

## Gerçek Zamanlı İnsan Davranışı Anlamaya Doğru: Optimal-Altı Bir Şekil Tanımlayıcı

Güray SONUGÜR<sup>1\*</sup>, Elif Ebru ÇAKI<sup>2</sup>, Simge Ayşe AKAN<sup>3</sup>, Celal Onur GÖKÇE<sup>4</sup>

<sup>1</sup>Afyon Kocatepe Üniversitesi ANS Kampüsü Mekatronik Mühendisliği Bölümü

<sup>2</sup>Afyon Kocatepe Üniversitesi ANS Kampüsü Fen Bilimleri Enstitüsü Mekatronik Mühendisliği Ana Bilim Dalı

<sup>3</sup>Mustafapaşa Mah. İbrahim Ağa Cd., No:9 K:3 Gebze/KOCAELİ

<sup>4</sup>Afyon Kocatepe Üniversitesi ANS Kampüsü Yazılım Mühendisliği Bölümü

Sorumlu yazar, e-posta: gsonugur@aku.edu.tr

elif\_ebru1999@hotmail.com

simgeakan1@gmail.com

cogokce@aku.edu.tr

ORCID ID: <http://orcid.org/0000-0003-1521-7010>

ORCID ID: <http://orcid.org/0000-0002-2225-5675>

ORCID ID: <http://orcid.org/0000-0002-9319-1330>

ORCID ID: <http://orcid.org/0000-0003-3120-7808>

Geliş Tarihi: 07.04.2022

Kabul Tarihi: 07.08.2022

### Öz

#### Anahtar kelimeler

İnsan davranışı  
anlama;  
Görüntü işleme;  
Optimal-altı yöntem;  
Weizmann

Bu çalışmada insan davranışı anlama (İDA) probleminin çözümünde kullanılmak üzere özgün optimal ve optimal-altı şekil tanımlayıcıları önerilmiştir. Bu şekilde en az veri kullanımıyla en fazla davranış bilgisini sınıflandırabilmek amaçlanmıştır. Optimal şekil tanımlayıcısı başarısı yüksek olmakla beraber algoritmik karmaşıklığı yüksek olduğu için oldukça yavaş çalışmaktadır. Bu sorunu gidermek için daha hızlı çalışan bir optimal-altı tanımlayıcı önerilmiştir. Optimal-altı tanımlayıcının başarısı optimal tanımlayıcıya çok yakın olmakla beraber çok daha düşük algoritmik karmaşıklığa sahip olup çok daha hızlıdır. Sonuçlar Weizmann veri setinde denenmiş ve şekiller ve video bağlantıları ile gösterilmiştir. Veri setinden elde edilen siluet görüntü akışlarından 12 adet istatistiksel öznitelik çıkarılıp sınıflandırmada kullanılmıştır. Sınıflandırmada kullanılan Öklid uzaklığı yöntemi sayesinde oldukça hızlı sonuçlar üretilerek %92 doğruluk oranına ulaşılmıştır.

## Towards Real-Time Human Behavior Understanding: A Suboptimal Shape Descriptor

#### Keywords

Human behavior  
understanding;  
Image processing;  
Suboptima method;  
Weizmann

#### Abstract

In this study, two novel shape descriptors are proposed to be used in human behavior understanding problem. First is optimal shape descriptor, which has high performance but works very slow due to high algorithmic complexity. Second is suboptimal shape descriptor, performance of which is very close to optimal one, but works much more faster. Optimal means using minimum data to represent maximum knowledge. Algorithms are run on Weizmann dataset and results are shown both as figure and video link. Classification was performed using 12 statistical features extracted from the data sets' human silhouettes. An accuracy rating of 92 percent was obtained by using Euclidean distance in classification.

© Afyon Kocatepe Üniversitesi

### 1. Giriş

Günümüzde bilgisayarlı görme çalışmalarının en önemlilerinden biri de insan davranışlarını anlamadır. Çok farklı kullanım alanları bulunan bu konuda çeşitli çalışmalar yapılmaktadır. Örneğin, spor müsabakalarının hareketli görüntülerinden sporcuların yaptığı eylemlerinin neler olduğunun tanınması, bale koreograflerinin algılanması, yaşlı bakımında düşme tespiti, dudak okuma, jest ve

mimikleri anlama gibi konular bu tip çalışmaların uygulama alanını oluşturmaktadır.

Gözlem videolarından insan davranışı tanıma işlevi, görüntü işleme ve bilgisayarla görmenin aktif kullanıldığı bir alandır. Otogar, tren garı, hava alanı, alışveriş merkezi, halka açık merkezler, meydanlar, konser alanları, bankalar, alışveriş merkezleri, hükümet binaları, askeri alanlar gibi alanlarda; bagaj bırakma, kavgalar, saldırılar, silah varlığı gibi şüpheli

durumları gözlemek ve önlemek amaçlı insan davranışları algılama faaliyetleri gerçekleştirilmektedir. Bu tip faaliyetleri belirleyebilmek için video karelerinin analiz edilmesi gerekir. Bunun yanında koşmak, yürümek, el sallamak gibi davranışlar ise olağan davranışlar olarak sınıflandırılır. İnsan davranışı anlama çalışmalarında görüntü ve sinyal işleme, öznitelik çıkarma, makine öğrenimi, 2-3 boyutlu geometri gibi pek çok disiplin bir arada kullanılmaktadır. İnsanları algılama ve otomatik tespit etme, davranış ve hareketlerini anlama yeteneği için öğrenme bazlı yöntemlere sıklıkla başvurulmaktadır.

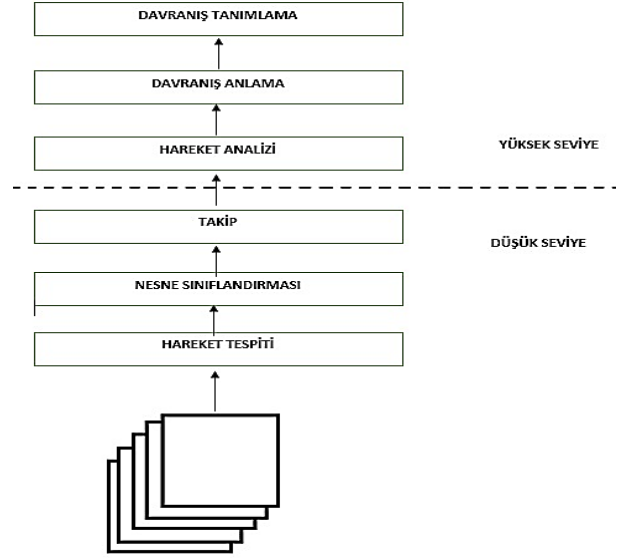
Bu çalışmada hareketli görüntülerdeki insan eylemlerinin algılanması ana hedeftir. Bu hedef doğrultusunda uygulama alanı olarak Weizmann veri setinde bulunan insan eylemlerinin analizi seçilmiştir. Veri kümesinde bulunan 10 farklı eylem ve 9 farklı şahsın hareketlerini inceleyerek medyan arka plan çıkarma (Median Background Subtracting) ve Canny kenar bulma yöntemleri kullanılarak insan silüetleri yüksek başarı oranı ile elde edilmiştir. İnsan davranışlarının en az miktarda veri kullanılarak temsil edilebilmesi için optimal ve optimal-altı yöntemler geliştirilerek davranışlar nokta bulutu kümeleri ile gösterilmiş ve klasik vektörel uzaklık yöntemi ile davranış sınıflandırması gerçekleştirilmiştir.

İkinci bölümde çalışmada kullanılan yöntemler detaylı bir şekilde anlatılmış, üçüncü bölümde hareket bilgilerini kullanarak eylem tanıyan sistemin detayları ve deneysel sonuçları verilmiştir. Beşinci bölümde elde edilen sonuçların diğer çalışma sonuçları ile karşılaştırılmasına ve elde edilen sonuçların değerlendirilmesine yer verilmiştir. Son olarak altıncı bölümde elde edilen sonuçlara yer verilmiş olup çalışmanın devamı ele alınmıştır.

İnsan davranışı anlama ile ilgili literatürde yapılan çalışmalar video gözetim süreçlerinin bir parçası olarak tanımlanmaktadır. Bu konuda pek çok başlık altında çalışmalar yapılmıştır. Bunlardan başlıklar: İnsan etkinliği tespiti ve etiketlenmesi (Dhulekar, Chitte ve Pardeshi, 2017), Anormal insan davranışı tanıma (Oluwatoyin and Kejun 2012), Mimik/jest tanıma sistemleri (Siddharth and Anupam 2015),

Kalabalık içinde hareket tanıma (de Campos 2014) olarak gösterilebilir.

Antonakaki (2011) tarafından video gözetim süreçleri Şekil-1'de gösterildiği gibi düşük seviye ve yüksek seviye olarak iki sınıfta tanımlanmıştır.



Şekil 1. Video gözetim süreçleri (Antonakaki et al. 2011)

İnsan davranışlarının anlaşılması, basitçe, zamanla değişen öznitelik verilerinin sınıflandırılması, yani bilinmeyen bir test dizisinin tipik davranışları temsil eden bir grup etiketli referans dizisiyle eşleştirilmesi olarak düşünülebilir. İnsan davranışı anlama süreçleri (insan etkinliği tanıma ve etkinlik örüntülerinin bulunması), bilgisayarlı görme ve makine öğrenimi disiplinlerini bir arada kullanmayı gerektirmektedir (Popoola ve Wang, 2012). Bu nedenle, bir derin öğrenme süreci ve öğrenmenin iyileştirilmesi için doğru ve etkili özniteliklere ihtiyaç duyar. İnsan davranışlarını görüntü dizilerinden anlamadaki temel zorluklar: (a) eğitim için toplanan örnek görüntülerden referans davranış dizilerini öğrenmek, (b) her bir hareket modelindeki öznitelik verilerinin küçük varyasyonlarını da anlayabilmek için hem eğitim hem de eşleştirme yöntemleri geliştirmek, (c) görüntü arka planındaki ve nesne görünümündeki değişikliklerden etkilenmeyecek öznitelikleri bulmaktır (Wang ve Maybank, 2004; Mabrouk ve Zagrouba, 2018). Literatürde davranış anlama süreçlerinde pek çok öznitelik kullanılmıştır. Örnek olarak; optik akış temelli öznitelikler (Tao et

al. 2017), Uzak-zaman temelli hacimsel, küboid ve blob öznelikleri (Yogameena and Priya 2015), HOG ve siluet gibi şekilsel öznelimler (Aslan vd. 2015) ve MoSIFT (Xu et al. 2014) gibi ölçekten bağımsız öznelimler örnek olarak verilebilir.

İnsan davranışı anlama süreçleri belirli bir zaman aralığındaki etkinliklerin tanınmasını gerektirir ve genellikle güvenlik amaçlı olarak anormal davranışların tespit edilmesine dayanır. Anormal davranışları tespit etmek için araştırmacılar tarafından genel olarak üç modelde çalışmalar yapılmıştır. İlk yöntemde yarı-eğitici bir model kullanılmaktadır. Bu yöntemde öğrenme sistemleri öncelikle normal davranışlar için eğitilmekte ve normal davranışların dışında kalan etkinlikler anormal olarak tespit edilmektedir (Chianese et al. 2008). İkinci modelde eğitici öğrenme sistemi kullanılmaktadır. Bu yöntemde öğrenme sistemi hem normal hem de anormal davranışlar için etiketli veriler ile kural tabanlı olarak eğitilmektedir (Park et al. 2010). Üçüncü model olan eğitici öğrenme modelinde normal ve anormal davranışlar etiketsiz veriler üzerinden öğrenilmektedir (Weiya et al. 2015).

Eğitim modelleri ile birlikte sınıflandırma yöntemleri de sonuçlar üzerinde etkilidir. En çok kullanılan yöntemlerin başında Gauss Karışımı ve derin öğrenme destekli Gauss Karışımı temelli yöntemler gelmektedir. (Zhao and Li 2014, Feng et al. 2017), Kullanılan diğer yöntemler arasında; eğitici öğrenme modellerinden bulanık mantık temelli (Acampora et al. 2015) ve Öz Örgütlenmeli Sinir Ağları (Johnson and Hogg 1996), Evrimsel Sinir Ağı (ESA) modelleri (Shifu et al. 2016), Yapay Sinir Ağları modelleri (Gökçe and Sonugür 2022), Zaman Gecikmeli Sinir Ağları (Yang and Ahuja 1998) ve Gizli Markov Model tabanlı yöntemler (Acharya and Gantayat 2015) sayılabilir.

İnsan hareketini algılama konusunda, Akdağ (2015) 2 boyutlu hareket algılama tekniğini kullanmıştır. 3-B derinlik haritası veri dizileri için siluet akı olarak adlandırdığımız yeni bir insan davranışı tanıma yöntemini önermektedir. Gökçe (2016), insan davranışı anlama çalışmalarına hiyerarşik bir yapı önermiştir. Futbol video oyunu (FVO) veri seti isimli yeni bir veri seti üretip, diğerlerinden farklı olarak daha karmaşık davranışları içermektedir. Futbol

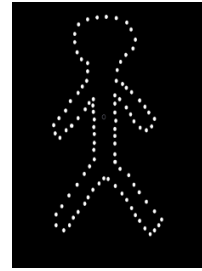
topu bayır inişi (gradient descent) metoduyla eğitilmiş ve sorgularla öğrenme tekniğiyle geliştirilmiş çok katmanlı yapay sinir ağı ile tespit edilmektedir. Topun etrafındaki ilgili alan bu yöntemle çıkarıldıktan sonra atomik aksiyon iki yöntemden birisiyle tanınmaktadır.

Bu çalışmada kullanılan optimal ve optimal-altı yöntem, şekillerin tanımlanmasında farklı bir algoritma sunmaktadır. Kenar çıktılarında elde edilen veriler üzerinden hesaplamalar ile belirlenen noktalar şekil tanımlamasında kullanılmaktadır. Bu yaklaşım ile hızlı ve etkin sonuç çıktılarında ulaşım hedeflenmektedir.

## 2. Materyal ve Metot

### 2.1 Uyarlanan Yöntemler

Önerilen yöntemlerde öncelikle video karelerinin arka planları ile ön planları ayrılmaktadır. Ön planların her biri ayrı ayrı eşikleme işlemine tabii tutulmaktadır. Elde edilen görüntülerin Canny kenar tespit algoritması ile kenar görüntüleri oluşturulmaktadır. Kenarların baz alınması ile nesne silüetlerinin kenarları 100 adet eşit aralıklı nokta koordinatları ile temsil edilecektir. Oluşturulan her bir nokta için x ve y koordinatları belirlenmektedir. Silüetler üzerindeki nokta dağılımları Şekil 2'de verilmiştir.

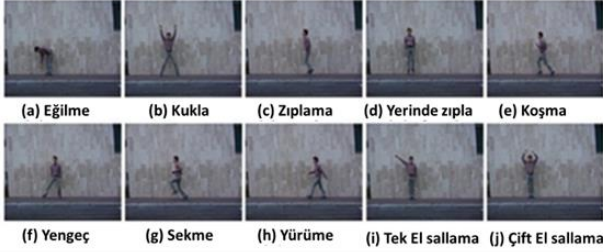


Şekil 2. Eşit Aralıklı Kenar Çıktı Görseli

#### 2.1.1 Weizmann Veri Kümesi

Weizmann Veri Seti literatürde sıkça kullanılan bir veri setidir. Hareketli görüntülerde insan eylemlerinin tanınması alanında, önerilen farklı algoritmaların karşılaştırılmasında kullanılan Weizmann veri kümesinde 90 tane 180x144 çözünürlükte hareketli görüntü bulunmaktadır. Bu hareketli görüntülerde dokuz farklı şahıs (Daria, Denis, Eli, Moshe, Odo, Irai Lena, Lyova ve Shahar), on farklı eylemi gerçekleştirmektedir. 10 farklı eylem: koşma, yürüme, tek ayak sekerek ilerleme,

eğilme, tek elini sallama, iki elini sallama, bir uçtan bir uca yengeç yürüyüşü, zıplayarak ilerlemek, yerinde sabit kalarak zıplamak, kolları ve bacaklarını açıp kapayarak yerinde zıplama şeklinde sıralanmıştır. Arka plan sabit olduğu için başlangıç çalışmalarına uygundur. Şekil 3'de bu veri kümesinden alınan birkaç örnek görüntü verilmiştir.



Şekil 3. Weizmann Veri Kümesinden Alınan Örnek Çerçeveler

### 2.1.2 Arka Plan Çıkarma

Arka plan çıkarma, görüntüdeki arka planı ortadan kaldırmanın bir yoludur. Birçok görüntü tabanlı uygulamada kullanılan ön işleme adıdır. Bu işlem için çeşitli yöntemlerle bir arka plan modeli hazırlanmakta ve ardışık olarak video karelerinden çıkarılmaktadır.

Arka Plan Çıkarmasının günlük hayatta çeşitli kullanım durumları vardır. Örnek olarak, trafik kameralarının araçlar hakkında bilgi toplaması veya girilmesi yasak alanlara girenleri tespit etmesi verilebilir. Arka plan hareketsiz ise işlemler kısmen kolaylaşmaktadır. Ancak çoğu durumda arka planda bir kısım hareketler oluşmaktadır ve bu nedenle elimizdeki görüntüler ile bir arka plan modeli oluşturmamız gerekmektedir. Gölge, ışık ve titreşim gibi çevresel faktörler olduğunda bu durum daha karmaşık hale gelmektedir.

Arka plan çıkarma işlemi için çoğunlukla Gauss Karışım tabanlı (Zhao and Su 2017) ve Çerçeve Çıkarımı tabanlı (Akilan *et al.* 2018) yöntemler kullanılmaktadır. Bu çalışmada Medyan arka plan çıkarımı (Sakpal and Sabnis 2018) yaklaşımı kullanılmıştır.

### 2.1.3 Canny Kenar Tespit Edici

Canny kenar algılama (Canny 1986), görüntülerdeki çok çeşitli kenarları algılamak için kullanılan çok adımlı bir algoritmadır. Gradyanların yoğunluğunu hesaplamak için Gauss'un türevini temel alan bir filtre kullanır. Gauss filtresi, görüntü gürültüsünün etkisini azaltır. Ardından, gradyan büyüklüğünün

maksimum olmayan pikselleri silinerek, olası kenarlar 1 piksellik eğrilere indirgenir. Son olarak, gradyan büyüklüğüne histerezis eşiklemesi uygulayarak, kenar pikselleri korunur veya elenir. Avlash and Kaur (2013)'e göre Canny optimal bir kenar bulucudur yani kenar olabilecek tüm piksel gruplarını işaretler. Yapılan çalışmada çok sayıda kenar noktasına ihtiyaç duyulduğu için Canny yöntemi seçilmiştir. Ayrıca literatürdeki kenar bulucuların karşılaştırıldığı çalışmalar incelendiğinde en yüksek doğrulukta sonuçlara Canny yöntemi ile ulaşıldığı gözlenmiştir (Jain *et al.* 2014, Acharya *et al.* 2012).

### 2.1.4 Eşikleme

Seçilen eşik değerlere göre piksel değerlerinin değiştirilmesi için kullanılan bir tekniktir. Bu teknikte her piksel değeri eşik değeri ile karşılaştırılmaktadır. Piksel değeri eşikten küçükse 0, aksi takdirde maksimum değere (genel olarak 255) ayarlanmaktadır. Eşikleme, ön plan olarak kabul edilen bir nesneyi arka planından ayırmak için kullanılan çok popüler bir bölütleme tekniğidir. Eşik değerinin adaptif olarak seçilmesi görüntüdeki gürültülerin azaltılması açısından gerekli bir yaklaşımdır. Bu çalışmada her görüntü çerçevesinde ışık oranları farklı olduğu için ayrı ayrı eşik değerler Otsu yöntemi kullanılarak hesaplanmıştır. Bu şekilde video sinyali içinde bir uyarlanma sağlanmıştır.

## 2.2 Önerilen Özgün Algoritmalar

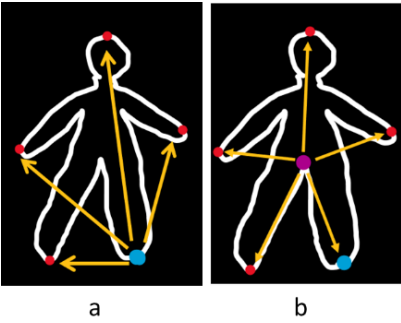
### 2.2.1 Optimal Algoritma

Çalışmamızda öncelikle tarafımızca geliştirilen Optimal algoritma uygulanmıştır. Optimal algortmada Canny yöntemi ile elde edilen nesne silüetlerinin 100 adet sınır koordinatı kullanılmaktadır. Bu koordinatlarının (x,y) eksenlerindeki en düşük ve en yüksek değerlerini kullanarak doğu, batı, kuzey ve güney olmak üzere 4 adet nokta belirlenmiştir. Geriye kalan sınır koordinatları içerisinden belirlenen 4 noktaya en uzak olan 20 sınır noktası hesaplanarak insan davranışını temsil edecek olan öznitelik vektörü oluşturulmaktadır. Söz konusu dört noktanın şekilsel gösterimi Şekil 4'de verilmiştir.



Şekil 4. Optimal-altı algoritmada seçilen sınır koordinatlarının temsili gösterimi

Örnek olarak 5. nokta koordinatları hesaplanırken başta belirlenen 4 sabit koordinata en uzak olan koordinat belirlenmiştir. Sonrasında 6. koordinat ilk beş koordinata en uzak olan koordinat olarak belirlenmiştir. Sınır koordinatlarının yerlerinin seçilmesi Şekil 5’de gösterilmiştir.



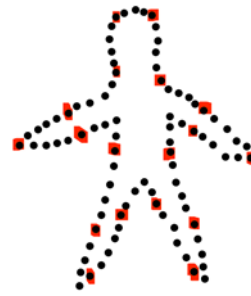
Şekil 5. Beşinci (a) ve Altıncı (b) Koordinatların Temsili Gösterimi

Bu işlem parametrik olarak belirlenen toplam sınır noktası sayısına ulaşıncaya kadar devam etmektedir. Tüm uzaklık hesaplamalarında Öklid uzaklığı kullanılmıştır. Optimal algoritmanın hesapsal yükünün oldukça yüksek miktarda bilgisayar kaynağı ve zaman gerektirmesi nedeniyle davranışı belirleyen özneliklerin oluşturulması oldukça uzun sürmektedir. Bu nedenle Optimal-altı algoritma geliştirilmiştir.

### 2.2.2 Optimal-altı Algoritma

Optimal-altı algoritmada Canny yöntemi ile elde edilen 100 adet nesne sınır koordinatı önce x koordinatına göre sıralanmaktadır. Ardından nesne silüetinin temsilinde kullanılacak nokta sayısı kadar (x,y) koordinatı eşit aralıklı olacak şekilde seçilmektedir. Örneğin, bir nesne silüeti 20 nokta ile temsil edilecekse her beş koordinattan birisi, 5

nokta ile temsil edilecek ise her 20 koordinattan birisi seçilmektedir. Aynı işlem adımları y koordinatları için yapılarak bulunan tüm koordinatlar kaydedilir. Toplamda elde edilen tüm koordinatlar için Optimal algoritma prensipleri uygulanır. Optimal-altı algoritma, toplam sınır koordinatı sayısından çok daha az koordinat sayısında uygulandığı için daha hızlı çalışmakta ve Optimal algoritmaya oldukça yakın doğrulukta sonuçlar vermektedir. Optimal-altı algoritmada eşit aralıklarla gerçekleştirilen koordinat seçimleri Şekil 6’da gösterilmiştir.



Şekil 6. Optimal-Altı Algoritmada Seçilen Sınır Koordinatlarının Temsili Gösterimi

### 3. Bulgular

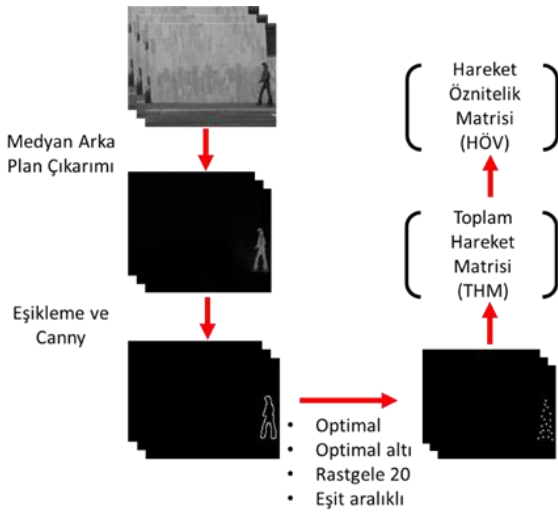
Geliştirilen yöntemlerdeki amaç; en küçük boyutlu öznelik vektörü kullanılarak insan davranışlarını en yüksek oranda temsil edebilmektir.

Yapılan deneysel çalışmalarda Weizmann veri setinde bulunan ve her biri dokuz farklı kişi tarafından gerçekleştirilen yürüme (walk), koşma (run), zıplayarak ilerleme (jump), yanal yürüme (side walk) ve el sallama (wave) hareketi videoları kullanılmıştır. Her bir videonun süresi farklı olduğu için toplam çerçeve sayısı da farklıdır.

Önerilen Optimal ve optimal-altı yöntemde videolardaki insan silüetleri 20’şer sınır noktası ile temsil edilmiştir. Bu temsildeki başarıyı karşılaştırmak üzere rastgele 20 ve eşit aralıklı 20 sınır noktası ile temsil edilen silüetler kullanılmıştır. Öncelikle başarılı temsil ölçütleri oluşturulmuştur. Bunun için seçilen silüet sınır noktaları hareketin toplam süresi boyunca çerçeve çerçeve kaydedilmiş ve bir toplam hareket matrisi (THM) oluşturulmuştur. Her bir hareket videosu için bu matris farklı boyutlarda oluşacağı için bire bir karşılaştırma yapılamayacaktır. Bu durumu



düzeltebilmek amacıyla her THM normalize edilmiş ve istatistiksel öznitelikleri çıkarılmıştır. Bu öznitelikler: Aritmetik ortalama, Geometrik ortalama, Standart sapma, Medyan, Değişim katsayısı, Değişim aralığı, Çarpıklık, Basıklık, Pearson çarpıklık katsayısı, Bowley çarpıklık katsayısı, 3.moment ve 4.moment olarak seçilmiştir. Bu şekilde her bir hareketin tamamını temsil eden ve silueti temsil eden nokta sayısından bağımsız [1x12] boyutlarında bir hareket öznitelik vektörü (HÖV) oluşturulmuştur. Gerçekleştirilen işlem adımları Şekil 7’de gösterilmiştir.



Şekil 7. Uygulamada Gerçekleştirilen İşlem Adımları

Böylece tüm hareketlerin karşılaştırılabileceği bir eşit boyutta bir ölçüt meydana getirilmiştir. Öncelikle her bir hareket videosu için kesin referans (ground truth) bir HÖV oluşturulmuştur. Burada hareket eden insan silüetleri 100’er adet sınır noktası kullanılarak THM oluşturulmuş ve bu matristen de HÖV çıkarılmıştır. Yapılan işlemin doğru olarak kabul edilebilmesi için birden fazla kişi tarafından gerçekleştirilen örneğin koşma hareketlerine ait HÖV’ler arası Öklid uzaklığı düşük iken farklı hareketler için (koşma ve yürüme) çıkarılan HÖV’ler arası Öklid uzaklığının daha fazla olması gerekir. Bu durumu ortaya koyan deneysel veriler Çizelge 1’de verilmiştir.

Çizelge 1. Kesin Referans HÖV Değerleri Arasındaki Öklid Uzaklıkları

	Daria Walk	Daria Run	Denis Walk	Denis Run
Daria Walk	0	40.3	19	44.1
Daria Run	40.3	0	40.9	18.2

Denis Walk	19	40.9	0	39.8
Denis Run	44.1	18.2	39.8	0

Çizelge 1’den anlaşılacağı üzere aynı tip hareketler arası uzaklıklar, farklı tipteki hareketler arasındaki uzaklıklardan çok daha düşük kalmaktadır. Bu nedenle HÖV’ler arası Öklid uzaklıklarının geçerli bir ölçüt olarak kabul edilmiştir.

Optimal ve optimal-altı yöntemlerin gerçek insan hareketlerini hangi doğrulukta temsil edebildiklerini görebilmek için bir karşılaştırma deneyi gerçekleştirilmiştir. Bu deneyde herhangi bir hareket videosunun kesin referans HÖV’ü ile optimal yöntem, optimal-altı yöntem, rastgele 20 sınır noktası ve eşit aralıklı 20 sınır noktası ile temsil edilen hareketlerin HÖV’leri arasındaki Öklid uzaklıkları hesaplanmıştır. Kesin referansa vektörel olarak en yakın olan hareket, en doğru temsil edilmiş hareket olarak kabul edilecektir. Yapılan deneylerin sonuçları Çizelge 2’de verilmiştir.

Çizelge 2. Kesin Referans HÖV’leri ile Hesaplanan HÖV’ler Arasındaki Öklid Uzaklıkları

Kesin Referans Seçilen Hareketler	Optimal	Optimal-altı	Rastgele 20	Eşit Aralıklı 20
Daria Walk	23.3	27.1	42.1	29.8
Denis Run	24.1	25.6	44.8	25.5
Moshe Jump	21.8	26.6	45	31.6
Eli Wave	20.9	24.8	45.2	30.1
Lena Side	22.7	23.1	52	30.1
Daria Run	27.1	30.8	44.1	33
Denis Jump	21	27.2	44.2	32.2
Moshe Wave	20.1	25.1	39.7	29.3
Eli Side	23.3	31.1	43.3	38.1
Lena Walk	24.2	32.6	45.9	40.7

Çizelge 2’den anlaşılacağı üzere kesin referanslara en yakın olan yöntemler optimal ve optimal-altı yöntemlerdir. Rastgele seçilen sınır noktaları oldukça kötü bir performans sergilemiştir. Eşit aralıklı noktalar ise performans olarak optimal-altı yöntemden daha düşük performans göstermiştir. Bu durumda; çalışmada baz alınan hareketlerden birisini sınıflandırabilmek için optimal veya optimal-altı yöntem ile oluşturulan HÖV ile kesin referans HÖV’leri arasındaki uzaklıkları hesaplayıp en düşük değeri seçmek yeterli olacaktır. Yukarıda deneyleri gerçekleştirilen ölçüt haricinde işlem süreleri de önemli bir ölçüttür. İşlem süreleri ölçülürken

hareketin başlangıcından sınıflandırılmasına kadar geçen süre sn olarak hesaplanmıştır. Elde edilen sonuçlar Çizelge 3’de verilmiştir.

**Çizelge 3.** Seçilen Bazı Hareket Tipleri için Gerçekleşen İşlem Süreleri (sn)

Kesin Referans Seçilen Hareketler	Optimal	Optimal-altı	Rastgele 20	Eşit Aralıklı
Daria Walk	5.2	3.9	3.5	3.8
Denis Run	4.5	2.5	2	2.2
Moshe Jump	3.4	2.5	1.9	2.4
Eli Wave	9.1	7.9	5.8	6.9
Lena Side	4.3	3.3	2.6	3
Daria Run	3.3	2.4	1.9	2.1
Denis Jump	3.9	2.8	2.1	2.6
Moshe Wave	4.3	3	2.5	2.8
Eli Side	3.8	2.6	1.9	2.2
Lena Walk	4.9	3.7	3	3

En düşük işlem süreleri rastgele 20 yöntemi ile seçilen sınır noktalarında hesaplanmıştır. Ancak Çizelge 2’deki verilere göre en başarısız sınıflandırma performansı yine rastgele 20’dir.

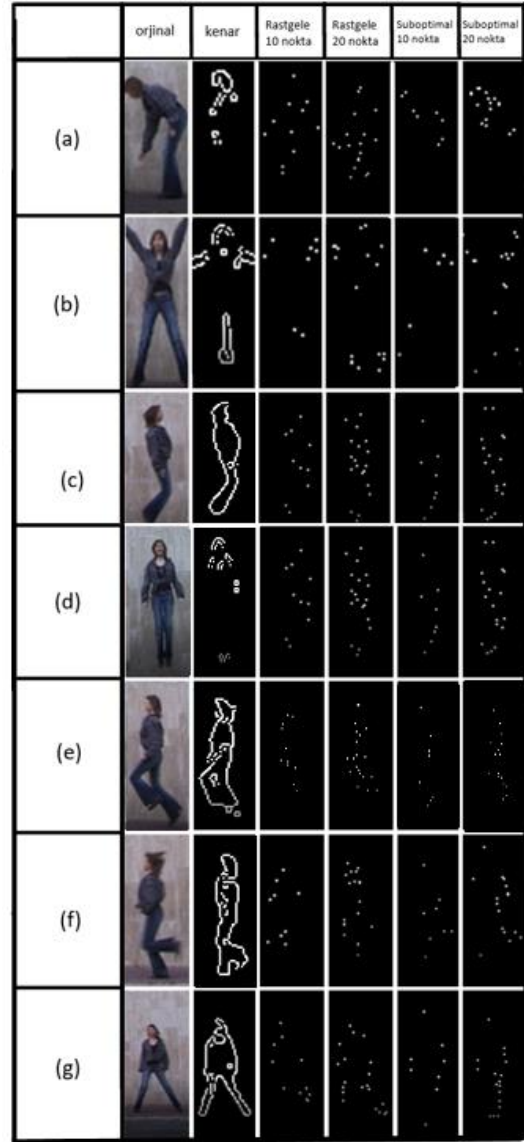
Davranış yani hareket tanımadaki toplam performansı ölçmek amacıyla eşik değer “Öklid uzaklığı=25” seçilerek deneysel çalışmalarda kullanılan tüm videolar test edilmiştir. Bu işlem sonucunda elde edilen doğru sınıflandırma oranları Çizelge 4’de verilmiştir.

**Çizelge 4.** Doğru Sınıflandırma Oranları (%)

Seçilen Hareketler	Optimal	Optimal-altı
Walk (Yürüme)	92	90
Run (Koşma)	90	90
Jump (Zıplama)	91	89
Wave (El sallama)	89	85
Side (Yanal yürüyüş)	90	90

El sallama hareketinde vücut hareketi yapılmayıp sadece el sallama hareketi yapıldığı için arka plan çıkarımı düşük başarıyla gerçekleştirildiği için doğruluk oranları da düşük çıkmıştır. Optimal yöntemde işlem sürelerinin uzunluğu dikkat çekmektedir. Optimal-altı yöntemde ise işlem süreleri ve sınıflandırma performansı bakımından diğer seçeneklere göre daha iyi bir seçenek olduğu gözlenmiştir. Kullanılan yöntemler neticesinde elde edilen tüm hareketlerden örnek sahneler Şekil 8’de verilmiştir. Ayrıca, temsil videolarının linkleri aşağıda verilmiştir.

- (a) Daria Bend - <https://youtu.be/546D60sjS3Y>,
- (b) Daria Jack - <https://youtu.be/7Dts1UFzt7w>,
- (c) Daria Jump - <https://youtu.be/0apS35dGupl>,
- (d) Daria P-Jump - <https://youtu.be/tevPMZaFc4U>,
- (e) Daria Run - <https://youtu.be/H1osp7ct1Ao>,
- (f) Daria Side - <https://youtu.be/y6AQ0UDNqko>,
- (g) Daria Skip - <https://youtu.be/TPUIXifBdKg>



**Şekil 8.** Örnek Hareket Çıktıları

#### 4. Tartışma ve Sonuç

Optimal-altı şekil tanımlayıcı yönteminin videolar üzerinde yapılan işlemler sonucunda performans ve süre ilişkisi açısından en verimli sonucu ürettiği gözlenmiştir. Çalışmadaki amaç; en az nokta ile en başarılı hareket temsilini gerçekleştirmektir. Optimal yöntem ile en başarılı sınıflandırma

performansı elde edilmiştir. Ancak bu yöntemde noktaların belirlenmesi işlemi oldukça zaman almaktadır. Sınır noktalarının optimal-altı olarak belirlenmesi en ideal ve hızlı çıktıyı oluşturmaktadır.

Çalışmamız kapsamında, günümüzde ileri teknoloji kullanımı ile yasa dışı faaliyet ve olaylar, kaza, kavga, hırsızlık, terör gibi durumlarda algılayabilen sistem tasarımının belirli bir kısmı gerçekleştirilmiştir. Kullanılan veri kümesi Optimal-altı yöntem ile işlenmiştir. Optimal-altı yöntem optimal yöntemle göre daha düşük işlem sürelerinde sonuç ürettiği gözlenmiştir. Doğruluk performans sonuçlarının da yakın değerlerde olması ve hızlı işlemi sonucu nedeniyle tercih edilmiştir.

Çalışma kapsamında uzaysal düzende bulunacak noktalar belirlenmiş, kayıt altına alınmıştır. Klasik bir yaklaşım olan vektörel uzaklıklar kullanılarak insan hareket ve davranışlarını içeren videolar hareket tipilerine göre sınıflandırılmıştır. Yapılan gözlemler sonucunda insan davranışlarının çok daha düşük boyutlu veri kümeleri kullanılarak tanınabileceği gözlenmiştir. Herhangi bir makine öğrenmesi yöntemi kullanılmadığı için verilerin eğitime ve etiketlenmesine ihtiyaç duyulmamıştır. Projenin bir sonraki aşamasında verilerin Temel Bileşen Analizi (Principal Component Analysis-PCA) kullanılarak boyutunun daha da azaltılması sağlanacaktır. Ayrıca Weizmann veri kümesi harici bir veri kümesinde deneysel çalışmalar gerçekleştirilecektir. Makine öğrenmesi yöntemleri, özellikle Evrimsel Sinir Ağları ve Yapay Sinir Ağları kullanılarak sınıflandırma başarısının yükseltilmesi hedeflenmektedir. YSA ağ yapısının basitliği göz önüne alındığında işlem süresi olarak oldukça düşük sürelerle ulaşılabileceği öngörülmektedir. Çalışmanın anlık kayıt alabilen ve anlık analiz sağlayabilen bir sisteme dönüştürülmesi hedeflenmektedir. Hedefler doğrultusunda projenin başarı ile sonuçlanması sistemin birçok alanda faaliyete geçmesine olanak tanıyacaktır. Sistemin kullanımı ile birçok alanda üst düzey güvenlik sağlanabilecektir.

## 5. Kaynaklar

Acampora, G., Foggia, P., Saggese, A., Vento, M. 2015. A hierarchical neuro-fuzzy architecture for human behavior analysis, *Information Sciences*, **310**, 130-148.

Acharya, B. R. ve Gantayat, P. K. 2015. Recognition of human unusual activity in surveillance videos. *International Journal of Research and Scientific Innovation (IJRSI)*, **2(5)**, 18-23.

Acharjya, P. P., Das, R., & Ghoshal, D. 2012. Study and comparison of different edge detectors for image segmentation. *Global Journal of Computer Science and Technology*.

Akdağ E. 2015. Human Behavior Understanding Through 3D Data, Yüksek Lisans Tezi, Fen Bilimleri Enstitüsü, ODTÜ, Ankara, 87.

Akılan, T., Wu, Q. J., Yang, Y. 2018. Fusion-based foreground enhancement for background subtraction using multivariate multi-model Gaussian distribution. *Information Sciences*, **430**, 414-431.

Antonakaki, P., Kosmopoulos, D., & Perantonis, S. J. 2011. Detecting abnormal human behaviour using multiple cameras. *Signal Processing*, **89(9)**, 1723-1738.

Aslan, M., Sengur, A., Xiao, Y., Wang, H., Ince, M.C., Ma, X. 2015. Shape feature encoding via fisher vector for efficient fall detection in depth-videos, *Applied Soft Computing*.

Avlash, M., & Kaur, L. 2013. Performances analysis of different edge detection methods on road images. *International Journal on Recent and Innovation Trends in Computing and Communication*, **2(6)**, 27-38.

Blank M., Gorelick L., Shechtman E., Irani M. ve Basri R. 2005. Actions as Space-Time Shapes, The Tenth IEEE International Conference on Computer Vision (ICCV), Beijing, China, 1395-1402.

Chianese, A., Moscato, V., ve Picariello, A. 2008. Detecting abnormal activities in video sequences. In Proceedings of the 2008 Ambi-Sys workshop on Ambient media delivery and interactive television, 1-8.

De Campos, T. 2014. A survey on computer vision tools for action recognition, crowd surveillance and suspect retrieval, XXXIV congresso da sociedade brasileira de computacao (CSBC) 1123-1132.

Dhulekar, P., Gandhe, S. T., Chitte, H., ve Pardeshi, K. 2017. Human action recognition: An overview. In Proceedings of the international conference on data engineering and communication technology, Springer, Singapore, 481-488.

Feng, Y., Yuan, Y. ve Lu, X. 2017. Learning deep event models for crowd anomaly detection. *Neurocomputing*, **219**, 548-556.



- Gökçe C.O., 2016. Human Behavior Understanding Using Video Analysis, Doktora Tezi, Fen Bilimleri Enstitüsü, ODTÜ, Ankara, 106.
- Gökçe, B., ve Sonugür, G. 2022. Recognition of dynamic objects from UGVs using Interconnected Neural network-based Computer Vision System. *Automatika*, **63(2)**, 244-258.
- Jain, A., Gupta, M., & Tazi, S. N. 2014. Comparison of edge detectors. In 2014 International Conference on Medical Imaging, m-Health and Emerging Communication Systems (MedCom), 289-294. IEEE.
- Johnson, N., ve Hogg, D. 1996. Learning the distribution of object trajectories for event recognition. *Image and Vision computing*, **14(8)**, 609-615.
- Mabrouk, A. B., ve Zagrouba, E. 2018. Abnormal behavior recognition for intelligent video surveillance systems: A review. *Expert Systems with Applications*, **91**, 480-491.
- Oluwatoyin, P.P. ve Kejun, W. 2012. Video-based abnormal human behavior recognition – a review, *IEEE Transactions on Systems, Man, and Cybernetics, Part C*, **42 (6)**. 865-878.
- Park, K., Lin, Y., Metsis, V., Le, Z., ve Makedon, F. 2010, June. Abnormal human behavioral pattern detection in assisted living environments. In Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments, 1-8.
- Siddharth, R. ve Anupam, A., 2015. Vision based hand gesture recognition for human computer interaction: A survey, *Artificial Intelligence Review*, **43(1)**.
- Sakpal, N. S. ve Sabnis, M., 2018. Adaptive background subtraction in images. In 2018 International Conference on Advances in Communication and Computing Technology (ICACCT), 439-444. IEEE.
- Wang, L. ve Maybank, S., 2004. A survey on visual surveillance of object motion and behaviors. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, **34(3)**, 334-352.
- Weiya, R. L. Guohui, S., Boliang ve Kuihua, H. 2015. Unsupervised kernel learning for abnormal events detection, *The Visual Computer*, **31**, 245-255, 10.1007/s00371-013-0915-0.
- Xu, L., Gong, C., Yang, J., Wu, Q., & Yao, L., 2014. Violent video detection based on MoSIFT feature and sparse coding. In 2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), 3538-3542. IEEE.
- Yang, M. H. ve Ahuja, N. 1998. Extraction and classification of visual motion patterns for hand gesture recognition. In *Proceedings. 1998 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (Cat. No. 98CB36231)* 892-897. IEEE.
- Yogameena, B., ve Priya, K. S., 2015. Synoptic video based human crowd behavior analysis for forensic video surveillance. In 2015 Eighth International Conference on Advances in Pattern Recognition (ICAPR), 1-6. IEEE.
- Zhang, T., Jia, W., Yang, B., Yang, J., He, X., ve Zheng, Z. 2017. MoWLD: a robust motion image descriptor for violence detection. *Multimedia Tools and Applications*, **76(1)**, 1419-1438.
- Zhao, F. ve Li, J., 2014. Pedestrian motion tracking and crowd abnormal behavior detection based on intelligent video surveillance. *Journal of Networks*, **9(10)**, 2598.
- Zhao, Y. ve Su, Y. 2017. Vehicles detection in complex urban scenes using Gaussian mixture model with FMCW radar. *IEEE Sensors Journal*, **17(18)**, 5948-5953.
- Zhou, S., Shen, W., Zeng, D., Fang, M., Wei, Y. ve Zhang, Z., 2016. Spatial-temporal convolutional neural networks for anomaly detection and localization in crowded scenes. *Signal Processing: Image Communication*, **47**, 358-368.