

Derin Öğrenme ve Görüntü İşleme Kullanılarak Yayın Akışı Üzerinden Yol İzleme Analitiği Yaklaşımı

Path Tracking Analytics Approach on Streaming Using Deep Learning and Image Processing

Alihan Öztürk^{1,2}, Elif Yavuz³, Nedime Gaye Nur Balkanlı³, Şeref Naci Engin^{1,2}, S. Levent Kuzu³

¹Kontrol ve Otomasyon Mühendisliği Bölümü
Yıldız Teknik Üniversitesi, Davutpaşa Kampüsü, İstanbul
alihan.ozturk1@std.yildiz.edu.tr nengin@yildiz.edu.tr

²RoboGör Robot ve Görme Teknolojileri
Yıldız Teknopark, İstanbul
alihan.ozturk@robogor.com serefnaci.engin@robogor.com

³Çevre Mühendisliği Bölümü
Yıldız Teknik Üniversitesi, Davutpaşa Kampüsü, İstanbul
elify@yildiz.edu.tr gayenur.balkanli@yildiz.edu.tr skuzu@yildiz.edu.tr

Özetçe

Bu bildiri, çeşitli problemlerin çözümünde değerlendirilmek amacıyla güvenlik kameralarının otomatik izlenerek analitiğinin çıkarıldığı ön çalışmalar sunulmuştur. Yüksek seviyeli bilgiler bulunduran yol izleme analitiği için görüntü işleme ve derin öğrenme yöntemleri kullanılmıştır. Elde edilen bilgiler, güvenlik kameralarının görüş açısındaki her ayrı yol için geçen araçların sınıfını, sayısını ve hızlarını içermektedir. Bu verileri tespit edebilmek amacıyla önceden belirlenmiş nesnelere tespit eden ve eş zamanlı olarak çoklu nesne takibi yapan bir görüntü işleme algoritması geliştirilmiştir. Nesne tespit algoritmasını eğitmek için "uym.ibt.gov.tr" sitesindeki kameralar taranmış ve bir veri seti oluşturulmuştur. Elde edilen derin öğrenme eğitim grafikleri, kullanılan yöntemler, geliştirilen algoritmalar, konunun zorlukları ve bu bilgilerin kullanılabileceği alanlar açıklamalarla birlikte sunulmuştur.

Abstract

In this paper, we present preliminary work on automatic monitoring and analytics of surveillance cameras to be used in solving various problems. Image processing and deep learning methods are used for road monitoring analytics containing high-level information. The acquired information includes the class, number and speed of passing vehicles on each individual road in the angle of view of the surveillance cameras. In order

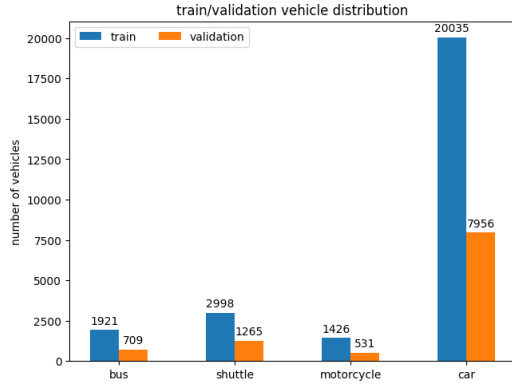
to detect this data, an image processing algorithm was developed that detects predefined objects and performs simultaneous multi-object tracking. To train the object detection algorithm, the cameras on the "uym.ibt.gov.tr" website were scanned and a dataset was created. The deep learning training graphs obtained, the methods used, the algorithms developed, the challenges of the subject and the areas where this information can be used are presented with explanations.

1. Giriş

Günümüzde güvenlik önlemlerinin artırılması amacıyla güvenlik kameralarının yaygın bir şekilde kullanıldığı görülmektedir. Bu kameraların geniş kullanım alanları, çeşitli problemlerin çözümünde otomatik izleme ve yüksek seviyeli bilgi çıkarımı gibi teknolojik yöntemlerin geliştirilmesine yol açmıştır. Güvenlik kameralarının görüş açısında kaydedilen görüntülerden elde edilecek yüksek seviyeli bilgiler trafiğin kontrolü, güvenlik önlemlerinin artırılması ve karbon ve hava kirlenimi emisyonlarının salımı tespiti için büyük önem taşımaktadır.

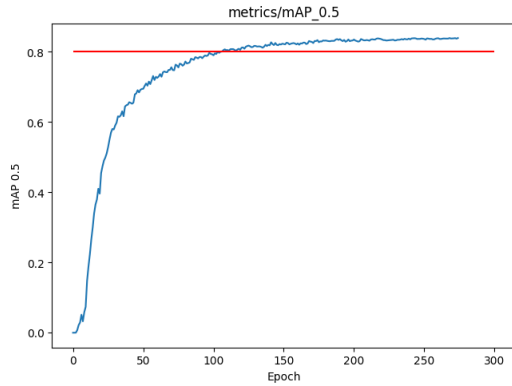
Bu bildiri, güvenlik kameralarının otomatik izlenmesi ve yüksek seviyeli bilgi çıkarımı çalışması hakkında detaylı bir inceleme sunulmuştur. Çalışmada, bilgisayarla görü ve derin öğrenme yöntemlerinden yararlanılarak nesne tespit algoritması ile çoklu nesne takip algoritmasını içeren görüntü işleme yöntemleri geliştirilmiştir. Bu sayede, kameraların görüş açısında yer alan araçların tipleri, sayıları ve hızlarını gerçek zamanlı olarak tespit eden bütüncül bir algoritma geliştirilmiştir. Ortaya çıkan yol izleme analitiği algoritmasının geliştirilme aşamaları aşağıda sırasıyla verilmiştir.

Kameradan günün farklı saatlerinde ve çeşitli hava koşullarında araç tespit modelini eğitmek amacıyla araç görüntüleri İstanbul Büyükşehir Belediyesinin trafik MOBESE (MOBil Elektronik Sistem Entegrasyonu) kamera görüntülerinin gerçek zamanlı olarak yayımlandığı “uym.ibb.gov.tr” adresinden toplanmıştır. Araçlar otomobil, servis (öğrenci veya personel mekik taşıtı), otobüs ve motosiklet olarak 4 sınıfa ayrılmıştır. Görüntülerde bulunan araçlar, ait oldukları sınıfı belirterek çevreleyen en küçük alana sahip dikkörtgenler içerisinde alınarak etiketlenmiştir. Araç tespit modelinin başarımını yükseltmek amacıyla görüntü çoklama teknikleri kullanılarak eğitimde kullanılan görüntü verilerinin sayısı artırılmıştır. Elde edilen veri seti ile nesne tespit modelinin eğitimi gerçekleştirilecektir. Şekil 1’de veri seti araç dağılımı sunulmuştur.



Şekil 1: Eğitim/Doğrulama veri seti araç dağılımı

Veri setindeki 1125 görüntünün yaklaşık % 70’i ile nesne tespit modeli eğitilmiş ve kalan 452’si (\approx %30’u) ile modelin tespit yeteneği sınanmıştır. Model performans kriteri olan mAP (ortalama doğruluk değeri) 0.5 %80 limit değeri olmuştur. Eğitim süreci Şekil 2 de verilmiş ve mAP 0.5 %80’i geçerek algoritmada kullanılmak için yeterli seviyeye gelmiştir.

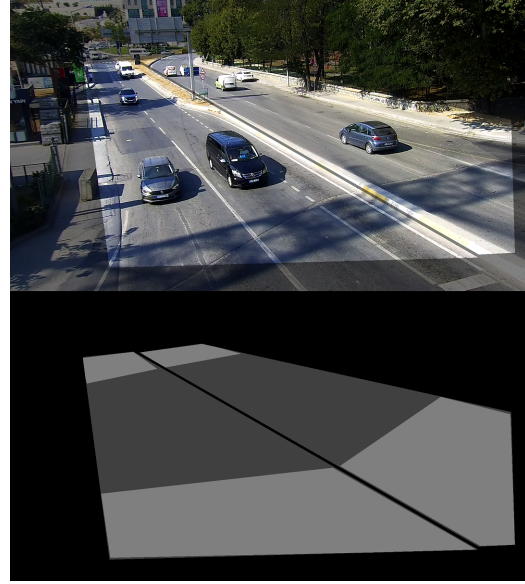


Şekil 2: Nesne tespit algoritması mAP 0.5 eğitim grafiği

Tespit edilen araçların hareketi takip edilecektir. Böylece aynı aracın sadece tek bir kimlik vermesi sağlanacak ve çoklu (mükerrer) sayımın önüne geçilecektir. Araçların hareket takibi

için Kalman tabanlı SORT (Simple Online and Real-time Tracking, Basit Çevrimiçi Gerçek Zamanlı Takip) algoritması kullanılmıştır.

Araç sayma ve hız tespiti algoritması için görüntü üzerinde yollar ve başlangıç/bitiş alanları elle maskelenmiştir. Araçlar başlangıç/bitiş alanlarından herhangi birinden geçtiklerinde nesne tespit ve çoklu nesne takibi algoritması ile izlemeye alınır ve herhangi bir alandan çıkışı yaptıklarında geçen süre ve rotası hesaplanır sonrasında hafızadan silinir. Harita üzerinden alınan başlangıç ve bitiş alanları arasındaki yol uzunluğu bilgisi ile geçen süre işlenerek aracın hızı tespit edilir. Şekil 3’te Alibeyköy Yavuz Selim Caddesi üzerinde bulunan gözetleme kamerasından alınan görüntü kullanılarak maskeleme işlemleri sunulmuştur.



Şekil 3: Alibeyköy Yavuz Selim Caddesi ile yol ve başlangıç/bitiş maskeleri

2. Yapay Sinir Ağları

Yapay sinir ağları tahminlemelerde genellikle karmaşık ve doğrusal olmayan ilişkileri yakalayabilmesinden dolayı seçilmektedir. Tahminde doğruluğu arttırmak amacıyla uygulamaya göre farklı mimariler kullanılmaktadır.

Yapay sinir ağları, görüntü işleme ve nesne tespit gibi görsel verilerin analizinde oldukça etkili ve başarılı sonuçlar elde edilen derin öğrenme modelleri olarak kullanılır. Bu modeller, görüntü üzerindeki karmaşık öznelikleri yakalama yeteneği sayesinde çok çeşitli uygulama alanları bulmaktadır.

2.1. Derin Öğrenme Tabanlı Nesne Tespit Algoritmaları

Yapay sinir ağları temelli nesne tespiti için, özellikle iki popüler yaklaşım ön plana çıkmıştır: CNN (Convolutional Neural Network – Evrişimli Sinir Ağı) tabanlı ve Dönüştürücü (Transformers) tabanlı modeller. Her birinin kendine özgü üstün ve geri kalan özellikleri vardır.

2.1.1. CNN Tabanlı Modeller

- ✓ Yapısal olarak evrişim katmanları sayesinde, görsel verilerin özelliklerini hiyerarşik olarak çıkarmada oldukça etkilidirler.
- ✓ Nesnelerin özgün özelliklerini yakalamak için yerel bağlantı ve paylaşılan ağırlıklar kullanılır, bu da verimli ve ölçeklenebilir öznetelik çıkarmaya yarar.
- ✓ Derin mimarileri, büyük veri setlerini öğrenmeye daha uygun hale getirir ve bu sayede yüksek doğruluk elde edilebilir.
- Tam bağlantılı katmanlar içerdiklerinden dolayı, nispeten büyük parametre sayısına sahiptirler ve eğitim için daha fazla hesaplama gücü gerektirebilirler.
- Çevre bağımlılığına dayalı olarak çalıştıklarından, nesnelerin farklı konumları, boyutları ve açıları gibi çeşitlilikleri ele almakta bazen zorlanabilirler.

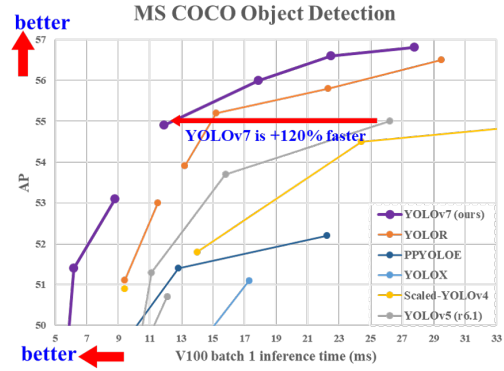
2.1.2. Dönüştürücü (Transformers) Tabanlı Modeller

- ✓ Daha az tabaka ve daha az parametreye rağmen, daha etkili modeller öğrenilebilirler ve daha düşük hesaplama maliyetine sahip olabilirler.
- ✓ Dikkat mekanizması sayesinde, nesneler arasındaki bağlantıları daha iyi anlayabilirler ve uzun mesafeli bağımlılıkları ele alabilirler.
- ✓ GPT-3 ve BERT gibi dil işleme alanında büyük başarılar elde ederek, transfer öğrenme için uygun bir yapıya sahiptirler.
- Geleneksel CNN tabanlı modellere kıyasla, görsel verilerin işlenmesinde henüz yeni ve daha az araştırılmıştır.
- Görsel veriler için dikkat mekanizmasını kullanmak, dil verilerine göre bazen daha fazla işlem gücü gerektirebilir.
- Yeterli miktarda veri olmadığında, derin dönüştürücü (transformers) tabanlı modellerin eğitimi daha zor olabilir ve aşırı öğrenme riski taşır.

Sonuç olarak, CNN tabanlı modeller güçlü ve yaygın olarak kullanılan bir temele sahiptir, ancak büyük veri setleri ve hesaplama gücü gerektirebilirler. Dönüştürücü tabanlı modeller ise görsel veriler için henüz yeni olmalarına rağmen, dikkat mekanizması ve transfer öğrenme yetenekleri nedeniyle gelecekte daha da önemli hale gelebilirler. Hangi modelin tercih edileceği, uygulanacak senaryo, veri seti büyüklüğü ve hesaplama kaynakları gibi çeşitli faktörlere bağlı olacaktır. [1, 2, 3]

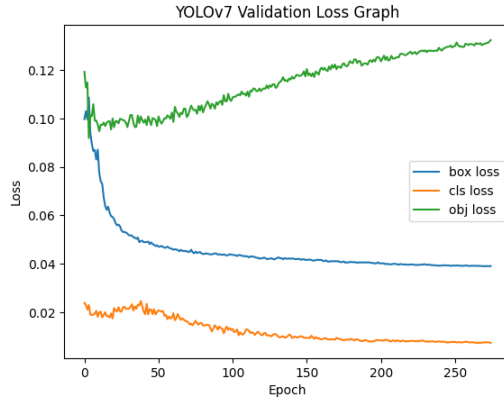
Modelleri karşılaştırmak amacıyla 2 metrik yeterli olacaktır. Bunlar mAP ve birim çerçevede çalışma süresidir. COCO veriseti nesne tespiti kıyas ölçütüne bakarak canlı olarak okunan verinin (10-30 fps) çalışmada kullanılan grafik kartında sorun oluşturmadan ve diğer işlemlere zaman kalacak şekilde çalıştıracak CNN tabanlı YOLOv7 modeli seçilmiştir. Modellerin karşılaştırma grafiği Şekil 4'te verilmiştir.

Yolo modeli eğitiminde 3 kayıp fonksiyonu birlikte kullanılarak model eğitilir. Bunlar tahmin edilen kutunun gerçek kutuya göre konum ve boyut açısından ne kadar yakın olduğuyla ilgili kayıp fonksiyonu, tahmin edilen sınıf etiketleri ile gerçek



Şekil 4: Modellerin birim çalışma sürelerine göre AP değerleri [4]

etiketler arasındaki farkı niteleyen kayıp fonksiyonu ve sınırlayıcı kutuda bir nesnenin varlığını doğru bir şekilde algılamak için ne kadar iyi olduğuyla ilgili kayıp fonksiyonudur. Eğitim boyunca doğrulama verisi üzerinde test edilen bu kayıp fonksiyonu değerleri aşağıda Şekil 5'te grafik halinde verilmiştir.



Şekil 5: YOLOv7 eğitim süreci

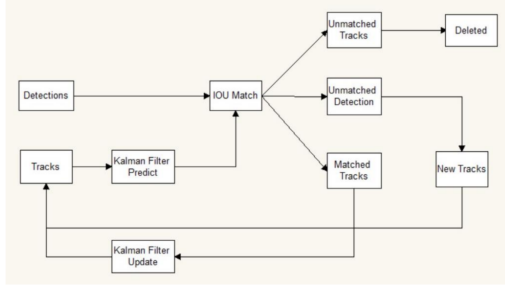
3. Çoklu Nesne Takibi

Çoklu nesne takibi, bir video veya görüntü akışında birden fazla nesnenin sürekli olarak izlenmesi işlemidir. Çoklu nesne takibi, otonom sürüş, güvenlik sistemleri, izleme sistemleri, video analitiği ve oyunlar gibi birçok alanda kullanılabilir ve önemli bir araştırma konusudur. Ancak bu konu bazen karmaşık ve gerçek zamanlı hesaplama gerektiren bir süreç olabilir. Bu nedenle, yüksek performanslı bilgisayar donanımları ve optimize edilmiş algoritmalar kullanılarak uygulanır. Görüntü bilgisi yerine genellikle nesnelerin zaman içindeki konumlarını ve sınıfını kullanarak takibini gerçekleştirir.

3.1. SORT Algoritması

Basit Çevrimiçi Gerçek Zamanlı Takip Algoritması (SORT), nesnelere algıladıktan sonra onların anlık konumlarını güncelle-

meye yönelik temel bir yaklaşımdır. Söz konusu yöntem, anlık konum bilgisine dayanarak nesnelere hareketini tahmin eder ve ardından takip etmek için güncellenmiş konumları kullanır. SORT, tahmin için nesnelere merkez piksel lokasyonlarını bir önceki merkez pikseller ile karşılaştırarak ve Kalman filtresinden geçirerek tahminde bulunur. Şekil 6’da SORT Algoritmasının işleyiş şeması sunulmuştur. [5]



Şekil 6: SORT algoritması işleyiş şeması [5]

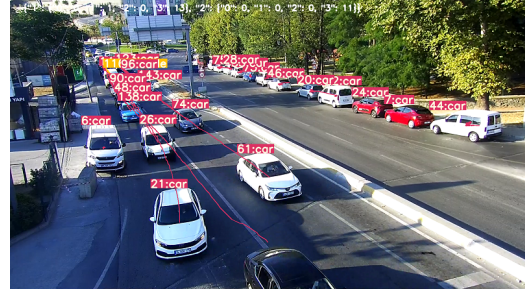
3.2. Çok Yollu Araç Sayma ve Hız Takip Algoritması

Literatürde incelenen [6, 7] nesne sayma algoritmaları tespit edilen nesnelere tek yönlü giriş-çıkışları ya da tek yol için tasarlandığı görülmüştür. İstanbul’daki gözetleme kameraları ele alındığında izlenen yolların birden fazla giriş-çıkışı olmasından ve bir kameranın birden fazla yola bakmasından dolayı araç sayma ve hız takip algoritması geliştirilmiştir.

Kamera görüntüsünden elle belirlenen yolların giriş-çıkışlarından geçen araçlar bir sonraki giriş-çıkışa kadar takip edilerek süresi ve giriş-çıkış noktaları kayıt edilmektedir. Bunun için 2 farklı tablo oluşturulmuştur. Tablolardan biri izlenmeye başlayan araçların eklendiği geçici tablodur. Bu tabloya araçlar giriş-çıkış alanlarına ilk defa gelmişse ya da aynı alana gelmişse takip algoritmasının belirlediği isim, aracın sınıfı, başlangıç noktası, başlangıç zamanı eklenerek güncellenir. Ayrıca bellek sızıntısı olmaması için her araç izlenmeye başladığı anda kullanıcının belirlediği değere göre ömür süresi oluşur. Algoritmanın her döngüsünde araçların ömür süreleri 1 birim azalır. Ömür süresi sıfırlanan araçlar izlenme listesinden çıkarılır. Diğer tablo ise geçmiş tablosudur. Araç farklı bir giriş-çıkış bölgesinden geçtiğinde geçici tablodan alınan bilgiler ile ulaşma süresi hesaplanarak geçmiş tablosuna eklenir ve geçici tablodan bu araç silinir. Aracın alanın içinde bulunup bulunmadığı, nesne tespit modelinin çıktısı olan sınırlayıcı kutunun merkezi oluşturulan maskelerde dizin olarak kullanılarak tespit edilir.

4. Sonuç

Çalışmanın amacı olan yüksek seviyeli bilgi çıkarımı, literatürden sunulan yöntemlerden yararlanarak geliştirilen algoritmalarla sağlanmıştır. Aracın sınıfı, hızı, bulunduğu yol, nereden nereye gittiği bilgileri hem canlı yayın akışından hem de video üzerinden tespit edilebilmektedir. Şekil 7’de canlı yayın akışından alınan görüntü üzerinde araçların sınıfları, atanan isimleri ve izleri görselleştirilmiştir.



Şekil 7: İzlenen araçların görselleştirilmesi

Geliştirilen algorithmada sırasıyla nesne tespit, nesne takip, nesne sayma ve hız takip döngüleri çalışmaktadır. Her döngü kendinden önceki döngünün çıktılarını kullanarak çalıştığından başlangıçtaki bir hata sonuca katlanarak etki etmektedir. Gözlemlenen hatalardan en önemlilerinden biri servis araçlarının çeşitli kamera açılarından araba olarak tespit edilmesidir. Diğer bir hata ise trafik hızının değişken ve yoğun olması durumunda kullanılan SORT algoritmasının araçları izlemede zorlanmasıdır. Bunun sebebi takip algoritmasının görüntü bilgisi yerine nesne tespit modelinin çıktısı kullanılarak elde edilen araçların merkez konumlarını kullanmasıdır. Takip algoritmasının görevi araçları Kalman filtresi kullanarak bir sonraki adımda bulacağı noktaya göre izlemesidir. Trafikteki hız değişimi sebebiyle tahmin edilen yere uzaklığı eşik değeri altında kalan araç farklı bir kimlik ile ilişkilendirileceğinden performansda düşüş meydana gelmektedir.

Burada sunulan çalışmaların çıktıları "Taşıt Kaynaklı Emisyonların Modellemesinde Derin Öğrenme ve Hesaplamalı Akışkanlar Dinamiği Yöntemleri Kullanarak Yüksek Çözünürlüklü Hava Kirliliği Modelinin Geliştirilmesi" başlıklı TÜBİTAK 1001 projemizde kullanılacak algoritmanın geliştirilmesinde değerlendirilecektir. Bir diğer gelecek çalışma "Trafik Sıkışıklığını Azaltmak İçin Pekiştirmeli Öğrenmeye Dayalı Kontrol Yaklaşımının Geliştirilmesi" için de aynı algorithmadan yararlanılacaktır.

5. Teşekkür

Bu bildiri, TÜBİTAK tarafından 1001 Bilimsel ve Teknolojik Araştırma Projelerini Destekleme Programı kapsamında desteklenen 122Y267 numaralı projede gerçekleştirilen çalışmalardan faydalanılarak hazırlanmıştır.

6. Kaynakça

- [1] Li, Zheng, vd. "Deep learning-based object detection techniques for remote sensing images: A survey." *Remote Sensing*, 14.10 (2022): 2385.
- [2] Zaidi, Syed Sahil Abbas, vd. "A survey of modern deep learning based object detection models." *Digital Signal Processing*, 126 (2022): 103514.
- [3] Sharma, Vipul, Roohie Naaz Mir. "A comprehensive and systematic look up into deep learning based object detec-

tion techniques: A review.” *Computer Science Review*, 38 (2020): 100301.

- [4] Wang, Chien-Yao, vd. “YOLOv7: Trainable bag-of-freebies sets new SOTA for real-time object detector” *Proceedings of the IEEE/CVF Conference on CV and PR*, 2023.
- [5] Bewley, Alex, vd. “Simple online and realtime tracking.” *IEEE ICIP*, 2016.
- [6] Liang, Mingpei, vd. “Counting and classification of highway vehicles by regression analysis.” *IEEE Transactions on Intelligent Transportation Systems*, 16.5 (2015): 2878-2888.
- [7] Dai, Zhe, vd. “Video-based vehicle counting framework” *IEEE Access* 7, (2019): 64460-64470.