

# **Ölçeklenebilir Hibrit Nesne Tanıma Sistemi**

**Program Kodu: 3501**

**Proje No: 113E496**

Proje Yürütücüsü

**Yrd. Doç. Dr. Mustafa ÖZUYSAL**

Bursiyerler:

Ali KÖKSAL

Furkan Eren UZYILDIRIM

EKİM 2016

İZMİR

# ÖNSÖZ

Mobil donanımlar gündelik hayatın bir parçası olurken Bilgisayar Bilimi ve Mühendisliği alanlarında da yeni uygulama sahaları açılmaktadır. Bunlardan önemli bir tanesi mobil cihaz kamerası yardımıyla nesnelere otomatik olarak tanınması ve uygulama bağlamına göre kullanıcının yönlendirilmesidir. Geçtiğimiz son on senelik dönemde Bilgisayarlı Görü alanındaki gelişmeler nesne tanımanın sunucu destekli olarak mobil cihazlarda rahatlıkla gerçekleştirilmesini sağlamıştır.

Bu raporda sonuçları verilen proje bu sistemlerin sunucu desteği olmadan, doğrudan mobil cihaz üzerinde çalışmasını amaçlamaktadır. Bunun önündeki en önemli engel mobil nesne tanıma uygulamalarında sıkça kullanılan ikilik betimleyici eşleştirme yöntemlerinin sunucuda kullanılan tekniklere göre daha düşük başarımlar göstermesidir. Projenin ana amacı da ikilik betimleyiciler ile daha yüksek eşleştirme ve nesne tanıma başarımlarını gösteren yöntemlerinin geliştirilmesidir.

Proje kapsamında ikilik betimleyicilerin zaafı incelenmiş ve betimleyicilerdeki kamera bakış açısından doğan değişimler istatistiksel olarak modellenerek hem nokta eşleştirme hem de yüzün üzerinde nesneden oluşan veri tabanlarındaki nesne tanıma başarımlarını artıran yeni bir yöntem geliştirilmiştir. Bu yöntemin mobil donanımlarda nesne tanıma konusunda avantajları, sınırları ve gelecekte yapılması planlanan çalışmalar raporda yer almaktadır.

Bu projede yer alan çalışmalar TÜBİTAK tarafından 3501 programı kapsamında 113E496 numaralı proje ile desteklenmiştir.

# İçindekiler

<b>1 GİRİŞ</b>	<b>1</b>
1.1 İkilik Betimleyiciler ile Gerçek Zamanlı Nokta Eşleştirme . . . . .	1
1.2 İkilik Betimleyiciler için Ölçeklenebilir Eşleştirme . . . . .	3
1.3 İkilik Betimleyici Kümelerinin Sıkıştırılması . . . . .	5
1.4 Literatüre Katkı . . . . .	5
<b>2 İLGİLİ LİTERATÜR</b>	<b>7</b>
2.1 İkilik Betimleyici Eşleme Yaklaşımları . . . . .	7
2.2 Anahtar Nokta Eleme Yaklaşımları . . . . .	8
<b>3 YÖNTEM</b>	<b>9</b>
3.1 Betimleyici Değişimlerinin İstatistiksel Modellenmesi . . . . .	9
3.2 İkili Betimleyiciler için Olasılığa Dayalı Karşılaştırma Yöntemi . . . . .	11
3.3 İkili Betimleyiciler için Ölçeklenebilir Eşleme . . . . .	11
3.4 Betimleyici Kümesinin Boyutunun Düşürülmesi . . . . .	13
3.4.1 Tanınma Başarımı Düşük Betimleyicilerin Elenmesi . . . . .	13
3.4.2 Anahtar Noktalar Arasında Betimleyici ve İstatistik Paylaşımı . . . . .	14
<b>4 DENEYLER</b>	<b>16</b>
4.1 Olasılığa Dayalı Karşılaştırma Yönteminin Testleri . . . . .	16
4.1.1 Oxford Veri Kümesi ile Testler . . . . .	16
4.1.2 Üç Boyutlu Nesne Veri Kümesi ile Testler . . . . .	20
4.1.3 Hesaplama Süresi . . . . .	21
4.2 Ölçeklenebilirlik Testleri . . . . .	22
4.2.1 Oxford Veri Kümesi ile Testler . . . . .	22
4.2.2 Orsay Müzesi Veri Kümesi ile Testler . . . . .	22
4.3 Veri Boyutunun Düşürülmesi Testleri . . . . .	25

**5 TARTIŞMA VE SONUÇ**

**26**

**Kaynakça**

**28**

## Şekil Listesi

1	Graffiti veri kümesi bir ve üçüncü görüntüleri için doğru eşlemelerin en yakın komşu sıralarının dağılımı . . . . .	2
2	Önerilen ölçeklenebilir ve hibrit yöntemin mevcut yöntemler ile karşılaştırması . .	4
3	Doğru tanınma oranının $K$ en yakın komşu liste boyutu ile ilişkisi . . . . .	12
4	Referans betimleyicileri için olasılıksal dağılımların uzaklıklarının dağılımı . . . .	15
5	Tanınma oranının farklı iki boyutlu veri kümeleri ve betimleyici türleri için değişimi	17
6	3B veri kümesinde geometrik olarak anahtar nokta eşleme yöntemi . . . . .	20
7	Orsay Müzesi veri kümesi referans imge örnekleri . . . . .	23
8	Orsay Müzesi veri kümesi test koşulları örnekleri . . . . .	24

## Tablo Listesi

1	Farklı düzlemsel nesne veri kümelerinin üçüncü test imgesi için doğruluk oranlarının beş farklı betimleyici için değerleri . . . . .	19
2	<i>Dog, Horse ve Oil</i> veri kümeleri için 3B test sonuçları. . . . .	21
3	Sekiz referans imge ve yerelliğe duyarlı karma ile elde edilen doğruluk oranları . . . . .	23
4	Orsay veri kümesinde nesne tanıma test sonuçları . . . . .	24

## ÖZET

Gerçek zamanlı mobil nesne tanıma uygulamalarında anahtar noktaların eşlenmesi için genellikle ikilik betimleyiciler kullanılır. Ancak bu betimleyiciler kamera bakış açısı benzeri etmenlerden oldukça etkilendiğinden nokta eşleme başarımları daha karmaşık betimleyicilere göre düşüktür. Proje kapsamında ikilik betimleyicilerin eşlenmesi için iki aşamalı yeni bir yöntem geliştirilmiştir. Bu yöntemin ilk adımı bilinen nesnelerin imgelerinden çıkarılan betimleyicilerin arasında  $K$  en yakın betimleyiciyi tespit eder. İkinci adım ise bu betimleyicilerden doğru eşleme olasılığı en yüksek olanı söz konusu her bir betimleyiciye özel istatistiksel bir yaklaşımla seçer. Bu istatistikler bir öğrenme adımında sentetik görüntülerden hesaplanır. Yöntem düzlemsel ve üç boyutlu nesnelere üzerindeki anahtar noktaların doğru eşlenme oranını sadece Hamming uzaklığı ile en yakın komşunun belirlenmesine göre arttırmaktadır. Ayrıca önerilen yöntemin ilk adımı yaklaşık en yakın komşu yöntemleriyle uyumlu olduğundan bu iki adımlı hibrit yöntem benzerlerinin aksine büyük betimleyici kümeleriyle çalışmasını sağlar. Projede ayrıca nesne tanıma katkısı az olacağı öngörülen betimleyicilerin tespiti için bir yöntem geliştirilmiş, bu betimleyiciler çıkarılsa dahi nesne tanıma başarımında büyük bir kayıp yaşanmadığı gösterilmiştir.

**Anahtar Kelimeler:** Bilgisayarla Görü, Nesne Tanıma, İkilik Betimleyiciler, Yerelliğe Duyarlı Karma

## ABSTRACT

Real-time mobile object identification approaches usually employ binary descriptors for keypoint matching. However, these descriptors can not reach the keypoint matching performance of more complex alternatives since they are more susceptible to variations due to camera viewpoint. As part of the project, a novel two step approach to descriptor matching has been developed. The first step of this approach determines the  $K$  nearest neighbors among the descriptors computed from the images of the known objects. Then the second step picks the descriptor that is most likely to be the correct match by employing statistics specific to each descriptor under consideration. These statistics are computed from a set of synthetic images in a training phase. The approach has been shown to increase the ratio of correct matches of keypoints detected on images of planar and three dimensional objects as compared to using just the nearest neighbor according to the Hamming distance. Moreover, since the first stage of the proposed approach is compatible with the approximate nearest neighbor algorithms, unlike comparable approaches, this two step hybrid approach allows the usage of large descriptor sets. Also, an approach for the determination of descriptors with low contribution to the object identification performance has been developed during the project and it has been shown that the object identification performance is not greatly affected even if such descriptors are left out of the descriptor set.

**Keywords:** Computer Vision, Object Identification, Binary Descriptors, Locality Sensitive Hashing



# 1. GİRİŞ

Nesne tanıma problemi daha önceden bilinen ve sınırlı sayıda bakış açısından görüntüleri toplanmış nesnelerin içeriği bilinmeyen bir sorgu görüntü içerisinde yer alıp almadığının ve eğer yer alıyorsa konumlarının tespiti olarak özetlenebilir. Bu problemin çözümü için yaygın olarak kullanılan yöntem, sorgu görüntü ile daha önceden çekilmiş görüntülerin eşleştirilmesidir. Bu eşleştirme doğrudan görüntüler üzerinden değil, görüntülerde tespit edilen anahtar noktalar (Mikolajczyk vd., 2005; Rosten vd., 2010) üzerinden gerçekleştirilir.

Anahtar nokta eşleştirme amacı ile her bir anahtar nokta etrafındaki dokuyu tanımlayan bir betimleyici (Mikolajczyk ve Schmid, 2004; Calonder vd., 2012) hesaplanır ve bu betimleyiciler arasındaki uzaklıklar hesaplanarak yakın betimleyicilere karşılık gelen anahtar noktaların doğru eşler olduğu varsayılır.

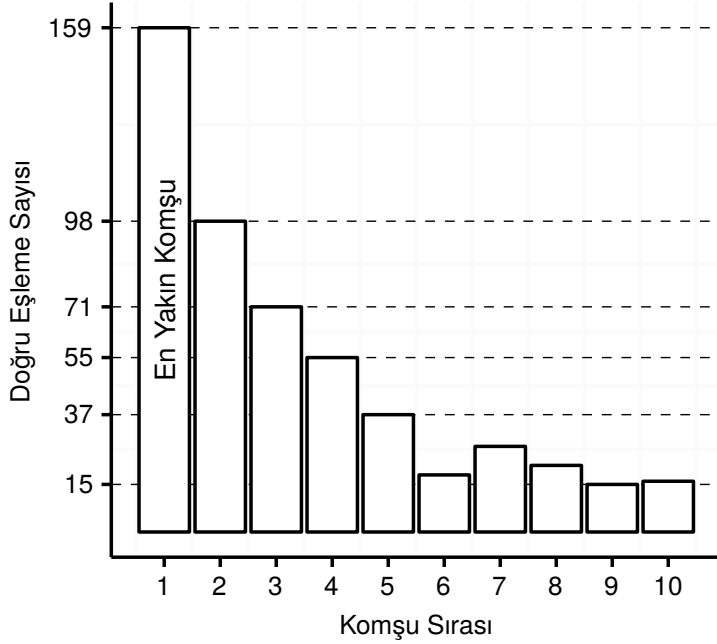
Nesne tanıma probleminin doğru yapılabilmesi için hem anahtar noktaların farklı bakış açılarında tekrar tespit edilebilmesi hem de bakış açısı değiştikçe hesaplanan betimleyicilerin ilk bakış açısında hesaplanan betimleyiciye uzaklığının çok artmaması gerekir. Bu nedenle nesne tanıma uygulamalarında genellikle hesaplaması daha çok zaman alan ama eşleşme başarımı yüksek anahtar nokta çıkarımı ve betimleyici hesaplama yöntemleri kullanılmaktadır. Bu da genellikle nesne tanıma işleminin bir sunucu desteğiyle yapılmasını gerektirmektedir.

## 1.1 İkilik Betimleyiciler ile Gerçek Zamanlı Nokta Eşleştirme

Gerçek zamanlı nesne tanıma uygulamalarında kullanılan yöntemin saniyede yaklaşık yirmi–yirmi beş kere çalışabilmesi gerekmektedir. Bu nedenle sunucu ve mobil donanımda benzer de olsa aynı tekniklerin kullanılması zordur. Hesaplama gücünün sınırlı olduğu durumlarda eşleme adımında genellikle ikilik betimleyiciler (Calonder vd., 2012; Rublee vd., 2011; Leutenegger vd., 2011; Alahi vd., 2012; Trzcinski vd., 2015; Levi ve Hassner, 2015; Oszust, 2016) tercih edilmektedir. Bu betimleyicilerin hem hesaplanması kolaydır hem de iki betimleyici arasındaki uzaklık sadece bir kaç makine komutu ile hesaplanabilir.

Ancak bu betimleyiciler görüntünün elde edildiği bakış açısından SIFT (Lowe, 2004) ben-

Graffiti 1-3 Görüntüleri için Doğru Eşlemelerin Kaçınıcı Sırada Olduğu



Şekil 1. Graffiti veri kümesi bir ve üçüncü görüntülerinde hesaplanan BRIEF betimleyicileri için doğru eşlemelerin en yakın komşu sıralarının dağılımı. Hamming uzaklığına göre hesaplanan en yakın komşular 848 olası doğru eşlemeden sadece 159'unu içerir. İlk on en yakın komşu ise 517 doğru eşlemeyi içerir ki bu da yaklaşık üç kat bir iyileştirmeyi olası kılar. Pratikte önerilen yöntem ile bunların 385 tanesi yakalanmaktadır (Şekil 5) ki bu da iki kattan fazla bir iyileştirme sağlamaktadır.

zere betimleyicilere göre daha çok etkilenirler. Burada temel etmen bu betimleyicilerin ikilik sayı sistemini kullanmasıdır. Sonuç olarak bu betimleyiciler kullanılarak yapılan anahtar nokta eşlemesinde görece daha çok hata yapılır. Çünkü sorgu görüntüdeki anahtar noktanın betimleyicisinin, ait olduğu nesne bilinen betimleyiciler arasındaki en yakın komşusu doğru nesneye ait olmayabilir. Aslında büyük nesne kümeleri ile çalışırken en yakın komşu eşsiz dahi olmayabilir. Bu durumda da eşleme hataları ortaya çıkar ve nesne tanıma başarımı düşer.

Bu durumu önlemek için en yakın tek bir komşu yerine en yakın  $K$  komşunun hesaplanması öngörülebilir. Sorgu görüntüde yer alan bir betimleyiciye olan uzaklıklara göre sıralama yapılarak en yakın  $K$  betimleyici kolaylıkla hesaplanabilir. Doğru eş betimleyicinin  $K$  elemanlı bu betimleyici kümesinde bulunma olasılığı, sadece en yakın komşuya denk gelmesine göre daha olasıdır. Ancak nesne eşleme uygulamalarında genellikle sadece en yakın komşu kullanılmaktadır. Bunun temel nedeni  $K$  elemanlı yakın komşu listesinden doğru olanı seçecek bir yöntemin betimleyici uzaklıkları dışında bir bilgiye dayanması gerekliliğidir.

Bu projede önerilen ise  $K$  en yakın komşu listesinden iki adımlı bir eşleştirme yöntemi kul-

lanılarak yararlanılmasıdır. İlk adımda her bir sorgu betimleyici için bilinen nesnelere ait betimleyicilerden Hamming uzaklığına göre en yakın  $K$  tanesi bulunarak bir listede saklanır. İkinci adımda ise bu betimleyicilerden en uygunu olasılığa dayanan ve her bir anahtar nokta için özel olarak hesaplanan bir değer yardımıyla seçilir. Bu değer hesaplanması için betimleyicilerdeki değişimleri modelleyen ve bilinen görüntülerden daha önce hesaplanmış olan veriden faydalanılır.

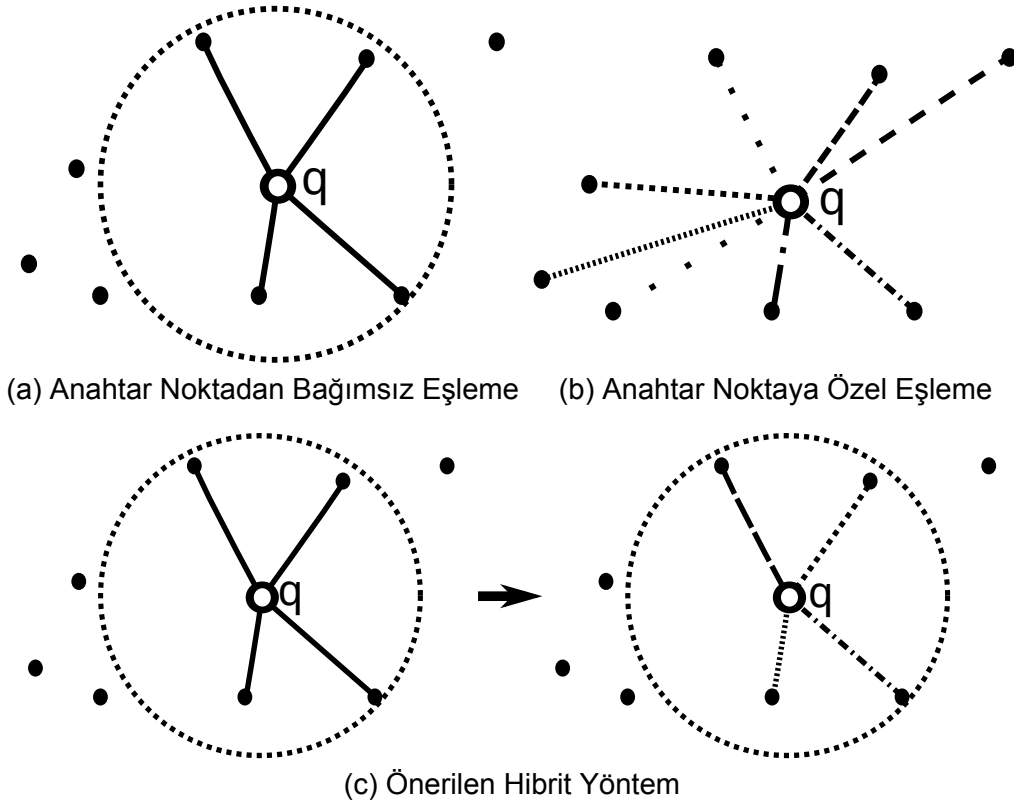
Önerilen yöntemin başarılı olması için ilk koşul  $K$  en yakın betimleyici listesinin doğru eşleşmeyi sadece en yakın komşuya göre daha yüksek oranda içermesidir. Şekil 1 örnek bir görüntü ikilisi için doğru eşleşmelerin yakınlık sıralamasına göre dağılımını göstermektedir.  $K = 10$  değeri için bile doğruluk oranının üç kata kadar artırılması olasıdır. Ancak aşağı sıradaki doğru eşleşmelere ulaşabilmek için Hamming uzaklığı dışında ve daha ayırt edici bir kritere ihtiyaç vardır.

Bu ikincil eşleme kriterini tasarlarırken şu gözlemden faydalanabiliriz: Neredeyse tüm ikilik betimleyiciler her anahtar nokta için aynı betimleme yöntemini kullanırlar. Ancak betimlemenin her bir anahtar noktaya özel şekilde gerçekleşmesi o anahtar noktanın eşleşme oranını artırabilir (Balntas vd., 2015). Mevcut literatürde de her bir anahtar noktaya özel öznelikler kullanan yöntemler vardır (Gupta ve Mittal, 2008; Lepetit ve Fua, 2006; Özuysal vd., 2010; Balntas vd., 2015). Fakat bu yöntemler bilinen nesne görüntülerinden çıkarılan her bir betimleyiciye tek tek uzaklık hesaplanmasını gerektirirler. Bu nedenle de büyük nesne veri tabanları için uygun değildirler. Bu yöntemlerin aksine bu projede önerilen yöntem için sadece  $K$  adet anahtar nokta eşleşmesi için hesaplama yapılması gerekir. Böylelikle önerilen yöntem, bilinen nesne veri tabanı büyüklüğünden bağımsızdır ve büyük veri tabanlarıyla da çalışılabilir.

Önerilen yöntemde ikilik betimleyicilerin kamera bakış açısına göre değişimleri her bir anahtar nokta için ayrı ayrı istatistiksel olarak modellenmiş ve elde edilen olasılık değerlerini kullanan bir betimleyici eşleşme skoru belirlenmiştir. Eşleşme anında ilk önce Hamming uzaklığına göre  $K$  en yakın komşu belirlenmekte ve bu listeden en uygunu  $K$  anahtar nokta eşleşmesinin herbiri için kendilerine özel bir skor ile değiştirilmektedir. Bunun sonucunda eşleşme performansı da küçük bir ek hesaplama adımı yardımıyla artırılmaktadır.

## 1.2 İkilik Betimleyiciler için Ölçeklenebilir Eşleştirme

Büyük nesne veri tabanlarıyla çalışırken önemli bir nokta eşleşme sırasında bir sorgu betimleyici ile bilinen her betimleyici arasındaki uzaklığın ayrı ayrı hesaplanmasının pratikte mümkün olmamasıdır. Bu nedenle yaklaşık en yakın komşu yöntemleri ile sorgu betimleyicinin eş olma



Şekil 2. (a) Eğer her bir betimleyici için betimleme aynı şekilde yapılırsa Yerelliğe Duyarlı Karma yöntemi ile sorgulanan betimleyici  $q$  ile Hamming uzaklığı hesaplanacak betimleyici sayısını sınırlamak mümkündür. (b) Bazı yöntemler her bir anahtar noktaya özel bir betimleme yaparlar ve bu yöntem daha ayırt edici olduğundan eşleme daha gürbüz yapılabilir. (c) Önerilen hibrit ve iki kademeli yöntem, anahtar noktaya özel betimlemeyi sadece ikinci adımda kullandığından bu iki farklı yaklaşımın avantajlarını birlikte içerir.

olasılığı yüksek betimleyiciler seçilere sadece bunlara olan uzaklık hesaplanır. İkilik betimleyiciler için yaygın olarak kullanılan yaklaşık en yakın komşu yöntemi Yerelliğe Duyarlı Karma (YDK)'dır (Andoni ve Indyk, 2008).

Ancak YDK ve benzeri yaklaşımlar tek bir betimleyici uzayında çalıştığından her bir anahtar nokta için betimleyicilerin aynı şekilde hesaplanmasını gerektirir. Bu nedenle anahtar noktaların etrafındaki dokuya özel betimleme yapan yöntemler (Gupta ve Mittal, 2008; Lepetit ve Fua, 2006; Özuysal vd., 2010; Balntas vd., 2015) YDK ile birlikte doğrudan kullanılamazlar. Pratikte bu yaklaşımlar daha ayırt edici olmalarına rağmen büyük betimleyici kümelerinde kullanılamamaktadırlar.

Bu projede önerilen iki kademeli yöntem önce  $K$  adet en yakın komşuyu belirleyip bundan sonra bu listedeki aday anahtar noktalar için özel bir betimleme yaptığından hem ölçeklenebilir hem de daha ayırt edici ve gürbüz bir yöntem elde edilmiştir. Önerilen bu yöntem ile literatürdeki mevcut yaklaşımlar arasındaki ilişki Şekil 2'de verilmiştir.

### 1.3 İkilik Betimleyici Kümelerinin Sıkıştırılması

Nesne tanıma amacıyla bilinen nesnelerin görüntüleri bir araya getirilirken bu görüntülerden çıkarılan anahtar noktaların betimleyicilerinin tümünün saklanması veri boyutunu olumsuz yönde etkiler. Bu nedenle bu betimleyicilerin seçilmesi ya da bir şekilde sıkıştırılması gerekmektedir.

Bu amaçla projede iki farklı yöntem denenmiştir. Bunlardan ilki performansı düşük betimleyicilerin ve bunlara ait anahtar noktaların elenmesi yöntemidir. Bu yöntemde önemli olan nokta elemeye esas olan betimleyici performansının doğru belirlenmesidir. Çünkü betimleyicilerin elenmesi nesne tanıma performansını görece fazla etkilememelidir.

İkinci denenen yaklaşım ise betimleyiciler ve bir önceki kısımda bahsedilen betimleyici istatistikleri arasındaki benzerlikler kullanılarak farklı nesnelere ait anahtar noktalar için tek bir betimleyici/betimleyici istatistikleri verisinin tutulmasıdır. Pratikte betimleyici istatistikleri arasındaki farkın bu tür bir paylaşım için çok büyük olduğu gözlemlendiğinden bu yaklaşım terk edilmiştir. Ancak referans olması için Bölüm 3'te elde edilen uzaklık verileri listelenmiş ve Bölüm 5'te bazı öneriler verilmiştir.

Sonuç olarak betimleyicilerin elemesine dayalı sıkıştırma yöntemi ile projede önerilen hibrit betimleyici eşleme yöntemi birlikte test edilmiş ve nesne tanıma performansını sadece %2 azaltarak yüksek sıkıştırma oranları elde edilmiştir.

### 1.4 Literatüre Katkı

Özetle projede ikilik betimleyicilerdeki bakış açısına bağlı değişimlerin istatistiklerine dayanan ve eşleme sırasında sadece en yakın değil  $K$  adet en yakın komşunun kullanılmasını sağlayan bir yöntem geliştirilmiştir. Geliştirilen yöntem her bir anahtar noktaya özel bir betimleme yapmasına rağmen, benzerlerinin (Gupta ve Mittal, 2008; Lepetit ve Fua, 2006; Özuysal vd., 2010; Balntas vd., 2015) aksine büyük veri tabanlarında da çalışabilir niteliktedir. Ayrıca önerilen yöntem kullanılan betimleyiciden görece bağımsız olduğundan BRIEF (Calonder vd., 2012), ORB (Rublee vd., 2011), BRISK (Leutenegger vd., 2011), FREAK (Alahi vd., 2012), LATCH (Levi ve Hassner, 2015) ve benzeri pek çok ikilik betimleyici ile birarada kullanılabilir.

Önerilen eşleme yöntemi,  $K$  en yakın komşunun belirlenmesi için kullanılan Yerelliğe Duyarlı Karma yöntemi ile birlikte de çalışabildiğinden bu yöntemle elde edilen eşleme kalitesini de arttırmaktadır. Bu sayede hem  $K$  en yakın komşu hızlı bir şekilde hesaplanmakta hem de eşleme kriteri kısmen eşlenen anahtar noktalara özel bir betimleme yapmaktadır. Bu da büyük veri tabanlarında eşleme doğruluğunu veri tabanı büyüklüğünden etkilenmeden ölçeklenebilir

bir şekilde arttırmaktadır (Uzyıldırım ve Özuysal, 2016).

Raporun kalanında önerilen yöntemin detayları, yerelliğe duyarlı karma ile birlikte kullanımı, yüzden fazla tablodan oluşan bir nesne veri tabanında test sonuçları ve bu sonuçların sıkıştırma amacıyla nesne veri tabanından betimleyiciler atıldığında değişimi yer almaktadır. Ayrıca Bölüm 5'te önerilen yöntemin mobil nesne tanıma uygulamaları için avantaj ve dezavantajları, öneriler ile birlikte verilmiştir.

## 2. İLGİLİ LİTERATÜR

Nesne tanımanın farklı aşamalarıyla ilgili pek çok çalışma mevcuttur. Projeye konu olan anahtar nokta eşlemesi amacıyla geliştirilen ikilik betimleyiciler ile ilgili çalışmalar BRIEF (Calonder vd., 2010) betimleyicisi ile başlamıştır. Rublee vd. (2011) BRIEF betimleyicisini kamera  $z$  eksenine etrafındaki dönmelerden bağımsız hale getiren ORB betimleyicisini önermiştir. Alahi vd. (2012) ise BRIEF tarafından kullanılan rastgele testleri eniyileme yoluyla seçerek FREAK betimleme yöntemini geliştirmiştir. Leutenegger vd. (2011) ise rastgele test konumları yerine belirli bir ızgara üzerindeki noktaların karşılaştırılmasına dayalı bir betimleyici ve bu betimleyici ile birlikte çalışan bir anahtar nokta tespit yöntemi geliştirmiştir. Levi ve Hassner (2015) ise iki yerine üç noktanın karşılaştırılmasına dayanan bir betimleyici geliştirmiştir. Tüm bu betimleyiciler her bir anahtar nokta için aynı betimleme yöntemini kullanırlar ve betimlenecek dokuya has özellikleri dikkate almazlar.

### 2.1 İkilik Betimleyici Eşleme Yaklaşımları

Genel olarak anahtar nokta etrafındaki dokuya özel bir betimleyici hesaplanmasının nokta eşleme başarımını arttırdığı gözlemlenmiştir (Balntas vd., 2015). Gupta ve Mittal (2008) küçük dönme ve ötelemelerden daha az etkilenen testleri anahtar noktanın etrafındaki dokuya bakarak seçen ve bunları betimleme amacıyla kullanan bir yöntem geliştirmiştir. Yakın zamanda ise Balntas vd. (2015) BRIEF benzeri betimleyicilerin küçük dönme ve ilgin dönüşümler ile değişmeyen ikilerini gösteren maskeler hesaplamıştır. İki betimleyici arasındaki uzaklık hesaplanmasında her iki betimleyiciye ait maskeler de hesaba katılır. Her iki yöntem de eşleme başarımını arttırmakla beraber eşleme sırasında her bir betimleyiciye olan uzaklığın ayrı ayrı hesaplanmasını gerektirir ve büyük nesne kümelerinde gerçek zamanlı eşleme için uygun değildirler.

Benzer şekilde Lepetit ve Fua (2006); Özuysal vd. (2010) anahtar nokta eşlemesi için olasılığa bağlı sınıflandırmaya dayalı anahtar nokta eşleme yöntemleri geliştirmişlerdir. Bu yöntemler de her bir anahtar nokta için farklı bir olasılık dağılımı hesaplar ve kullanır. Bu nedenle de büyük anahtar nokta kümeleri için uygun değildirler. Ayrıca kullandıkları olasılıksal dağılımlar anahtar

nokta başına bu projede önerilen yöntemle göre çok daha fazla hafıza kullanılmaktadır. Çünkü önerilen yöntemde kullanılan olasılıklar sadece ilk  $K$  en yakın komşuyu birbirinden ayırt etmekte kullanılır.

Önerilen iki adımda eşleme yöntemi imge erişim yaklaşımlarında (Chum vd., 2011; Li vd., 2015) kullanılan genel sorgu ardından görüntüye has geometrik kıstaslar ile eleme yöntemiyle benzerlikler taşımaktadır. İmge erişimde ikinci adımdaki eleme sorgu imge ile aday imge arasındaki geometrik uyuma bağlıdır. Önerilen yöntemde ise geometrik uyum yerine olasılıksal uyum kullanılmaktadır.

## 2.2 Anahtar Nokta Eleme Yaklaşımları

Anahtar nokta veri kümelerinin sıkıştırılması konusunda robotik alanında çalışmalar mevcuttur. Bir ortamda bulunan tüm anahtar noktaların saklanması gerçek zamanlı robotik uygulamaları için de ortam büyüklüğü arttıkça sorun olmaktadır. Li ve Kosecka (2006) nokta kümelerinin azaltılması için olasılıksal bir yaklaşım geliştirmiştir. Dymczyk vd. (2015) ise nokta kümeleri ve sahneler üzerinden bir tam sayı eniyileme problemi çözümlenerek noktalar arasındaki ilişkiyi de kullanarak bazı noktaları eler. Projede kullanılan yaklaşım bu yaklaşımlardan daha basit ama istatistiklerinin hesaplanması için kullanılan eğitim aşamasıyla daha uyumludur. Ayrıca veri tabanındaki nesne görüntüleri birbiri ile örtüşmediğinden robotik uygulamalarından farklı bir veri tipine sahiptir.



## 3. YÖNTEM

İkilik betimleyicilerin eşleştirilmesi için genel yaklaşım betimleyici uzayında sorgu betimleyiciye en yakın komşunun bulunmasıdır. Giriş bölümünde belirtildiği gibi projede tek adımda eşleme yerine öncelikle  $K$  en yakın komşunun Hamming uzaklığına bağlı olarak bulunması ve ilk aday betimleyici listesinin elde edilmesi önerilmektedir. Bu aday listeden doğru eşin seçimi ise her bir adayın istatistiksel özellikleri göz önünde bulundurularak yapılır.

Bu ikinci adımda kullanılmak üzere bilinen anahtar noktaların betimleyicilerinin ikil değerleri bir öğrenme aşamasında pek çok ilgin dönüşüm altında incelenir. İkil değerlerdeki değişimin nasıl modellendiği aşağıda detaylarıyla verilmektedir.

### 3.1 Betimleyici Değişimlerinin İstatistiksel Modellenmesi

İkilik betimleyicilerin ikil değerleri perspektif ve ışık değişimlerinden tamamen bağımsız değildir. Anahtar nokta etrafındaki dokuya göre bazı ikil değerleri değişkenlik gösterirken bazıları daha gürbüz bir davranış sergilerler. Örneğin eğer ikil değer hesaplanmasında kullanılan parlaklık değerleri civarında görüntünün gradyanı yüksek bir değere sahipse bu ikil daha fazla değişkenlik gösterecektir. İkiller ve bakış açısı arasındaki karmaşık ilişkinin modellenmesi için anahtar noktaya bağımlı bir olasılıksal model uygundur:

$$P(D | C = k_i) = P(d_1, d_2, \dots, d_S | C = k_i). \quad (3.1)$$

Denklemden  $D$  ve  $C$  betimleyici değerine ve anahtar noktanın kimliğine karşılık gelen iki rastgele değişkeni göstermektedir.  $C = k_i$  ifadesi ise dağılımın  $i$  sıralı anahtar nokta için hesaplandığını belirtir.  $d_j$  betimleyicinin  $j$  sıralı ikil değerini gösteren rastgele değişkendir.  $S$  ise betimleyicideki toplam ikil sayısıdır.

Bu yaklaşım  $2^S$  parametreye sahiptir ve  $S$  genellikle 64'ten çok daha büyük olduğundan tüm ikil değerlerin ortak dağılımının doğrudan modellenmesi pratikte mümkün değildir. [Özuyal vd. \(2010\)](#) tarafından önerilen yaklaşımı kullanarak betimleyici ikillerini  $M$  ikil içeren  $N$  grupta

toplayıp, gruplar arasında bağımsızlık varsayımını uygularsak

$$P(D | C = k_i) = \prod_{j=1}^N P(D_j | C = k_i), \quad (3.2)$$

istatistiksel modelini elde ederiz. Burada  $D_j$ ,  $j$  sıralı gruptaki ikil değerlerini göstermektedir. Bu yaklaşım  $N \times 2^M$  parametreye sahiptir ve pratikte kullanımı mümkündür. Hafıza kullanımı da göz önünde bulundurularak  $M$  için 4, 6 ve 8 değerlerini kullanabiliriz.  $N$  değeri ise  $M$  sabit tutularak ve  $N = \frac{\text{betimleyici boyutu}}{M}$  hesabıyla elde edilir. 256 ikillik betimleyiciler için  $N$  değeri 32 ve 64 arasında elde edilmiş olur. Daha yüksek  $M$  değerleri hafıza ihtiyacını artırır ve  $M > 12$  için eşleme başarımı da deneysel verilere göre düşmektedir. Bunun sebebinin ihtiyaç duyulan öğretim kümesi boyutundaki artış olduğu düşünülmektedir.

Betimleyici gruplarının belirlenmesi için pek çok yol kullanılabilir. En basiti ardışık ikilerin aynı grupta toplanmasıdır. BRIEF benzeri rastgele testler kullanan betimleyiciler için bu doğal bir yaklaşımdır. BRISK benzeri daha sabit yapıdaki betimleyiciler için ise gruplamalar üzerinden yapılacak bir eniyileme göz önünde bulundurulabilir. Pratikte ardışık gruplamaların başarımı tüm betimleyici türleri için yeterli bulunduğundan projedeki deneylerde bu yaklaşım kullanılmıştır.

Denklem 3.2 içindeki parametrelerin hesaplanması için öğrenme adımında her bir anahtar nokta için Morel ve Yu (2009) tarafından önerilen ilgin dönüşüm modeli ile eğitim kümesi sentezlenmiştir. Anahtar noktalar etrafındaki doku yerel olarak ilgin modellenebileceği için bu yaklaşım çok sınırlayıcı değildir. Bölüm 4.1.2'de aynı model üç boyutlu nesnelere için de test edilmiştir. Tüm deneylerde eğitim kümesinin sentezlenmesi için aynı parametreler kullanılmıştır. Ölçekleme parametresi  $\frac{1}{\sqrt{2}}$  ile  $\sqrt{2}$  arasında, düzlemsel dönme açısı  $-30$  ve  $+30$  derece arasında, eğme miktarı (Morel ve Yu, 2009, için  $\theta$ )  $0$  ve  $60$  derece arasında, eğme açısı (Morel ve Yu, 2009, için  $\phi$ ) ise  $0$  ve  $180$  derece arasında değişmektedir.

Eğitim kümesi için yaklaşık 200000 imge sentezlenmiştir ki bu da anahtar nokta başına aynı sayıda betimleyiciye karşılık gelir. Her bir ikil grubu için Denklem 3.2'de verilen parametreler eğitim betimleyicilerinin aldığı değerler göz önünde bulundurularak hesaplanır. Örneğin grup boyutu 4 ise grubun alabileceği toplam 16 farklı değer vardır. Her bir değere denk gelen betimleyici sayısı sayılarak toplam değere bölünür.

Eğitim setinin boyutundan kaynaklı olarak ortaya çıkan ve sıfır değerine karşılık gelen olasılıklar hesaplamaya zarar vereceğinden sayma işlemine birden başlanır. Bu da grupların farklı değerlerine eşit dağılan  $2^M$  sanal betimleyici örneğine karşılık gelmektedir. Bu yaklaşım genellikle Laplace Düzleştirilmesi (Manning vd., 2008) olarak anılır ve Denklem 3.2'deki olasılıklar için Bayes tahmini yapılırken bir Dirichlet ön dağılımı kullanılmasına karşılık gelir (Bishop, 2006).

## 3.2 İkili Betimleyiciler için Olasılığa Dayalı Karşılaştırma Yöntemi

Eşleme için sorgu beimleyiciye en yakın  $K$  komşu Hamming uzaklığı ile bulunduktan sonra bu liste içinden gerçek eşin seçilmesi gerekir. Bunun için  $K$  en yakın komşu listesinin her bir aday anahtar nokta için yukarıda modellenen olasılıklar kullanılarak tekrar sıralanması gerekir.

Bu sıralama için kullanılacak skor değeri için projede farklı fonksiyonlar değerlendirilmiş ve Hamming uzaklığı ile betimleyici istatistikleri farklı şekillerde birleştirilmiştir. En iyi sonuca negatif Hamming uzaklığı ile sorgu betimleyicinin Denklem 3.2 ile verilen gözlenme olasılığının toplamı ile ulaşılmıştır:

$$\text{skor}(Q, k_i) = -|Q - D^i|_H + \log \prod_{j=1}^N P(Q_j | C = k_i) . \quad (3.3)$$

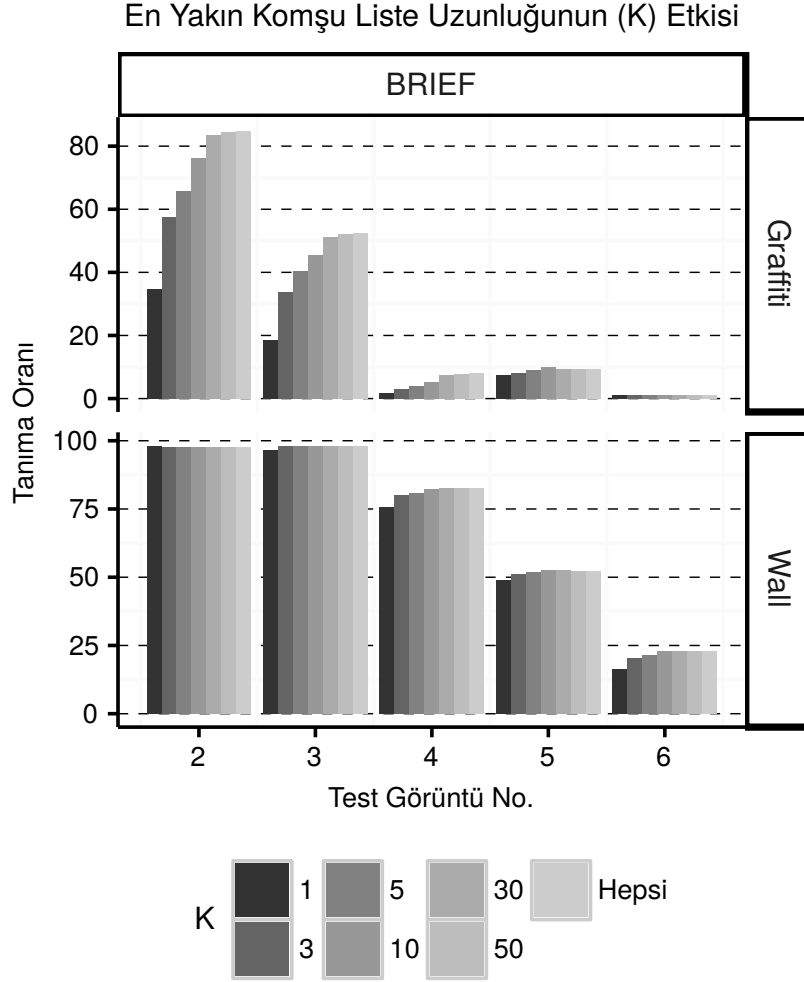
Burada  $Q$  sorgu betimleyiciyi,  $D^i$   $i$  sıralı referans betimleyiciyi ve  $|X - Y|_H$  ise  $X$  ve  $Y$  betimleyicileri arasındaki Hamming uzaklığını belirtmektedir.

İkinci terim  $k_i$  değerini içerdiğinden bu değer doğrudan hesaplanma zamanı referans betimleyici kümesinin boyutuyla doğrusal artacaktır. Projede önerilen yöntemin ana çıkış noktası bu terimin sadece  $K$  en yakın komşu için hesaplanmasının yeterli olacağıdır. Bu sayede sadece en yakın komşudan çok daha fazla doğru eşleme yapmak mümkün olacaktır.

Bu durumun basit bir gösterimini yapmak ve  $K$  değerinin doğru tanıma oranına etkisini ölçmek amacıyla *Graffiti* ve *Wall* veri kümeleriyle (Mikolajczyk ve Schmid, 2004; Mikolajczyk vd., 2005) ilksel bir deney yapılmıştır. Deney için referans imgelerde yaklaşık 1000 adet anahtar nokta tespit edilmiş ve bunların koordinatları veri kümesinde verilen gerçek perspektif dönüşüm ile diğer test imgelerine aktarılmıştır. Hem referans hem de test imgelerdeki anahtar noktalar etrafında BRIEF betimleyicileri çıkarılmıştır. Her bir test betimleyici için  $K$  en yakın komşu hesaplanmış ve Denklem 3.3 ile verilen skoru en yüksek betimleyici ile eşleme yapılmıştır. Her bir test imgesi için doğruluk oranı doğru eşlenen test nokta sayısının tüm test noktalarına oranı olarak hesaplanmış ve farklı  $K$  değerleri için kaydedilmiştir. Deneyde ikil gruplarının boyutu  $M = 8$  olarak alınmıştır. Sonuçlar Şekil 3'te verilmiştir. Betimleyici tipinden bağımsız olarak  $K$  değerinin büyümesi sonuçları iyileştirmiş ve  $K > 10$  için artış görece daha azdır.

## 3.3 İkili Betimleyiciler için Ölçeklenebilir Eşleme

Bilinen nesne sayısının yüksek olduğu durumlarda en yakın komşu ya da komşuların belirlenmesi için bilinen nesnelere ait referans betimleyici kümesindeki her bir betimleyiciye ayrı ayrı uzaklık hesaplanması uzun zaman almaktadır. Bu nedenle büyük betimleyici veri kümelerinde



Şekil 3. BRIEF betimleyicisi için doğru tanıma oranı  $K$  en yakın komşu (EYK) liste boyutu arttıkça artmaktadır. Bu durum özellikle  $K = 1$  için tanıma oranı düşük olan ve hem ölçek hem de dönme açısı değişimleri içeren *Graffiti* veri kümesi için doğrudur. İlk on en yakın komşu ile tüm betimleyicilerin değerlendirilmesi arasındaki fark görece azdır.  $K = 10$  değeri için doğru eşlemelerin sayısı *Graffiti*-3 için 385'e yükselir. Bu maksimum potansiyel rakam olan 517'den azdır (Şekil 1), ama sadece en yakın komşuyu kullanılmasıyla elde edilen 159'dan çok daha yüksektir.

eşleme için gerçek en yakın komşu yerine yaklaşık en yakın komşu yöntemleri kullanılır. Burada amaç tüm referans betimleyicileri yerine bu betimleyici kümesinin bir alt kümesini seçip sadece bu alt küme içinde uzaklık hesabı yapılmasıdır. Gerçek en yakın komşular seçilen alt küme içinde yüksek olasılıkla bulunuyorsa eşleme sonucu bu varsayımdan kötü etkilenmez.

İkilik betimleyiciler için en sık kullanılan yaklaşık en yakın komşu yöntemi doğrudan ikilik sistemde uygulanabilmesi nedeniyle Yerelliğe Duyarlı Karma'dır (YDK). Bu yöntemde  $P$  adet karma fonksiyonu seçilir ve bu karma fonksiyonların izdüğümlerine göre referans betimleyiciler gruplanır. Sorgu sırasında arama yapılacak betimleyici alt kümesi sorgulanan betimleyici ile en az bir karma fonksiyonu için aynı değere sahip referans betimleyiciler kümesidir.

Projede hem YDK yöntemine alternatif yaklaşımlar araştırılmış hem de geliştirilen olasılığa dayalı betimleyici eşleme yöntemi YDK için uyarlanmış ve test edilmiştir. Denenen yöntemlerden en başarılısı önerilen olasılığa dayalı eşleme yönteminin YDK ile birlikte kullanımı olmuştur. Bu yöntemde YDK ile  $K$  en yakın komşu listesi elde edilmiş ve bu liste içinden eşleme sonuncu Denklem 3.3 ile verilen skor ile seçilir. Aslında bu skor eşlenmeye çalışılan anahtar noktaya özel olarak hesaplandığından YDK ile birlikte doğrudan kullanılması mümkün değildir. Ancak önerilen yaklaşımın iki kademeli bir eşleme yapması ve YDK'nın sadece ilk aşamada kullanılması nedeniyle birlikte kullanılmaları mümkün olmuştur.

Bölüm 4.2'de verilen deneysel veriler  $K$  en yakın komşu listesinin yaklaşık olarak seçilmesinin yöntemin başarımını fazlaca etkilemediğini ve YDK ile sadece en yakın komşunun kullanılmasına göre nesne tanıma başarımını arttırdığını göstermektedir.

### 3.4 Betimleyici Kümesinin Boyutunun Düşürülmesi

Projede bilinen nesnelere tanımlamak amacıyla kullanılan referans betimleyici kümesinin boyutunun azaltılması için iki farklı yöntem önerilmiştir. Bunlardan ilki betimleyicilerden tanıma oranı düşük olanların elenmesine dayanır. İkincisi ise farklı nesnelere üzerinde olsalar dahi birbirine benzer dokuya sahip anahtar noktalar için tek bir betimleyicinin depolanmasıdır. Aşağıda bu yöntemlerin detayları verilmiştir.

#### 3.4.1 Tanınma Başarımı Düşük Betimleyicilerin Elenmesi

Bilinen nesnelere çıkarılan her anahtar noktanın referans betimleyici listesine katkı sağlaması pratikte betimleyici küme listesini büyük oranda arttırmaktadır. Bazı anahtar noktalar etrafındaki doku ayırt edici özellikler taşımadığında bu dokulardan hesaplanan betimleyiciler ile de sağlıklı

bir eşleme yapmak mümkün değildir. Bunlara örnek tekrar eden pencere gibi dokular ya da ağaç benzeri yüksek frekanslı ama ayırt edici olmayan dokulardır.

Betimleyicilere ait istatistikler zaten bir öğrenme aşamasında hesaplandığından bu aşama sırasında betimleyicilerin eşleme başarımı da ölçülebilir. Bu amaçla referans imgeler ( $\mathcal{I}_i$ ) yapay ilgin dönüşümler ile dönüştürülerek sentetik test imgeleri ( $\mathcal{I}'_i^j, j = 1, \dots, N_j$ ) oluşturulur.  $\mathcal{I}'_i^j$  üzerindeki her bir sentetik test betimleyicisi ( $q_{il}^j, l$  sıralı anahtar nokta etrafında hesaplanmış betimleyici) için tüm referans imgeler içinde arama yapılır ve eşlenen betimleyicinin doğru referans imgeden gelme oranı hesaplanır. Eğer  $q_{il}^j$  test betimleyicisine eşlenen betimleyici  $k$  imgesindeki  $q_{kt}^{*j}$  ise,  $i$  sıralı referans betimleyicisi için doğru tanınma oranı

$$R_i = \frac{\sum_j \delta(i - k)}{N_j}$$

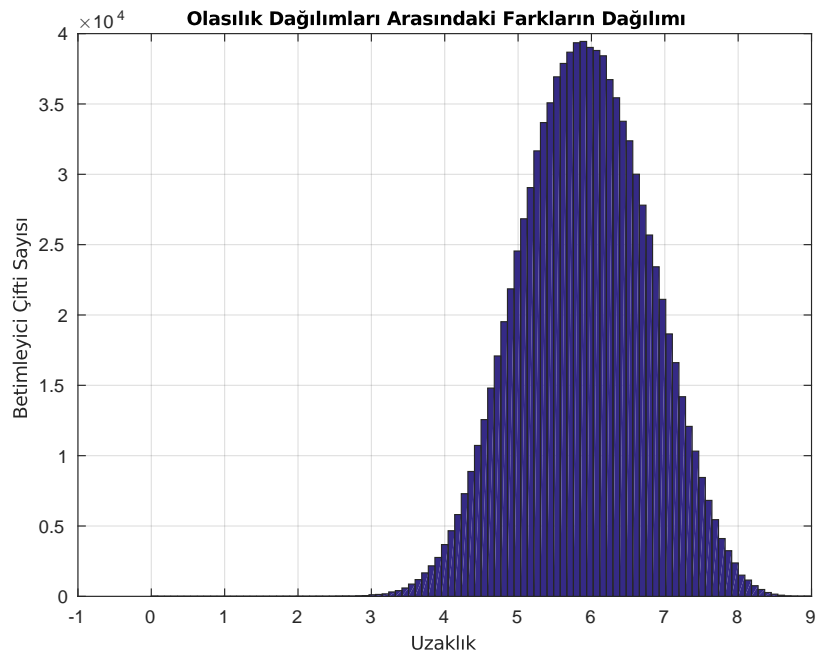
ile verilir. Burada önemli bir nokta gerçek imge sırası  $i$  ile eşlenen betimleyicinin imge sırasının ( $k$ ) aynı olmasının yeterli oluşudur,  $l$  ve  $t$  değerlerinin aynı olması beklenmez. Yani eşlenen betimleyici doğru imgeye ait olduğu sürece eşleme doğru kabul edilmiş olur. Bu sayede bir referans imge üzerinde yakın noktalardan çıkarılan betimleyici benzerliklerinden doğacak hatalar doğru imge seçildiği için ihmal edilmiş olur. Bu da nesne tanıma yönteminin çalışmasıyla daha uyumludur.

Pratikte  $N_j = 300$  kadar sentetik test görüntüsü ile hesaplama yapılmış ve referans betimleyiciler hesaplanan  $R_i$  değerlerine göre sıralanarak en düşük değere sahip olanlar elenmiştir. Deneysel sonuçlar Bölüm 4.3'te verilmiştir.

### 3.4.2 Anahtar Noktalar Arasında Betimleyici ve İstatistik Paylaşımı

Projede denenen ikinci bir yaklaşım da birbirine benzer betimleyiciler arasında veri paylaşımıdır. Özellikle betimleyiciler için hesaplanan istatistiksel veriler betimleyicilerden daha fazla yer kapladığı için her bir betimleyiciye ait olasılık dağılımlarının paylaşımına dayalı bir sıkıştırma yöntemi öngörülmüştür. Bu amaçla referans betimleyicilerden rastgele çiftler alınarak bunların olasılık dağılımları arasındaki uzaklıklar ölçülmüş ve bu uzaklıkların dağılımı incelenmiştir. Bu dağılım Şekil 4'te verilmiştir.

Deney sonuçları göstermiştir ki hesaplanan olasılık dağılımları arasındaki farklar paylaşımına dayalı bir sıkıştırma yönteminin başarılı olmasını önlemektedir. Bu nedenle betimleyici elenmesi yoluyla veri kümesinin azaltılması yöntemi daha başarılı bulunmuştur.



Şekil 4. Betimleyiciler için hesaplanan olasılık dağılımlarının paylaşımı için birbirine benzer dağılımların belli bir oranda olması gerekmektedir. Ancak şekilde görülebileceği gibi olasılık dağılımları arasındaki uzaklıklar yüksek bir ortalama değer etrafında toplanmıştır. Olasılık dağılımları arasındaki uzaklık sıfıra yakın olan çok az sayıda betimleyici çifti bulunmaktadır.

## 4. DENEYLER

### 4.1 Olasılığa Dayalı Karşılaştırma Yönteminin Testleri

Önerilen yöntemin  $K$  en yakın komşu içerisinde doğru eş seçme yeteneğini test etmek amacıyla iki farklı veri kümesinde toplam üç farklı deney tasarlanmıştır. Bunların ilk ikisinde iki boyutlu Oxford veri kümesi ikincisinde ise üç boyutlu nesnelere içeren bir veri kümesi kullanılmıştır.

#### 4.1.1 Oxford Veri Kümesi ile Testler

Düzlemsel nesnelere ile ilk deneyler için Şekil 3 ile sonuçları verilen deney düzeneği kullanılmış ve gerçek eşlemeler üzerinden tanınma oranı hesaplanmıştır. Burada test görüntülerinde anahtar nokta çıkarımı yapılmamış, referans görüntülerdeki nokta konumları bilinen gerçek dönüşüm kullanılarak test imgelerine aktarılmıştır.

Deneyde dört farklı referans imge ve beş betimleyici tipi için  $K = 10$  değeri kullanılmış, deneyler farklı ikil grup büyüklüklerinde tekrarlanmıştır ( $M \in \{4, 6, 8\}$ ). Sonuçlar<sup>1</sup> Şekil 5'te verilmiştir.

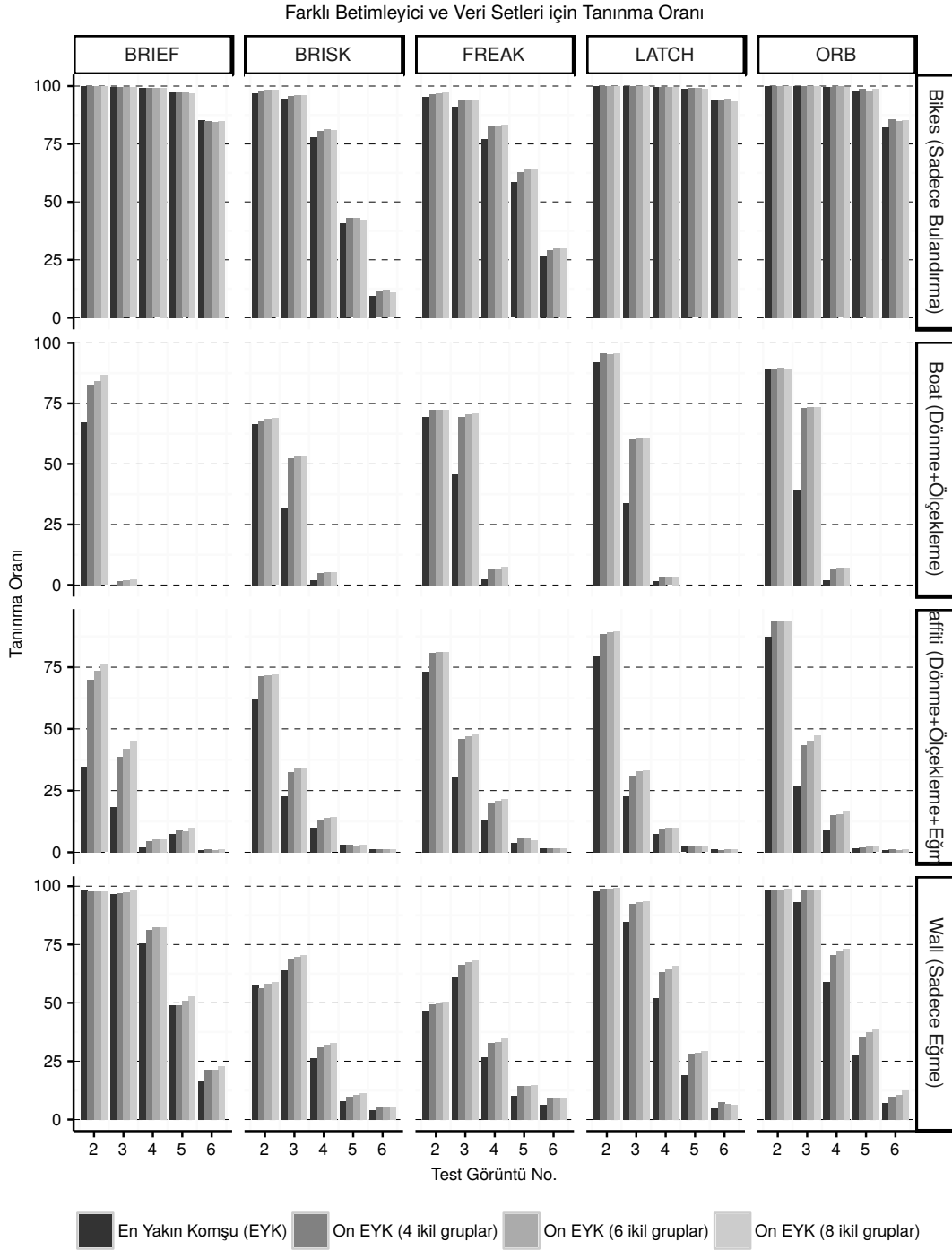
Projede önerilen yöntem tanınma oranını özellikle aşırı perspektif değişimlerin olduğu durumlarda oldukça arttırmaktadır. *Bikes* veri kümesi sadece bulanıklaştırma içerdiğinden artış daha düşüktür. Hesaplanan olasılıksal dağılımlar sadece perspektif değişimler altında ikilerin davranışını modellediğinden bu beklenen bir durumdur. Eğitim kümesine bulanıklaştırma örnekleri artırıldığında ise sonuçlarda belirgin bir iyileşme olmamıştır. Bunun sebebi olarak ayırt ediciliğin bu durumda azalması öngörülmektedir.

BRIEF betimleyicisi düzlemsel dönüş açısından fazlaca etkilendiğinden *Boat* veri kümesinin ikinci imgesinden sonra ilk on en yakın komşu dahi yeterli doğru eşleme içermediğinden önerilen yöntem bir iyileşme sağlayamamaktadır.

---

<sup>1</sup>Her bir betimleyici tipi için farklı bir anahtar nokta bulma yöntemi ve betimleyici parametreleri (ya yazarın önerdiği ya da OpenCV kütüphanesinde yer alan) kullanılmıştır. Bu nedenle elde edilen değerler farklı betimleyiciler arasında karşılaştırma yapmak için değil önerilen yöntemin farklı betimleyiciler için verdiği sonuçları incelemek için kullanılmalıdır.





Şekil 5. Tanınma oranı sadece en yakın komşu yerine ilk on en yakın komşu kullanıldığında artar. Artış özellikle tanınma oranının düşük olduğu imgeler (3–6) için önemlidir çünkü önerilen yöntem kullanılmadığında başarılı tanıma için yeterli doğru eşleme yapılamamaktadır.  $M$  arttıkça başarımlar da artar fakat  $M > 10$  için başarımların azaldığı gözlemlenmiştir. Bu nedenle deneyler  $M = 4$  ve  $M = 8$  değerleri için devam ettirilmiştir.

Düzlemsel dönüş açısını tahmin eden ORB benzeri betimleyiciler için ise sonuçlar özellikle test imgesi üç için büyük oranda iyileştirme içermektedir. En büyük artış ölçek ve perspektif değişimleri içeren *Graffiti* veri kümesinde olmuştur.

İlk deney ile gerçek eşleşmeler kullanılarak sadece anahtar nokta eşleme başarımı ölçülmüştür. İkinci deneyde doğruluk oranı eşleşen anahtar nokta sayısına bağlı olarak verilmiştir. Bu deney düzeneği hem referans hem de test imgelerinde anahtar nokta çıkarıldığından gerçek eşleme için daha uygun bir modeldir. Bu amaçla test imgelerinde tespit edilen noktaların betimleyicileri için eşleme yapılarak bu eşleşmeler negatif betimleyici uzaklığına göre sıralanmıştır. Ölçülen değer doğru eşleşmelerin sayısının toplam eşleşen nokta sayısına bölünmesiyle elde edilir. Bu değer *doğruluk oranına* karşılık gelir ve PROSAC (Chum ve Matas, 2005) benzeri gürbüz geometrik eşleme yöntemlerinin çalışma hızını doğrudan etkiler.

Anahtar noktaya özel skorlamanın etkisini göstermek için üç farklı değer ölçülmüştür. Öncelikle karşılaştırma amacı ile sadece en yakın komşunun eşlemede kullanılması ve Hamming uzaklığı ile sıralanmasıyla doğruluk değeri hesaplanmıştır. Daha sonra ise sadece en yakın komşu kullanılarak elde edilen eşleşmeler Denklem 3.3 ile verilen skora göre sıralanmıştır. Bu anahtar noktaya özel skorlamanın etkisini toplam doğruluk oranını değiştirmeden gösterir. Son olarak ise hem on en yakın komşu içinden eşleme yapılmış hem de Denklem 3.3'e göre sıralama yapılmıştır.

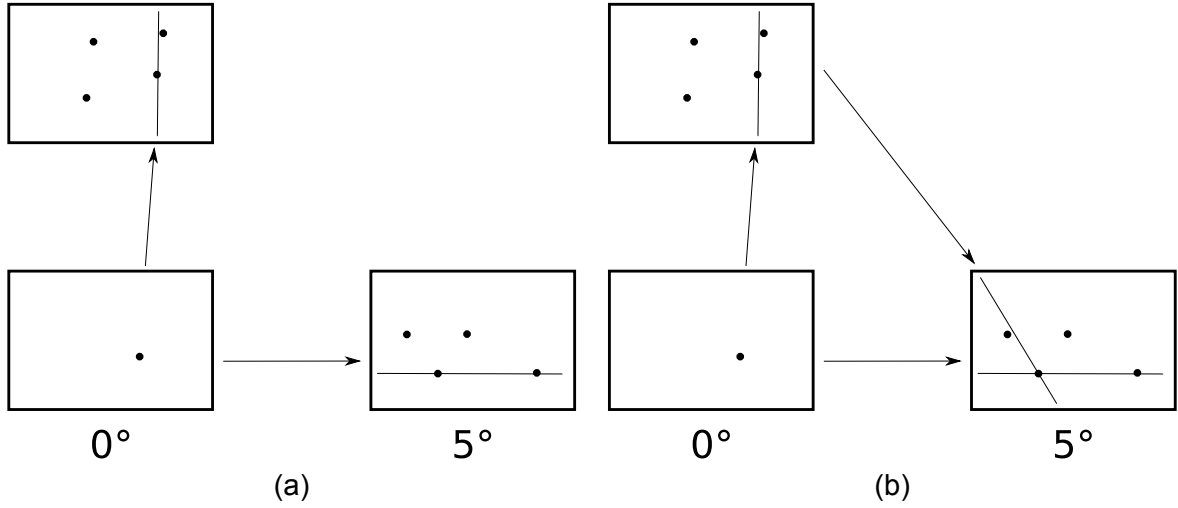
Tablo 1 elde edilen sonuçları göstermektedir. Ağırlıklandırılmış en yakın komşu (A-EYK) değerleri sadece Denklem 3.3'e göre hesaplanan skorun etkisini göstermektedir. Yaklaşık 500 eşleme yapıldığında elde edilen son doğruluk değeri aynı kalsa da genelde daha az sayıda eşleme için yüksek doğruluk oranları elde edilmiştir. K-EYK değerleri ilk on en yakın komşu içinde aramanın getirdiği ek katkıyı göstermektedir. Hemen hemen her durumda  $K$  en yakın komşunun kullanılması 500 eşleme için son doğruluk oranını arttırmaktadır.

Özellikle *Graffiti* veri kümesi için, görece zayıf başarıma sahip olan BRIEF için değerler BRISK ve FREAK betimleyicilerinin üzerine çıkmıştır. Diğer veri kümeleri için BRIEF başarımı ya yükseltilemeyecek kadar düşük (*Boat*) ya da başarımda gözlenebilir bir fark yaratmayacak kadar yüksektir (*Bikes* ve *Wall*).

*Boat* veri kümesi için, 250 eşleme için EYK doğruluk oranları BRISK ve FREAK için sırasıyla %44 ve %45'tir. On en yakın komşunun kullanılmasıyla doğruluk oranı neredeyse iki katına, %86 ve %83'e yükselmektedir. Bu oranlar PROSAC benzeri yöntemlerin sadece 5 tekrarda sonuca varması için yeterlidir ve bu da EYK'ya göre 20 kat daha hızlıdır.

Tablo 1. Farklı düzlemsel nesne veri kümelerinin üçüncü test imgesi için doğruluk oranlarının beş farklı betimleyici için değerleri. Sonuçlar sadece en yakın komşunun Hamming uzaklığı ile sıralanmasıyla (EYK), Denklem 3.3'e göre sıralanmasıyla (Ağırlıklandırılmış EYK  $\rightarrow$  A-EYK) ya da  $K$  en yakın komşudan seçilen eşlemelerin Denklem 3.3'e göre sıralanmasıyla ( $K$ -EYK,  $K \in \{5, 10, 20\}$ ) elde edilmiştir. Her bir veri kümesi ve betimleyici için doğruluk oranları ilk 100, 250 ve 500 anahtar nokta eşi için ayrı ayrı verilmiştir. Genel olarak anahtar noktaya özel skorları kullanmak ilk birkaç yüz eşleme için doğruluk oranlarını arttırmakta ve ilk on en yakın komşuyu değerlendirmek 500 eşleme için doğruluk oranlarını yükseltmektedir. Daha yüksek bir doğruluk oranı PROSAC benzeri eşlemelerin skorlarını kullanan yöntemlerin daha kısa sürede sonuca ulaşmasını sağlar.

Betimleyici Eşleme Sayısı	BRIEF			BRISK			FREAK			LATCH			ORB		
	100	250	500	100	250	500	100	250	500	100	250	500	100	250	500
Bikes-NN	<b>0.97</b>	<b>0.97</b>	<b>0.88</b>	0.78	0.61	0.46	<b>0.86</b>	0.72	0.53	<b>0.99</b>	0.95	0.70	<b>0.99</b>	0.97	-
A-EYK-4 ikil	0.93	0.90	0.85	<b>0.84</b>	0.73	0.51	0.85	0.79	0.58	0.96	0.95	0.70	0.98	0.97	-
A-EYK-8 ikil	0.94	0.91	0.87	0.81	0.70	0.51	<b>0.86</b>	0.79	0.59	0.98	<b>0.96</b>	0.70	0.98	0.97	-
5-EYK-4 ikil	0.94	0.91	0.86	<b>0.84</b>	0.74	0.59	0.84	0.80	0.62	0.96	<b>0.96</b>	<b>0.72</b>	<b>0.99</b>	0.97	-
5-EYK-8 ikil	0.95	0.92	<b>0.88</b>	0.80	0.76	<b>0.60</b>	0.85	<b>0.81</b>	<b>0.63</b>	0.98	<b>0.96</b>	<b>0.72</b>	<b>0.99</b>	<b>0.98</b>	-
10-EYK-4 ikil	0.94	0.91	0.86	<b>0.84</b>	0.74	<b>0.60</b>	0.84	0.80	0.62	0.96	<b>0.96</b>	<b>0.72</b>	<b>0.99</b>	0.97	-
10-EYK-8 ikil	0.95	0.92	<b>0.88</b>	0.80	<b>0.77</b>	<b>0.60</b>	0.85	<b>0.81</b>	<b>0.63</b>	0.98	<b>0.96</b>	<b>0.72</b>	<b>0.99</b>	<b>0.98</b>	-
20-EYK-4 ikil	0.94	0.91	0.86	<b>0.84</b>	0.74	<b>0.60</b>	0.84	0.80	<b>0.63</b>	0.96	<b>0.96</b>	<b>0.72</b>	<b>0.99</b>	0.97	-
20-EYK-8 ikil	0.95	0.92	<b>0.88</b>	0.80	<b>0.77</b>	<b>0.60</b>	0.85	<b>0.81</b>	<b>0.63</b>	0.98	<b>0.96</b>	<b>0.72</b>	<b>0.99</b>	<b>0.98</b>	-
Boat- NN	0.00	0.00	0.00	0.65	0.44	0.29	0.60	0.45	0.37	0.40	0.22	0.13	0.13	0.14	0.11
A-EYK-4 ikil	0.00	0.00	0.00	0.94	0.63	0.35	0.87	0.75	0.48	0.54	0.27	0.14	0.51	0.28	0.15
A-EYK-8 ikil	0.00	0.00	0.00	0.96	0.65	0.35	<b>0.90</b>	0.77	0.49	0.61	0.27	0.14	0.55	0.28	0.15
5-EYK-4 ikil	0.00	0.00	0.01	0.96	0.78	0.52	0.85	0.80	0.65	0.66	0.44	0.23	0.64	0.46	0.29
5-EYK-8 ikil	0.00	0.00	0.01	<b>0.98</b>	0.82	0.53	0.89	<b>0.83</b>	0.67	0.71	0.44	0.24	0.72	0.50	0.29
10-EYK-4 ikil	0.00	0.00	0.02	0.96	0.80	0.56	0.85	0.81	0.67	0.68	0.50	0.27	0.67	0.50	0.34
10-EYK-8 ikil	0.01	0.01	0.03	<b>0.98</b>	<b>0.86</b>	0.59	0.89	<b>0.83</b>	0.70	0.74	0.52	0.29	<b>0.76</b>	0.54	0.36
20-EYK-4 ikil	0.00	0.00	0.01	0.96	0.80	0.59	0.85	0.81	0.67	0.69	0.51	0.29	0.67	0.50	0.37
20-EYK-8 ikil	0.01	0.02	0.03	<b>0.98</b>	<b>0.86</b>	<b>0.63</b>	0.89	<b>0.83</b>	<b>0.71</b>	<b>0.76</b>	<b>0.56</b>	<b>0.32</b>	0.75	<b>0.55</b>	<b>0.39</b>
Graffiti- NN	0.28	0.21	0.16	0.61	0.36	0.23	0.53	0.40	0.30	0.43	0.27	0.17	0.37	0.26	0.18
A-EYK-4 ikil	0.49	0.32	0.20	0.70	0.44	0.26	0.70	0.54	0.34	0.54	0.36	0.19	0.55	0.38	0.22
A-EYK-8 ikil	0.54	0.34	0.20	<b>0.72</b>	0.44	0.26	0.72	0.54	0.35	0.53	0.37	0.19	<b>0.59</b>	0.39	0.22
5-EYK-4 ikil	0.54	0.44	0.34	0.70	0.50	0.33	0.72	0.57	0.43	0.51	0.37	0.22	0.52	0.40	0.27
5-EYK-8 ikil	0.67	0.52	0.37	<b>0.72</b>	0.53	0.35	0.73	0.61	0.43	0.54	0.39	0.24	0.54	0.42	0.29
10-EYK-4 ikil	0.54	0.44	0.35	0.70	0.50	0.34	0.73	0.58	0.44	0.51	0.40	0.23	0.52	0.40	0.27
10-EYK-8 ikil	0.68	0.54	0.41	<b>0.72</b>	0.53	0.36	<b>0.74</b>	<b>0.62</b>	0.46	<b>0.55</b>	<b>0.41</b>	0.26	0.53	0.42	0.30
20-EYK-4 ikil	0.54	0.46	0.38	0.70	0.50	0.35	0.73	0.59	0.45	0.51	0.39	0.24	0.52	0.40	0.28
20-EYK-8 ikil	<b>0.69</b>	<b>0.56</b>	<b>0.45</b>	<b>0.72</b>	<b>0.54</b>	<b>0.38</b>	<b>0.74</b>	<b>0.62</b>	<b>0.47</b>	<b>0.55</b>	<b>0.41</b>	<b>0.27</b>	0.53	<b>0.43</b>	<b>0.31</b>
Wall-NN	0.99	<b>0.99</b>	0.88	<b>1.00</b>	0.96	0.67	0.97	0.86	0.56	0.90	0.78	0.47	<b>1.00</b>	0.95	0.62
A-EYK-4 ikil	0.99	<b>0.99</b>	0.90	<b>1.00</b>	0.97	0.73	<b>0.98</b>	0.91	0.61	0.89	0.84	0.52	<b>1.00</b>	<b>0.96</b>	0.64
A-EYK-8 ikil	0.99	<b>0.99</b>	<b>0.92</b>	<b>1.00</b>	<b>0.98</b>	0.74	0.97	0.91	0.61	0.91	0.85	0.52	<b>1.00</b>	<b>0.96</b>	0.64
5-EYK-4 ikil	<b>1.00</b>	<b>0.99</b>	0.89	<b>1.00</b>	0.97	0.75	<b>0.98</b>	0.90	0.64	0.90	0.84	0.56	<b>1.00</b>	<b>0.96</b>	0.65
5-EYK-8 ikil	<b>1.00</b>	<b>0.99</b>	<b>0.92</b>	<b>1.00</b>	<b>0.98</b>	0.78	0.97	<b>0.92</b>	0.66	<b>0.92</b>	<b>0.87</b>	0.56	<b>1.00</b>	<b>0.96</b>	<b>0.67</b>
10-EYK-4 ikil	<b>1.00</b>	<b>0.99</b>	0.89	<b>1.00</b>	0.97	0.75	<b>0.98</b>	0.90	0.65	0.90	0.84	0.55	<b>1.00</b>	<b>0.96</b>	0.65
10-EYK-8 ikil	<b>1.00</b>	<b>0.99</b>	<b>0.92</b>	<b>1.00</b>	<b>0.98</b>	<b>0.79</b>	0.97	<b>0.92</b>	<b>0.67</b>	<b>0.92</b>	<b>0.87</b>	<b>0.57</b>	<b>1.00</b>	<b>0.96</b>	<b>0.67</b>
20-EYK-4 ikil	<b>1.00</b>	<b>0.99</b>	0.89	<b>1.00</b>	0.97	0.76	<b>0.98</b>	0.90	0.65	0.90	0.84	0.55	<b>1.00</b>	<b>0.96</b>	0.65
20-EYK-8 ikil	<b>1.00</b>	<b>0.99</b>	<b>0.92</b>	<b>1.00</b>	<b>0.98</b>	<b>0.79</b>	0.97	<b>0.92</b>	<b>0.67</b>	<b>0.92</b>	<b>0.87</b>	0.56	<b>1.00</b>	<b>0.96</b>	<b>0.67</b>



Şekil 6. (a) İki görüntü arasındaki kalibrasyon bilgisi kullanılarak noktaları bir görüntüden diğer bir görüntüdeki bir doğruya taşımak mümkündür. Ancak bu anahtar noktalar arasında kesin bir eşlemeye varmayabilir. (b) Üçüncü bir görüntü kullanılarak belirsizlikler ortadan kaldırılabilir ve görüntüler arasında noktadan noktaya bir ilişki kurulabilir.

#### 4.1.2 Üç Boyutlu Nesne Veri Kümesi ile Testler

Önerilen yöntemde betimleyicilerin kamera bakış açısına bağlı değişimleri istatistiksel olarak modellenirken ilgin dönüşümler ile sentezlenen eğitim verisinden yararlanılmıştır. Bilinen nesneler düzlemsel olduğunda bu gerçekçi bir yaklaşımdır. Nesneler üç boyutlu olduğunda ise tek bir görüntüden eğitim verisi çıkarılması gerçekçi değildir. Önerilen yöntemin üç boyutlu nesneler için de genellenmesi ve test edilmesi için üç boyutlu nesnelere oluşan bir veri kümesi kullanılmıştır.

Bu veri kümesi [Moreels ve Perona \(2006\)](#) tarafından anahtar nokta bulma ve betimleme yöntemlerinin test edilmesi için geliştirilmiştir. Üç boyutlu nesnelere dönen bir tablaya yerleştirilmiş ve iki adet kamera ile farklı dönme açılarında görüntüleri kaydedilmiştir. Kameraların tablaya ve birbirlerine göre konumları sabit olduğundan aralarındaki dönme ve öteleme dönüşümlerinin hesaplanması mümkündür. Ayrıca iç kalibrasyon parametrelerinin hesaplanması için bir damalı düzlemsel kalibrasyon nesnesinin görüntüleri de veri kümesinde yer almaktadır.

Projede bu veri kümesindeki görüntüler için hem iç hem de dış kalibrasyon hesaplanmış, bu sayede görüntülerdeki anahtar noktaların konumları diğer görüntülere taşınabilmektedir. Görüntüler üç boyutlu nesnelere ait olduğundan bir görüntüdeki konumlar diğer görüntülerde bir doğru parçası üzerine taşınır. Ancak Şekil 6'da gösterildiği şekilde üç görüntü kullanılarak bir noktanın diğer bir görüntüdeki karşılığını bulmak mümkündür. Testlerde kullanılmak amacıyla geometrik yöntemle 0° referans imgesindeki anahtar noktalar diğer görüntülere taşınmıştır.

Tek bir imgeden hesaplanacak olasılık dağılımları üç boyutlu değişimleri doğru modelleye-

Tablo 2. *Dog*, *Horse* ve *Oil* veri kümeleri için 3B test sonuçları.

<i>Dog</i>	-45	-35	-25	-15	-5	5	15	25	35	45	<i>Ortalama</i>
EYK	2.3	9.4	11.9	15.8	24.0	<b>33.7</b>	13.2	5.6	8.2	3.4	12.8
10-EYK	<b>8.0</b>	<b>15.1</b>	<b>16.0</b>	<b>20.1</b>	<b>24.8</b>	33.2	<b>15.0</b>	<b>10.8</b>	<b>14.3</b>	<b>6.1</b>	<b>16.3</b>
<i>Horse</i>	-45	-35	-25	-15	-5	5	15	25	35	45	<i>Ortalama</i>
EYK	7.6	11.5	16.7	24.6	44.6	<b>44.1</b>	29.9	15.7	10.5	3.5	20.9
10-EYK	<b>13.6</b>	<b>18.1</b>	<b>21.2</b>	<b>27.8</b>	<b>44.8</b>	<b>44.1</b>	<b>32.1</b>	<b>24.9</b>	<b>15.2</b>	<b>7.0</b>	<b>24.9</b>
<i>Oil</i>	-45	-35	-25	-15	-5	5	15	25	35	45	<i>Ortalama</i>
EYK	5.5	5.6	9.7	<b>18.3</b>	<b>29.4</b>	34.9	<b>37.9</b>	29.5	20.4	12.2	20.3
10-EYK	<b>6.6</b>	<b>7.4</b>	<b>11.7</b>	<b>18.3</b>	29.1	<b>36.2</b>	37.7	<b>30.4</b>	<b>22.0</b>	<b>14.9</b>	<b>21.4</b>

meyeceğinden 10° aralıklarda eğitim imgeleri alınmış bunlar zayıf ilgin dönüşümlerle bükülerek eğitim verisi oluşturulmuş ve bu eğitim verisinden de betimleyici davranışlarının olasılıksal dağılımları hesaplanmıştır.

Test için eğitim verisinde yer almayan ve 5°'li imgeden başlayıp 10° aralıklarla devam eden imgeler kullanılmıştır. Tüm bu imgelerde betimleyiciler hesaplanarak 0° referans imgeye eşleme yapılmış buradan da tanınma oranı hesaplanmıştır.

Sonuçlar Tablo 2'de verilmiştir. Hesaplanan betimleyici istatistikleri kullanılarak ilk on en yakın komşu içerisinde eşleme yapıldığında tanınma oranları da artmaktadır. Bu durum özellikle *Dog* ve *Horse* veri kümelerinin uç aç değerleri için geçerlidir. Bazı durumlarda doğru eşleme sayısını iki-üç kat arttırmak mümkün olmuştur.

#### 4.1.3 Hesaplama Süresi

Betimleyici eşleşme başarısının artmasının hesaplama süresine yansımalarının ölçülmesi için testler yapılmıştır. Yapılan tüm testlerde ilk on en yakın komşu ile aramanın neredeyse hesaplama süresini hiç değiştirmedeği gözlenmiştir. Burada temel etmen olasılıkların sadece on betimleyici adayı için kullanılması ama bilinen tüm betimleyiciler için Hamming uzaklığının hesaplanmasıdır. Betimleyici sayısı arttıkça olasılıkların kullanıldığı ek aşamanın hesaplama süresine katkısı yüzdesel olarak azalacaktır.

Örneğin 1000 referans betimleyicinin 928 sorgu betimleyici için Hamming uzaklığına bağlı olarak kaba kuvvet eşlenmesi 4.0 milisaniye alırken ilk on en yakın komşu için Denklem 3.3 ile seçim yapıldığında toplan süre 4.2 milisaniye olmaktadır. Pek çok gerçek zamanlı uygulama için bu ek süre fazlaca bir anlam ifade etmemektedir.

## 4.2 Ölçeklenebilirlik Testleri

### 4.2.1 Oxford Veri Kümesi ile Testler

Referans veri kümesi birden fazla imgeden oluştuğunda betimleyici sayısı da buna bağlı olarak artar ve betimleyici eşlemesi için her bir referans kümesinde kaba kuvvet eşleme yapmak zorlaşır. Önerilen yöntemin bu gibi durumlardaki başarımı ölçmek için referans imgeleri olarak Oxford veri kümesindeki *Graffiti*, *Boat*, *Bikes*, *Wall* imgeleri ve diğer dört imge bir araya getirilmiştir. Toplam betimleyici sayısı yaklaşık 8000'dir. Bu küme için de doğruluk değerleri hesaplanmış ancak  $K$  en yakın komşunun tespiti için Yerelliğe Duyarlı Karma (YDK) yöntemi kullanılmıştır. YDK hesaplaması FLANN (Muja ve Lowe, 2014) kodlarının OpenCV kütüphanesi aracılığıyla çağrılması ile gerçekleştirilmiştir. YDK parametreleri olarak 12 tablo, 20 karma anahtar uzunluğu ve çoklu sorgu seviyesi 2 kullanılmıştır. Bu yaklaşık %90 keskinlik değerine karşılık gelmektedir (YDK ile hesaplanan en yakın komşular kaba kuvvet yaklaşımıyla bulunanlar ile on seferin dokuzunda örtüşmektedir).

Tablo 3'te gösterildiği gibi sonuçta elde edilen doğruluk oranları Tablo 1'e göre daha düşüktür çünkü referans betimleyici sayısı sekiz kat daha fazladır. Ancak her durumda projede önerilen eşleme yöntemi sadece en yakın komşuyla eşlemede kaçırılacak pek çok doğru eşiy yakalayabilmektedir. Artış genellikle iki kattan fazladır ve özellikle *Boat* ve *Graffiti* veri kümeleri için gözlenmiştir.

### 4.2.2 Orsay Müzesi Veri Kümesi ile Testler

Geliştirilen yöntemlerin gerçekçi bir nesne tanıma uygulamasında test edilmesi için bir müzedeki tablolara ait imgelerden oluşan bir veri kümesi oluşturulmuştur. Orsay müzesindeki 112 tablonun ön cepheden çekilmiş fotoğrafları bilinen nesnelere kümesini oluşturmaktadır. Referans imge örnekleri Şekil 7 ile gösterilmektedir.

Daha sonra bu veri kümesinde nesne tanıma başarımını ölçmek için 11 adet imge seçilmiş ve bunlar basılarak farklı arkaplan ve açılar ile test veri kümesi oluşturulmuştur. Aynı test imgesi için farklı test koşulları Şekil 8 ile gösterilmektedir. Test imgelerindeki görüntülerde referans nesne konumları el ile işaretlenerek test ve referans imgeler arasındaki iki boyutlu perspektif dönüşümler hesaplanmıştır. Bu dönüşümler kullanılarak yapılacak eşlemelerin doğruluk ya da yanlışlığı tespit edilebilir.

Nesne tanıma başarımının ölçülmesi için her bir test imgesi için betimleyiciler çıkarılarak tüm nesnelere hesaplanan betimleyiciler arasında arama yapılmıştır. Bu aramanın sonuç-



Tablo 3. Sekiz referans imge ile üçüncü test imgeleri için YDK ile en yakın komşu listesi hesaplandığında elde edilen doğruluk oranları. Tablo 1'e benzer şekilde eşlemelerin Denklem 3.3'e göre sıralanması sonuçları iyileştirmektedir. İlk on en yakın komşu içerisinde arama yapılması ise sonuçta elde edilen doğruluk oranlarını yükseltmiştir. Pek çok durumda önerilen yöntem ile elde edilen doğruluk oranları sadece en yakın komşuya göre iki kattan daha iyidir.

Betimleyici Eşleme Sayısı	BRIEF			BRISK			FREAK			LATCH			ORB		
	100	250	500	100	250	500	100	250	500	100	250	500	100	250	500
Bikes – NN	<b>0.97</b>	<b>0.96</b>	<b>0.87</b>	0.53	0.37	0.27	0.81	0.59	0.40	<b>0.99</b>	0.95	0.69	<b>0.99</b>	0.94	-
A–EYK-4 ikil	0.93	0.90	0.84	0.66	0.49	0.33	0.84	0.68	0.46	0.96	0.95	0.69	0.98	0.96	-
A–EYK-8 ikil	0.94	0.91	0.86	0.64	0.47	0.33	<b>0.86</b>	0.70	0.47	0.98	<b>0.96</b>	0.69	0.98	0.96	-
10–EYK-4 ikil	0.94	0.91	0.83	<b>0.70</b>	0.52	0.40	0.83	0.71	0.51	0.96	<b>0.96</b>	<b>0.71</b>	<b>0.99</b>	<b>0.97</b>	-
10–EYK-8 ikil	0.95	0.92	0.86	<b>0.70</b>	<b>0.56</b>	<b>0.41</b>	0.84	<b>0.72</b>	<b>0.53</b>	0.98	<b>0.96</b>	<b>0.71</b>	<b>0.99</b>	<b>0.97</b>	-
Boat – NN	0.01	0.01	0.00	0.40	0.25	0.15	0.35	0.28	0.20	0.21	0.13	0.08	0.04	0.04	0.04
A–EYK-4 ikil	0.00	0.01	0.00	0.72	0.34	0.17	0.74	0.48	0.27	0.41	0.18	0.09	0.25	0.12	0.06
A–EYK-8 ikil	0.01	0.00	0.00	0.71	0.34	0.17	0.76	0.49	0.27	0.45	0.18	0.09	0.26	0.12	0.06
10–EYK-4 ikil	0.00	0.00	0.00	0.88	0.59	0.35	0.76	0.68	0.50	0.59	0.31	<b>0.17</b>	0.46	0.28	<b>0.17</b>
10–EYK-8 ikil	0.00	0.00	0.00	<b>0.91</b>	<b>0.64</b>	<b>0.36</b>	<b>0.82</b>	<b>0.74</b>	<b>0.51</b>	<b>0.64</b>	<b>0.32</b>	<b>0.17</b>	<b>0.49</b>	<b>0.30</b>	<b>0.17</b>
Graffiti – NN	0.14	0.10	0.07	0.59	0.34	0.21	0.47	0.31	0.22	0.39	0.21	0.13	0.32	0.20	0.15
A–EYK-4 ikil	0.27	0.16	0.10	0.65	0.39	0.23	0.64	0.42	0.27	0.46	0.26	0.14	0.46	0.31	0.17
A–EYK-8 ikil	0.29	0.18	0.10	0.69	0.41	0.23	0.66	0.43	0.27	0.47	0.26	0.14	0.48	0.31	0.17
10–EYK-4 ikil	0.35	0.24	0.19	0.66	0.47	0.31	0.67	0.49	0.34	0.42	0.26	<b>0.17</b>	0.42	0.30	0.20
10–EYK-8 ikil	<b>0.46</b>	<b>0.33</b>	<b>0.23</b>	<b>0.70</b>	<b>0.48</b>	<b>0.32</b>	<b>0.68</b>	<b>0.52</b>	<b>0.37</b>	<b>0.49</b>	<b>0.29</b>	<b>0.17</b>	<b>0.46</b>	<b>0.33</b>	<b>0.21</b>
Wall–NN	0.99	<b>0.98</b>	0.80	<b>1.00</b>	0.88	0.56	0.92	0.68	0.45	0.90	0.69	0.41	0.98	0.82	0.52
A–EYK-4 ikil	0.99	0.96	0.86	<b>1.00</b>	0.97	0.60	<b>0.96</b>	0.80	0.52	0.89	0.81	0.45	0.98	<b>0.94</b>	0.56
A–EYK-8 ikil	0.99	<b>0.98</b>	<b>0.88</b>	<b>1.00</b>	<b>0.98</b>	0.61	0.95	0.83	0.52	0.91	0.82	0.45	<b>0.99</b>	<b>0.94</b>	0.56
10–EYK-4 ikil	0.99	0.96	0.86	<b>1.00</b>	0.95	0.66	<b>0.96</b>	0.81	0.59	0.90	0.83	0.50	0.98	0.92	<b>0.59</b>
10–EYK-8 ikil	<b>1.00</b>	<b>0.98</b>	<b>0.88</b>	<b>1.00</b>	0.97	<b>0.68</b>	0.95	<b>0.84</b>	<b>0.60</b>	<b>0.92</b>	<b>0.85</b>	<b>0.52</b>	<b>0.99</b>	<b>0.94</b>	<b>0.59</b>



Şekil 7. Orsay Müzesi veri kümesi referans imge örnekleri



Şekil 8. Orsay Müzesi veri kümesi test koşulları örnekleri

Tablo 4. Orsay veri kümesinde nesne tanıma test sonuçları

Nesne Tanıma Oranı	En Yakın Komşu	On En Yakın Komşu
Tüm Betimleyiciler	%79.09	<b>%88.18</b>
En Başarılı Betimleyiciler (%50'si elenmiş)	%80.55	<b>%86.36</b>

ları hem en yakın komşu için hem de 10 en yakın komşu için YDK yöntemiyle hesaplanmıştır. Bu büyüklükteki veri kümelerinde betimleyicilerin büyük çoğunluğu yanlış eşleşeceğinden doğrudan nesne tanıma yapılması mümkün değildir. Bunun yerine eşlenen betimleyicilerin hangi nesnelere ait olduğuna bakılarak en çok betimleyicinin eşlendiği ilk üç nesne belirlenir. Bu ilk üç nesne içinden her birine ayrı ayrı eşleme yapılarak geometrik sınırlamaları da kullanan gürbüz bir yöntemle (PROSAC) eşlemeler kontrol edilir. En çok eşleme yapılan nesne seçilmiş olur.

Betimleyici eşlemesi sadece en yakın komşu ile ve ilk on en yakın komşu ile her bir test imgesinde çalıştırılmış ve seçilen nesnelerin doğru olup olmadığı tespit edilmiştir. Sonuçlar Tablo 4 ile verilmiştir. Sadece en yakın komşu kullanılarak 110 test imgesinin 87 tanesi doğru tanınırken projede önerilen yöntem ile 97 test imgesi için başarılı tanıma yapılmıştır.



### 4.3 Veri Boyutunun Düşürülmesi Testleri

Betimleyiciler için istatistiksel verilerin tutulması başarıyı artırırken ek hafıza kullanımını da zorunlu kılmaktadır. Bu nedenle nesne tanıma fazlaca fayda sağlamayan betimleyiciler için istatistiklerin tutulması gereksizdir. Bir önceki bölümde anlatıldığı gibi nesne tanıma başarımına katkı eğitim aşamında her bir betimleyici için sentetik test imgeleri üzerinden hesaplanabilir. Böylelikle betimleyiciler nesne tanıma olası katkılarına göre sıralarak en düşük katkıya sahip olanlar elenebilir.

Yapılan testlerde nesne tanıma başarısında %2 kayıpla %50'ye kadar veri boyutunda azalma sağlamanın mümkün olduğunu göstermiştir. Sonuçlar Tablo 4'ün ikinci sırasında yer almaktadır.

## 5. TARTIŞMA VE SONUÇ

Proje kapsamında ikili betimleyicilerin eşlenmesi için Hamming uzaklığından daha iyi çalışan ve betimleyicilerin farklı bakış açılarındaki değişim istatistiklerinden faydalanan bir yöntem geliştirilmiştir. Yöntemin başarısındaki ana etmen her bir anahtar nokta için özel bir betimleme yapılmış olmasıdır. Bu da tüm anahtar noktalar için aynı özneliklerin kullanılmasına göre daha başarılı olmuştur.

Anahtar noktaların etrafındaki dokuya has betimleme yapan mevcut yöntemlerin aksine proje kapsamında büyük betimleyici kümelerine ölçeklenebilirlik temel bir hedef olduğundan iki aşamalı bir yöntem geliştirilmiştir. En yakın komşunun doğrudan eşlenmesi yerine  $K$  en yakın komşu anahtar nokta dokusundan bağımsız hesaplanmış, bu liste içerisindeki eleme dokuya has olacak şekilde tasarlanmıştır.

Geliştirilen yöntem yüzden fazla nesne içeren bir veri kümesinde test edilmiş ve nesne tanıma başarımının arttığı gözlenmiştir. Buna ilave olarak betimleyicilerden nesne tanımaya katkısı az olanlar atılarak nesne tanıma başarımında sadece %2 kayıpla hafıza kullanımı yarıya indirilmiştir.

Projede geliştirilen yöntem mobil cihazlarda gerçek zamanlı uygulamalar için uygun olmakla beraber betimleyici istatistiklerinin öğrenilmesi gerektiğinden ancak daha önceden bilinen nesnelerin tanınması mümkündür. Mevcut durumda eşlenecek imgelerin her ikisinin de uygulama çalışırken elde edildiği durumlar için kullanılamaz. Bu durumda betimleyici istatistikleri yerine [Balntas vd. \(2015\)](#) tarafından önerilen yaklaşımın önerilen iki adımda eşleme yöntemiyle harmonlanması düşünülebilir.

Bir başka geliştirme olanağı da nesnelerin mevcut nesne veri tabanına eklenme işleminin gerçek zamanlı olmasa da hızlı yapılabilmesidir. Betimleyici istatistiklerinin hesaplanması için çok sayıda görüntü sentezlenmesi gerekmektedir. Ancak bu işlemin grafik kartı kullanılarak yapılması işlem hızını arttırabilir. Bu sayede sisteme yeni bir nesnenin tanıma işlemi hızlandırılabilir.

Betimleyici veri kümelerinin daha da büyük olduğu durumlarda YDK ile elde edilen sonuçlar yeterli olmayabilir. Bu durumda yeni geliştirilmekte olan çizge temelli yaklaşım en yakın komşu

yöntemleri (Harwood ve Drummond, 2016) bir alternatif sunabilir. Ancak bu yöntemler ikilik betimleyiciler ile çalışmadığından bunların gerçek zamanlı uygulamalar için nasıl uyarlanabileceği de açık bir araştırma konusudur.

Projede kullanılan betimleyicilerin başarımlarına göre elenmesi yaklaşımı da geliştirilmeye açıktır. Önerilen yaklaşım basit ve hızlı olmasına karşın daha kapsamlı bir hedef fonksiyon tanımlanarak bunun eniyilenmesiyle elenecek betimleyicilere karar verilebilir. Örneğin önerilen eleme yönteminde her bir betimleyici ile ilgili karar birbirinden bağımsız olarak verilmektedir. Ancak bir betimleyicinin atılması aynı imgedeki diğer betimleyicilerin atılıp atılmama kararını etkileyebilir. Bu tip problemler için bir tam sayı eniyileme problemi tanımlanıp bunun çözülmesi önerilebilir (Dymczyk vd., 2015).

Eklenecek nesnelere imgelerinden çıkarılan anahtar noktalar için betimleyici ve istatistik paylaşımı yaklaşımı betimleyici elenmesine göre başarısız olmuştur. Öngörülen sebep, betimleyicilerin ayırt edici dokulardan gelmeseler bile birbirlerinden farklı oluşudur. Bu da bu betimleyicilerin doğrudan elenmesini zorlaştırmaktadır. Projede sentetik test görüntüleri yardımıyla nesne tanıma başarımı tahmin edilmiştir ama bu işlem yoğun bir yöntemdir. Anahtar nokta etrafındaki dokudan ya da betimleyici içeriğinden ayırt ediciliğin doğrudan tahmin edilmesi faydalı olacaktır.

Sonuç olarak eğitim aşamasının mümkün olduğu, betimleyici istatistiklerinin depolanmasının sorun olmadığı uygulamalar için iki aşamalı ve sadece en yakın komşu değil  $K$  en yakın komşunun tamamını kullanabilen, büyük nesne veri kümelerine ölçeklenebilen bir betimleyici eşleme yöntemi geliştirilmiştir. Nesne başına düşen betimleyici sayısını az bir başarımla düşüren bir yaklaşım da denenmiştir. İlerleyen aşamalarda nesnelere anahtar noktaların seçimi ve eğitim zamanının düşürülmesi konularının önem taşıdığı düşünülmektedir.

## Kaynakça

- Alahi, A., Ortiz, R. ve Vandergheynst, P. 2012. “Freak: Fast Retina Keypoint”, *Conference on Computer Vision and Pattern Recognition*, 510–517. [1](#), [5](#), [7](#)
- Andoni, A. ve Indyk, P. 2008. “Near-optimal Hashing Algorithms for Approximate Nearest Neighbor in High Dimensions”, *Communications of the ACM*, 51, 1, 117–122. [4](#)
- Balntas, V., Tang, L. ve Mikolajczyk, K. 2015. “BOLD-Binary Online Learned Descriptor For Efficient Image Matching”, *Conference on Computer Vision and Pattern Recognition*, 2367–2375. [3](#), [4](#), [5](#), [7](#), [26](#)
- Bishop, C. 2006. “Pattern Recognition and Machine Learning”, Springer. [10](#)
- Calonder, M., Lepetit, V. ve Fua, P. 2010. “Pareto-optimal Dictionaries for Signatures”, *Conference on Computer Vision and Pattern Recognition*. [7](#)
- Calonder, M., Lepetit, V., Özuysal, M., Trzcinski, T., Strecha, C. ve Fua, P. 2012. “BRIEF: Computing a Local Binary Descriptor Very Fast”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 34, 7, 1281–1298. [1](#), [5](#)
- Chum, O. ve Matas, J. 2005. “Matching with PROSAC - Progressive Sample Consensus”, *Conference on Computer Vision and Pattern Recognition*, 220–226, San Diego, CA. [18](#)
- Chum, O., Mikulik, A., Perdoch, M. ve Matas, J. 2011. “Total recall II: Query expansion revisited”, *Conference on Computer Vision and Pattern Recognition*, 889–896. [8](#)
- Dymczyk, M., Lynen, S., Bosse, M. ve Siegwart, R. 2015. “Keep it brief: Scalable creation of compressed localization maps”, *IEEE/RSJ International Conference on Intelligent Robots and Systems, IROS*, 2536–2542. [8](#), [27](#)
- Gupta, R. ve Mittal, A. 2008. “Smd: A locally stable monotonic change invariant feature descriptor”, *European Conference on Computer Vision*, 265–277. [3](#), [4](#), [5](#), [7](#)

- Harwood, B. ve Drummond, T. 2016. “FANNG: Fast Approximate Nearest Neighbour Graphs”, *Conference on Computer Vision and Pattern Recognition*. 27
- Lepetit, V. ve Fua, P. 2006. “Keypoint Recognition using Randomized Trees”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28, 9, 1465–1479. 3, 4, 5, 7
- Leutenegger, S., Chli, M. ve Siegwart, R. Y. 2011. “BRISK: Binary robust invariant scalable keypoints”, *International Conference on Computer Vision*, 2548–2555. 1, 5, 7
- Levi, G. ve Hassner, T. 2015. “LATCH: Learned Arrangements of Three Patch Codes”, *CoRR*, abs/1501.03719. 1, 5, 7
- Li, F. ve Kosecka, J. 2006. “Probabilistic Location Recognition using Reduced Feature Set”, *International Conference on Robotics and Automation*. 8
- Li, X., Larson, M. ve Hanjalic, A. 2015. “Pairwise Geometric Matching for Large-Scale Object Retrieval”, *Conference on Computer Vision and Pattern Recognition*. 8
- Lowe, D. 2004. “Distinctive Image Features from Scale-Invariant Keypoints”, *International Journal of Computer Vision*, 20, 2, 91–110. 1
- Manning, C., Raghavan, P. ve Schütze, M. 2008. “Introduction to Information Retrieval”, Cambridge University Press. 10
- Mikolajczyk, K. ve Schmid, C. 2004. “A Performance Evaluation of Local Descriptors”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 27, 10, 1615–1630. 1, 11
- Mikolajczyk, K., Tuytelaars, T., Schmid, C., Zisserman, A., Matas, J., Schaffalitzky, F., Kadir, T. ve Van Gool, L. 2005. “A Comparison of Affine Region Detectors”, *International Journal of Computer Vision*, 65, 1/2, 43–72. 1, 11
- Moreels, P. ve Perona, P. 2006. “Evaluation of Features Detectors and Descriptors based on 3D Objects”, *International Journal of Computer Vision*. 20
- Morel, J.-M. ve Yu, G. 2009. “ASIFT: A new framework for fully affine invariant image comparison”, *SIAM Journal on Imaging Sciences*, 2, 2, 438–469. 10
- Muja, M. ve Lowe, D. G. 2014. “Scalable Nearest Neighbor Algorithms for High Dimensional Data”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 36. 22
- Oszust, M. 2016. “An optimisation approach to the design of a fast, compact and distinctive binary descriptor”, 1–8. 1

- Özuysal, M., Calonder, M., Lepetit, V. ve Fua, P. 2010. “Fast Keypoint Recognition Using Random Ferns”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32, 3, 448–461. 3, 4, 5, 7, 9
- Rosten, E., Porter, R. ve Drummond, T. 2010. “Faster and Better: A Machine Learning Approach to Corner Detection”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 32, 1, 105–119. 1
- Rublee, E., Rabaud, V., Konolige, K. ve Bradski, G. 2011. “ORB: an Efficient Alternative to SIFT or SURF”, *International Conference on Computer Vision*, 2564–2571. 1, 5, 7
- Trzcinski, T., Christoudias, M. ve Lepetit, V. 2015. “Learning Image Descriptors with Boosting”, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 37, 3, 597–610. 1
- Uzyıldırım, F. ve Özuysal, M. 2016. “Instance Detection by Keypoint Matching beyond the Nearest Neighbor”, *Signal, Image and Video Processing*, 10, 8, 1527–1534. 6

**TÜBİTAK  
PROJE ÖZET BİLGİ FORMU**

Proje Yürütücüsü:	Yrd. Doç. Dr. MUSTAFA ÖZUYSAL
Proje No:	113E496
Proje Başlığı:	Ölçeklenebilir Hibrit Nesne Tanıma Sistemi
Proje Türü:	3501 - Kariyer
Proje Süresi:	30
Araştırmacılar:	
Danışmanlar:	
Projenin Yürütüldüğü Kuruluş ve Adresi:	İZMİR YÜKSEK TEKNOLOJİ ENS.
Projenin Başlangıç ve Bitiş Tarihleri:	01/03/2014 - 01/09/2016
Onaylanan Bütçe:	159313.9
Harcanan Bütçe:	61951.38
Öz:	<p>Gerçek zamanlı mobil nesne tanıma uygulamalarında anahtar noktaların eşlenmesi için genellikle ikilik betimleyiciler kullanılır. Ancak bu betimleyiciler kamera bakış açısı benzeri etmenlerden oldukça etkilendiğinden nokta eşleme başarımları daha karmaşık betimleyicilere göre düşüktür. Proje kapsamında ikilik betimleyicilerin eşlenmesi için iki aşamalı yeni bir yöntem geliştirilmiştir. Bu yöntemin ilk adımı bilinen nesnelerin imgelerinden çıkarılan betimleyicilerin arasında K en yakın betimleyiciyi tespit eder. İkinci adım ise bu betimleyicilerden doğru eşleme olasılığı en yüksek olanı söz konusu her bir betimleyiciye özel istatistiksel bir yaklaşımla seçer. Bu istatistikler bir öğrenme adımında sentetik görüntülerden hesaplanır. Yöntem düzlemsel ve üç boyutlu nesneler üzerindeki anahtar noktaların doğru eşlenme oranını sadece Hamming uzaklığı ile en yakın komşunun belirlenmesine göre arttırmaktadır. Ayrıca önerilen yöntemin ilk adımı yaklaşık en yakın komşu yöntemleriyle uyumlu olduğundan bu iki adımlı hibrit yöntem benzerlerinin aksine büyük betimleyici kümeleriyle çalışılmasını sağlar. Projede ayrıca nesne tanıma katkıları az olacağı öngörülen betimleyicilerin tespiti için bir yöntem geliştirilmiş, bu betimleyiciler çıkarılsa dahi nesne tanıma başarımında büyük bir kayıp yaşanmadığı gösterilmiştir.</p>
Anahtar Kelimeler:	Bilgisayarla Görü, Nesne Tanıma, İkilik Betimleyiciler, Yerelliğe Duyarlı Karma
Fikri Ürün Bildirim Formu Sunuldu Mu?:	Hayır
Projeden Yapılan Yayınlar:	<ol style="list-style-type: none"><li>1- A detailed analysis of MSER and FAST repeatability (Bildiri - Uluslararası Bildiri - Sözlü Sunum),</li><li>2- Instance detection by keypoint matching beyond the nearest neighbor (Makale - İndekli Makale),</li><li>3- Accelerating Mobile Object Tracking (Bildiri - Ulusal Bildiri - Sözlü Sunum),</li><li>4- KEYPOINT MATCHING BASED ON DESCRIPTOR STATISTICS (Tez (Araştırmacı Yetiştirilmesi) - Yüksek Lisans Tezi),</li></ol>