



Türk Doğa ve Fen Dergisi Turkish Journal of Nature and Science

<http://www.bingol.edu.tr/dergiler/turk-doga-ve-fen-dergisi.aspx>



Nesne tespit ve takip metotları: Kapsamlı bir derleme

Kazım HANBAY^{*1}, Hüseyin ÜZEN¹

Özet

Görüntü işleme dijital bir görüntü içerisindeki önemli bilgilerin okunması, çıkartılması ve işlenmesi için kullanılan bir yöntemdir. Görüntü içerisinde bulunan bir nesne ya da bir ortam hakkında insan görme sistemine benzer şekilde nitel bilgiler edinilmesi ve kullanılması görüntü işlemenin temel amaçlarındandır. Görüntülerde bulunan nesnelerin tespiti, tanımlanması, sınıflandırılması ve takibi gibi ihtiyaçları karşılayacak birçok yöntem geliştirilmiştir. Özellikle görüntülerdeki hedef nesnenin bulunması ve ileriki zaman dilimlerinde bu nesnenin kaybedilmemesi birçok alandaki uygulamalarda sıklıkla kullanılmaktadır. Takip edilecek nesnenin değişken bir ortam içinde bulunması nesne takibi ve analizini zorlaştıran temel problemdir. Bu problemleri çözmek ve nesnenin başarılı bir şekilde takip edilmesi için birçok farklı yöntem geliştirilmiştir. Bu çalışmada nesne takibi için güncel ve yaygın kullanılan yöntemler ele alınmıştır. İncelenen yöntemler güçlü/zayıf yönleri ile irdelenmiştir.

Anahtar Kelimeler: Nesne izleme, arka plan çıkarma, nesne algılama, performans analizi, nesne sınıflandırması, derin öğrenme

Object detection and tracking methods: A comprehensive review

Abstract

Image processing is a method for reading, extracting, and processing important data in a digital image. One of the fundamental part of the image processing is collecting knowledge acquisition about recognized objects and environments in the image like human vision system. Many methods are developed to meet the needs of detection, identification, classification and tracking of objects in images. Especially, finding the target object in the images and following this object in the future time periods is frequently used in many applications. One of the main problems which makes object tracking and analysis difficult is tracking the object in a changing environments. Several effective methods are developed in order to solve these kinds of problems. In this paper, current and widely used methods for object tracking are discussed and those methods were examined with their strengths and weaknesses.

Keywords: Object tracking, background subtraction, object detection, performance analysis, object classification, deep learning

1. Giriş

Günümüzde gelişen teknoloji sayesinde kamera ve gelişmiş bilgisayarların edinilmesi oldukça kolaylaşmış ve bunun sonucunda her an bir görüntü ya da bir video kaydedilmesi olağan bir durum haline gelmiştir. Bu dijital ortama geçiş ile elde edilen görüntü verisinin analizi ve sonuçların çıkartılması ayrı bir öneme sahip olmuştur. Görüntü işleme teknikleri ile ham görüntü verilerinden anlamlı sonuçlar çıkartılabilmektedir. Örneğin akıllı telefonlardaki kameralar sayesinde yüz tanımlama [1] veya trafikteki araçların plaka tespiti [2] gibi günlük uygulamalar görüntü işleme alanının en belirgin uygulamalarındandır.

Görüntü işleme uygulamalarının önemli konulardan biri de nesne takibidir [3]. Nesne takibi bir görüntü dizisi ya da bir videodaki önceden belirlenmiş ya da belirlenmemiş [4] nesnelerin konum, hız veya doğrultu gibi bilgilerinin edinilmesidir. Nesne takip yöntemleri literatürde oldukça çeşitlilik göstermektedir. Bir video içerisindeki arka arkaya

gelen iki imgenin birbirinden çıkarılması gibi basit uygulamalardan başlayarak günümüzde popüler olan derin öğrenme [5] yöntemlerine kadar pek çok yöntem bulunmaktadır [6]. Mevcut video içerisinde nesnenin bulunduğu ortamın arka planının çıkartılmasıyla nesne hareketi takip edilebilir [7]. Yine bir görüntü dizisindeki takip edilecek nesnenin arka plan ve nesne olarak sınıflandırılmasıyla nesne takibi sağlanabilmektedir [8].

Günümüzde dinamik ortam ve değişen ışık koşullarında bile gerçek zamanlı olarak nesne takibi yapan derin öğrenme yöntemleri oldukça önem kazanmıştır. Derin öğrenme, yapay sinir ağları tabanlı doğrusal olmayan dönüşümlerden oluşan sinir katmanları kullanarak verilerdeki üst düzey soyutlamaları modellemeye dayanan bir makine öğrenmesi tekniğidir. Derin yapay sinir ağları giriş çıkış katmanları, bu katmanlar arasında bulunan gizli katmanlardan ve katmanlar arasında geçişi sağlayan bağlantılardan oluşmaktadır. Derin öğrenme yöntemleri birçok görevde en iyi performans sağlayabilmektedir [5]. Son zamanlarda artan işlem gücü ve grafik işlemcilerdeki gelişmelere paralel olarak, derin öğrenme yöntemleri, büyük veri analizinde, konuşma

¹ Bingöl Üniversitesi, Mühendislik ve Mimarlık Fakültesi, Bilgisayar Mühendisliği Bölümü, 12000 Bingöl, Türkiye

*Sorumlu yazar E-posta: kazimhanbay@gmail.com

tanımlama [9], görüntü sınıflandırması [10] ve nesne takibi gibi çeşitli alanlarda kullanılmaya başlamıştır. Bazı şirketler (ör. Google [11] ve Facebook [12]) ayrıca büyük miktarda veriyi günlük olarak toplamak ve analiz etmek yoluyla derin öğrenme ile ilgili projeleri başlatmıştır [5]. Eğitimlerinin zaman alıcı olmasına rağmen sonuçları oldukça başarılıdır.

1.1. Nesne takip yöntemlerinin sonuçlarının değerlendirmesinde kullanılan yöntemler

Yapılan nesne takip çalışmalarında sonuçların değerlendirilmesi için en yaygın kullanılan metotlardan biri hassasiyettir (Precision). Bu metot video üzerinde çalıştırılan yöntemin belirli olan tüm nesnelerden bulduğu nesnelerin doğru bulunma oranı şeklinde tanımlanabilir. Bir diğer başarı ölçme metodu ise anımsamadır (Recall). Bu metot algoritmamızın bulduğu nesnelerin kaçının doğru nesne olduğu şeklinde açıklanabilir. Bu metotların hesaplanmasında True Positive (TP), False Positive (FP), False Negative (FN), True Negative (TN) değerleri kullanılmaktadır. Değerler Tablo 1’de açıklanmıştır[13].

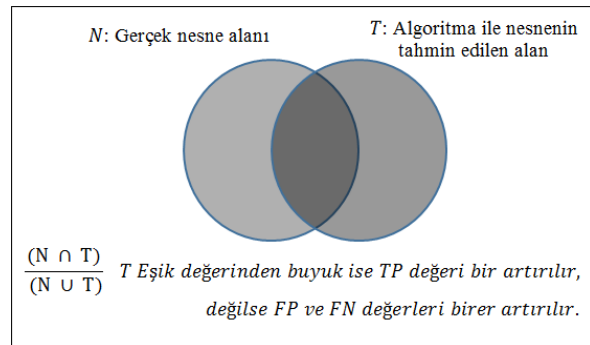
Tablo 1. True Positive (TP), False Positive (FP), False Negative (FN), True Negative (TN) değerlerin hesaplanması

	Gerçek değer: Nesne var	Gerçek değer: Nesne yok
Algoritma tahmini: Nesne var	TP	FP
Algoritma tahmini: Nesne yok	FN	TN

TP: Algoritmamızın nesne olarak bulduğu ve nesnenin olduğu durum sayısı, FN: Algoritmamızın nesne olarak bulamadığı fakat nesne olduğu durum sayısı, FP: Algoritmamızın nesne olarak bulduğu fakat nesne olmadığı durum sayısı, TN: Algoritmamızın nesne olarak bulamadığı ve nesnenin de olmadığı durum sayısı olarak hesaplanır.

Nesne takip algoritmaları nesneyi birebir örtüşecek şekilde bulamayabilir. Bu durumda nesne bulunma durumu kabul edilip edilmeme işlemi 4 adım ile hesaplanır.

1. Algoritmada nesne olarak bulunan alan ile nesnenin gerçekten bulunduğu alanın kesişimi sonucu elde edilen alanın piksel sayısı hesaplanır ($N \cap T$).
2. Algoritmada nesne olarak bulunan alan ile nesnenin gerçekten bulunduğu alanın birleşimi ile elde edilen alanın piksel sayısı hesaplanır ($N \cup T$).
3. Kesişen alan birleşim alanına bölünür.
4. Elde edilen oran, daha önceden belirlenmiş olan T eşik oranı ile kıyaslanır. Eğer T değerinden büyük ise tespit etme durumu kabul edilir (TP artırılır), değilse ret edilir (FP , FN değerleri artırılır) (Şekil 1).



Şekil 1.Kesişme durumunda TP , FP ve FN değerlerin hesaplanması

Recall ve Precision değerleri TP , FP , FN , TN değerlerinden faydalanılarak aşağıdaki matematiksel ifadeler ile hesaplanmaktadır:

$$Recall = \frac{TP}{TP+FN} \quad Precision = \frac{TP}{TP+FP} \quad (1)$$

Yukarıda verilen performans ölçme kriterleri nesne takip algoritmalarının son aşamada elde ettikleri doğrulukları ifade etmektedir. Literatürde bu ölçütlere benzer farklı ölçütler de bulunmaktadır. Bir nesne takip yönteminin performansı içerdiği işlem adımlarına bağlıdır. Özellikle ham veriden anlamlı özniteliklerin çıkartılması ve nesneyi ifade eden ayırt edici özelliklerin fark edilmesi problemin en önemli aşamasıdır.

1.2. Materyal ve metot

Bu çalışmada nesne tespit ve takip yöntemi adı altında yapılan çalışmalar 3 farklı başlık altında toplanmıştır. Bu başlıklar sırasıyla; 1) 2007-2017 yılları arasında literatürde bulunan ve çok kullanılan yöntemler, 2) Zor test koşullarında başarı değeri yüksek çıkan yöntemler, 3) Son zamanlarda önem kazanıp geliştirilmeye açık olan yöntemler şeklinde gruplanmıştır. Bu başlıklar göz önünde bulundurularak literatürde bulunan çalışmalar ön eleme işleminden geçirilmiştir. Ön eleme işleminden sonra çalışmalar nesne tespiti, nesne sınıflandırma ve nesne takibi konularında ilgili başlık altında incelenmiştir. Nesne takip yönteminde yeni yöntemlere ışık tutabilecek çalışmalara da yer verilmiştir.

2. Literatürde Nesne Tespiti ve Takibi için Kullanılan Yöntemler

Literatür çalışmaları incelendiğinde nesne takibi yapan çalışmaların genel olarak 4 farklı aşama içerdikleri anlaşılmaktadır. Bu aşamalar ön işlemler, nesne tespiti, nesne sınıflandırma ve nesne takibi olarak sıralanabilmektedir [3](Şekil 2).

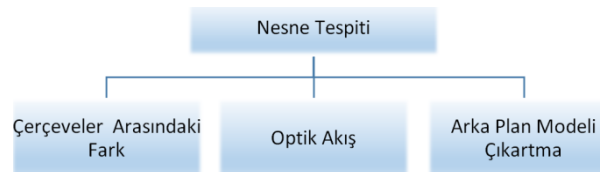


Şekil 2.Nesne takip algoritmalarının genel diyagramı

Bu adımlardan özellikle nesne tespiti temel ve önemli bir yere sahiptir. Çünkü ileriki adımlardaki işlemlerin başarısı bu adımın performansına bağlıdır. Bu işlem genel olarak video imgelerinde nesnenin belirlenmesi ve işlenecek olan nesnenin arka plandan ayrılması olarak tanımlanabilir [14].

2.1. Nesne tespiti

Nesne tespiti ile ilgili olarak literatürde basit veya karmaşık yapılara sahip olarak oldukça fazla yöntem bulunmaktadır. Bu yöntemler genel olarak Şekil 3'teki gibi gruplandırılabilir [3], [6].



Şekil 3.Nesne tespit yöntemlerinin genel başlıkları

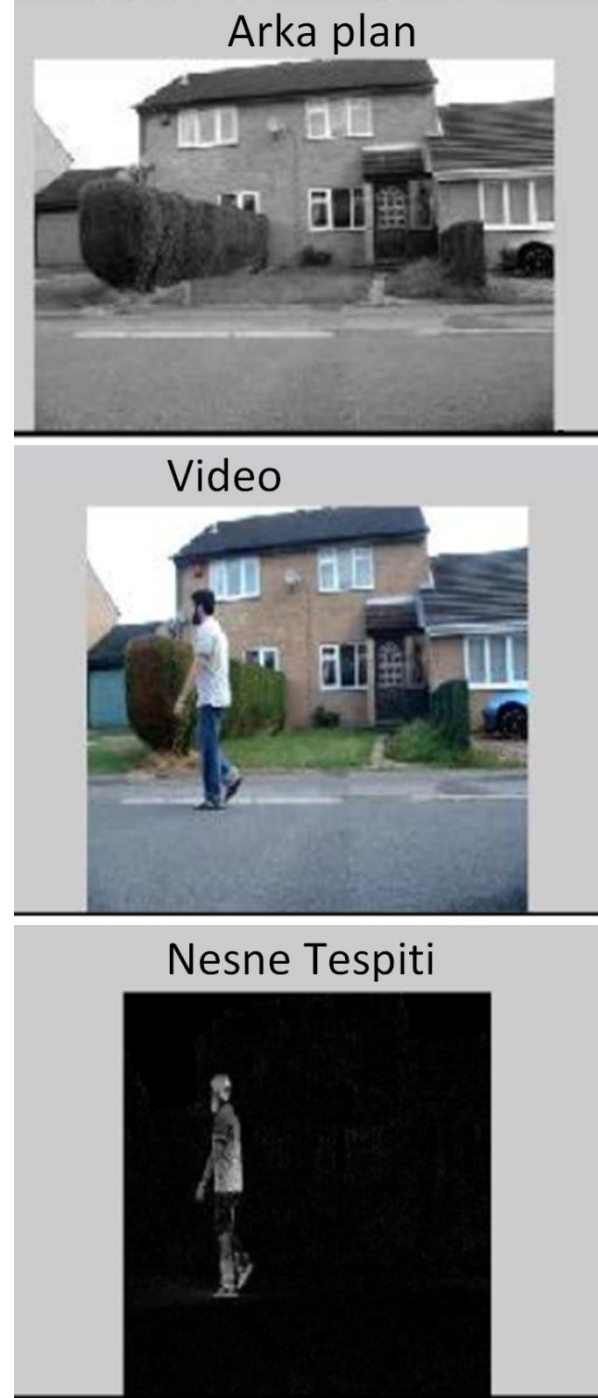
Bir videodaki veya görüntüdeki nesnenin tespiti için iki temel bilgi kullanılır. Bunlar görsel öznitelik (renk, doku, şekil gibi) ve hareket bilgileridir. Bu bilgilerin edinilmesi veya kullanılması için belli bir işlem gücü ve işlem zamanı harcanmaktadır. Bu bilgileri tek başına kullanan yöntemler olduğu gibi özellikle nesne takibi ve tespitinin zor olduğu durumlarda farklı öznitelikleri ve hareket bilgilerini birlikte kullanan yöntemlerde bulunmaktadır [6]. Nesne tespiti için kullanılan bilgilerin çeşitliliği başarıyı arttırmakla beraber işlem zamanında bir artışa sebep olmaktadır. Bu durum gerçek zamanlı uygulamalarda nesne takibi için oldukça önemli bir problem oluşturmaktadır [6], [7].

Kullanılan başlıca yöntemlerden birisi videoda arka arkaya gelen iki görüntüyü arasındaki geçici değişiklikleri bulma yöntemidir (temporal frame differencing). Bir video içerisindeki arka arkaya gelen iki video çerçevesinin birbirinden çıkartılması ile elde edilen bilgilerden oluşur (Şekil 4) [15]. Bu yöntem oldukça basit ve hızlı olmakla beraber ışık değişimi ve gürültüye karşı oldukça hassastır.



Şekil 4. Geçici değişiklikler yöntemi, a: $t - 1$ zamanındaki görüntü (I_{t-1}), b: t zamanındaki görüntü (I_t), c: t zamanındaki fark görüntüsü $N_t = (I_t) - (I_{t-1})$ [15]

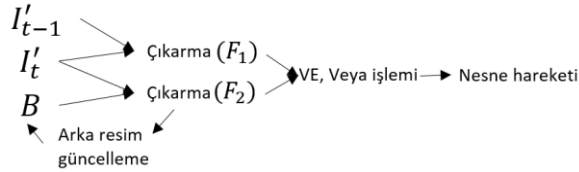
Bir görüntüde nesne tespit işlemi için kullanılan geçici değişiklikler yönteminin hata oranı yüksek çıkabilmektedir. Bundan dolayı daha kullanışlı ve başarılı olan diğer yöntemlerden birisi ise belirlenmiş nesnenin olduğu görüntüden sadece arka plan görüntüsünün bulunduğu çerçevenin çıkarılmasıdır. Sonraki adımda ise elde edilen fark görüntüsü belli bir eşik değerden geçilir ve istenmeyen nesneler elenir [16], [17]. Bu yöntemle elde edilen yeni görüntüden gürültüleri temizlemek için morfolojik işlemler yapılır ve belirli oranlarda istenen nesne ortaya çıkarılmış olur (Şekil 5).



Şekil 5. Bir video çerçevesindeki nesne tespiti için arka plan görüntüsü çıkartma yöntemi [16]

Li ve arkadaşlarının [18] yaptığı çalışmada iki imge arasındaki fark yöntemi ile gerçek zamanlı nesne takibi yapılmıştır. Bu çalışmada kullanılan yöntemin temel

prensibi 3 adım ile sıralanabilir: 1) Yeni gelen video çerçevesi ile bir önceki çerçevenin birbirinden çıkartılması, 2) Arka plan görüntüsü ve yeni gelen çerçevenin birbirinden çıkartılması, 3) Birinci ve ikinci adımlarda elde edilen iki görüntünün birbirleriyle ve/veya mantıksal kapı işlemine sokulmasıdır. Kullanılan yöntemin algoritması Şekil 6'da gösterilmiştir. Bu adımların işlem zamanı çok düşük olduğundan gerçek zamanlı uygulamalar için uygundur. Fakat çoklu nesne takibi ve dinamik arka plana sahip video görüntülerinde başarısız olmaktadır.



Şekil 6. Li ve arkadaşların nesne hareketi tespitine yönelik algoritmalarının diyagramı [18]

Nesne tespiti için oluşturulacak iyi bir arka plan modelinin değişebilen ışık ortamında başarı gösterebilmesi gerekmektedir. Arka plan görüntü modeli için tek bir arka plan görüntüsü kullanıldığında inşa edilen bu model ışık değişimine karşı oldukça hassas olmaktadır. Yaklaşık Medyan Metodu (Approximate Median Method) yöntemi ile referans alınan tek bir görüntü yerine n sayıda hafızada saklanan arka plan görüntüleriyle ile elimizdeki görüntünün belli bir algoritma kullanılarak kıyaslanıp hedef nesne ortaya çıkartılmaya çalışılmaktadır [16]. Bu yöntemde fazladan belli bir hafızaya ihtiyaç duyulmaktadır. Ayrıca bu yöntem arka plan görüntü modelinin çıkartılması için n adet nesne içermeyen görüntüye ihtiyaç duymaktadır.

Gaussian Ortalama (Gaussian Average) arka görüntü çıkartma yönteminde ise, arka plan görüntüsü çıkartmak için yine mevcut arka plan görüntüleri kullanılarak her bir pikselin olasılık yoğunluk dağılımları çıkartılmaktadır. İçerisinde nesne barındıran bir video çerçevesi geldiğinde, gelen görüntünün her bir pikselinin ilgili olasılık yoğunluk dağılımına bakılarak pikselin arka plana mı yoksa nesneye mi ait olduğu belirlenmektedir. Balaji ve arkadaşlarının [3] yaptığı çalışmada Gaussian ortalama yöntemi 3 aşamada incelenmiştir: 1) Arka plan görüntüsünün olasılık yoğunluk modelinin oluşturulması, 2) Yeni gelen görüntüdeki arka plan ve nesnenin birbirlerinden ayrılması, 3) Arka plan görüntü modelinin güncellenmesi. Bu çalışmada arka plan görüntü güncelleme işlemi şu şekilde yapılmaktadır:

Her olasılık yoğunluk dağılımının, ortalama (μ_t) ve varyans (σ_t^2) olmak üzere iki parametresi bulunmaktadır. Bu değerlerin güncellenmesi olasılık yoğunluk fonksiyonunun (pdf) dağılımını güncelleyecektir. Ortalama değerlerin güncellenmesi için formül 2 uygulanır.

$$\mu_t = \alpha x_t + (1 - \alpha)\mu_{t-1} \quad (2)$$

$$\sigma_t^2 = \alpha(x_t - \mu_t)^2 + (1 - \alpha)\sigma_{t-1}^2 \quad (3)$$

Formülde σ_t^2 güncel varyans değerini, σ_{t-1}^2 eski varyans değerini, α eğitim katsayısı, μ_t güncellenmiş ortalama ve x_t ise mevcut görüntünün piksel değeridir.

Literatürde Gaussian ortalama yöntemini farklı renk kanallarında ayrı ayrı kullanarak nesne tespiti yapan çalışmalar bulunmaktadır. Martin'in yaptığı çalışmada [19] kırmızı, yeşil ve mavi renk kanalları ayrı ayrı ele alınmış ve nesne takip performansları incelenmiştir. Bu yöntem gri seviye görüntü ile karşılaştırılmıştır. Uçakların havaalanında alınmış gün içindeki görüntüleri kullanılarak test edilmiştir.

Yapılan uygulamalı test çalışmalarında RGB kanallarının hepsinin birlikte kullanımı sonucu başarı oranı artmıştır.

Manipriya ve arkadaşların [20] yaptığı çalışmada görüntüler öncelikle RGB renk uzayından HSV, CIE LAB, LUV ve YCrCb renk uzayındaki görüntülere dönüştürülmüştür. Bu elde edilen farklı renk uzayları kullanılarak nesne tespiti yapılmaya çalışılmıştır. Bu çalışmada HSV renk uzayının nesne tespiti için daha başarılı sonuçlar verdiği tespit edilmiştir.

Nesne tespiti için arka plan çıkartma yöntemlerinde referans alınan arka plan görüntüsünün dinamik olarak güncellenmemesi, yöntemlerin ışık gibi oluşabilecek dış ortam değişikliklerine karşı oldukça hassas olmalarına sebep olduğu gözlemlenmiştir. Bu yüzden arka plan modelinin güncel tutulması oldukça önemlidir [15]. Stauffer ve arkadaşları [21] çalışmalarında 16 ay boyunca günün her saatinde ve farklı hava koşullarında gerçek zamanlı olarak nesne takibi yapmışlardır. Bu çalışmada arka plan çıkartma işlemlerinde yöntemin farklı ışık ve hava durumlarına uyum sağlanması için arka plan çıkartma yöntemlerinden Gaussian karışım yöntemi kullanılmıştır. Arka plan modeli için oluşturacak her bir piksellin Gaussian karışım yoğunluk dağılım grafiği oluşturulmakta ve gelen yeni bir görüntünün piksel değerlerine bakılarak ilgili dağılımdan nesne tespit edilebilmiştir. Geride kalan arka plan ise arka plan modelinin güncellenmesi için kullanılmıştır. Bu çalışmada arka plan modelini güncelleyen ve güncellemeyen yöntemlerde kıyaslanmıştır. Arka planı dinamik olarak güncellenen yöntemlerin yapılan testler sonucunda daha başarılı olduğu tespit edilmiştir.

Arka plan ve nesnenin ortaya çıkartılması ile takip etme yöntemlerinden bir tanesi de optik akış yöntemidir. Bu yöntemde arka arkaya gelen görüntülerden piksel hareketliliğine bakılarak arka plan çıkartılmakta ve bu şekilde nesne takibi yapılabilmektedir. Doyle ve arkadaşlarının [22] yaptığı çalışmada gerçek zamanlı olarak hareketli kamera ile nesne tespit ve takibi yapılmıştır. Bu çalışmada işlemler grafik kart üzerinde yapılarak işlem hızının yüksek olabileceği belirtilmiştir. Denemelerin yapıldığı ortamın arka planının homojen olması başarı oranının yüksek kalmasını sağlamıştır. Bu çalışmada kullanılan yöntemin başarısı dış ortam görüntülerinde ve homojen olmayan arka plan durumlarında düşecektir.

Tablo 2'de bir video görüntüsü içerisindeki nesnelerin arka plan çıkartma yöntemleri ile tespit edildiği çalışmalar özetlenmiştir. Tiwari ve arkadaşları [23] bu yöntemleri tablo halinde kıyaslamıştır.

2.2. Nesne sınıflandırma

Görüntü işlemede sınıflandırma, bir veri kümesindeki belirli nesnelerin özelliklerine göre gruplara ayrılması şeklinde açıklanabilmektedir. Bu ayrıştırma yapılırken nesnenin kendine has öznelikleri kullanılmaktadır. Kullanılan öznelik nesnenin tanımlanması için kullanılan sayısal değerlerden oluşmaktadır. Bu değerler bir pikselin değeri veya bir görüntüdeki ortalama yoğunluk değeri gibi anlamlı ifadeler olabilmektedir. Nesne takibi veya sınıflandırması yöntemlerinde öznelik seçimi ve kullanımı performansı doğrudan etkileyen işlemlerdir. Bir sınıflandırma veya takip algoritmasında kullanılan öznelikler seçilirken başarıyı arttırmasının yanı sıra getirdiği işlem yüküne de bakılması gerekmektedir. Gerçek zamanlı uygulamalarda daha basit öznelikler kullanılırken zaman problemi olmayan uygulamalarda ise çok daha farklı ve zor öznelikler kullanılmaktadır [30].

Sınıflandırma algoritmaları nesne takip yöntemlerinde önemli bir yere sahiptir. Özellikle çoklu nesne takip yöntemlerinde takip edilen her bir nesnenin diğer nesneler ile karıştırılmaması için doğru bir şekilde sınıflandırılması

oldukça önemli olmaktadır. Ayrıca bazı nesne takip yöntemlerinde sınıflandırma algoritmaları kullanılarak arka plan ve hedef nesne iki sınıfta tanımlanarak nesne takibi yapılmaktadır [5], [8].

Literatürde nesne takibi için en yaygın kullanılan öznitelikler renk, kenar, doku, derinlik, süper piksel, hareket ve optik akış şeklinde sıralanabilir. Literatürde yapılan bazı çalışmalarda bu öznitelikler birlikte de kullanılmaktadır. Literatürde yaygın olarak kullanılan öznitelikler Tablo 3’de verilmiştir [31].

Balaji ve arkadaşlarının [3] yaptığı çalışmada sınıflandırma algoritmaları 4 temel başlık altında toplanmıştır. Bunlar sırasıyla hareket tabanlı, doku tabanlı, şekil tabanlı ve renk tabanlı algoritmalarıdır.

Hareket tabanlı sınıflandırma yönteminde, nesne hareket bilgileri kullanılarak sınıflandırma yapılmaktadır. Nesne takip çalışmalarında hareket öznitelikleri kullanılarak, bir sonraki gelecek video çerçevesinde nesnenin potansiyel olarak varacağı konum tahmin edilmektedir. Bu ön tahmin nesnenin konumunu bulmak için aranacak alanı da azaltmaktadır. Bu yöntemlerde genellikle nesnenin belli bir

ortalama hızla hareket edildiği varsayılmaktadır [31]. Belli bir şekli olmayan nesnelerin takibinde kenar, şekil ve renk bilgileri seçici olmadığından, hareket tabanlı sınıflandırıcıların başarı oranları yüksek olabilmektedir. Optik akış özniteliği bu başlık altında ayrı bir öneme sahip olmaktadır. Bu yöntem ön eğitim yapılmadan nesne takibi üzerinde yapılan çalışmalarda önemlidir [3]. Optik akış yöntemi hareket tabanlı sınıflandırmada çoklu nesne takibi yapılan çalışmalarda nesne tespitini kolaylaştırmaktadır [31].

Renk ve doku tabanlı sınıflandırmalar çok farklı alanlarda kullanılabilir. Doku tabanlı sınıflandırma yardımıyla nesne sınıflandırma işlemlerinde nesnenin yapısına ait öznitelikler kullanılmaktadır. Özniteliklerin çıkartılması ve kullanılması işlem maliyetini yükseltir. Bununla birlikte yüksek başarılar elde edilmektedir. Renk tabanlı sınıflandırma yöntemleri en yaygın kullanılan yöntemlerdir. Renk özniteliği kullanımı işlem zamanını azaltır. Bununla birlikte gürültü ve ışık değişimi içermeyen videolarda başarı oranını yükseltir [31]. Bu sınıflandırma başlıkların kıyaslanması Tablo 4’te verilmiştir [23].

Tablo 2. Nesne tespiti yöntemlerin kıyaslanması [23]

Nesne Tespit Yöntemi	Temel Prensipleri	İşlem Gücü	Başarı Oranı	Yorumlar
Geçici Değişlikler Yöntemi	İki imge arasındaki fark	Düşük	Yüksek	Kolay uygulanabilir [24, 25] Ortam değişikliklerine karşı hassas [26]
Arka Görüntü Modül Çıkartma	Görüntüler arasındaki fark	Düşük-Orta	Orta-Yüksek	En yaygın kullanılan yöntemdir [16, 26] Gerçek zamanlı uygulamalara uygun değildir [25]
	Yaklaşık medyan	Düşük-Orta	Orta-Yüksek	Yeterli arka plan imgesine ihtiyaç duymamaktadır [16] Bir hafıza alanına ihtiyaç duyar [16]
	Gaussian ortalama	Orta-Yüksek	Orta-Yüksek	Gerçek zamanlı uygulamalarına uygundur [16] İşlem gücü yüksek
	Gaussian karışım	Orta-Yüksek	Orta-Yüksek	Düşük hafıza gereksinimi [27] Gürültüye karşı hassas [28]
Optik Akış	Piksellerin optik akış hareketleri	Orta-Yüksek	Yüksek	İşlem gücü yüksek olmasına rağmen tüm hareketli nesneleri ortaya çıkartır [29]

Tablo 3. Nesne sınıflandırma çalışmalarında kullanılan özniteliklerin karşılaştırılması [31]

Öznitelik Tipi	Kullanımı biçimi	Avantajı	Dezavantajı
Renk	Renk Histogramı	Etkili	Işık değişikliğine hassas
Eğim	Hog, Shift	İnsan takibi için kullanışlı	Gürültüye karşı hassas
Doku	Lbp, Shift	Yüksek doğruluk oranı	İşlem zamanı yüksek
Hareket	Hog, Shift	Nesne takip sırasında oluşan kesilmeler	İşlem zamanı yüksek
Zamansal Değişimler	Süper Piksel, Süper Voxel	Nesne takip sırasında oluşan kesilmeler ve nesnenin şekil durum çeşitliliği	İşlem zamanı yüksek Yüksek öznitelik boyutu

Tablo 3. Nesne sınıflandırma çalışmalarında kullanılan özniteliklerin karşılaştırılması [31]

Öznitelik Tipi	Kullanımı biçimi	Avantajı	Dezavantajı
Renk	Renk Histogramı	Etkili	Işık değişikliğine hassas
Eğim	Hog, Shift	İnsan takibi için kullanışlı	Gürültüye karşı hassas
Doku	Lbp, Shift	Yüksek doğruluk oranı	İşlem zamanı yüksek
Hareket	Hog, Shift	Nesne takip sırasında oluşan kesilmeler	İşlem zamanı yüksek
Zamansal Değişimler	Süper Piksel, Süper Voxel	Nesne takip sırasında oluşan kesilmeler ve nesnenin şekil durum çeşitliliği	İşlem zamanı yüksek Yüksek öznitelik boyutu

Tablo 4. Sınıflandırma yöntemlerin kullandığı özneliteliklere göre kıyaslanması

Nesne Sınıflandırma Yöntemi	Hesaplama Zamanı	Başarı Oranı	Yorum
Şekil tabanlı	Düşük	Orta-Yüksek	Model uydurma algoritmaları barındırır. Değişen şekle sahip olan nesnelerde zayıftır [32].
Hareket tabanlı	Yüksek	Orta	Hareketli nesneleri tanımayı amaçlar [32].
Renk tabanlı	Yüksek	Yüksek	Yüksek başarılar yanında işlem zamanı artmaktadır [33].
Doku tabanlı	Yüksek	Yüksek	Öznitelik çıkartmada işlem maliyeti yükselmekte.

2.3. Nesne takibi

Nesne takibi genel olarak bir video veya sıralı gelen görüntü dizisinde bulunan nesnenin takip edilmesidir. Nesne takibi yöntemleri temel olarak 3 kategoriye ayrılır. Bunlar sırasıyla nokta tabanlı, çekirdek tabanlı ve silüet tabanlı yöntemlerdir.

Nokta takip yönteminde takip edilecek nesne noktalar ile ifade edilir. Bu noktaların bir sonraki imgedeki konumları ve birbirlerine olan uzaklıkları gibi verilerin sonraki gelen video çerçevesinde de birbirine paralel olması beklenir. Bu bilgilerden yola çıkılarak nesne takibi sağlanır. Bu yöntemde temel amaç nesnenin video çerçevesi içinde tespit edilmesi ve bir önceki çerçevede kullanılan nokta benzerliklerin hesaplamasıdır. Bu yöntem deterministik ve istatistiksel yöntemler olarak ta kendi içinde alt sınıflandırmalar içermektedir [34], [35].

Nokta takibi yöntemlerinde en yaygın kullanılan yöntemlerden birisi kalman filtresidir. Bu yöntem nesnenin Gaussian dağılıma sahip durum değişkenleri yardımıyla videodaki bir sonraki gelen çerçevede nesne konumunu tahmin etmektedir. Kalman filtresi basit ve hızlı olma açısından gerçek zamanlı nesne takip uygulamalarında kullanıma uygundur [34, 36]. Durum değişkenleri Gaussian dağılımına sahip olmayan sistemlerde kalman filtresi başarısız olabilmektedir. Bu tür problemlerin giderilmesi için parçacık filtresi yöntemi geliştirilmiştir. Parçacık filtresi olasılıksal yöntemeye dayanmaktadır. Bu yöntemin en büyük

avantajı doğrusal olmayan ve çoklu dağılıma sahip sistemlerde çalışabilmesidir [31].

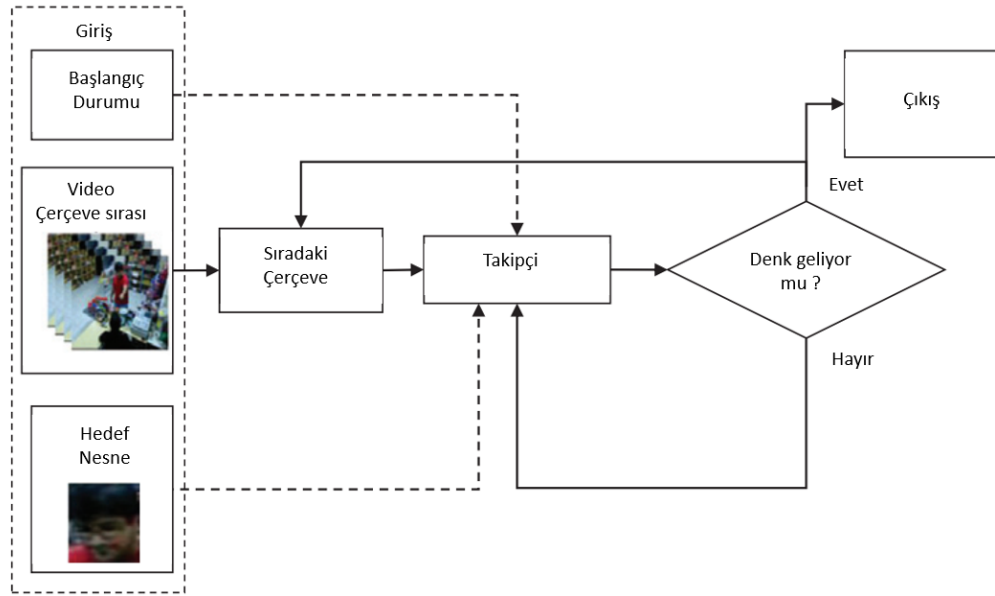
Çekirdek tabanlı yöntemlerde bir geometrik şekil yardımıyla takip edilecek nesne çerçevelenir. Bu çerçeve içerisinde bulunan nesne parçasının anlamlı bilgileri hesaplanarak başlangıçtaki şekil yardımıyla nesne takip edilir. Bu yöntemde nesnenin şeklinden ziyade kullanılan geometrik şeklin içerisinde bulunan nesne bilgilerinin çıkarılması yeterli olabilmektedir. Bu şekil içinde bulunan piksellerin hesaplanan olasılık yoğunluk bilgileri veya histogram özellikleri gibi bilgileri sonraki video çerçevelerinde takip edilebilmektedir.

Silüet tabanlı yöntemler genellikle takip edilen nesnenin insan ya da hayvan gibi belli bir geometrik şekille ifade edilemediği durumlarda kullanılır. Bu yöntemin temel amacı nesneyi tanımlayacak kenar bilgisi ya da şekil bilgisi çıkartılarak sonraki imgelerde bu bilgiyi aramaktır. Bu yöntem şekil değişikliğine karşı oldukça hassas olmaktadır.

Çekirdek ve silüet tabanlı yöntemler kıyaslandığında, çekirdek tabanlı yöntemlerin daha düşük işlem zamanına ve daha yüksek başarı oranlarına sahip oldukları görülmektedir. Bu sebepten dolayı çalışmalarda çekirdek tabanlı yöntemler geniş bir kullanım alanına sahiptir. Nokta tabanlı yöntemler diğer yöntemlere oranla daha düşük işlem zamanına sahip olmakla birlikte daha düşük başarı oranına sahiptirler. Literatürde nesne takibi için yaygın olarak kullanılan yöntemler Tablo 5'te gösterilmiştir [23].

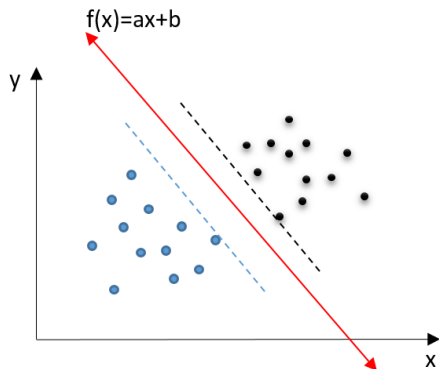
Tablo 5. Literatürde nesne takibi için yaygın olarak kullanılan yöntemler

Nesne Takip Yöntemi	Kullanımı	İşlem Zamanı	Başarı Oranı
Nokta tabanlı	Kalman filtresi	Kalman filtre algoritmaları	Düşük-Orta
	Parçacık filtresi	Öz yinelemeli bayes filtresi	Orta-Yüksek
	Çoklu hypothesis takip	MHT algoritması	Düşük
Çekirdek tabanlı	Temel şablon uydurma	Video içinde uyum aram algoritmaları	Düşük-orta
	Destek vektör makinesi (dvm)	Görüntü içindeki piksellerin nesne ve arka plan olarak sınıflandırması	Orta
	Uydurma tabanlı sınıflandırma	Şekil içindeki piksel yoğunluklara bakılır	Orta
Silüet tabanlı	Kenar kesiştirme	Gradient descent algoritmaları	Orta
	Şekil uydurma	Hough dönüşümü	Yüksek



Şekil 8. KLT takip sistemi [37]

Avidan [8] yaptığı çalışmada Destek Vektör Makinesi (DVM) sınıflandırıcıdan yola çıkarak oluşturulan Destek Vektör İzleme yöntemini önermiştir. DVM algoritması verilen öznitelikler kullanılarak iki sınıf arasında bir lineer doğru bularak farklı sınıfları birbirinden ayırmayı hedeflemektedir (Şekil 9). DVM yöntemi sınıfları ayırmak için kullanılan fonksiyonun sadece iki parametre içermesi açısından uygun işlem zamanına sahiptir. Bu düşük işlem zamanı nesne takip yönteminde büyük avantaj sağlamaktadır. Çalışmanın önerdiği yöntemde videodaki belli görüntüler kullanılarak DVM sınıflandırıcı ile eğitim yapılmıştır. Eğitim sonunda video çerçevesini nesne ve arka plan olarak ayırmak için en uygun parametreler bulunmuş olmaktadır. Test aşamasında yeni video çerçevesi daha küçük görüntü parçalarına ayrılır ve bu parçalar sırasıyla DVM algoritması kullanılarak çerçevenin alt parçalarının nesne ya da arka plan görüntüsü olup olmadığı belirlenmektedir. Son olarak nesne olarak tahmin edilen ve başarı değeri yüksek olan parçalar birleştirilir ve doğru nesne konumu belirlenir. Bu şekilde algoritma tekrarlanarak nesne takibi yeni video çerçevelerinde de tekrar edilmektedir.

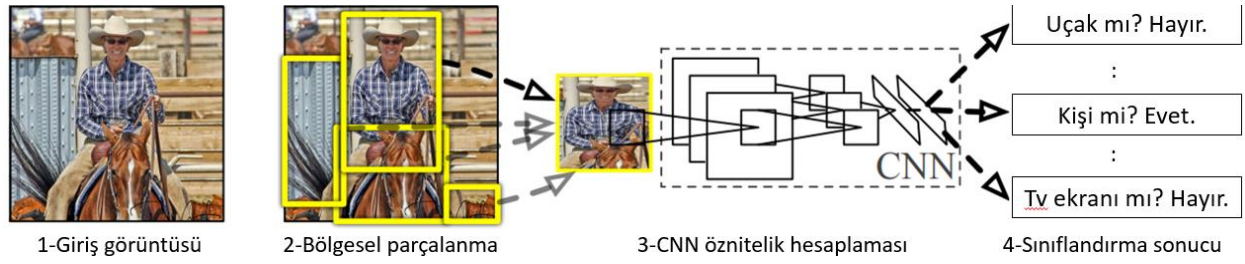
Şekil 9. DVM algoritması kullanılarak iki sınıfı ayıran $f(x)$ fonksiyonu

Şekil 10. Öğrenilen SDAE'nin ilk katmanındaki bazı filtreler [38]

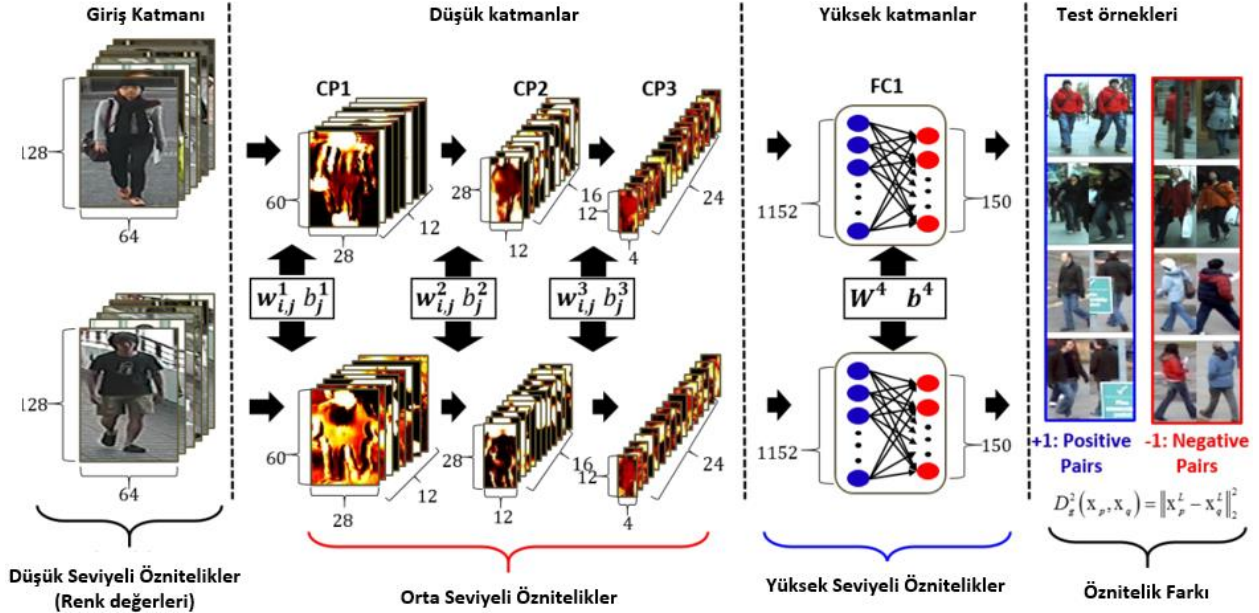
Son zamanlarda derin öğrenme yaklaşımları ile nesne takibi yapan çok sayıda çalışma yapılmıştır. Wang ve arkadaşları [38] yaptıkları çalışmada literatürde bulunan 7 farklı nesne takip yöntemi ile derin öğrenme tabanlı nesne takip yöntemini karşılaştırmışlardır. Yürütülen karşılaştırmalı test çalışmalarında gerçek zamanlı nesne takibi yapılmıştır. Kullanılan veri setleri ani nesne kayıpları, dinamik arka plan ve değişen ışık ortamı gibi karmaşık ve zor arka planlar içermektedir. Karşılaştırılan yöntemler genel olarak piksel yoğunluk değerleri, Haar özellikleri ve histogram gibi özniteliklerden faydalanmaktadırlar. Derin öğrenme algoritmasında ise büyük veri setlerinden oluşan görüntüler Stacked Denoising AutoEncoder (SDAE) algoritmasından geçirilerek yüksek seviyeli öznitelikler elde edilmiştir (Şekil 10). Seyrek kodlamaya dayalı önceki yöntemlerin yaptığı gibi derin öğrenmede de temsili görüntüler kullanılarak algoritmada optimizasyon problemi çözümüne gereksinim duyulması önlenmiştir. Çalışmanın elde ettiği sonuçlara bakıldığında derin öğrenme yönteminin diğer yöntemlerden genel olarak daha yüksek başarılar elde ettiği görülmüştür.

Zhou [39] çalışmasında bir önceki [38] çalışması ile yakın yöntem kullanmıştır. Dinamik hareketliliğe sahip karmaşık arka planlarda nesne izleme bilgisayar görme alanında zor bir problem olmuştur. Derin öğrenme mimarisini içeren yeni yaklaşımlar sayesinde bu sorunun üstesinden daha kolay gelinebilmektedir.

Zhang [40] yaptığı çalışmada nesne takibi için derin öğrenme tabanlı Deep RL Tracker (DRLT) yöntemini geliştirmiştir. Bu çalışmada derin öğrenme tabanlı nesne takip yöntemlerinde kullanılan Convolutional Neural Network (CNN) (Şekil 11) ve Recurrent Neural Network (RNN) yöntemlerinden de bahsedilmiştir. DRLT yöntemi literatürde yaygın olarak kullanılan 25'ten fazla video veri seti üzerinde test edilmiştir. Bu testlerde DRLT yöntemi başarı sıralamasında genel olarak ilk iki sırada yer almaktadır.



Şekil 11. CNN sınıflandırıcının genel yapısı [41]



Şekil 12. Discriminative Deep Appearance modelinin genel ağ yapısı [44]

Behrendt ve arkadaşları [42] yaptıkları çalışmada, gerçek zamanlı olarak trafik ışıklarının tespiti ve takibi için derin öğrenme tabanlı bir yöntem geliştirmiştir. Bu yöntem YouOnlyLookOnce (YOLO) yöntemine dayanarak trafik ışıklarını tespit ve takip için 2 farklı sinir ağı kullanmıştır. Sinir ağından bir tanesi trafik ışığı tespiti için eğitilmiştir. Diğer sinir ağı ise trafik ışıklarının yanlış tespit edildiği durumları (FP) belirlemek için eğitilmiştir. Önerilen yöntem gerçek zamanlı testlerde test edilmiştir. Sonuçlara bakıldığında zorlu koşullarda yüksek doğruluk elde edildiği görülmüştür. Bu çalışmada trafik ışık tespiti ve takibi için Bosch Small Traffic Lights Dataset adı altında güçlü bir veri tabanı oluşturulmuş ve yayınlanmıştır.

Gordon ve arkadaşları [43] yaptıkları çalışmada Recurrent Neural Network (RNN) tabanlı nesne takip yöntemi geliştirmiştir. Bu yöntemde diğer derin öğrenme tekniklerinden farklı olarak nesnenin görünümünü ve hareketini temsil etmek için bir RNN sinir ağı kullanılmıştır. RNN sinir ağ yapısı, test öncesi farklı birçok örnekten eğitilebilir ve belirli bir nesneyi hızlı bir şekilde güncelleyebilir özelliklerine sahiptir. Bunun sayesinde gerçek zamanlı olarak yapılan testlerde nesnenin bir kısmının kaybolması veya dinamik arka planlarda bile güçlü başarılar sağlanmıştır. Bu çalışma sayesinde nesne takibi için RNN yöntemlerinin yüksek başarılar verebileceği öngörülmüştür.

Çoklu nesne takip çalışmalarında başarı oranını düşürecek önemli zorluklardan bir tanesi de nesnelerin görünüşleri arasındaki benzerliklerin olduğu durumlardır. Bae ve arkadaşları [44] bu problemten yola çıkarak derin öğrenme tabanlı Discriminative Deep Appearance Learning yöntemini geliştirmiştir. Bu yöntem nesneler arasındaki

farkı arttıracak güçlü öznitelikleri kullanarak nesnelerin sınıflandırılması ve takip edilmesini amaçlamıştır. Nesne içeren görüntü parçalarının RGB renk kanalları ile elde edilen düşük seviye öznitelikleri, geliştirilen yapay sinir ağı katmanlarından geçirilerek sınıfları ayıracak güçlü öznitelikler elde edilmiştir (Şekil 12). Bu öznitelikler kullanılarak nesnelerin tespit ve takibi yapılmıştır. Geliştirilen yöntem birçok zorlu veri kümeleri ile test edilmiştir. Testler sonucunda çalışmadaki diğer yöntemlere kıyasla belirgin bir performans artışı gösterilmiştir.

3. Sonuçlar

Bu çalışmada nesne takibi ve sınıflandırma problemleri için kullanılan basit ve karmaşık yöntemler hakkında bir derleme çalışması sunulmuştur. Nesne tanımlama, sınıflandırma ve takibi için kullanılan yöntemlerin avantaj ve dezavantajları incelenmiştir. Kolay ve homojen arka planlarda çalışan basit ve hızlı yöntemlerden oldukça karmaşık ve zor problemlere karşı kullanılan derin öğrenme yöntemlerine kadar birçok farklı çalışma incelenmiş ve değerlendirilmiştir. Geliştirilen yöntemlerin farklı karakteristiklere sahip veri setleri üzerindeki davranışları gözlemlenmiş ve olumlu/olumsuz yönleri tartışılmıştır.

Son yıllarda geliştirilen derin öğrenme yöntemlerin eğitim süreleri uzun olmasına rağmen test aşamasında elde edilen başarı oranları derin öğrenme yöntemlerine olan güveni arttırmıştır. Fakat basit problemlerde işlem zamanı ve hesapsal karmaşıklığı az olan algoritmaların kullanılması çok daha uygun olacaktır. Varılan bir diğer sonuç ise Kalman ve parçacık filtre yöntemlerinin bulunmasının nesne takip algoritmalarına farklı bir perspektif kazandırdığıdır. Bununla birlikte derin öğrenme algoritmalarının farklı

türelerinin nesne takibi ve sınıflandırılmasında uzun bir süre kullanılmaya devam edeceği öngörülmüştür.

Kaynaklar

- [1] Hjelmås E., Kee L.B., Face Detection: A Survey, Computer vision and image understanding, 83, 236–274., 2001
- [2] Anagnostopoulos C.-N.E., Anagnostopoulos I.E., Psoroulas I.D., Loumos, V., Kayafas, E., License Plate Recognition From Still Images and Video Sequences: A Survey, IEEE Transactions on intelligent transportation systems, 9, 3:377–391, 2008
- [3] Balaji S.R., Karthikeyan S., A survey on moving object tracking using image processing, 2017 11th International Conference on Intelligent Systems and Control (ISCO), IEEE, pp. 469–474, 2017
- [4] Wang D., Unsupervised video segmentation based on watersheds and temporal tracking, IEEE Transactions on Circuits and Systems for video Technology, 8.5:539–546, 1998
- [5] Chen Y., Yang X., Zhong B., Pan S., et al., CNNTracker: Online discriminative object tracking via deep convolutional neural network, Applied Soft Computing, 38:1088–1098, 2016
- [6] Luo W., Xing J., Milan A., Zhang X., Multiple Object Tracking: A Literature Review, arXiv Prepr. arXiv1409.7618, 2014
- [7] Risha K.P., Kumar A.C., Novel Method of Detecting Moving Object in Video, Procedia Technology, 24:1055–1060, 2016
- [8] Avidan S., Support vector tracking, IEEE transactions on pattern analysis and machine intelligence, 26.8: 1064–1072, 2004
- [9] Hinton G., Deng L., Yu D., Dahl G., Deep Neural Networks for Acoustic Modeling in Speech Recognition: The Shared Views of Four Research Groups, IEEE Signal Processing Magazine, 29.6: 82–97, 2012
- [10] Krizhevsky A., Sutskever I., Hinton G.E., ImageNet Classification with Deep Convolutional Neural Networks, In: Advances in neural information processing systems, p. 1097–1105, 2012
- [11] Szegedy C., Liu W., Jia Y., Sermanet P., Going Deeper With Convolutions, In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 1–9, 2015
- [12] Taigman, Y., Yang, M., Ranzato, M., Wolf, L., DeepFace: Closing the gap to human-level performance in face verification, In: Proceedings of the IEEE conference on computer vision and pattern recognition, p. 1701–1708, 2014
- [13] Wu Y., Lim J., Yang M.-H., Online Object Tracking: A Benchmark, In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp. 2411–2418, 2013
- [14] Karasulu B., Videolardaki Hareketli Nesnelerin Tespit Ve Takibi İçin Uyarlanabilir Arkaplan Çıkarımı Yaklaşımı Tabanlı Bir Sistem, Uludağ Üniversitesi Mühendislik-Mimarlık Fakültesi Dergisi, 18, 2013
- [15] Shaikh S.H., Saeed K., Chaki N., Moving Object Detection Approaches, Challenges and Object Tracking, In: Moving Object Detection Using Background Subtraction. Springer International Publishing, p. 5–14, 2014
- [16] Aldhaheri A.R., Edirisinghe E.A., Detection and Classification of a Moving Object in a Video Stream, In: Proc. of the Intl. Conf. on Advances in Computing and Information Technology-ACIT, 2014.
- [17] Hardas A., Vibha M., Moving Object Detection using Background Subtraction Shadow Removal and Post Processing, Int. J. Comput. Appl., 975–8887, 2015
- [18] Li G., Wang Y., Shu W., Real-Time Moving Object Detection for Video Monitoring Systems, In: Intelligent Information Technology Application, 2008. IITA'08. Second International Symposium on, IEEE, pp. 163–166, 2008
- [19] Martin C., Background Subtraction Using Running Gaussian Average: a Color Channel Comparison, In: Seminar aus Bildverarbeitung und Mustererkennung, 2014
- [20] Manipriya S., Mala C., Mathew S., Performance Analysis of Spatial Color Information for Object Detection Using Background Subtraction, IERI Procedia, 10:63–69, 2014
- [21] Stauffer, C., Grimson W.E.L., Adaptive background mixture models for real-time tracking, Proceedings, 1999. IEEE Computer Society Conference on, IEEE, pp. 246–252, 1999
- [22] Doyle D.D., Jennings A.L., Black J.T., Optical flow background estimation for real-time pan/tilt camera object tracking, Measurement, 48:195–207, 2014
- [23] Tiwari, M., Singhai, R., A Review of Detection and Tracking of Object from Image and Video Sequences, International Journal of Computational Intelligence Research, 13, 973–1873, 2017
- [24] Chate M., Amudha S., Gohokar V., Object Detection and tracking in Video Sequences, ACEEE International Journal on signal & Image processing, 3, 2012,
- [25] Mohan A.S., Resmi R., Video image processing for moving object detection and segmentation using background subtraction, In: Computational Systems and Communications (ICCSC), 2014 First International Conference on, IEEE, pp. 288–292, 2014
- [26] Haritaoglu I., Harwood D., Davis L.S., W/sup 4/: real-time surveillance of people and their activities, IEEE Transactions on pattern analysis and machine intelligence, 22, 809–830, 2000
- [27] Zhiqiang W., Xiaopeng J., Peng W., Real-time moving object detection for video monitoring systems, Journal of Systems Engineering and Electronics, 17, 731–736, 2006
- [28] Zhang T., Liu Z., Lian X., Wang X., Study on moving-objects detection technique in video surveillance system, Chinese Control and Decision Conference, IEEE, pp. 2375–2380, 2010
- [29] Krishna M.T.G., Ravishankar M., Babu D.R.R., Automatic detection and tracking of moving objects in complex environments for video surveillance applications, In: Electronics Computer Technology (ICECT), 2011 3rd International Conference on, IEEE, pp. 234–238, 2011
- [30] Due Trier., Jain A.K., Taxt T., Feature extraction methods for character recognition-A survey, Pattern Recognition, 29, 641–662, 1996.
- [31] Fan L., Wang Z., Cail B., Tao C., A survey on multiple object tracking algorithm, In: Information and Automation (ICIA), 2016 IEEE International Conference on, IEEE, pp. 1855–1862, 2016
- [32] Javed O., Shah M., Tracking and object classification for automated surveillance, In: European Conference on Computer Vision. Springer, Berlin, Heidelberg, pp. 343–357, 2002
- [33] Hu, W., Tan, T., Wang, L., Maybank, S., 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, 2004
- [34] Talu M.F., Nesne Takip Yöntemlerinin Sınıflandırılması, İstanbul Ticaret Üniversitesi Fen

- Bilim. Derg., 18, 45–63, 2010
- [35] Yilmaz A., Javed O., Shah M., Object Tracking: A Survey, *ACM computing surveys*, 38, 2006
- [36] Parekh H.S., Thakore D.G., Jaliya U.K., A survey on object detection and tracking methods, *International Journal of Innovative Research in Computer and Communication Engineering*, 2.2: 2970-2979. 2014
- [37] Bagherpour P., Cheraghi S.A., Mokji M.B.M., Upper Body Tracking Using KLT and Kalman Filter, *Procedia Computer Science*, 13, 185–191, 2012
- [38] Wang, N., Yeung, D.-Y., Learning a deep compact image representation for visual tracking, *Advances in neural information processing systems*, p. 809-817, 2013.
- [39] Zhou X, Xie L, Zhang P, Zhang Y. An ensemble of deep neural networks for object tracking, In: *Image Processing (ICIP), 2014 IEEE International Conference on*, IEEE, p. 843-847, 2014
- [40] Zhang D., Maei H., Wang X., Wang Y.-F., Deep Reinforcement Learning for Visual Object Tracking in Videos, *arXiv Prepr. arXiv1701.08936*, 2017
- [41] Girshick R., Donahue J., Darrell T., Malik J., Rich feature hierarchies for accurate object detection and semantic segmentation, In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, pp.580–587, 2014
- [42] Behrendt K., Novak L., Botros R., A deep learning approach to traffic lights: Detection, tracking, and classification, In: *Robotics and Automation (ICRA), 2017 IEEE International Conference on*, pp. 1370–1377, 2017
- [43] Gordon D., Farhadi A., Fox D., Re3: Real-Time Recurrent Regression Networks for Object Tracking, *arXiv preprint arXiv:1705.06368*, 2017.
- [44] Bae S.-H., Yoon K.-J., Confidence-Based Data Association and Discriminative Deep Appearance Learning for Robust Online Multi-Object Tracking, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2017