

## Fotokapan Görüntülerinde Yerel Öznitelikler ile Nesne Tespiti

### Object Detection on Camera-Trap Images with Local Features

Emrah ŞİMŞEK<sup>1,2,a</sup>, Barış ÖZYER<sup>\*2,b</sup>, Gülşah TÜMÜKLÜ ÖZYER<sup>2,c</sup>

<sup>1</sup>Recep Tayyip Erdoğan Üniversitesi, Mühendislik Fakültesi, Bilişim Sistemleri Mühendisliği Bölümü, 53020,Rize

<sup>2</sup>Atatürk Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü,25240, Erzurum

• Geliş tarihi / Received: 09.01.2019 • Düzeltilecek geliş tarihi / Received in revised form: 22.05.2019 • Kabul tarihi / Accepted: 19.06.2019

#### Öz

Fotokapanlar doğal ortamda yaşayan canlıların davranışlarını izlemek amacıyla yaygın olarak kullanılan cihazlardır. Fotoğraf ve video kaydı yapan bu cihazlar ile doğal görüntülerde yapılan nesne (hayvan veya insan) tespiti işlemi, arka planın karmaşık yapıda olması, ışık şiddeti yetersizliği, ışık şiddeti değişimi, nesnenin parçalı bulunması gibi nedenlerden dolayı zor bir problemdir. Ayrıca nesnenin hareketli olması, görüntü içerisinde bulunduğu konumun tespit edilmesini zorlaştırmaktadır. Son yıllarda kullanılan yerel öznitelikler konum bilgisi içerdiğinden, hem konumlandırma probleminde çözüm olmakta hem de yerel öznitelik dönüşüm yöntemlerinin içerdiği ölçek, dönme, afin dönüşümü, aydınlatma değişimi gibi zorluklara karşı değişmezlikler sayesinde daha başarılı tespit işlemi yapılabilmektedir. Bu çalışmada foto-kapan görüntülerinde yerel öznitelik dönüşüm yöntemleri olan Ölçek Değişmez Öznitelik Dönüşümü (Scale Invariant Feature Transform-SIFT), Hızlandırılmış Sağlam Öznitelikler (Speeded Up Robust Features-SURF), İkili Sağlam Bağımsız Temel Öznitelikler (Binary Robust Independent Elementary Features-BRIEF), Yönlendirilmiş Hızlı ve Sağlam Brief Öznitelikleri (Oriented Fast And Robust Brief-ORB), öznitelik eşleştirme yöntemlerinde kullanılarak nesne tespiti gerçekleştirilmiştir. Hatalı yerel öznitelik eşleşmelerinin elenmesi için yüzdelik ve medyan tabanlı aykırılık tespiti ile k-en yakın komşu öznitelik eleme yöntemleri kullanılmıştır. Çalışmada öznitelik dönüşüm yöntemleri ile elde edilen nesne tespit başarıları, eşleşen öznitelik sayıları, sınırlayıcı kutu büyüklükleri, elenen öznitelik sayıları ve bunların nesne tespit başarısına olan etkileri incelenmiştir.

**Anahtar kelimeler:** Fotokapan, Nesne Tespiti, ORB, SIFT, SURF, BRIEF, Yerel Öznitelikler

#### Abstract

Camera-traps are the devices that commonly used to monitor the behavior of living creatures in the natural environment. Object (animal or human) detection in the natural image or video by recorded these devices has difficulties such as cluttered background, inefficient light intensity, light intensity change or partial object presence. Furthermore, the fact that the object is moving makes it difficult to determine the position in the image. Since the local features used in recent years contain location information, it is a solution to the problem of localization as well as more successful detection can be made by the invariance of the scale, rotation, affine transformation, lighting change included in the local feature transformation methods. In this study, local feature description methods are used in camera-trap images, such as Scale Invariant Feature Transform (SIFT), Speeded-Up Robust Features-SURF, Binary Robust Independent Elementary Features (BRIEF), Oriented Fast and Robust Brief-ORB was performed with feature matching methods. Percentile and median based outlier detection methods and k nearest neighboring feature elimination methods were used to eliminate incorrect feature matches. In this study, the effect of feature description methods on object detection accuracies, number of matching features, bounding box sizes, number of eliminated features and their effects on object detection success were analyzed.

**Keywords:** Camera Trap, Object Detection, ORB, SIFT, SURF, BRIEF, Local Features

<sup>\*b</sup> Barış Özyer; baris.ozyer@atauni.edu.tr; Tel: 0 442 231 60 51; <https://orcid.org/0000-0003-0117-6983>

<sup>a</sup> <https://orcid.org/0000-0002-1652-9553> <sup>c</sup> <https://orcid.org/0000-0002-0596-0065>

## 1. Giriş

Foto-kapanlar doğal ortamdaki canlı yaşamı hakkında bilgi toplanmasını sağlayan görüntü ve video kaydı yapan cihazlardır (Kays vd., 2010). Foto-kapan görüntüleri üzerinde bölgede canlıların tespit edilmesi, türlerinin ve nüfuslarının belirlenmesi gibi çalışmalar yapılmaktadır. Bu çalışmalarda, görüntülerde bulunan canlıların otomatik olarak tespit edilebilmesi için bilgisayarlı görme ve örüntü tanıma yöntemleri kullanılmaktadır. Bu yöntemler kullanılarak görüntüdeki hayvanlar ve insanlar otomatik olarak tespit edilmeye veya tanınmaya çalışılmaktadır. Foto-kapan görüntülerinde yapılan çalışmalarda, nesne tespit etme ve tanıma işlemlerinin yapılmasını zorlaştıran karmaşık ve hareketli arka plan, ışık yoğunluğu değişimi ve cismin kısmi görünümü gibi önemli problemler bulunmaktadır (Yu vd., 2013; Nguyen vd., 2017; Norouzzadeh vd., 2018; Meek vd., 2014; Andavarapu ve Vatsavayi, 2017). Foto-kapan görüntüleri üzerinde yapılan çalışmalarda, nesne tespit etme ve nesne tanıma olmak üzere iki temel konu hedeflenmektedir. Görüntüde nesne tespiti işlemi, nesnenin görüntüde bulunması ve eğer var ise görüntüdeki konumunun belirlenmesi olarak ifade edilebilmektedir (Chen vd., 2015). Nesne tanıma ise tespit edilen nesnelerin tür veya farklı şekilde sınıflandırılması anlamına gelmektedir. Nesne tespiti amacıyla yapılan çalışmalarda genellikle nesnelere ait model görüntülerin öznelikleri, foto-kapan görüntülerinde kayan pencere tabanlı yöntemle aranarak nesne tespit işlemi gerçekleştirilmektedir (Yu vd., 2013; Andavarapu ve Vatsavayi, 2017).

Nesne tespit etme ve konumlandırma için bir pencereyi görüntü üzerinde çoklu ölçekler şeklinde kaydırmak en yaygın yaklaşımdır. Bu yerel penceredeki görüntüler aranan nesne veya arka plan şeklinde sınıflandırılarak tespit işlemi gerçekleştirilebilmektedir. Bu yaklaşım, görüntülerde yüz, yaya ve araba gibi nesnelere tespit etmek için başarıyla kullanılmaktadır (Murphy vd., 2006). Bunun dışında son yıllarda kullanılan bir yaklaşım da, görüntüden yerel özneliklerini çıkarmak ve sonra tüm olası alt pencerelere bakmak yerine bu noktaların etrafındaki bölgelerin her birini sınıflandırmaktır (Murphy vd., 2006). Görüntüdeki nesnelerin tespit edilmesi ve nesnelerin tanınması, özellikle sınıf içi çeşitlilik, karmaşıklık, tıkanıklık ve poz değişikliklerinin olmasından dolayı bilgisayarlı görüdeki en zorlu problemlerden biridir (Zhang vd., 2007). Görüntüde nesne tespit etme veya nesne tanıma amacıyla kullanılacak temelde

iki tip öznelik çıkarılmaktadır. Bunlar global ve yerel özneliklerdir. Global öznelikler görüntüyü bir bütün olarak tanımlamakta ve tüm bir nesneyi tek bir vektörle genelleştirmektedir. Yerel öznelikler ise, görüntüde belirlenen birden çok anahtar nokta ve bu anahtar noktaların tanımlayıcılara dönüştürülmesi ile elde edilmektedir. Bu tanımlayıcılara dönüştürme işlemi, özneliğe tıkanma ve karmaşıklığa karşı dayanıklılık sağlamaktadır. Ancak, her nesne veya görüntü için değişken sayıda öznelik vektörünün bulunduğu durumları ele almak için özel sınıflandırma algoritmaları gerekebilmektedir (Lisin vd., 2005). Görüntüdeki nesnelerin tespit edilmesi veya tanınması için kullanılacak bir yerel öznelik, yakın komşuluklarından farklılık ifade eden bir görüntü deseni olarak tanımlanmaktadır. Yerel öznelikler noktalar, kenarlar veya küçük görüntü yamalarıdır. Tipik olarak, bazı ölçümler yerel bir öznelik üzerinde bulunan bir bölgeden alınmakta ve daha sonra çeşitli uygulamalar için kullanılacak tanımlayıcılara dönüştürülmektedir. Bu yerel öznelik tanımlayıcıları, parçalara ayırma gereği duymadan nesnelere veya sahneleri tanımaya yarayan sağlam bir görüntü temsili olarak kullanılabilirler için sıklıkla tercih edilebilmektedir (Tuytelaars ve Mikolajczyk, 2008).

Gerçek görüntülerde nesne tespit etme ve nesne tanıma, görüntüdeki karmaşıklık veya kısmi tıkanıklıktan etkilenmeyen yerel görüntü öznelikleri kullanmayı gerektirmektedir. Yerel özneliğin, tekrarlanabilirlik, ayırt edilebilirlik, yerellik, yeterli miktar, doğruluk ve verimlilik gibi özelliklere sahip olması gerekmektedir (Tuytelaars ve Mikolajczyk, 2008). Bu öznelikler, aydınlatma değişikliğine, 3B dönüşümlere ve ortak nesne varyasyonlarına karşı değişmezlik sahibi olmalıdır. Öte yandan, öznelikler farklı nesnelere tanımlayabilmeleri için ayırt edici olmalıdır (Lowe, 1999). Yerel değişmez öznelikler, yalnızca görüntüleme koşullarındaki değişimlere, oklüzyon veya görüntü dağınıklığına rağmen görüntüdeki benzerliklerin bulunmasını sağlamakla kalmamakta, aynı zamanda görüntü içeriği ile ilgili tanımlama yapabilmektedir (Tuytelaars ve Mikolajczyk, 2008).

Foto-kapan görüntülerinde yapılan nesne tespit etme ve tanıma çalışmalarında farklı yerel ve global öznelikler kullanılmıştır. Bu çalışmalarda, nesnelerin tespit edilme ve kırılma işleminin insanlar tarafından elle yapılmış görüntülerin yerel veya genel öznelikleri, sınıflandırma yöntemleri ile birlikte kullanılabilirliği gibi (Yu

vd., 2013; Nguyen vd., 2017; Andavarapu ve Vatsavayi, 2017), bu öznelikler kayan pencere veya sinir ağları tabanlı yöntemlerde kullanılarak nesne tespiti yapılabilmektedir (Nguyen vd., 2017; Norouzzadeh vd., 2018; Zhang vd., 2016; Khorrani vd., 2012). Yapılan çalışmalar incelendiğinde, nesne görüntülerinin insanlar tarafından tespit edildiği ve arka plandan el ile kırıldığı çalışmada (Yu vd., 2013), nesnelere SIFT ve cLBP (cell-structured Local Binary Pattern) öznelikleri kullanılarak çok sınıflı SVM (Support Vektör Machine) ile sınıflandırılmıştır. Tespit işleminin otomatik olarak kayan pencere tabanlı yöntemle yapıldığı (Andavarapu ve Vatsavayi, 2017) diğer bir çalışmada ise 5 farklı vahşi hayvan türünün bulunduğu 5000 görüntüde vahşi hayvanların tespit edilmesi ve tanınması amacıyla W-CoHOG (Weighted Co-occurrence Histogram Of Gradients) öznelikleri ve çok sınıflı doğrusal SVM sınıflandırıcı kullanılmıştır. Sinir ağları kullanılan bir çalışmada ise (Nguyen vd., 2017) piksel yoğunluklarının 0-1 arasına normalize edildiği, 224x224 piksel boyutlarına dönüştürülmüş eğitim verisi ile eğitilen derin öğrenme tabanlı sinir ağları kullanılmıştır. 1.2 milyon görüntünün eğitim için kullanıldığı bir diğer sinir ağları tabanlı çalışmada ise (Norouzzadeh vd., 2018), derin öğrenme ile nesnelere tanıma, sayma, davranış belirleme gibi işlemler gerçekleştirilmiştir ve bu çalışmada 3 gizli katman kullanılmış olup, bu katmanlarda sırasıyla kenarlar, köşeler ve motifler, nesne parçaları işlenmiştir. Çapraz doğrulama kullanılan (Zhang vd., 2016) bir çalışmada ise, görüntüdeki 20x20 piksel büyüklüğündeki parçaların arka plan olup olmadığının tespit edilmesi amacıyla derin öğrenme tabanlı sinir ağı ve adaptif pencere büyüklüğünde kayan pencere tabanlı HOG öznelikleri ve Fisher vektör kodlaması kullanılmıştır. Çok sayıda hayvan türünün tespit edildiği (Khorrani vd., 2012) çalışmada, görüntüler üzerinde öncelikle Sağlam Temel Bileşen Analizi (RPCA-Robust Principal Component Analysis) ile arka plan analizi yapılmış ve daha sonra yerel entropi ve bağlı bileşen analizi ile nesnelere tespit edilmiştir. Bu çalışmada geniş yer değiştirme optik akış yöntemi ile nesne içeren bölgeler tespit edilmiştir. Sonraki görüntülerde ise nesnenin takip edilmesi ve hareketi ile ilgili bilgi elde edilmesi amacıyla optik akış yöntemi ile görüntüler arasındaki büyük değişikliklerden yararlanılmıştır.

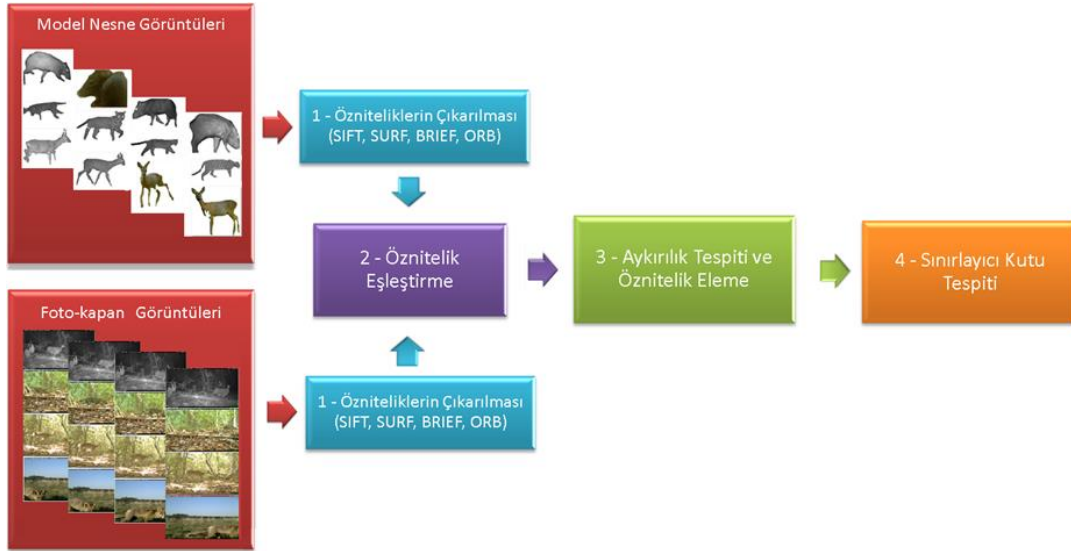
Literatürde yerel öznelikler ile yapılan çalışmalarda, genellikle yerel öznelik dönüşümleri birebir eşleştirilmek yerine, bunların

istatistiki bilgilerini içeren görsel kelime çantası yaklaşımları tercih edilmektedir. Böyle uygulamalarda görüntüdeki öznelikler görsel kelime çantası yöntemi ile gruplandırılıp, k-en yakın komşu, destek vektör makineleri ve rastgele orman gibi sınıflandırıcılarda kullanılarak, kayan pencere tabanlı yöntemle, her pencerede ilgili nesnenin olup olmadığına bakılarak tespit işlemi gerçekleştirilmektedir. Bizim bu çalışmamızda, kırılmış ve arka plandan ayrılmış nesne görüntülerinin SIFT, SURF, BRIEF ve ORB yerel öznelik dönüşümleri, kaba kuvvet (brute-force) ve k-en yakın komşu öznelik eşleştirme yöntemlerinde kullanılarak foto-kapan görüntülerinde otomatik nesne tespit işlemi gerçekleştirilmiştir. Hatalı öznelik eşleşmelerinin elenmesi amacıyla yüzdelik ve medyan tabanlı aykırılık tespiti ve k-en yakın komşu öznelik eleme işlemi uygulanmıştır. Çalışmada öznelik dönüşüm yöntemleri ile elde edilen nesne tespit başarıları, eşleşen öznelik sayıları, sınırlayıcı kutu büyüklükleri, elenen öznelik sayıları ve bunların nesne tespit başarısına olan etkileri incelenmiştir.

Çalışmanın 2. bölümünde kullanılan veri seti, yerel öznelik dönüşüm yöntemleri, öznelik eşleştirme yöntemleri ve öznelik eleme işlemleri anlatılmış olup, 3. bölümde çalışmadan elde edilmiş bulgular ifade edilmiş ve yorumlanmıştır, 4. bölümde ise sonuçlar tartışılmış ve sonraki aşamalarda yapılması planlanan çalışmalar ifade edilmiştir.

## 2. Materyal ve Metot

Fotokapan görüntülerinde yerel öznelikler kullanılarak nesne tespit için kullandığımız dört aşamadan oluşan sistem modeli Şekil 1'de gösterilmiştir. Öncelikle Şekil 2'de örnek görüntüleri gösterilen veri setinde tespit edilecek nesnelere ait model görüntüler, Şekil 3'te görüldüğü üzere kırılarak ve arka plandan ayrılarak elde edilmiştir. Birinci aşamada veri setindeki görüntülerin SIFT, SURF, BRIEF ve ORB yöntemleri ile öznelikleri çıkarılmaktadır. İkinci aşamada görüntülerden elde edilmiş yerel öznelikler kaba kuvvet (brute-force) ve k-en yakın komşu tabanlı yöntem ile eşleştirilmektedir. Üçüncü aşamada yüzdelik veya ortalama tabanlı aykırılık tespiti ile k-en yakın komşu öznelik eleme yöntemleri kullanılarak hatalı eşleşmelerin giderilmesi ve elenmesi sağlanmaktadır. Dördüncü aşamada eleme işlemi sonucunda kalan özneliklerin konum bilgileri kullanılarak tespit edilen nesneye ait sınırlayıcı kutular belirlenmektedir.



Şekil 1: Sistem modeli

## 2.1. Veri Seti

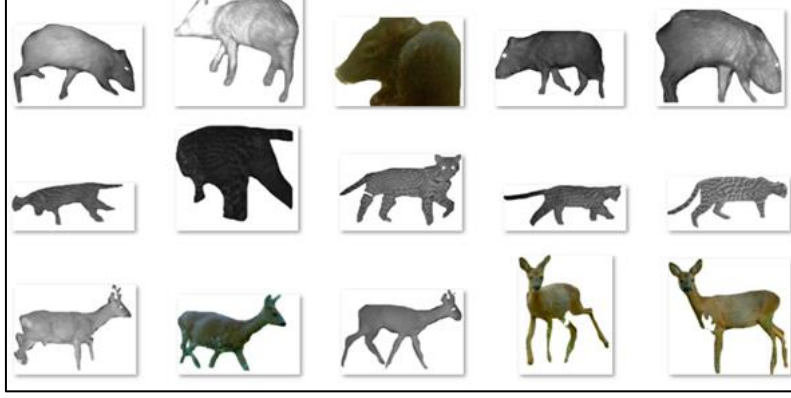
Çalışmamızda Kays ve arkadaşlarının 2010 yılında yaptığı çalışmada (Kays vd., 2010) elde edilen veri seti kullanılmıştır. Veri setinde 20 farklı hayvan türüne ait 2048x1536 piksel boyutlarında toplam 1016 görüntü bulunmaktadır. Şekil 2’de veri setinden rasgele seçilmiş gece ve gündüz kaydedilmiş

görüntüler gösterilmektedir. Gece görüntüleri kızıl ötesi flaş kullanılarak kaydedilmiştir. Veri setindeki görüntülerde nesnelere farklı ortam, uzaklık ve açılardan elde edilmiş olup tam veya parçalı olabilmektedir. Nesne tespit işleminde her tetikleme için 1 model olmak üzere 20 farklı türe ait 79 model nesne görüntüsü kullanılmıştır.



Şekil 2: Veri setine ait örnek görüntüler





Şekil 3: Nesnelere ait kırılmış ve arka plandan ayrılmış örnek görüntüler

## 2.2. Yerel Öznitelik Çıkarma

Öznitelik tespiti, görüntü bilgisinin anlamlandırılması ve bu noktada belirtilen tipte bir görüntü özelliğinin olup olmadığını görmek için her bir görüntü noktası için yerel bir karar verme işlemidir. Yerel öznitelikler renk, doku, şekil gibi düşük seviyeli özniteliklerden farklı olarak görüntüdeki dönme, ölçek, aydınlatma, gürültü ve afin dönüşümü için dayanıklı olmalıdır (Karami vd., 2017). Bir dizi yerel öznitelik, nesnelere ve sahneleri tanıyabilmek için bölütleme işlemine ihtiyaç duyulmadan sağlam bir görüntü temsili olarak kullanılabilir (Tuytelaars ve Mikolajczyk, 2008). Çalışmamızda bu yerel öznitelik dönüşüm yöntemlerinden SIFT, SURF, BRIEF ve ORB kullanılmıştır.

### 2.2.1. SIFT (Scale Invariant Feature Transform)

Ölçek Değişmez Öznitelik Dönüşümü (SIFT), 1999 yılında Lowe tarafından geliştirilen bir öznitelik detektörüdür (Lowe, 1999). SIFT detektörü ve tanımlayıcısı nesne tanıma (Lowe, 1999), görüntü birleştirme (Snavely vd., 2008), görsel haritalama (Se vd., 2002), vb. görsel özellikleri kullanan bir dizi uygulamada oldukça başarılı olduğu kanıtlanmıştır (Rublee vd., 2011). SIFT'in nesne tanıma uygulamalarında çok etkili olduğu kanıtlanmış olsa da, büyük bir hesaplama yükü gerektirmesi gibi özellikle gerçek zamanlı uygulamalar için dezavantaja sahiptir (Karami vd., 2017; Lowe, 1999; Rublee vd., 2011; Karami vd., 2017).

Her SIFT özniteliği, görüntü dönüştürme, ölçekleme ve rotasyona göre değişmeyen ve aydınlatma ve yerel görüntü deformasyonlarındaki değişikliklere kısmen değişmeyen bir şekilde yerel görüntü ölçümlerinin bir vektörünü temsil

etmektedir (Lowe, 2001). SIFT algoritmasının 4 temel adımı vardır: 1) tepe noktası seçimi, 2) anahtar nokta yerelleştirilmesi ve elenmesi, 3) yönlendirme ataması ve 4) anahtar noktaların tanımlanmasıdır. Birinci aşamada, Gaussian'ın Farkı (DoG) kullanılarak bir ölçek uzayı hesaplanmaktadır. Bu aşamada, Gaussian'ların farkının maksimum veya minimumu olan yerleri arayarak ölçek alanındaki anahtar konumları tanımlanmaktadır. İkinci aşama, anahtar noktaların (keypoint) yerleştirildiği ve adayların düşük karşıtlığa sahip noktalarının, iyi betimleyici özellikleri olmadığı için elendiği işlemidir. Üçüncü aşama, yerel görüntü eğimi temeline dayanan anahtar noktalara bir yönelim değeri atanmasıdır. Yönelim histogramlarına ait en yüksek bir veya birkaç tepelik öznitelik noktası için yönelim vektörü olarak atanmaktadır. Bu aşamada gradyan yönelimlerinin histogramı kullanılarak dönmeye karşı değişmezlik kazandırılmaktadır. Son aşama ise, görüntü eğim büyüklüğü ve yön belirlenmesine dayanan her bir anahtar noktası için yerel görüntü tanımlayıcısının hesaplanmasıdır. Her bir öznitelik noktası için, etrafındaki 4x4'lük bölgede 8 bölmeli bir histogram elde edilerek 128 (4x4x8) boyutlu bir tanımlayıcı belirlenmektedir. Bu 128 boyutlu vektör, birim uzunluğa dönüştürülerek aydınlatma değişimine karşı kısmen dayanıklılık elde edilmektedir. Ayrıca SIFT öznitelikleri, görüntü gradyanı konumlarını bulanıklaştırarak, afin veya 3D görünüm gibi yerel çeşitliliklere kısmi değişmezlik sağlamaktadır (Lowe, 1999).

### 2.2.2. SURF (Speeded Up Robust Features)

SURF, Hessian matrisine dayanan bir ölçek-değişmez öznitelik detektörüdür. SIFT'in hızlandırılmış bir yaklaşımı olan SURF tekniği, tespit edilen noktaların kalitesini düşürmeden SIFT'den daha hızlı çalışmaktadır (Bay vd., 2006;

Karakuş ve Karabörk, 2014). Konum ve ölçeği seçmek için farklı bir ölçü kullanmak yerine, her ikisi için de Hessian detektörü kullanılmaktadır. Hessian matrisi sayısal bir alanın ikinci dereceden kısmi türevlerinden oluşan kare matristir ve kabaca bir dizi kutu tipi filtrenin kullanılmasına benzemektedir (Tuytelaars ve Mikolajczyk, 2008). LoG (Laplacian of Gaussian) kullanılarak, ikinci dereceden türevi alınan görüntüde, ölçek uzayında değişmeyen anahtar noktalar elde edilmektedir. SURF'te SIFT'ten farklı olarak, çok işlem gerektiren ikinci dereceden Gauss türevleri yerine, bir kutu filtre kullanılmaktadır ve bu kutu filtredeki hesaplama işlemi, bütünleşik görüntü kullanılarak kolayca yapılabilmektedir (Tuytelaars ve Mikolajczyk, 2008). Öznitelik tanımlayıcısı, ilgi noktasını tanımlamak için 20x20 piksellik komşu bölgelerin yatay ve dikey yönde Haar dalgacık tepkilerinin dağılımı kullanılmaktadır. Baskın yönelim, 60 derecelik bir kaydırma penceresindeki tüm tepkilerin toplamı hesaplanarak tahmin etmektir. Dalgacık tepkisi, herhangi bir ölçekte çok kolay bir şekilde bütünleşik görüntüler kullanılarak bulunabilmektedir. Bütünleşik görüntü, görüntüdeki piksel değerlerinin dikdörtgen ızgaralar şeklinde hızlı ve etkili olarak toplanabilmesi için kullanılan veri yapısıdır. Ayrıca SURF yönteminde, Laplacian'ın işaretine dayanan eşleşme hızını azaltan ve sağlamlığı arttıran bir indeksleme yöntemi kullanılmaktadır (Bay vd., 2006). Bu indekslemede, temel ilgi noktası için Laplacian (Hessian Matrix) izi kullanılmaktadır. Bu değerler algılama sırasında zaten hesaplandığından ek hesaplama maliyeti getirmemektedir. Laplacian'ın işareti, koyu arka plandaki parlak lekeleri ters durumlardan ayırmaktadır. Eşleştirme aşamasında, ilgi noktasının yalnızca aynı türde kontrast özellikleri varsa özellikleri karşılaştırılmaktadır. Bu minimum bilgi, tanımlayıcının performansını düşürmeden daha hızlı eşleştirilmesine olanak tanımaktadır. SURF, bulanıklaştırma ve döndürmeye karşı dayanıklı olmasına rağmen, bakış açısı değişikliği ve aydınlatma değişimine karşı dayanıklılık içermemektedir (Bay vd., 2006).

### 2.2.3. BRIEF (Binary Robust Independent Elementary Features)

BRIEF (İkili Sağlam Bağımsız Temel Öznitelikler), öznitelik detektörleriyle birleştirilebilen genel amaçlı bir öznitelik tanımlayıcısıdır. Tipik fotometrik ve geometrik görüntü dönüşümlerine karşı dayanıklıdır. BRIEF, SIFT için neredeyse benzer eşleştirme performansına sahip SIFT'den daha az

karmaşıklık gerektiren başka bir alternatiftir (Calonder vd., 2010). BRIEF yönteminde, yumuşatılmış görüntü yamalarındaki pikseller arasında basit ikili testleri kullanılmaktadır. Performansı, aydınlatma değişimine karşı dayanıklılık, bulanıklık ve perspektif bozulma gibi birçok açıdan SIFT'e benzemektedir. Bununla birlikte, düzlem içi rotasyona karşı dayanıksızdır (Rublee vd., 2011; Calonder vd., 2010). BRIEF içerisinde bir öznitelik veya anahtar nokta detektörü içermemektedir, yalnızca tespit edilmiş öznitelik veya anahtar noktaların tanımlanması için kullanılmaktadır. Çalışmamızda BRIEF öznitelik tanımlayıcısı, yıldız öznitelik detektörü (star feature dedector) ile birlikte kullanılmıştır.

### 2.2.4. ORB (Oriented and Rotated BRIEF)

ORB (Yönlendirilmiş FAST ve Döndürülmüş BRIEF), FAST (Features from Accelerated Segment Test) anahtar noktası dedektörünün ve BRIEF tanımlayıcısının birtakım modifikasyonları ile birleşmesidir

(Karami vd., 2017; Rublee vd., 2011; Bay vd., 2006). Her iki teknik de iyi performansları ve düşük işlem maliyeti nedeniyle caziptir (Rublee vd., 2011). FAST, gerçek zamanlı uygulamalarda kullanılabilen, diğer köşe detektöründen hızlı olan, gürültüye dayanıklı bir detektördür (Karami vd., 2017). SIFT ve SURF anahtar nokta detektörleri bir yönelim operatörü içermesine rağmen, FAST'de bu işlem bulunmamaktadır (Rublee vd., 2011). FAST algoritması yalnızca köşelerin tespit edileceği dairenin yarıçapı bilgisini parametre olarak almakta ve ORB yönteminde FAST ile birlikte tespit edilen köşelerin ölçülendirilebilmesi için Harris Köşe Ölçüsü, köşe yöneliminin belirlenmesi için ise yoğunluk ağırlık merkezi (intensity centroid) yöntemi kullanılmaktadır. Yoğunluk ağırlık merkezi yönteminde standart sabitler kullanılarak köşe yönelimleri tespit edilmektedir. Tespit edilen anahtar noktalar, BRIEF kullanılarak tanımlanmaktadır (Rublee vd., 2011).

### 2.3. Yerel Öznitelik Çıkarma

Foto-kapan görüntülerinde nesnelere tespit edilmesi için kullanılan öznitelik eşleştirme algoritması Algoritma 1'de gösterilmiştir. Bu algoritmada model nesne görüntülerine ait çıkarılmış yerel öznitelikler, foto-kapan görüntülerindeki öznitelikler ile eşleştirilmektedir. Eşleştirilme işleminde, görüntülerdeki bütün yerel öznitelikler arasında Mahalanobis, Manhattan (denklem 1), Öklid (denklem 2) veya Hamming (denklem 3) uzaklığının hesaplanması ile elde

edilen mesafe değerleri kullanılmaktadır. Elde edilen bu mesafe değerleri karşılaştırılarak en uygun eşleştirmeye karar verilmektedir.

$$d(f_a, f_b) = \sum |f_a - f_b| \quad (1)$$

$$d(f_a, f_b) = \sum (f_a - f_b)^2 \quad (2)$$

$$d(f_a, f_b) = \sum \text{XOR}(f_a, f_b) \quad (3)$$

Kaba Kuvvet eşleştirme işlemi (Brute Force Matcher), ilk görüntüdeki bir öznitelik tanımlayıcısının, ikinci görüntüdeki tüm özniteliklerle olan mesafesi denklem (1), denklem (2) ve denklem (3)'e göre hesaplanmış ve en yakın olan eşleşme tespit edilmiştir. Bu eşleştirici iki opsiyonel parametre gerektirmektedir. Birincisi, kullanılacak mesafe ölçümünü

belirleyen Manhattan, Öklid ve Hamming değerleridir. Varsayılan olarak uygulanan Öklid, SIFT ve SURF yöntemleri için kullanılmıştır. ORB, BRIEF gibi ikili tabanlı tanımlayıcılar için Hamming mesafesi tercih edilmiştir. İkinci parametre ise, varsayılan olarak kapalı olan çapraz kontrol değişkenidir. Bu özellik kullanılarak, birinci görüntüdeki özniteliklerin, ikinci görüntüdeki en yakın öznitelik ile eşleştirdiği gibi, ikinci görüntüdeki özniteliklerin birinci görüntüdeki en yakın eşleştirme olup olmadığını da kontrol edilmiştir. k-en yakın komşu eşleştiricisinde ise bir tanımlayıcıya karşılık girilen parametre kadar en yakın komşuluğa sahip 2 tanımlayıcı geriye döndürülmektedir. Bu iki tanımlayıcı arasındaki mesafe değeri öznitelik eleme işleminde kullanılmıştır.

#### **Algoritma 1:** Öznitelik eşleştirme algoritması

*Veri Setindeki Her Görüntü İçin:*

*Bütün Model Nesne Görüntüleri İçin:*

*Tanımlayıcı 1 = yerel öznitelikleri tanımla (Model Nesne Görüntüsü)*

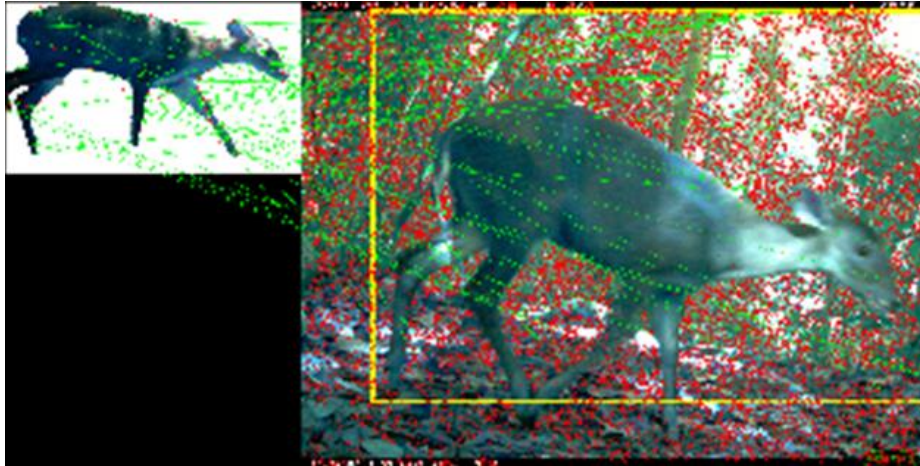
*Tanımlayıcı 2 = yerel öznitelikleri tanımla (Foto-kapan Görüntüsü)*

*Eşleşmeler = kaba kuvvet eşleştirici(Tanımlayıcı 1, Tanımlayıcı 2)*

#### **2.4. Aykırılık Tespiti ve Öznitelik Eleme**

Öznitelik eşleştirme işlemi görüntülerde benzer iki nesne arasında mutlak doğrulukta bir eşleşme ve tespit sağlayamamaktadır. Şekil 4'te gösterildiği üzere kırılmış görüntü ile foto-kapandan elde edilen görüntü eşleşmesinde yeşil renkli çizgiler eşleştirilmiş öznitelik noktalarını,

kırmızı renkli pikseller ise eşleşmemiş öznitelik noktalarını temsil etmektedir. Görüntüdeki en benzer özniteliklerin eşleştirilmesi işlemi, bazen nesneye ait olmayan kısımların hatalı bir biçimde nesneye ait olarak etiketlenmesine, bu nedenle tespit işleminin hatalı yapılmasına neden olmaktadır.



**Şekil 4:** Hatalı ve hatasız öznitelik eşleştirme sonucuna ait örnek görüntü

Şekil 4'teki sarı sınırlayıcı kutunun nesnenin üzerindeki arka plan bölgesini de kapsadığı, yani bazı özneliklerin hatalı olarak arka plandaki özneliklerle eşleştiği anlaşılmaktadır. Bu hatalı özneliklerin giderilmesi amacıyla çalışmamızda kaba kuvvet eşleştirici ile birlikte yüzdelik ve ortalama tabanlı aykırılık tespiti yöntemleri kullanılmıştır. Ayrıca k-en yakın komşu

eşleştirici, farklı eşik değerleri ile birlikte kullanılarak öznelik eleme işlemleri uygulanmıştır. Algoritma 2'de gösterildiği üzere, k-en yakın komşu eşleştiricisinde, tanımlayıcılardan eşleştirilmiş her m ve n eşleşme çiftinin uzaklıkları oranlanmış ve bu oran belirli bir değerin üzerinde ise eşleşme kabul edilmiştir.

### Algoritma 2: k-en yakın komşu öznelik seçme ve eleme algoritması

Veri Setindeki Her Görüntü İçin:

Bütün Model Nesne Görüntüleri İçin:

Tanımlayıcı 1 = yerel öznelikleri tanımla (Model Nesne Görüntüsü)

Tanımlayıcı 2 = yerel öznelikleri tanımla (Foto-kapan Görüntüsü)

Eşleşmeler = k-en yakın komşu eşleştirici(Tanımlayıcı 1, Tanımlayıcı 2, k=2)

Her m, n eşleşmesi için:

Eğer m'nin mesafesi < [0.80, 0.85, 0.90, 0.95, 0.99]\*n'nin mesafesi ise

Eşleşmeyi kabul et

Değil ise

Eşleşmeyi ele

Ortalama tabanlı aykırılık tespitinde kullanılan ortanca mutlak sapmanın (Median Absolute Deviation - MAD) nasıl hesaplandığı denklem (4)'te ifade edilmektedir. Hesaplanan ortanca mutlak sapma değerlerinden, parametrede girilen eşik değerden büyük olanlar aykırılık olarak tespit edilmiştir.

$$MAD = median ( |x - median(x)| ) \quad (4)$$

Yüzdelik tabanlı aykırılık tespitinde ise sıralı bir dizide girilen parametredeki konumlara karşılık gelen elemanlarının değerlerini kullanarak maksimum ve minimum eşik değerleri belirlenmiştir. Dizide bu minimum ve maksimum değerlerin dışındaki elemanlar aykırılık olarak tespit edilmiştir.

### 2.5. Sınırlayıcı Kutu Tespiti

Sınırlayıcı kutu tespiti, aykırılık tespiti veya öznelik eleme işlemi sonrasında kalan eşleştirilmiş özneliklerin yerel (x,y) konum bilgilerinin kullanılması ile elde edilmiştir. Eşleşen özneliklerin nesneyi temsil ettiği varsayıldığından, bu eşleşen özneliklerin en küçük ve en büyük (x,y) noktalarının nesneyi kapsadığı kabul edilmiştir. Sınırlayıcı kutu tespitinde başarı ölçümü amacıyla denklem (5)'te gösterilen birleşimlerin kesişimi yöntemi (IoU-Intersection Over Union) (Zitnick ve Dollar, 2014) 0.5 eşik değer parametresi ile kullanılmıştır.

İki sınırlayıcı kutu için konumlandırma işlemi, bu iki kutunun kesiştiği bölgenin piksellerinin, iki kutunun birleşimi olan bölgenin piksellerine oranı 0.5 ve üzerinde ise başarılı kabul edilmiştir. Denklem (5)'te IoU sınırlayıcı kutunun başarı değerini,  $B_P$  tahmin edilmiş sınırlayıcı kutuyu,  $B_G$  ise gerçek sınırlayıcı kutuyu ifade etmektedir.

$$IoU = \frac{B_P \cap B_G}{B_P \cup B_G} \quad (5)$$

### 3. Bulgular

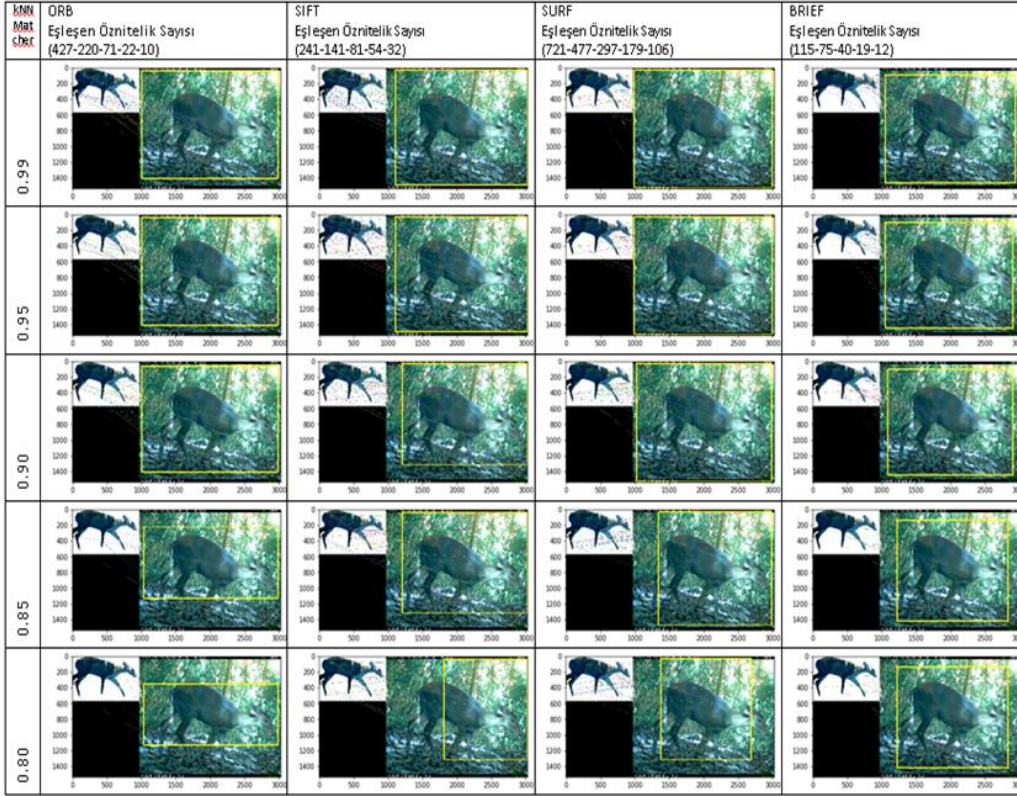
Çalışmamızda SIFT, SURF, BRIEF ve ORB yerel öznelikleri, 2 farklı öznelik eşleştirme yönteminde kullanılarak foto-kapan görüntülerinde nesne tespiti gerçekleştirilmiş olup, kaba kuvvet eşleştirici ile elde edilen sonuçlar Tablo 2 ve Tablo 5'te, k-en yakın komşu eşleştirici ile elde edilen sonuçlar ise Tablo 8'de gösterilmiştir. Aykırılık tespiti ve öznelik eleme işlemleri, hatalı eşleşmelerin giderilmesi ve sınırlayıcı kutu tespitinin daha doğru yapılmasını sağlamıştır. Hatalı öznelik eşleşmelerinin elenmesinden sonra, tespit edilen sınırlayıcı kutuların boyutları azalmış ve daha yüksek tespit başarısı elde edilmiştir. Eşleştirilen öznelik sayısı ile sınırlayıcı kutu büyüklüğünün ilişkili olduğu ve tespit başarısını etkilediği gözlemlenmiştir. Tablo 1'de öznelik tanımlayıcılarının boyutları gösterilmektedir. Tanımlayıcıların, özneliklerin eşleştirilmesinde



etkili olduğu ve bu nedenle bu boyutların tespit sonuçlarını arttırdığı anlaşılmaktadır.

Şekil 5'te k-en yakın komşu eşleştiricide, birinci görüntüdeki öznitelik tanımlayıcısı ile ikinci görüntüde bu öznitelik ile eşleşen en yakın 2 öznitelik tanımlayıcısının mesafelerinin birbirine olan oranlarının 0.99, 0.95, 0.90, 0.85, 0.80 eşik

değerlerine göre özniteliklerin elenmesi ve sınırlayıcı kutudaki değişiklikler gösterilmektedir. Bu oran azaldıkça sınırlayıcı kutular gerçek değerlerine yaklaşmakta ve tespit başarısı artmaktadır. k-en yakın komşu eşleştiricisinde, en yakın 2 komşunun elde edilmesi amacıyla K=2 parametresi kullanılmıştır.



Şekil 5: k-en yakın komşu eşleştirme işleminde eşik değere göre öznitelik eleme sonrası sınırlayıcı kutu değişimi

Tablo 1: Özniteliklerin tanımlayıcı boyutları

	SIFT	SURF	BRIEF	ORB
Tanımlayıcı Uzunluğu	128	64	256	256
Tanımlayıcı Boyutu (Byte)	512 Byte	256 Byte	32 Byte	32 Byte

0.65	88.87	87.55	97.22	92.50
0.60	90.60	88.34	98.68	94.22
0.55	92.98	90.07	99.60	95.41
0.50	95.23	91.52	99.66	96.60

Tablo 1: Yüzdeler tabanlı aykırılık tespiti sonrası nesne tespit başarı oranları

Yüzdeler Parametre	ORB	BRIEF	SIFT	SURF
0.95	80.00	78.41	89.80	81.37
0.90	81.59	79.47	90.99	83.36
0.85	83.05	80.66	92.19	86.54
0.80	84.64	82.91	93.38	87.86
0.75	85.70	83.97	94.57	88.79
0.70	87.02	85.83	95.76	89.72

Yüzdeler tabanlı öznitelik eleme işleminde, Tablo 2, Tablo 3 ve Tablo 4'te 1. sütunda gösterilen öznitelik eleme parametre değeri azaldıkça, Tablo 4'teki ortalama eşleşen öznitelik sayılarının azaldığı, buna bağlı olarak Tablo 3'teki ortalama sınırlayıcı kutuların gerçek değerlere yaklaştığı ve bunun sonucunda Tablo 2'de nesne tespit başarısında artış elde edildiği anlaşılmaktadır. Tablo 2, Tablo 3 ve Tablo 4'te görüldüğü üzere, yüzdeler tabanlı aykırılık tespiti işleminde, eşleşmelerin ortalama olarak yaklaşık yarısının elendiği ve bu durumun nesne tespit başarısını %10-%15 arası arttırdığı anlaşılmaktadır.

**Tablo 3:** Yüzdeler tabanlı aykırılık tespiti sonrası ortalama sınırlayıcı kutu boyutları (piksel)

Yüzdeler Parametre	ORB	BRIEF	SIFT	SURF
0.95	473	418	498	544
0.90	423	383	455	486
0.85	382	354	416	443
0.80	350	326	381	404
0.75	322	303	347	372
0.70	294	280	313	344
0.65	270	259	280	319
0.60	249	235	280	295
0.55	228	213	249	269
0.50	209	193	218	245

**Tablo 2:** Yüzdeler tabanlı aykırılık tespiti sonrası ortalama eşleşen öznitelik sayıları

Yüzdeler Parametre	ORB	BRIEF	SIFT	SURF
0.95	1182	197	471	1490
0.90	1121	187	447	1422
0.85	1062	177	422	1354
0.80	1000	167	397	1286
0.75	939	158	373	1220
0.70	877	148	348	1151
0.65	816	137	323	1084
0.60	754	128	323	1008
0.55	693	117	298	919
0.50	633	107	273	839

Bütün sonuçlarda en yüksek nesne tespit başarısı bütün öznitelik eleme yöntemlerinde SIFT özniteliği ile elde edilmiş olup, ona en yakın sonuçlar ise SURF özniteliğinde elde edilmiştir. Tablo 1’de görüldüğü üzere, en yüksek boyutlu öznitelik tanımlayıcısı olan SIFT ile en yüksek tespit başarısı elde edilmiştir. Diğer dönüşüm yöntemleri, öznitelik dönüştürme veya eşleştirme işlemini hızlandırmayı veya öznitelikleri daha az boyutta temsil etmeyi amaçladığından daha düşük tespit başarısı elde edilmiştir.

Tablo 7’den anlaşıldığı üzere ortalama tabanlı aykırılık tespitinde, öznitelik eşleşmelerinin %35-%40’ı elenmiştir. Yüzdeler tabanlı eleme yöntemine göre daha az öznitelik elenmiş ve buna bağlı olarak Tablo 5’teki gösterilen nesne tespit başarısında Tablo 2’deki yüzdeler tabanlı yöntemle göre daha düşük tespit başarısı elde edilmiştir.

Tablo 6’da görüldüğü üzere ortalama tabanlı aykırılık tespitinde elde edilen sınırlayıcı kutuların yüzdeler tabanlı aykırılık tespitinde elde edilenlere göre daha küçük boyutlu olduğu anlaşılmaktadır.

**Tablo 5:** Ortalama tabanlı aykırılık tespiti sonrası nesne tespit başarı oranları

Ortalama Parametresi	ORB	BRIEF	SIFT	SURF
2.5	78.15	76.69	90.07	82.65
2.0	81.85	81.99	94.30	87.15
1.5	84.90	85.43	96.29	91.26
1.0	87.95	87.81	97.88	93.11

**Tablo 3:** Ortalama tabanlı aykırılık tespiti sonrası sınırlayıcı kutu boyutları (piksel)

Ortalama Parametresi	ORB	BRIEF	SIFT	SURF
2.5	389	351	470	423
2.0	340	314	428	376
1.5	277	260	351	314
1.0	201	187	229	228

**Tablo 4:** Ortalama tabanlı aykırılık tespiti sonrası eşleşen öznitelik sayıları

Ortalama Parametresi	ORB	BRIEF	SIFT	SURF
2.5	1060	177	462	1300
2.0	985	167	440	1216
1.5	872	149	393	1074
1.0	682	116	299	851

k-en yakın komşu eşleştirme işleminde elde edilen sonuçlar incelendiğinde, Tablo 9’daki sınırlayıcı kutu boyutlarının, olması gereken nesne boyutlarından daha küçük olduğu ve bu nedenle düşük tespit başarısı elde edildiği anlaşılmaktadır.

**Tablo 5:** k-en yakın komşu tabanlı öznitelik eleme sonrası nesne tespit başarı oranları (%)

kNN Parametresi	ORB	BRIEF	SIFT	SURF
0.95	64.50	55.76	73.91	64.64
0.90	73.64	69.80	80.79	71.13
0.85	75.36	70.86	80.66	75.23
0.80	75.63	72.05	82.38	77.88

**Tablo 6:** k -en yakın komşu tabanlı öznelik eleme sonrası sınırlayıcı kutu boyutları (piksel)

kNN Parametresi	ORB	BRIEF	SIFT	SURF
0.95	207	191	255	224
0.90	203	185	250	220
0.85	196	177	237	204
0.80	166	139	184	169

Tablo 10'da gösterildiği üzere, k-en yakın komşu öznelik eleme işleminde 1. sütundaki öznelik yakınlık mesafe parametresindeki her 0.05'lik azalmaya karşılık öznelikler oldukça yüksek oranda elenmişlerdir. Bu işlemde özneliklerin %40 ile %50'sinin elenmesine karşılık, Tablo 9'daki sınırlayıcı kutularda %7 ile %20 oranlarında küçülme gerçekleşmiş, bunun sonucunda ise Tablo 8'deki nesne tespit başarıları %5 ile %17 arasında artmıştır.

**Tablo 7:** k-en yakın komşu tabanlı öznelik eleme sonrası eşleşen öznelik sayısı

kNN Parametresi	ORB	BRIEF	SIFT	SURF
0.95	1237	206	494	1559
0.90	608	125	274	977
0.85	238	66	133	548
0.80	32	20	34	173

Eşleşen öznelikler sayıları ile nesne tespit başarı oranı karşılaştırıldığında, en iyi sonuç Tablo 11'de görüldüğü üzere BRIEF öznelik tanımlayıcısında elde edilmiştir. Ayrıca k-en yakın komşu eşleştirme ve eleme işleminin, kaba kuvvet eşleştiricisi yüzdelik ve ortalama aykırılık tespitine göre, eşleşen öznelik sayıları ile nesne tespit başarılarına oranına göre daha iyi olduğu Tablo 11'den anlaşılabilir.

**Tablo 8:** Nesne tespit başarı oranları ile eşleşen öznelik sayıları ilişkisi (tespit başarı oranı / eşleşen öznelik sayısı \* 100)

Aykırılık Tespiti ve Öznelik Eleme	ORB	BRIEF	SIFT	SURF
Yüzdelik	7.3	42.6	20.1	6.2
Ortalama	6.1	34.4	16.3	4.8
kNN	8.3	49.6	21.2	7.2

#### 4. Sonuç ve Tartışma

Bu çalışmada SIFT, SURF, BRIEF ve ORB yerel öznelikleri, kaba kuvvet ve k-en yakın komşu öznelik eşleştirme yöntemleri kullanılarak foto-kapan görüntülerinde nesne tespit işlemi gerçekleştirilmiştir. Bu yerel öznelik tanımlayıcılarından elde edilen tespit sonuçları ile birlikte, eşleşen öznelik sayıları ve elde edilen sınırlayıcı kutuların boyutları değerlendirilmiştir. Öznelik eşleşme işlemindeki hataların, nesne tespitindeki sınırlayıcı kutuların büyük olmasına neden olduğu ve nesne tespit başarısını düşürdüğü anlaşılmaktadır. Ayrıca eşleşen öznelik sayıları ile nesne tespit başarılarının ilişkisi çıkarılarak, öznelik tanımlayıcılarının eşleşen özneliklerin nesne tespit başarısına etkisi incelenmiştir.

Yerel öznelik tanımlayıcılarının nesne tespit etmede, çok sayıda görüntü barındıran veri setleri ile eğitime ihtiyaç duymadan, birkaç model görüntü kullanılarak başarılı sonuçlar elde edilebileceği gözlemlenmiştir.

Sonraki çalışmalarda yerel öznelik tanımlayıcılarının işlem ve bellek gereksinimleri incelenerek dağıtık ve gömülü sistemler üzerinde uygulanabilirliğinin araştırılması planlanmaktadır.

#### Teşekkür

Bu çalışma Atatürk Üniversitesi Bilimsel Araştırma Projelerini destekleme programı kapsamında FYL-2018-6623 proje kodu ile desteklenmiştir.

#### Kaynaklar

- Andavarapu, N., & Vatsavayi, V. K. (2017). Wild-Animal Recognition in Agriculture Farms Using W-COHOG for Agro-Security. *International Journal of Computational Intelligence Research*, 13(9), 2247-2257.
- Bay, H., Tuytelaars, T., & Van Gool, L. (2006, May). Surf: Speeded up robust features. In *European conference on computer vision* (pp. 404-417). Springer, Berlin, Heidelberg.
- Calonder, M., Lepetit, V., Strecha, C., & Fua, P. (2010, September). Brief: Binary robust independent elementary features. In *European conference on computer vision* (pp. 778-792). Springer, Berlin, Heidelberg.
- Chen, Q., Song, Z., Dong, J., Huang, Z., Hua, Y., & Yan, S. (2015). Contextualizing object detection and classification. *IEEE transactions on pattern analysis and machine intelligence*, 37(1), 13-27.

- Karakuş, P., & Karabörk, H. "Surf Algoritması Kullanılarak Uzaktan Algılama Görüntülerinin Geometrik Kaydı". 5. Uzaktan Algılama-Cbs Sempozyumu (Uzal-Cbs 2014), 14-17 Ekim 2014, İstanbul.
- Karami, E., Prasad, S., & Shehata, M. (2017). Image matching using SIFT, SURF, BRIEF and ORB: performance comparison for distorted images. arXiv preprint arXiv:1710.02726.
- Karami, E., Shehata, M., & Smith, A. (2017). Image Identification Using SIFT Algorithm: Performance Analysis against Different Image Deformations. arXiv preprint arXiv:1710.02728.
- Kays, R., Tilak, S., Kranstauber, B., Jansen, P. A., Carbone, C., Rowcliffe, M. J., ... & He, Z. (2010). Monitoring wild animal communities with arrays of motion sensitive camera traps. arXiv preprint arXiv:1009.5718.
- Khorrani, P., Wang, J., & Huang, T. (2012, November). Multiple animal species detection using robust principal component analysis and large displacement optical flow. In Proceedings of the 21st International Conference on Pattern Recognition (ICPR), Workshop on Visual Observation and Analysis of Animal and Insect Behavior (pp. 11-15).
- Lisin, D. A., Mattar, M. A., Blaschko, M. B., Learned-Miller, E. G., & Benfield, M. C. (2005, June). Combining local and global image features for object class recognition. In Computer vision and pattern recognition-workshops, 2005. CVPR workshops. IEEE Computer society conference on (pp. 47-47). IEEE.
- Lowe, D. G. (1999). Object recognition from local scale-invariant features. In Computer vision, 1999. The proceedings of the seventh IEEE international conference on (Vol. 2, pp. 1150-1157). IEEE.
- Lowe, D. G. (2001). Local feature view clustering for 3D object recognition. In Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on (Vol. 1, pp. I-I). IEEE.
- Meek, P. D., Ballard, G., Claridge, A., Kays, R., Moseby, K., O'brien, T., & Townsend, S. (2014). Recommended guiding principles for reporting on camera trapping research. *Biodiversity and conservation*, 23(9), 2321-2343.
- Murphy, K., Torralba, A., Eaton, D., & Freeman, W. (2006). Object detection and localization using local and global features. In *Toward Category-Level Object Recognition* (pp. 382-400). Springer, Berlin, Heidelberg.
- Nguyen, H., Maclagan, S. J., Nguyen, T. D., Nguyen, T., Flemons, P., Andrews, K., ... & Phung, D. (2017, October). Animal recognition and identification with deep convolutional neural networks for automated wildlife monitoring. In *Data Science and Advanced Analytics (DSAA), 2017 IEEE International Conference on* (pp. 40-49). IEEE.
- Norouzzadeh, M. S., Nguyen, A., Kosmala, M., Swanson, A., Palmer, M. S., Packer, C., & Clune, J. (2018). Automatically identifying, counting, and describing wild animals in camera-trap images with deep learning. *Proceedings of the National Academy of Sciences*, 201719367.
- Rublee, E., Rabaud, V., Konolige, K., & Bradski, G. (2011, November). ORB: An efficient alternative to SIFT or SURF. In *Computer Vision (ICCV), 2011 IEEE international conference on*(pp. 2564-2571). IEEE.
- Se, S., Lowe, D., & Little, J. (2002). Mobile robot localization and mapping with uncertainty using scale-invariant visual landmarks. *The international Journal of robotics Research*, 21(8), 735-758.
- Snavely, N., Seitz, S. M., & Szeliski, R. (2008, June). Skeletal graphs for efficient structure from motion. In *CVPR (Vol. 1, p. 2)*.
- Tuytelaars, T., & Mikolajczyk, K. (2008). Local invariant feature detectors: a survey. *Foundations and trends® in computer graphics and vision*, 3(3), 177-280.
- Yu, X., Wang, J., Kays, R., Jansen, P. A., Wang, T., & Huang, T. (2013). Automated identification of animal species in camera trap images. *EURASIP Journal on Image and Video Processing*, 2013(1), 52.
- Zhang, J., Marszałek, M., Lazebnik, S., & Schmid, C. (2007). Local features and kernels for classification of texture and object categories: A comprehensive study. *International journal of computer vision*, 73(2), 213-238.
- Zhang, Z., He, Z., Cao, G., & Cao, W. (2016). Animal detection from highly cluttered natural scenes using spatiotemporal object region proposals and patch verification. *IEEE Transactions on Multimedia*, 18(10), 2079-2092.
- Zitnick, C. L., & Dollár, P. (2014, September). Edge boxes: Locating object proposals from edges. In *European conference on computer vision* (pp. 391-405). Springer, Cham.