştir. Bu davranış biçimleri köpeklerin günlük hayatta çok sık yaptığı davranışlardan olmasına özen gösterilmiştir. Bunun asıl nedeni ise, davranış tanıma ve sınıflandırma işlemlerinin başarısını artmasındaki temel etken verilerin olabildiğince çok olmasına dayanmaktadır. Çalışma için oluşturulan video veriseti 4K-12 MP GoPro Hero 5 giydirilebilir Aksiyon Kamerası ve köpeklerin sahipleri tarafından kullanılabilen 4K-14.2 MP Sony Video Kamera kullanılarak elde edilmiştir. Köpek üzerinden aksiyon kamera ile video toplama ve gerçek zamanlı takip için hazırlanmış olan düzenek Şekil 1'de gösterilmiştir.

**Tablo 1.** Köpeklerin altı farklı davranışı gösterdikleri eylemlerin anlamlandırılması  
(The meaning of the actions of dogs in six different behaviors)

Davranış Tipi	Köpek Vücut Diline Göre Anlamı	Davranış Görüntüsü
Ağız Açma (AA)	Ağız açma hareketini köpek, sinirli olduğunda, kızdığında, yemek yeme isteğini göstermek için, sahibiyile oyun oynama isteğini göstermek için kullanabilir.	
Dil Çıkarma (DC)	Eğlendiğinde, sahibini karşıladığı anda, yemek yeme eyleminden sonra temizlenmek için kullanabilir.	
Kulak Dikme (KD)	Bir uyararı aldığında, tehlike hissettiğinde, karşısındakine kendi kararlığını gösterirken kullanabilir.	
Kuyruk Sallama (KS)	Rahat anında, yürüme esnasında, sahibinden bir beklentisi olduğunda, sevindiğinde bu tip hareketler kullanabilir.	
Koklama (KO)	Arama kurtarma esnasında, yiyecek ararken, yemek istediği yiyeceğin güvenli olup olmadığını anlamak için sık koklama hareketi yapabilir.	
Oyun Oynama (OO)	Sahibiyile eğlenmek istediğinde, rahat olduğu anlarda, mutluluğunu göstermek için bu tip hareketler yapabilir.	

**Şekil 1.** Köpekten aksiyon kamera düzeneği ile video toplama ve takip sistemi  
(Video capture and tracking system with action camera for dog in proposed system)

## 2.2. Evrişimsel Sinir Ağları (Convolutional Neural Networks)

Derin öğrenme, çok büyük adetlerde verilerden özelliklerin çıkarılması için kullanılan ve çoklu katman yapısına sahip olan makine öğrenmesi sistemlerinin alt gruplarında incelenen ve de son zamanlarda bilgisayar bilimcilerinin yoğun olarak çalıştığı popüler bir alandır. Derin öğrenme yöntemlerinde kullanılan en popüler olanlarından biri hiç kuşkusuz ESA'dır. ESA, yapay sinir ağlarının ileriye dönük

işlem yapan ve diğer sinir ağlarından farklı olarak özellik çıkarıcı katmanlarıyla bir hiyerarşik yapıya sahip olan derin öğrenme yaklaşımıdır. ESA'yı diğer makine öğrenme yöntemlerinden ayıran en temel fark ham görüntüyü kullanarak uçtan uca öğrenme sağlayabilmesidir [36, 37]. Diğer makine öğrenmesi yöntemlerinde bir resim sınıflandırma işlemi yapılması için görüntünün içerdiği tüm sayısal değerler ağa aktarılması gerekmektedir, fakat evrişimsel bir ağda ilk katmanda sadece örüntüler belirlenir

ve bu özellikler ağa aktarılır. Bu örüntüler dikey ve yatay kenarlar gibi daha düşük özellikleri içeren bilgilerden oluşmaktadır.

Derin sinir ağların sahip olduğu yüksek kararlılık ve farklı katman sayısı/yapısı kombinasyonlarına sahip olması, sistemin eğitim aşamasında farklı yaklaşımlar getirilmesine yardımcı olmaktadır. Her bir katman yapısı, sistemin eğitimi aşamasında neyi aktifleştirdiğini görselleştirmek için uygun bir prosedür sunmaktadır. Sistem çalışma aşamalarında, gerektiğinde katman yapısını değiştirilebilir veya farklı başarılı iki model kombinasyonlarından yararlanılabilmektedir [26]. Derin öğrenme yapıları içerisinde birçok sınıflandırma modeli kullanılmaktadır. Görüntü tanıma ile ilgili olarak yapılan çalışmalarda ESA, en güçlü sınıflandırma modeli olarak karşımıza çıkmaktadır [38-42]. ESA, bilgisayarlı görü uygulamalarında kullanılmak üzere geliştirilmiş çok katmanlı yapay sinir ağlarının özel bir modelidir. Yapısında konvolüsyon katmanı, havuzlama katmanı, aktivasyon katmanı tam bağlı katman ve softmax gibi kendine özgü görevleri olan ayrı katmanları barındırır. Makineyle öğrenme yıllardır kullanılan bir yöntem olmasına rağmen, büyük veri ve derin öğrenme yöntemleri GPU'ların özelliklerin iyileştirilmesi ile birlikte çok daha fazla çalışma alanı bulmuştur ve bu çalışmalarda daha yüksek başarılar elde edilmeye başlanmıştır. Bu durum, farklı derin öğrenme modelleri tasarlanmasına da olanak sağlamıştır. Bu modellerden en yaygın kullanılanı ve bu çalışmada kullanılan modeli ESA metodudur. Bu metodu kullanarak birçok çalışma gerçekleştirilmiştir [26, 38-42]. Diğer makine öğrenmesi yöntemlerinde, sınıflandırmadan önce dikkatlice hangi özelliklerin kullanılması gerektiği iyi belirlenmeli ve bu özelliklere göre ağı tasarlanması gerekmektedir. Buna karşılık, ESA diğer makine öğrenmesi yöntemlerinde olduğu gibi özelliklerin manuel olarak belirlenmesi yerine, ham görüntülerden özelliklerin çıkarılmasını sağlar. Bu özelliklerin sırasıyla her evrişim katmanından geçtikten sonra kenarlar, köşeler, dokular, nesne parçaları ve sınıflar gibi anahtar noktalarının bilgi seviyesi artırılmaktadır.

Şekil 2'de genel mimarisi gösterilen ESA temel olarak giriş katmanı, konvolüsyon katmanı, örnekleme katmanı ve tam bağlantı katmanlarından oluşmaktadır [43]. ESA'de tam bağlantı katmanlara ek olarak çıkışı üretmesi için giriş katmanına konvolüsyon uygulanır. Bu, giriş katmanındaki bütün bölgelerin bir sonraki katmanda bulunan nöronlara

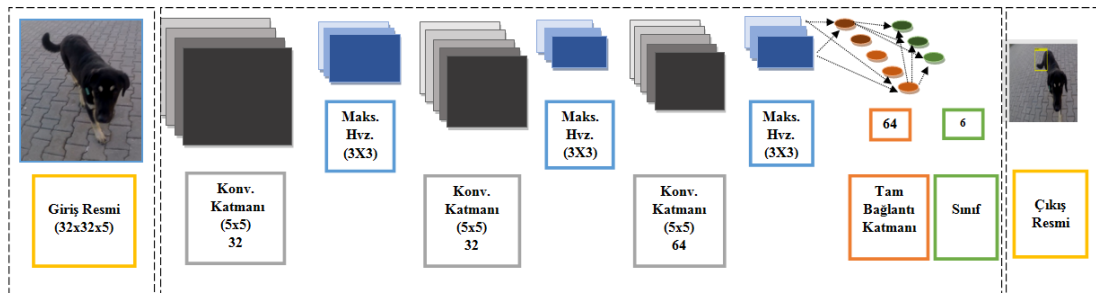
bağlı olduğu lokal bağlantıları oluşturur. Her katmanda farklı filtreler uygulanır ve sonunda bunların sonuçları birleştirilir. Eğitim boyunca ESA, filtrelerin içeriğini (öznitelikleri) gerçekleştirilmesi istenen işleme göre kendisi otomatik olarak öğrenir. Böylece baştan sona doğru her katmanda görüntüye ilişkin basitten karmaşığa öznitelikler öğrenilmiş olur. Filtre banka katmanında, değişik öznitelik çıkarılmak için kullanılan birçok çekirdek bulunmaktadır. Havuzlama katmanında elde edilen her öznitelik haritası ayrı ayrı ele alınarak her bir özellik için komşu değerinin ortalaması veya maksimum değerinin elde edilmesi sağlanmaktadır.

Bir ESA, esas olarak görüntü sınıflandırması için uygulamalarda yer bulurken, B-ESA ise nesne tespiti için tercih edilmektedir[44]. Görüntü sınıflandırma işlemi, genellikle bir görüntüdeki nesnenin tahmini üzerine çalışmaktadır. Nesnenin görüntünün neresinde olduğu ve kapladığı sınırların tespiti ise nesne tanımlama işlemidir. ESA mimarisinde kullanılan bir yapı olmakla birlikte, B-ESA giriş görüntüsüne ek olarak ilgilenilen bölgenin de giriş olarak verildiği bir yapı olarak düşünülebilir, aslında daha özelleştirilmiş ESA özellikli bölgeler denilebilir.

### 3. DH B-ESA İLE KÖPEK DAVRANIŞLARININ TANINMASI (RECOGNITION OF DOG BEHAVIORS USING FASTER R-CNN)

Köpeğin davranışını tespit etmek için ESA derin öğrenme uygulamalarında yaygın olarak kullanılan transfer öğrenme iş akışını kullanır. Transfer öğrenmesi, geniş bir veri seti ile eğitilmiş bir ağ modeli için yeni bir sınıflandırma veya tanıma işlemlerinde takip edilecek nesnenin başlangıç noktası olarak kullanılır. Bu öğrenme modeli ile çok farklı görseller üzerinde eğitim yapamaya olanak sağlayan zengin içeriğe sahip bir ön-eğitilmiş bir ağ elde edebilir. Bu öğrenme, ağa ince ayar yaparak yeni göreve aktarılabilir.

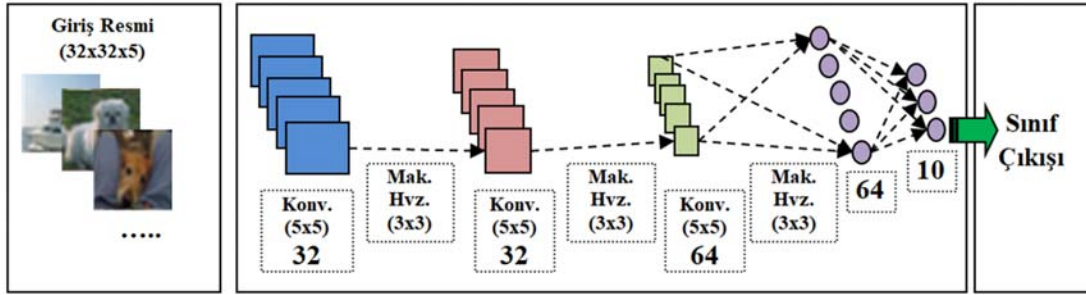
Bu çalışmadaki gibi köpek hareketlerinin algılanması için tüm video çerçevesine değil köpeklerin o hareketi gerçekleştirirken sergiledikleri bölgelere odaklanılmaktadır. Örneğin bir kuyruk sallama hareketi için tüm resme bakmak yerine köpeğin kuyruk bölgesi için ilgili alan tespit edilmektedir. Tam da bu noktada sistemin daha etkin ve hızlı çalışmasını sağlayan B-ESA devreye girmektedir. DH B-ESA, B-ESA'dan geliştirilmiş bir ESA tipidir [31, 32]. DH B-ESA, BÖA ve Hızlı ESA ağlarının birleşiminden oluşmaktadır. DH B-ESA ile BÖA tarafından belirlenen



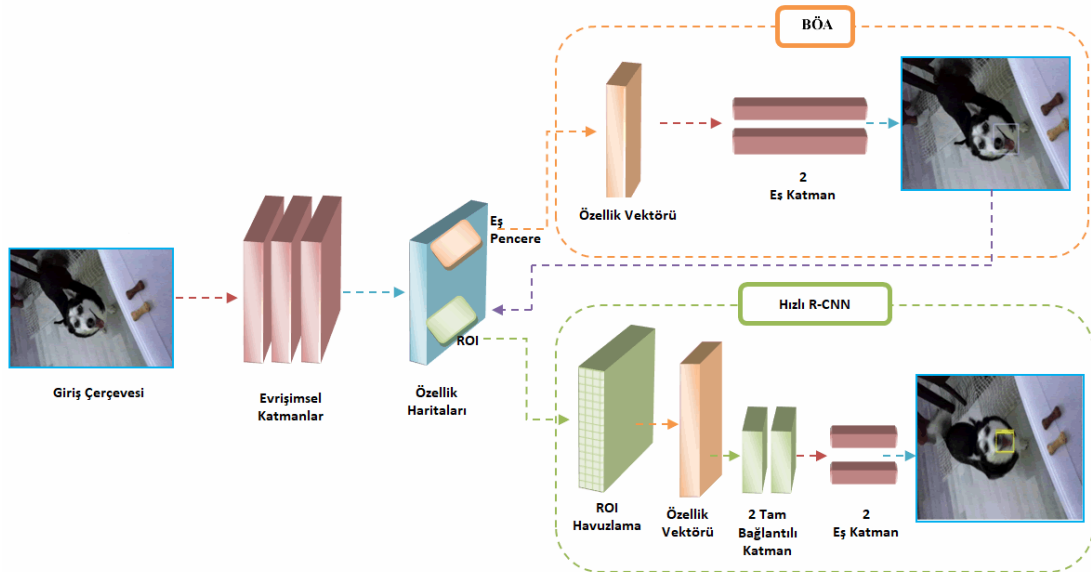
Şekil 2. ESA yapısının genel mimarisi (General architecture of CNN structure)

aday bölgelerden rastgele seçim yapılarak sınıflandırma işlemi gerçekleştirilir. DH B-ESA, ESA ile ortaya çıkan aynı sonuçların elde edilmesi problemini ortadan kaldırdığı için hesaplama süresini azaltmaktadır. Böylece DH B-ESA gerçek zamanlı tanıma işlemlerinde tercih edilmektedir. Köpeklerin davranışlarının karmaşıklığından dolayı, bilgisayar destekli tespit sistemleri tarafından tanınması kolay değildir. Ancak çok sayıda örnek veri ile yüksek hassasiyetli eğitim kümeleri oluşturularak sınıflandırma işlemlerinin başarımı artırılabilir. Bu nedenle ön- eğitilmiş ağlarla transfer öğrenme yöntemi problemi çözmek için kullanılır [45]. Çalışmada ön- eğitilmiş ağ olarak CIFAR-10 [46] veri setindeki görüntüler ile eğitilmiş bir ESA ağı kullanılmıştır. BÖA ve tespit ağlarını oluşturmak için genellikle ön eğitilmiş ağ modelleri kullanılmaktadır. CIFAR-10'daki [46] eğitim verileri (32x32 boyutunda 50.000 görüntü, 10 sınıf) ağ modelinin ön- eğitimi için kullanılmıştır. Ön- eğitim için kullanılan ESA ağının yapısı Şekil 3'te sunulmuştur. DH B-ESA modeline dayalı köpek davranışlarının tanıma işleminin mimarisi Şekil 4'te gösterilmiştir. Görüldüğü gibi, DH B-ESA, BÖA ve Hızlı B-ESA içeren alt iki ağdan oluşmaktadır. BÖA ve Hızlı B-ESA, ESA tarafından çıkarılan aynı girişe sahip özellik

haritalarını kullanırlar. Hedef bölgeleri oluşturmak için BÖA, bu bölgeleri sınıflandırmak için ise DH B-ESA kullanılır. BÖA, giriş özellik haritalarından tespit edilecek bölgeleri oluşturmak için kullanılır. 32 boyutlu özellik haritaları, tüm görüntüyü girdi olarak ön- eğitilmiş konvolüsyon ağı tarafından çıkarılır. DH B-ESA ağı, BÖA tarafından tespit edilen nesne bölgelerini sınıflandırmak için kullanılır. Hızlı B-ESA ağı, tüm görüntüyü ve tanınması amaçlanmış bölge veri seti kümesini bir giriş olarak alır. BÖA ile aynı şekilde, ön- eğitilmiş ağ üzerinde tüm görüntüler işlenir ve 32-boyutlu özelliğe sahip bir harita oluşturulur. Daha sonra, her bir hedef nesne özellik haritalarından ilgili bölge (ROI) işaretlenir. DH B-ESA'da tanıma işlemi dört adımda tamamlanır. DH B-ESA modelinde, öncelikle ön- eğitilmiş ağ ile BÖA başlatılır ve bölge adayları belirlenir. İkinci olarak BÖA ile belirlenen bölge adayları, ayrı bir ağ olan Hızlı B-ESA ile yeniden eğitilir. Üçüncü olarak, BÖA ağının konvolüsyon katmanları tespit ağına aktarılır ve BÖA ağına ince ayar yapılır. Son olarak ise, Hızlı B-ESA için de ince ayar yapılarak, konvolüsyon katmanları tespit ağı ile paylaşılır ve iki ağ birleştirilerek tek bir ağ üzerinde DH B-ESA modeli oluşturulur.



Şekil 3. CIFAR-10 veriseti üzerinde ön- eğitim için kullanılan ESA ağının yapısı  
(CNN structure used for pre-training on the CIFAR-10 dataset)



Şekil 4. Köpek davranışlarının tanıması için önerilen DH B-ESA modelinin mimarisi  
(The architecture of the proposed faster R-CNN model for the recognition of dog behavior)

### 3.1. Video Ön-İşleme (Video pre-processing)

Bu aşamada kameralar ile farklı zamanlarda alınan büyük video parçalarından köpeklerin anlamlı davranışlarının bir uzman eşliğinde videolardan çıkarılması sağlanmaktadır. Bu video özelleştirme işlemi eğitim sırasında, eğitim süresinin kısalmasını sağlamaktadır. Alınan görüntüler üzerinden gerekli incelemeler yapıldıktan sonra, eğitim ve sınıflandırması yapılacak olan davranışlara karar verilmiştir. Elde edilen videolardaki tüm bölümler izlenerek incelenecek olan davranışların sergilendiği bölümler tespit edilmiştir. Bu incelenecek davranışların olduğu kısa bölümler video kırpma programları kullanılarak daha küçük parçalara ayrılmışlardır. Bu sayede tüm altı hareketler için daha anlamlı bölümlerden oluşan özelleştirilmiş veri seti elde edilmiştir. Veri setinde her bir hareket için ortalama 5 saniyelik 30 farklı video parçası bulunmaktadır. Elde edilen videolardaki tüm bölümler izlenerek üzerinde çalışılacak olan davranışların sergilendiği bölümler tespit edilmiştir. Bu incelenecek davranışların olduğu kısa bölümler video kırpma programları kullanılarak daha küçük parçalara ayrılmıştır. Bu sayede tüm altı hareketler için daha anlamlı bölümlerden oluşan özelleştirilmiş veri seti elde edilmiştir. Minimum 3 saniye, maksimum 8 saniye davranış içeriklerinin olduğu kısa videolardan yeterli sayıda çerçeve JPG görüntü formatında alınmıştır. Bu video parçaları kullanılarak köpeklerin davranışlarının analizinin yapılması

ve harekete göre sınıflandırılması sağlanmıştır. Tüm veriseti kullanılarak elde edilen veri içerikleri Tablo 2'de gösterilmiştir.

### 3.2. Anahtar Çerçeve Seçimi (Determination of key frame)

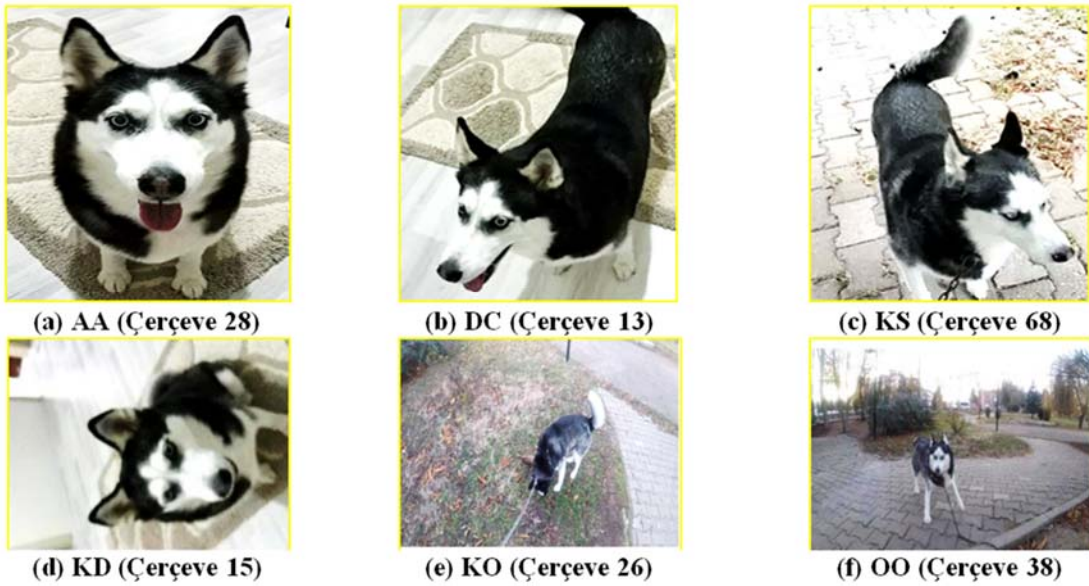
Bu aşamada anlamlı video parçalarından hayvanın davranışlarını gösteren anahtar video çerçeveleri seçilmektedir. Bu çerçeveler köpeğin davranışlarının analizi için video sahnesinde köpeğin olduğu bölümlerin yakalanması amacıyla kullanılmaktadır. Şekil 5'te tüm davranışlar için seçilmiş anahtar çerçeveler için örnekler gösterilmiştir.

### 3.3. İlgi Bölgelerin Etiketlenmesi (Labelling of ROI)

Önerilen çalışmada incelenmesine karar verilen altı farklı hareketler için video sahneleri üzerinde ilgili bölümün işaretlenmesi işlemi yapılmaktadır. Bunun için öncelikle köpeğin sergilediği hareket alanı içerisinde tanımlama yapılacak olan hareketin/davranışın olduğu bölge MATLAB ortamında geliştirilen bir uygulama ile işaretlenmekte ve ilgili her hareket için özellikleri çıkarılarak derin öğrenme ağına sunulacak sayısal verilere dönüştürülmektedir. Bu sayısal veriler her hareket için bölgenin matematiksel karşılıklarının çıkarıldığı 4x1'lik koordinat noktalarının olduğu matris dosyasına karşılık gelmektedir. Bu amaçla

**Tablo 2.** Oluşturulan video veri seti hakkında detaylar (Details about the generated video dataset)

Veri Seti Özelliği	Açıklama
Video adedi	150
Video süresi	10020 saniye (mp4 format, 1920x1080 boyut)
Çerçeve sayısı	6 farklı hareket için toplam:3000 (1920x1088 boyut, 96 dpi çözünürlük)
Eğitim çerçeve sayısı	2100 (Her bir hareket için: 350)
Test çerçeve sayısı	900 (Her bir hareket için: 150)



**Şekil 5.** Her bir davranış için anahtar çerçevelerin seçimi (Selection of key frames for each behavior)



köpeğin izlenecek olan altı farklı hareket için ilgili bölgelerinin belirlenmesi işlemi Şekil 6'da gösterilmiştir. Bu alan işaretleme işlemi, eğitim esnasında ağın tüm çerçeve üzerinden hareketi bularak öğrenmek yerine sadece ilgili hareketin bulunduğu alana odaklanarak ağın daha hızlı eğitilmesini sağlamaktadır.

### 3.4. Ağın Eğitimi (Training of Faster R-CNN)

Bu aşamada etiketlenen bölgelerin koordinatları alındıktan sonra bu koordinatlar önerilen DH B-ESA modeli ile eğitilmektedir. Ağın eğitiminin doğru yapıp yapılmadığını kontrol etmek için bazı aşamaların takip edilmesi gerekmektedir. Bu aşamalardan birincisi, köpeğin davranışının analiz edileceği bölgenin bulunmasıdır. Şekil 7'de gösterilen çerçeveler eğitim veri setinde her bir hareket için kullanılan videodan ağın aldığı ilk çerçevedir. Köpeğin altı farklı hareketi için kırmızı kutu ile çerçevelenmiş ROI bölgesinin (ilgili alanın) işaretlenmiş şekilde çıktılar elde edilmiştir. Bu aşamada eğitilen hareket için doğru alanı vermesi, hareket tanımlanması için doğru koordinat noktalarının tespit edildiğinin bir göstergesidir. Ayrıca bu işlem ile bir görüntüde tüm detaylarla ilgilenmek yerine sadece ilgili alan üzerinde odaklama yapılmaktadır. Eğitim aşaması başarılı bir şekilde tamamlanırsa, Şekil 8'de

görüldüğü gibi etiketlenen her bir davranışın ilgi bölgelerinin sarı kutu ile çerçevelenmektedir.

### 3.5. Köpek davranışlarının tanınması (Recognition of Dog Behaviors)

Bu aşamada, veri setindeki herhangi bir video kullanılarak köpeğin ilgili davranışı gösterip göstermediği test edilmektedir. Yapılan testler sonucunda köpek ilgili davranışı gösteriyorsa Şekil 9'da gösterildiği gibi sarı çerçeve ile işaretlenerek tespit edilen ilgili hareketin adı yazılmaktadır. Yapılan deneysel çalışmalar sonucunda ağız açma, dil çıkarma, koklama, kulak dikme, kuyruk sallama, oyun oynama davranışları eğitime tabi tutulmuş ve bu davranışlar için sırasıyla %100, %99,99,%99,28, %99,99, %95,99, %99,64 sınıflandırma başarımı elde edildiği Şekil 9'da sunulmuştur.

### 3.6. Davranışın video üzerinde takibi (Video tracking)

Tüm bu davranış tanıma ve sınıflandırma işlemlerine ek olarak video takip işlemi de ayrıca sisteme dahil edilmiştir. Bu aşamada ilgi bölgesinin DH B-ESA ile tanınmasından sonra köpeğin davranışı video üzerinde takip edilmektedir. Video takibi bilgisayarlı görü alanının en zorlayıcı



**Şekil 6.** Köpeğin altı farklı davranışları için ilgi bölgesinin etiketlenmesi a) ağız açma b) dil çıkarma c) kuyruk sallama d) kulak dikme e) koklama f) oyun oynama (Labelling of the region of interest for the dog's six different behaviors a) opening the mouth b) sticking out the tongue c) swinging the tail d) rearing the ear e) sniffing f) playing)



**Şekil 7.** Köpeğin altı farklı hareketi/davranışları için ROI bölgesinin bulunması a) ağız açma b) dil çıkarma c) kuyruk sallama d) kulak dikme e) koklama f) oyun oynama (Finding the ROI for six different behaviors of the dog a) opening the mouth b) sticking out the tongue c) swinging the tail d) rearing the ear e) sniffing f) playing)



**Şekil 8.** Eğitim sonrasında altı farklı davranışın sarı çerçeve ile gösterimi a) ağız açma b) dil çıkarma c) kuyruk sallama d) kulak dikme e) koklama f) oyun oynama (Six different behaviors with yellow frame after training a) opening the mouth b) sticking out the tongue c) swinging the tail d) rearing the ear e) sniffing f) playing)



**Şekil 9.** Test veri kümesindeki çerçeveler ile davranışların tanınması a) ağız açma b) dil çıkarma c) kuyruk sallama d) kulak dikme e) koklama f) oyun oynama (Recognition of behaviors with frames in the test dataset a) opening the mouth b) sticking out the tongue c) swinging the tail d) rearing the ear e) sniffing f) playing)

konularından biridir Çoklu hedef bulmadan tekli hedef bulmaya, çevrim içi-çevrim dışı gibi yaklaşımlarla pek çok sorunu ele alan ve teknolojik gelişmeler ile birlikte sürekli olarak yenilenen ve çoğalan alt alanları içermektedir. Bu çalışmada tek davranışı tek hedeften çevrim dışı olarak bulmaya çalışan bir uygulama gerçekleştirilmiştir. Burada teknik olarak, videodan alınan ilk anahtar çerçeve takip edilecek hareket için sisteme giriş olarak verilmektedir ve izleyici/takipçi (tracker) ile videonun devamında bu hareket çerçeveler boyunca takip edilmektedir. Daha önceden koordinatları belirlenen çerçevelerin eğitimi sonucunda, bu koordinatlara benzer olarak kare beyaz işaretçiler ile işaretlenerek elde edilen çıktılar Şekil 10'da görülmektedir. Eğitim videosunda hareketlerin izleyiciler ile izlemesine olanak sağlayan beyaz işaretçiler, hareketlerle ilgili bölgelere işaretler bırakarak video izlenmesinin daha kolay olmasını sağlamaktadır. Veri setindeki herhangi bir video kullanılarak köpeğin ilgili davranışı gösterip göstermediği tespit edildikten sonra köpeğin istenilen davranışı Şekil 11'de gösterildiği gibi yeşil işaretçiler vasıtasıyla video süresi boyunca takipçiler (tracker) ile izlenmektedir. Test işleminin sonucunda tüm hareketlerin videolarda başarılı olarak izlendiği tespit edilmiştir.

#### 4. DENEYSEL ÇALIŞMALAR (EXPERIMENTAL RESULTS)

Bu çalışmada oluşturulan bir veri seti için elde edilen köpek videolarından davranışların analizi, sınıflandırılması ve çevrimdışı olarak video üzerinden köpeklerin davranışlarının izlenmesi gerçekleştirilmiştir. Köpeğin ağız açma, dil çıkarma, koklama, kulak dikme, kuyruk sallama, oyun oynama gibi belirlenmiş altı farklı davranışının analizi ve sınıflandırılması ile ilgili deneysel çalışmalar yürütülmüştür. Önerilen ESA ağında eğitim algoritması, Momentum (SGDM) değeri 0,001'lik bir başlangıç öğrenme oranı ile Olasılıksal Dereceli Azalma kullanılmaktadır. Eğitim sırasında, ilk öğrenme oranı her sekiz adımda bir azaltılır. Olasılıksal dereceli azalma, verimlilik ve kod ayarlama için birçok uygulama kolaylığı sunmaktadır. Olasılıksal dereceli azalmada amaç, her bir adımda veri setindeki her bir örneği gezdikten sonra veri setini güncellemektir. Eğitim algoritması 50 epok için çalıştırılmıştır. Eğer örnekleme kümesi ne kadar büyükse geçiş o kadar zaman almaktadır. Örneğin, 1000 adet veriden oluşan bir veri seti için paket boyutu 500 seçilirse bir epoku tamamlamak için iki iterasyon gereklidir [47]. Veri toplama aşamasında, köpeklere giydirilen ve sahibi tarafından manuel olarak kullanılabilen



**Şekil 10.** Köpeğin davranışlarının takibi için ilgili alanın işaretlenmesi a) ağız açma b) dil çıkarma c) kuyruk sallama d) kulak dikme e) koklama f) oyun oynama (Marking the area of the dog to follow the behavior of the dog a) opening the mouth b) sticking out the tongue c) swinging the tail d) rearing the ear e) sniffing f) playing)



**Şekil 11.** Köpeğin davranışlarının video üzerinde yeşil işaretçiler ile izlenmesi. a) ağız açma b) dil çıkarma c) kuyruk sallama d) kulak dikme e) koklama f) oyun oynama (Tracking the behavior of the dog with green pointers on the video a) opening the mouth b) sticking out the tongue c) swinging the tail d) rearing the ear e) sniffing f) playing)

kameralar sayesinde köpeklerin doğal ortamlarında yaşarken belli aralıklarla videolar alınarak veri setleri oluşturulmuştur. Elde edilen videolardan altı farklı hareketlerin (ağız açma, dil çıkarma, koklama, kulak dikme, kuyruk sallama, oyun oynama) bulunduğu video parçaları çıkarılmıştır. Bu videoları kesmek için video kırma uygulamaları kullanılmıştır. Bu video parçacıkları en kısa 3 saniye en uzun 8 saniye olacak şekilde kısaltılmıştır. Bu saniye aralıklarındaki farklılık köpeklerin bu hareketleri yaparken hızlarından kaynaklanmaktadır. Bu sürelerin kısa olmasının iki ana sebebi vardır. Bunlardan birincisi, eğitim süresini hızlı ve etkili yapmak içindir. İkincisi ise çıkarılan video parçacıklarının her hareketten sadece birine odaklanmış olduğu anları yakalamaktır. Bu odaklanmış hareketlerin seçilmesi sınıflandırma aşamasında bir video parçacığında olan iki veya daha fazla hareket için çakışmayı önleyecektir ve ayrıca eğitimin daha yüksek oranlarda başarılı olmasını sağlayacaktır. Kırılan video veri seti her hareket için daha özelleştirilmiş bir içeriğe sahip olması sağlanmıştır. Daha sonra her hareketin video parçacığından çerçeve çıkarma uygulamaları ile altı hareketin çerçeveleri çıkarılmıştır. Veri setindeki farklı uzunluktaki 150 video içeriğinden her bir hareketten 500 olmak üzere toplamda 3000 anahtar çerçeve çıkarılmıştır. Bu çerçeve görüntülerinin %70'i (2100) eğitim

için ayrılırken %30'u (900) ise test verisi olarak kullanılmıştır. Verisi ile ilgili detaylı bilgiler Tablo 3'te sunulmuştur.

Çalışmada yapılan tüm uygulamalar MATLAB yazılımı ile oluşturulan uygulama üzerinde gerçekleştirilmiştir. Tüm deneysel çalışmalar Tablo 4'te özellikleri verilen bir bilgisayar kullanılarak yapılmıştır. Çalışmada ağız eğitim işlemleri adım adım takip edilerek yapılmaktadır. Eğitim ve test videolarında hareketlerin izleyici ile video üzerinde takip edilebilmesi durumunda hareketin tanınması işlemi başlamaktadır. Bu aşamadan sonra, her hareket için test veri setinde ki veriler ile kıyaslama yapılır. Bu kıyaslanma sonuçları sistemin sınıflandırma başarısının elde edilmesini sağlamaktadır. Elde edilen her bir başarılı veya hatalı eşleşmeler yorumlanarak ilgili hareket için veri seti değiştirilmiş veya eğitim tekrarlanmıştır.

Deneysel çalışmalar kapsamında 6 farklı hareket için sınıflandırma başarımını test etmek için, test verilerinden her bir hareket için 10 farklı anahtar çerçeve seçilmiştir ve sınıflandırmak için seçilen hareket diğer 5 hareket ile kıyaslanmıştır. Böylece önerilen sistemin eğitimin çapraz olarak test edilmesi sağlanmaktadır. Bu kıyaslamayı

yaparken sınıflandırma için seçilen hareket, örneğin ‘kulak dikme’ ise ‘kulak dikme’ için üretilmiş olan eğitim seti kullanılarak 150 test anahtar çerçevesi ‘kulak dikme’ verileri sınıflandırılmaya çalışılmaktadır. Çalışmada köpeklerin 6 farklı hareket için elde edilen sınıflandırma sonuçları Tablo 5’te verilmiştir. Tablodaki verilen ifadelerde DC ile DC kıyaslandığında 150 çerçeveden 147 tanesi başarılı etiketlenirken, 3 tane veri AA olarak hatalı sınıflandırma sonucu elde edilmiştir.

Tablo 5’te gösterilen köpeklerin her bir davranışı için elde edilen sınıflandırma sonuçlarına göre hesaplanan Doğruluk ve Kesinlik değerleri için sonuçlar Tablo 6’da sunulmuştur. Her bir davranış için Doğruluk oranı Doğru Pozitif ve Doğru Negatif değerlerinin ilgili davranış için toplam test veri sayısına bölümü ile hesaplanırken, Kesinlik, Doğru Pozitif değerinin Doğru Pozitif ile Yanlış Pozitif değerlerinin toplamına bölümü ile hesaplanmaktadır.

Test işleminde başarısız eşleşmelerin bazı farklı nedenleri vardır. Bunlardan birincisi, köpeklerin saniyeler içerisinde yüz ifadelerinde ve beden dillerindeki değişiklikler sistemin farklı hareketleri de yakalamasını sağlamaktadır. Bu nedenle

‘kulak dikme’ hareketi test çerçeveleri içerisinde anlık olarak köpek ağız açma hareketini gerçekleştirdiyse eğer sistem bu hareketi yakalayacaktır ve ‘ağız açma’ hareketi için eşleşme olacaktır. Böylelikle oluşan her bir eşleşme sınıflandırma başarımları olarak kaydedilmektedir. Sınıflandırma işlemleri sırasında gözlemlenen ikinci durum ise, benzer içeriklere sahip hareketlerin kıyaslanmasında gözlemlenmiştir. Bu benzer hareketler ‘ağız açma’ ve ‘dil çıkarma’ hareketleridir. ‘Ağız açma’ hareketi sergileyen köpeğin video açısı düşünüldüğünde bu hareketin ‘dil çıkarma’ hareketi olarak algılanabilmesi olasıdır. Bu nedenle, ‘ağız açma’ hareketi için seçilen on adet test veri setindeki çerçeveler ile ‘dil çıkarma’ çerçeveleri sınıflandırıldığında sarı işaretçi ile çerçevelenmiş ‘dil çıkarma’ imgesi gözlemlenebilmektedir. Diğer bir durum ise, çok hareketli davranışların çerçevelerinde davranışın takip edilememesinden kaynaklanmaktadır. Bu durumu en iyi örnekleyen hareket ‘kuyruk sallama’ hareketidir. Kuyruk sallama davranışı çok hareketli bir davranış yapısına sahip olduğundan dolayı veri setinde farklı yapılar da kuyruk boyutuna, pozisyonuna ve şekline sahip çerçeveleri veri setinde görmek mümkündür. Bu hareketlilik nedeniyle çerçevelerin bazılarında kuyruk sallama hareketi

**Tablo 3.** Veri setindeki her bir davranış için eğitim ve test verileri için detaylar  
(Details for training and test data for each behaviors in the dataset)

Veri Seti Detayları	Ağız Açma	Kulak Dikme	Oyun Oynama	Koklama	Kuyruk Sallama	Dil Çıkarma
Etiket	AA	KD	OO	KO	KS	DC
Toplam Veri Sayısı	500	500	500	500	500	500
Eğitim Veri Sayısı	350	350	350	350	350	350
Test Veri Sayısı	150	150	150	150	150	150

**Tablo 4.** Deneysel çalışmalar için kullanılan bilgisayarın konfigürasyon bilgileri  
(Configuration information of the computer used for experimental studies)

Donanım	Özellik
Merkezi İşlemci(CPU)	Intel Core i7-7700K @ 2,8 Ghz (8 CPUs)
Bellek(RAM)	16 GB (DDR4 2400 Mhz)
Anakart	ASUS X580VD
GPU	NVIDIA GeForce GTX 1050 (4 GB)
Harddisk	256 GB SSD HDD

**Tablo 5.** Test veri setinde sınıflandırma başarımları sonuçları (Classification performance results in test data set)

Test Edilen Davranış	AA	DC	KO	KD	KS	OO
AA	141	9	0	0	0	0
DC	3	147	0	0	0	0
KO	1	0	149	0	0	0
KD	0	0	1	149	0	0
KS	0	0	1	0	147	2
OO	0	0	0	0	2	148

**Tablo 6.** Önerilen sistem için Doğruluk ve Hassaslık oranları (Accuracy and Precision ratios for the proposed system)

	AA	DC	KO	KD	KS	OO
Doğruluk(%)	94,00	98,00	99,33	99,33	98,00	98,67
Kesinlik(DP/(DP+YP))	0,9724	0,9423	0,9867	1	0,9865	0,9867

izlenememesinden kaynaklı başarımların oranlarının düşmektedir. Ayrıca bu hareketlilik nedeniyle videolardan alınan çerçevelerin netliklerinde azalma da söz konusudur ve bu azalma yine başarımların oranını düşüren bir faktör olarak karşımıza çıkmaktadır. Bu gibi problemlerin üstesinden gelmek için, eğitim esnasında hareketi takip edilemediği çerçeveler yerine bu çerçevelere ek olarak yeni çerçeveler üretilerek ağız eğitimi yenilenmiştir. Çoğaltılmış çerçeve içeriğine sahip ağ daha yüksek başarımlar elde edilmektedir.

Önerilen DH B-ESA tabanlı köpek davranışlarını tanıma sisteminde Tablo 7’de dil çıkarma hareketlerinin sınıflandırma işlemi için ilk 10 çerçevelere ait sonuçlar sunulmuştur. Yapılan eğitimin sınıflandırma başarımları da ayrıca tespit etmek için, dil çıkarma çerçeveleri diğer beş davranış ile kıyaslanmıştır. Tablo 7’de DH B-ESA kullanılarak ilgili davranışın test edilmesi için seçilen çerçevenin sınıflandırma başarımları açısından sonuçları görülmektedir. İlk 10 çerçeveye bakıldığında büyük oranda doğru hareketin tespit edildiği anlaşılmaktadır. Şekil 12’de ise dil çıkarma test çerçevesi ile eşleşen veya eşleşmeyen sonuçların bulunduğu davranışlara ait görüntüler sunulmuştur. Dil çıkarma ve ağız açma hareketleri birbirine benzediği için bazı dil çıkarma davranışlarının sınıflandırma

sonucunda ağız açma olarak etiketlendiği ve bununla ilgili eşleşen çerçeve olduğu görülmektedir.

## 5. SONUÇLAR VE TARTIŞMALAR (RESULTS AND DISCUSSIONS)

Önerilen çalışmada, DH B-ESA modeline dayalı köpeklerin davranışlarının analiz edilmesi ve sınıflandırılması amacıyla literatür taraması, araç-gereçlerin temini, verilerin toplanması, video verilerinden anahtar çerçevelerin elde edilmesi, çerçeveler üzerinden özellik çıkarımı, ağız eğitilmesi, köpek davranışlarının tanınması ve köpeğin eş zamanlı takip edilmesi aşamalarından oluşan bütüncül bir sistem geliştirilmiştir. Tüm aşamalar birbirinden hem ayrılmaz bir bütün iken hem de çok farklı prosedürlerden oluşmaktadır. Yapılan deneysel çalışmalarda köpeklerin ağız açma, dil çıkarma, koklama, kulak dikme, kuyruk sallama, oyun oynama davranışları incelenmiş ve bu davranışlar için sırasıyla %94,00, %98,00, %99,33, %99,33, %98,00, %98,67 sınıflandırma başarımları elde edilmiştir. Köpeklere giydirilen ve sahibi tarafından manuel olarak kullanılabilen kameralar sayesinde köpeklerin doğal ortamlarında yaşarken belli aralıklarla videolar alınarak veri setleri oluşturulmuştur. Elde edilen videolardan altı farklı hareketlerin (ağız açma, dil çıkarma, koklama, kulak dikme, kuyruk sallama, oyun

**Tablo 7.** Dil çıkarma hareketi için B-ESA sınıflandırma sonuçları  
(Classification results with faster R-CNN for sticking out the tongue)

Test Edilen Davranış: DC	AA(%)	DC (%)	KO (%)	KD (%)	KS (%)	OO (%)
1	0	87,94	0	0	0	0
2	0	99,96	0	0	0	0
3	0	99,01	0	0	0	0
4	0	100	0	0	0	0
5	0	100	0	0	0	0
6	51,1	96,89	0	0	0	0
7	0	100	0	0	0	0
8	0	99,98	0	0	0	0
9	0	100	0	0	0	0
10	0	99,99	0	0	0	0
Ortalama(%)	5,11	98,38	0	0	0	0



**Şekil 12.** Köpeğin dil çıkarma davranışı sınıflandırma sonuçları için örnek çerçeveler  
(Sample frames of classification results for sticking out the tongue behavior of dog)

oyun) bulunduğu video parçaları çıkarılmıştır. Veri seti hazırlıklarının ardından video parçacıkları ile resim çerçeveleri kullanılarak bir dizi tanıma ve sınıflandırma işlem adımları gerçekleştirilmiştir.

Bu çalışma ile köpekler kameralarla izlenerek ve bilgisayarlı gözü destekli makine öğrenmesi yöntemleri ile davranışlarının analiz edilerek zaman ve personel problemlerin azaltılmasına katkı sağlaması hedeflenmiştir. Ek olarak, engelli bir insanın günlük yaşantısında rehberlik eden bir köpeğin izlenmesi davranışlarının takip edilmesi, hayati sonuçlar doğurabilecek problemler için bu kişileri koruyacağı düşünülmektedir. Bu nedenle, tüm özel amaçlar için yetiştirilmiş köpeklerin veya herhangi bir sokak köpeğinin davranışlarının anlaşılması, takip edilmesi insanların günlük yaşam standartlarını arttırabileceği gibi köpeklerin de günlük yaşam kalitelerini arttıracağı öngörülmektedir. Ayrıca bu çalışmada olduğu gibi güncel teknolojilerin kullanımı ile özel eğitilmiş köpeklerin günlük hayatımızda daha çok var olacağı ve insan-köpek fayda alanlarının daha da gelişeceği söylenebilir.

Davranışların eğitilmesi ve sınıflandırması için oluşturulan derin öğrenme modeli ağı eğitim başarısını etkileyen en önemli faktörlerden biridir. Konvolüsyon katmanı derin sinir ağının temel yapı taşıdır ve burada kullanılan filtrelerin özellikleri sayesinde sistem başarısı daha üst noktalara çıkarılabilmektedir. Derin ağların en önemli özelliklerinden biride sonsuz sayıda ağ modelleri oluşturmaya olanak sağlamasıdır. Bu sonsuzluğu sağlayan parametreler; katman sayısı, konvolüsyon, filtre boyutu/tipi/sayısı, havuzlama katmanı ve tam bağlı katmanların istenilen işlevlere uygun olarak evrilebilmesinden kaynaklanmaktadır. Çalışmada, eğitim setinin haricinde oluşturulan test video veri seti üzerinden köpeğin hangi davranışı sergilediğinin kıyaslanması yapılmıştır. Sınıflandırma yapılırken karşılaşılan problemlere çözümler bulunmuş ve sınıflandırma başarı oranları yükseltilmiştir. Bu sorunlardan bazıları, oluşturulan veri setlerindeki çerçevelerin iki veya üç farklı hareketi birden içeriyor olmasından kaynaklanmıştır. Bu sorunu gidermek için çerçevelerin çıkartıldığı video parçacıkları çok kısa sürelerde alınarak, köpek tarafından yapılan hareketin tam sergilenmiş olması ve başka hareketler içermemesi sağlanmıştır.

Önerilen çalışmadan sonra yapılacak olan benzer çalışmalarına ışık tutması açısından, derin öğrenme ağlarının eğitimi için çok fazla veri olması başarıyı artıran en önemli etken olmasının yanı sıra, transfer öğrenme ile ön-eğitilmiş ağlar kullanılmak sistemin başarısını artıran diğer önemli unsurlardandır. Önerilen çalışmada ağı ön-eğitimi için CIFAR-10 veri seti kullanılmıştır. Elde edilen video süreleri ve adetleri göz önüne alındığında, 5 dakikalık bir videodan alınacak çerçeve sayısı ile ortalama 5 saniyeye düşürülmüş video parçacığından alacağımız çerçeve sayısı arasında büyük bir fark vardır. Bu durumda veri setini geniş tutup çeşitlendirmek ağı eğitimi esnasında çok önemli olmaktadır. Ayrıca, ağı eğitimi sırasında resimlerin çözünürlüklerinin iyi olmasının ne kadar önemli olduğu da açıktır. Ek olarak,

derin öğrenme ağlarında işlem adımlarının daha hızlı gerçekleşmesi için, bilgisayarın donanımsal olarak güçlü olması gerekmektedir. Derin sinir ağlarının eğitiminde genellikle GPU kullanılmaktadır. Ancak GPU'ya ihtiyacınız olup olmadığı çalışılacak veri setine, modelin karmaşıklığına ve zaman kısıtına bağlı olarak değişmektedir. Bu çalışmada elde edilen başarı oranının yüksek olması ileride yapılacak olan diğer hayvan türleri üzerinde yapılacak çalışmalara ve altı farklı hareketin dışında davranışlarının belirlenmesi çalışmalarına altyapı oluşturacaktır. Daha sonraki çalışmalarda farklı sınıflandırma algoritmaları ile derin öğrenme sınıflandırma algoritmalarının birbirine entegre edilmiş hali ile eğitim başarısının oranı arttırılabileceği düşünülmektedir. Başarı oranını etkileyen diğer faktörler ise veri setinde kullanılan videoların sayısının çokluğudur. Önerilen çalışmada veri setindeki herhangi bir video parçacığı üzerinde çevrimdışı olarak sadece tek hedeften tek bir hareket izlenecek şekilde tasarlanmıştır. Bir sonraki çalışmalarda, köpeğin sergilediği birden fazla hareketi çevrim-içi veya çevrim-dışı izlemek ana çalışma konusu olması hedeflenmektedir.

#### TEŞEKKÜR (ACKNOWLEDGMENT)

Bu çalışmayı Proje No: 2017-01.BŞEÜ.03-09 ile destekleyen Bilecik Şeyh Edebalı Üniversitesi Bilimsel Araştırma Projeleri Koordinatörlüğüne ve video verilerinin toplanmasına katkı sunan köpeklerin sahibi Öğr. Gör. İbrahim Kılıç'a teşekkür ederiz.

#### KAYNAKLAR (REFERENCES)

1. Weisbord M. and Kachanoff K., Dogs with jobs: working dogs around the world: Simon and Schuster, 2000.
2. Prato-Previde E., Nicotra V., Pelosi A., and Valsecchi P., Pet dogs' behavior when the owner and an unfamiliar person attend to a faux rival, PloS one, 13, e0194577, 18 April, 2018.
3. Pan Y., Landsberg G., Mougeot I., Kelly S., Xu H., Bhatnagar S., et al., Efficacy of a therapeutic diet on dogs with signs of cognitive dysfunction syndrome (CDS): A prospective double blinded placebo controlled clinical study, Frontiers in Nutrition, 5 (127), 2018.
4. Lindsay S. R., Handbook of applied dog behavior and training, adaptation and learning, 1, John Wiley & Sons, 2013.
5. Peterson J. C., Soulos P., Nematzadeh A., and Griffiths T. L., Learning Hierarchical Visual Representations In Deep Neural Networks Using Hierarchical Linguistic Labels, arXiv preprint arXiv:1805.07647, 19 May, 2018.
6. Byosiere S.-E., Chouinard P. A., Howell T. J., and Bennett P. C., What do dogs (Canis familiaris) see? A review of vision in dogs and implications for cognition research, Psychonomic Bulletin & Review, 25 (5), 1798-1813, October, 2018.
7. Aenishaenslin C., Brunet P., Lévesque F., Gouin G. G., Simon A., Saint-Charles J., et al., Understanding the Connections Between Dogs, Health and Inuit Through a

- Mixed-Methods Study, *EcoHealth*, 1-10, 14 December, 2018.
8. Ladha C., Hammerla N., Hughs E., Olivier P., and Plotz T., Dog's Life: Wearable Activity Recognition for Dogs, 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing, Switzerland, 415-418, 2013.
  9. Leos-Barajas V., Photopoulou T., Langrock R., Patterson T. A., Watanabe Y. Y., Murgatroyd M., et al., Analysis of animal accelerometer data using hidden Markov models, *Methods in Ecology and Evolution*, 8 (2), 161-173, 2017.
  10. Gerencsér L., Vásárhelyi G., Nagy M., Vicsek T., and Miklósi A., Identification of behaviour in freely moving dogs (*Canis familiaris*) using inertial sensors, 8 (10), 77814, 2013.
  11. Brugarolas R., Loftin R. T., Yang P., Roberts D. L., Sherman B., and Bozkurt A., Behavior recognition based on machine learning algorithms for a wireless canine machine interface, 2013 IEEE International Conference on Body Sensor Networks, Cambridge-Ma-Usa ,1-5, 6-9 May, 2013.
  12. Sağıroğlu Ş. and Koç O., Büyük Veri Ve Açık Veri Analitiği:Yöntemler Ve Uygulamalar, Ankara: Gazi Üniversitesi Big Data Center, Ankara, Turkey, 2017.
  13. Huang H., Zhou H., Yang X., Zhang L., Qi L., and Zang A.-Y., Faster R-CNN for Marine Organisms Detection and Recognition Using Data Augmentation, *Neurocomputing*, 2019.
  14. Yang Q., Xiao D., and Lin S., Feeding behavior recognition for group-housed pigs with the Faster R-CNN, *Computers and Electronics in Agriculture*, 155, 453-460, 2018.
  15. Wang D., Tang J., Zhu W., Li H., Xin J., and He D., Dairy goat detection based on Faster R-CNN from surveillance video, *Computers and Electronics in Agriculture*, 154, 443-449, 2018.
  16. Zhao X., Wu Y., Song G., Li Z., Zhang Y., and Fan Y., A deep learning model integrating FCNNs and CRFs for brain tumor segmentation, *Medical image analysis*, 43,98-111, 2018.
  17. Sharma H., Zerbe N., Klempert I., Hellwich O., and Hufnagl P., Deep convolutional neural networks for automatic classification of gastric carcinoma using whole slide images in digital histopathology, *Computerized Medical Imaging and Graphics*, 61, 2-13, 2017.
  18. Havaei M., Davy A., Warde-Farley D., Biard A., Courville A., Bengio Y., et al., Brain tumor segmentation with deep neural networks, *Medical image analysis*, 35, 18-31, 2017.
  19. Kamnitsas K., Ledig C., Newcombe V. F., Simpson J. P., Kane A. D., Menon D. K., et al., Efficient multi-scale 3D CNN with fully connected CRF for accurate brain lesion segmentation, *Medical image analysis*, 36, 61-78, 2017.
  20. Pirsivash H. and Ramanan D., Detecting activities of daily living in first-person camera views, 2012 IEEE Conference on Computer Vision and Pattern Recognition, Providence-RI-USA, 2847-2854, 16-21 June, 2012.
  21. Hammerla N. Y., Halloran S., and Plötz T., Deep, convolutional, and recurrent models for human activity recognition using wearables, *IJCAI'16 Proceedings of the Twenty-Fifth International Joint Conference on Artificial Intelligence*, New York-USA, 1533-1540, 09-15 July, 2016.
  22. Cho Y., Nam Y., Choi Y.-J., and Cho W.-D., SmartBuckle: human activity recognition using a 3-axis accelerometer and a wearable camera, the 2nd International Workshop on Systems and Networking Support for Health Care and Assisted Living Environments, 7, 2008.
  23. Fathi A., Farhadi A., and Rehg J. M., Understanding egocentric activities, *IEEE International Conference on Computer Vision*, Barcelona-Spain , 407-414, 6-13 November, 2011.
  24. Iwashita Y., Takamine A., Kurazume R., and Ryoo M. S., First-person animal activity recognition from egocentric videos, 22nd International Conference on Pattern Recognition, Stockholm, Sweden, 4310-4315, 24-28 August, 2014.
  25. Dodge S. and Karam L., A Study And Comparison Of Human And Deep Learning Recognition Performance Under Visual Distortions, 26th International Conference On Computer Communication And Networks (ICCCN), Vancouver, BC, Canada,1-7, 31 July-3 August , 2017.
  26. Ehsani K., Bagherinezhad H., Redmon J., Mottaghi R., and Farhadi A., Who Let The Dogs Out? Modeling Dog Behavior From Visual Data, *IEEE/CVF Conference on Computer Vision and Pattern Recognition*, Salt Lake City-Ut-Usa , 4051-4060, 18-23 June, 2018.
  27. Zhou J., Li Z., Zhi W., Liang B., Moses D., and Dawes L., Using Convolutional Neural Networks And Transfer Learning For Bone Age Classification, *International Conference on Digital Image Computing: Techniques and Applications (DICTA)*, Sydney-NSW-Australia, 1-6, 29 November-1 December, 2017.
  28. Szegedy C., Liu W., Jia Y., Sermanet P., Reed S., Anguelov D., et al., Going Deeper With Convolutions, *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston-Usa, 1-9, 7-12 June, 2015.
  29. Ardö H., Guzhva O., Nilsson M., and Herlin A. H., Convolutional neural network-based cow interaction watchdog, *IET Computer Vision*,12 (2) , 171-177, 2018.
  30. Kang K., Li H., Yan J., Zeng X., Yang B., Xiao T., et al., T-cnn: Tubelets with convolutional neural networks for object detection from videos, *IEEE Transactions on Circuits and Systems for Video Technology*, 28 (10), 2896-2907, 2018.
  31. Girshick R., Donahue J., Darrell T., and Malik J., Rich Feature Hierarchies For Accurate Object Detection And Semantic Segmentation, 14 Proceedings of the 2014 IEEE Conference on Computer Vision and Pattern Recognition, Washington DC-Usa, 580-587, 23-28 June, 2014.
  32. Ren S., He K., Girshick R., and Sun J., Faster R-CNN: Towards Real-Time Object Detection With Region

- Proposal Networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39 (6) ,91-99, 2015.
33. Zheng C., Zhu X., Yang X., Wang L., Tu S., and Xue Y., Automatic recognition of lactating sow postures from depth images by deep learning detector, *Computers and Electronics in Agriculture*, 147, 51-63, 2018.
  34. Bulling A., Blanke U., and Schiele B., A tutorial on human activity recognition using body-worn inertial sensors, *ACM Computing Surveys (CSUR)*, 46 (3), 33, 2014.
  35. Modern Dog Magazine. The Lifestyle Magazine for Modern Dogs and Their Companies. Available: <https://moderndogmagazine.com/>. Yayın tarihi 2012. Erişim tarihi Mart 10, 2019.
  36. LeCun Y., Bottou L., Bengio Y., and Haffner P., Gradient-based learning applied to document recognition, *Proceedings of the IEEE*, 86 (11), 2278-2324, 1998.
  37. Krizhevsky A., Sutskever I., and Hinton G. E., Imagenet classification with deep convolutional neural networks, *Advances in neural information processing systems*, 60 (6), 84-90, 2017.
  38. Karpathy A., Toderici G., Shetty S., Leung T., Sukthankar R., and Fei-Fei L., Large-scale video classification with convolutional neural networks, 2014 *IEEE Conference on Computer Vision and Pattern Recognition*, Columbus-OH-Usa, 1725-1732, 23-28 June, 2014.
  39. Koppula H. S. and Saxena A., Anticipating human activities using object affordances for reactive robotic response, *IEEE Transactions on Pattern Analysis And Machine Intelligence*, 38 (1), 14-29, 2016.
  40. Kitani K. M., Ziebart B. D., Bagnell J. A., and Hebert M., Activity forecasting, *Proceedings of the 12th European conference on Computer Vision*, Florence-Italy, 201-214, 7-13 October, 2012.
  41. Lan T., Chen T.-C., and Savarese S., A hierarchical representation for future action prediction, *13th European Conference on Computer Vision*, Zurich, Switzerland, 6-12September, 689-704,2014.
  42. Liu Y. H., Feature Extraction and Image Recognition with Convolutional Neural Networks, in *Journal of Physics Conference Series*, 1087 (6), 062032, 2018.
  43. Lu Y., Yi S., Zeng N., Liu Y., and Zhang Y., Identification of Rice Diseases Using Deep Convolutional Neural Networks, *Neurocomputing*, 267, 378-384, 2017.
  44. Ali A., Hanbay D., Tumor detection in MR images of regional convolutional neural networks, *Journal of the Faculty of Engineering and Architecture of Gazi University*, 34 (3), 1395-1408, 2019.
  45. Weiss K., Khoshgoftaar T. M., and Wang D., A survey of transfer learning, *Journal of Big Data*, 3, 9, 2016.
  46. CIFAR-10. (11/03/2019). Available: <https://www.cs.toronto.edu/~kriz/cifar.html>
  47. Kurt F., Evrişimli Sınır Ağlarında Hiper Parametrelerin Etkisinin İncelenmesi, Yüksek Lisans Tezi, Hacettepe Üniversitesi, Eğitim Bilimleri Enstitüsü, Ankara, 2018.