



Derin Öğrenme Vasıtasıyla Masa Tenisi Topu Takibi

Erman Havuç¹, Şeyma Alpak², Gözde Çakırel³, Mehmet Kadir Baran^{4*}

^{1*} Marmara Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, İstanbul, Türkiye, (ORCID: 0000-0001-5006-2937), ermanhavuc@gmail.com

² Marmara Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, İstanbul, Türkiye, (ORCID: 0000-0003-0903-2140), seymaalpak@gmail.com

³ Marmara Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, İstanbul, Türkiye, (ORCID: 0000-0001-9612-2245), gozdecakirel@gmail.com

^{4*} Marmara Üniversitesi, Mühendislik Fakültesi, Bilgisayar Mühendisliği Bölümü, İstanbul, Türkiye, (ORCID: 0000-0002-7973-2794), crom.baran@gmail.com

(İlk Geliş Tarihi 24 Şubat 2021 ve Kabul Tarihi 8 Ağustos 2021)

(DOI: 10.31590/ejosat.885795)

ATIF/REFERENCE: Havuç, E. Alpak, Ş., Çakırel, G. & Baran, M. K. (2021). Derin Öğrenme Vasıtasıyla Masa Tenisi Topu Takibi. *Avrupa Bilim ve Teknoloji Dergisi*, (27), 629-635.

Özet

Bilgisayarlı görü teknolojisi 1960'lı yıllardan itibaren gelişmeye başladı ve günümüzde bu alanda oldukça ilerleme katedildi. Bugün gerçek zamanlı görüntüler üzerinde arabalar, insanlar gibi farklı hareket halinde olan nesnelere takip edilebilmektedir. Fakat doğrusal olmayan yörüngelerde ve çok hızlı hareket eden küçük cisimlerin tespiti ve takibi çoğu durum için daha doğrusal, normal hızlarda hareket eden büyük cisimlere göre çok daha zor olmaya devam ediyor. Bu tip nesnelere takibi için Kalman filtresi, parçacık filtresi, bulanık mantık ve Gaussian modellemesi gibi farklı metodlar uygulanmıştır. Fakat son on yıl içinde evrimsel sinir ağları kullanan yeni metodlar bu klasik metodlara alternatif olarak ortaya çıkmış ve birçok alanda büyük bir başarıyla uygulanmışlardır. Bu çalışmada, evrimsel sinir ağlarını kullanarak pinpon topu gibi doğrusal olmayan yönlerde ve yüksek hızlarda hareket eden küçük cisimlerin gerçek zamanlı tespiti ve yardımcı algoritmalarla nesne takibinin yapılabilmesini sağlayan bir sistem geliştirilmiştir. Evrimsel sinir ağı temelli bir nesne tespit algoritması olan YOLO, pinpon topunun farklı renk ve video üzerinde değişen biçimleriyle birlikte veri seti hazırlanarak eğitilmiş ve test edilmiştir. Büyük oranda başarı sağlandığı görülmüş, bu tür nesnelere evrimsel sinir ağı temelli algoritmalar ile gerçek zamanlı tespitinin ve takibinin mümkün olduğu görülmüştür. İleride yapılacak olan çalışmalarda masa tenisi oynayabilecek bir robot için araştırma yapılması planmaktadır. Bu nedenle, tespiti ve yardımcı algoritmalarla takibi yapılan pinpon topunu bir işaretçi ile sürekli olarak işaret eden 2 eksenli servo motor kullanan bir robot kol inşa edilmiştir.

Anahtar Kelimeler: Bilgisayarlı görü, Nesne tespiti, Nesne takibi.

Ping-Pong Ball Tracking Through Deep Learning

Abstract

Computer vision technology has been constantly evolving from 1960's on, and a lot of progress has been made in this field since that date. Today, different moving objects such as cars and people can be tracked on real-time images. However, the detection and tracking of small objects moving in nonlinear trajectories and at high speeds is still more challenging compared to large objects moving in more linear and slower speeds. Previously, different methods such as Kalman filters, particle filters, fuzzy logic and Gaussian modeling have been applied to the problem tracking such objects. But, in the last ten years, new methods using convolutional neural nets emerged as alternatives to these classical methods and they were applied to various problems with great success. In this study, a convolutional neural network-based system will be developed which enables the real-time detection and tracking of a fast-moving small object, such as a ping-pong ball. YOLO, which is a neural network-based object detection algorithm, is trained on the images of fast-moving ping-pong balls of various colors, where all the distortions accompanying fast moving objects like motion blur are present. For this purpose, a new training set is created. A high success rate is achieved in detecting and tracking the ping-pong ball, and it has been observed that real-time detection of such objects is possible with convolutional neural network-based algorithms. In future, this research is planned to be extended to a robot that can play table tennis. For this purpose, a two degree of freedom robot arm using 2 servos has been built, which continuously points the ping-pong ball with a pointer as it travels.

Keywords: Computer vision, Object detection, Object tracking.

* Sorumlu Yazar: crom.baran@gmail.com

1. Giriş

Bu çalışmada, pinpon topunun gerçek zamanlı olarak evrişimsel sinir ağları kullanılarak tespitini ve takibini yapan ve 2 eksenli bir robot kolu yardımıyla topun yerini işaret eden bir sistem geliştirilecektir. Bir kamera ile belirli bir açıdan pinpon masası kaydı canlı olarak bilgisayara aktarılacak, bilgisayar üzerinde topun koordinatları hesaplanacak ve sonuçlar anlık olarak robot koluna iletilip robot kolunun hareketi sağlanacaktır. Pinpon topunun tespiti ve takibi diğer çoğu senaryoya kıyasla topun doğrusal olmayan yönlerde ve yüksek hızlarda hareketinden dolayı ve oldukça küçük bir nesne olması dolayısıyla daha zordur.

Bilgisayarlı görü teknolojisi günümüzde birçok alanda kullanılmaktadır. Askeri alanda düşman askerlerinin ve araçlarının tespit edilmesi, tıbbi alanda hastaların sağlık durumlarının tespit edilmesi, film ve görüntü alanında otomatik restorasyon yapılması, sivil alanda polis gücüne yardımcı uygulamalarda yer alması, otonom araçlarda çeşitli görevler yüklenilip yürütülmesi gibi çeşitli alanlarda ve uygulamalarda yardımcı rol üstlenmektedir. İlk olarak 1960'lı yıllarda yapay zekâ alanı dahilinde geliştirilmeye başlanmıştır. Renk karşılaştırmaları, kenar bulma algoritmaları, çeşitli görüntü modellemeleri gibi temel dijital görüntü işleme algoritmaları kullanarak verilen görüntüdeki nesnelere ya da olayların anlamlandırılması üzerine çalışılmıştır (Szeliski, 2011). Günümüzde video üzerinden hareket tespiti, hareket takibi, nesne tanımlama, görüntü veya görüntüleri kullanarak 3 boyutlu modellemeler çıkarma gibi birçok farklı konuda geliştirilmiştir ve gelişmeye devam etmektedir.

Robotlar ile bilgisayarlı görü günümüzde birçok noktada beraber kullanılmaktadır. Uydulara ve gezegenlere yollanan arazi araçları, kayan şeritler, otonom araçlar, ameliyatlara için tasarlanmış robot kollar gibi bilimin ve endüstrinin temel noktalarında robotlar ve bilgisayarlı görü teknolojileri iş birliği içerisinde dirler (Corke, 2017).

Bilgisayarlı görüde, hareketli bir nesnenin tespit edilebilmesi ve hareket takibinin yapılabilmesi kullanılan birçok alan için oldukça önemlidir. Bu görev için farklı algoritmalar geliştirilmiştir. 1960'lı yıllarda geliştirilen Kalman Filtresi, bu alanda en yaygın algoritmalarından biridir. Sistemin önceki durumlarına göre bir Gaussian matematiksel model oluşturarak, sistemin sonraki hareketlerini öngörmeye çalışmaktadır (Simon, 2006). Takip edilen nesne ne kadar doğrusal olmayan bir yolda hareket ederse, tahmin o kadar zorlaşır. Bir başka yöntem olan parçacık filtresi, Gauss tipi olmayan süreçleri ve dağılımları içeren Monte Carlo yöntemlerini kullanır (Branko Ristic, 2003). Birçok gerçek yaşam senaryosunda Kalman Filtresi'nden daha iyi sonuç vermesine karşın, çoğu durumda daha fazla işlem gücü gerektirmektedir ve hareket takibinin başlangıcında, takip edilecek nesnenin tanıtılması gereklidir. Hareket takibini hedef alan bu algoritmaların takip edilen nesnenin farklı hızları ya da nesnenin doğrusal olmayan hareketleri nedeniyle hata payları ve dolayısıyla nesnenin izini kaybetme olasılıkları artmaktadır.

Pinpon topu şekli itibarıyla dairesel bir görüntü verir. Bilgisayarlı görü alanında şimdiye kadar kullanılan Hough dönüşümü gibi farklı daire tespit algoritmaları geliştirilmiştir (Rafael C. Gonzalez, 2017). Fakat bu algoritmalar ortamdaki ışığa bağlı olarak topun renginin arkaplandaki farklı nesnelere karışmasına ve daire tespit algoritmasının hatalı sonuç vermesine

yol açabilmektedir. (P. R. Kamble, 2019) Normal bir masa tenisi maçında top çok hızlı hareket edeceği için görüntü "yılanvari" bir şekil almakta, daire tespit algoritmasının yanılmasına sebep olmaktadır. Ayrıca çevredeki farklı dairesel nesnelere daire tespit algoritmasını yanıltmakta, topun olması gerektiğinden başka bir konumda bulunmasına yol açmaktadır. Bu hatalarla karşılaşmaması için arkaplanın düz bir renkle kaplanması, ortamdaki ışık seviyesinin ve ışık renginin sabit tutulması ve topun hızının görüntüde eliptik şekil alacak kadar arttırılmaması düşünülmüş, fakat gerçek hayat koşullarından oldukça uzaklaşıldığı için bu fikirden vazgeçilmiştir.

Masa tenisi topu takibi için daha önce yapılan çalışmalarda birçok değişik algoritma önerilmiştir. Literatürde Kalman filtresi (G. Chen, 2013) (P. R. Kamble, 2019), parçacık filtresi (Y. Huang, 2008), bulanık mantık algoritması (Hu Su, 2013), basit Gaussian modellemesi (J. Liu, 2014) gibi farklı algoritmalar kullanılmış ve bu farklı algoritmalar harmanlanıp optimum sonuç elde edilmeye çalışılmıştır.

Bu çalışmada yukarıda bahsedilen sorunların önüne geçilebilmesi için günümüzde oldukça popüler olan yapay sinir ağı bazlı algoritmalar göz önüne alınmıştır. Evrişimli sinir ağı kullanan bazı bilgisayarlı görü algoritmaları, hali hazırda bazı modifikasyonlarla nesne tespiti için de kullanılabilir. Bu projede, farklı nesne tespit algoritmaları karşılaştırılmış, gerçek zamanlı ve anlık nesne tespiti ve takibi yapılacağından evrişimli sinir ağı bazlı algoritmalar arasından literatüre ilk defa 2016 yılında giren YOLO adlı algoritma tercih edilmiştir (J. Redmon, 2016). YOLO yukarıda bahsedilen ve klasik algoritmaların performansını etkileyen arkaplan ve ışıklandırma gibi problemlerden daha az etkilenmektedir. Topun görüntüsünün alabileceği değişik formlar (yavaş hareket sırasında dairesel form, hızlı hareket sırasında bulanık yılanvari form) YOLO'ya ayrı ayrı öğretilbileceği için de algoritmanın güvenilirliği artmaktadır.

2. Materyal ve Metod

2.1. Sistem Bileşenleri

2.1.1. YOLO

Nesne tespiti ve takibi yapılabilmesi için yapay sinir ağı merkezli bir algoritmanın topun hızına bağlı olarak topun şeklinde ve ortam koşullarına bağlı olarak topun renginde meydana gelebilecek farklılıkları göz önüne alarak dijital görüntü işleme alanında kullanılan daire tespit algoritmalarından daha iyi sonuç vereceği anlaşılmıştır. Bu nedenle geliştirilmiş evrişimli sinir ağı (CNN) bazlı algoritmalar araştırılmıştır.

R-CNN algoritması verilen veri setine bağlı olarak görüntü üzerinde nesnenin bulunabileceği muhtemel bölgeleri belirleyerek bölgeler üzerinde CNN işlemleri yürütmektedir (Aggarwal, 2018). Birden fazla bölgede CNN algoritması yürütüldüğü için bu algoritma gerçek zamanlı senaryolar için oldukça yavaş kalmaktadır. Fast R-CNN ve Faster R-CNN algoritmaları her ne kadar R-CNN algoritmasının hızlandırılmış versiyonu olsalar da YOLO hız ve tespit oranıyla ön plana çıkmaktadır.

Bu çalışmada topun tespit edilebilmesi ve takip edilebilmesi için YOLO (You Only Look Once) algoritması kullanılmıştır. YOLO algoritması da R-CNN temelli algoritmalar gibi CNN temelli bir algoritma olsa da görüntü üzerinde birden çok CNN işlemi yapmak yerine tek seferde bu işlemi gerçekleştirip çok

hızlı sonuçlar alınmasını sağlamaktadır (J. Redmon, 2016). YOLO algoritmasının bir alt sürümü olan YOLO-tiny ise daha az evrişimli katman kullanarak daha hızlı fakat doğruluk payı daha az sonuç vermektedir.

Bu projede YOLO algoritmasının YOLOv3 versiyonunun YOLOv3-tiny sürümü kullanılacaktır. Bunun sebebi sadece pinpon topunun tespitinin yapılması ve daha kompleks nesnelere kıyasla pinpon topunun basit yapısı ve tek renkli olmasıyla daha kolay bulunabilmesidir. Masa tenisi maçlarında, top oldukça hızlı olabilmektedir. Bu nedenle algoritmanın gerçek zamanlı çalışması için doğruluk payından bir miktar fire verilmesi gerekmiştir.

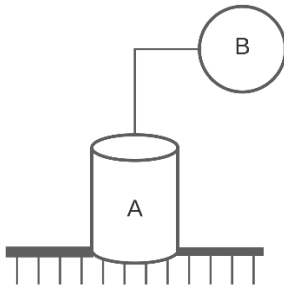
2.1.2. Kamera

Bu çalışmada YOLO algoritmasının yürütülmesi için özelleştirilmiş bir kamera kullanılmıştır. Bilgisayarlı görü ve dijital görüntü işleme için özelleştirilmiş olması, çözünürlüğü, görüntü kalitesi, renk skalası, görüş açısının değiştirilebilmesi gibi özelleştirilebilir bir yapıda olması görüntü üzerinde optimum sonucun bulunabilmesi için modifikasyonların önünü açarak kullanım kolaylığı sağlamıştır. Böylece, kamera üzerinde yapılacak gerekli parametre değişimleri ile YOLO performansı arttırılabilecek, ağ üzerinde sürekli olarak aktarım yapılan görüntünün çözünürlüğü darboğaz oranına göre değiştirilebilmiştir.

Çalışmada kullanılacak kameranın konumu, açısı ve görüş açısı, robot kolun hareketlerinin hesaplanması ile doğrudan bağlantılıdır. Algoritma topu kameranın algoritmayı beslediği görüntü üzerinde bulacak, robot kola iletilmesi gereken sinyaller bu görüntü üzerinde bulunan topun koordinatlarına göre hesaplanmıştır. Bu nedenle, kameranın bulunduğu topun konumu ile robot kolun işaret edeceği nokta arasındaki farkın az olması için kamera ile robot kol birbirine yakın yerleştirilmiştir.

2.1.3. Robot Kol

Bu çalışmada kullanılacak olan robot kol iki eksen üzerinde hareket edecektir. Şekil 1'de görülebileceği üzere eksenlerden biri masa tenisi masasına dik şekilde 180° içerisinde hareket etmektedir. Diğeri ise masa tenisi masasına paralel şekilde 180° açı içerisinde hareket etmektedir. Böylece masa tenisi masasının orta kenarına yerleştirilen robot kol, topun hareketlerini masa üzerinde takip edebilmektedir.

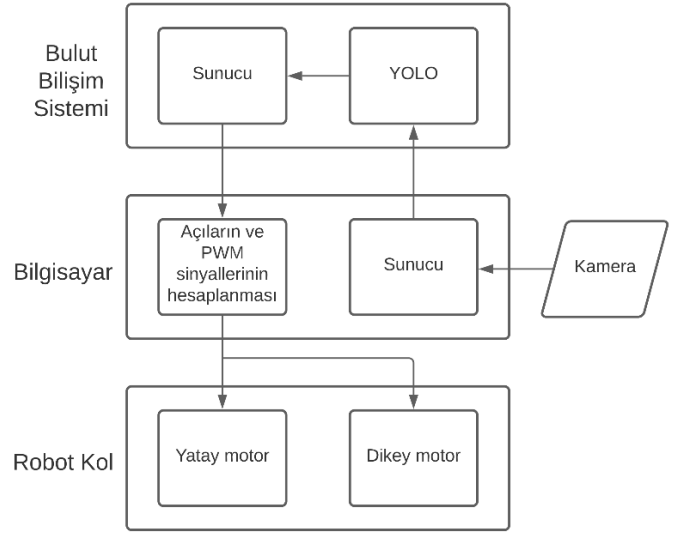


Şekil 1. Robot kolun şeması. Silindir ve daire servo motorları temsil etmektedir. A motoru robot kolun yatay eksendeki hareketini, B motoru dikey eksendeki hareketi sağlamaktadır.

Robot kolda eksenleri oluşturan ve hareket kabiliyetini sağlayan motorlar olarak servo motorlar tercih edilmiştir. Servo motorlar hareket kontrolünün yapılabildiği motor düzenekleridir. DC motorlardan farklı olarak, manyetik alan vektörü ile sargı akım vektörünü dikkate alarak -90° ile $+90^\circ$ arasındaki

hareketinin geri dönüş sinyalini kontrolcüsüne iletebilmektedir. Böylece motor istenen açıya hareket ettirilebilir. Motorun istenilen açıya hareket ettirilebilmesi için PWM (Sinyal Genişlik Modülasyonu) sinyallerine ihtiyaçları vardır (Paul Scherz, 2016). PWM sinyalleri kare dalgalar şeklindedir ve cihazlara verilen gücün ayarlanabilmesi için anahtar görevinde kullanılır. Bu çalışmada MG90S model servo motorlar kullanılmıştır.

Robot kolun eksenlerini servo motorlar oluşturmuştur. Servo motorlardan biri masa tenisi masasına yatay konumlandırılmış, böylece masa tenisi masasına paralel hareket kabiliyeti sağlanmıştır. Dik yerleştirilecek motorun üstünde bulunmakta olan dikey eksende hareket kabiliyetini sağlayacak olan servo motor ise masa tenisi masasına dik konumlandırılmıştır. İki servo motor birbirlerine hareket kabiliyetlerinin kısıtlanmadığı bir şase ile bağlıdır. Dikey eksende hareket kabiliyetini sağlayan servo motora bir işaretçi yerleştirilmiştir. Böylece robotun topun konumunu ne kadar doğru gösterdiği anlaşılabilmiştir.

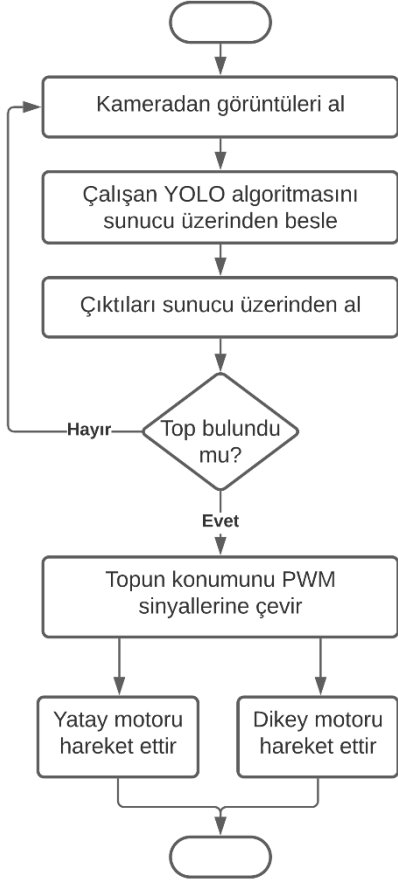


Şekil 2. Sistem bileşenlerinin bütün halinde görünümü ve birbirleriyle iletişimi

2.2. Sistem Mimarisi

Masa tenisi masasına robot kolun bulunduğu konumun üzerinde yerleştirilen kamera pinpon topunu anlık olarak kaydetmektedir. Alınan görüntü sürekli olarak Google COLAB adındaki bulut bilişim sisteminde çalışan ve uzak bir makinede yürütülen YOLO algoritmasına aktarılmaktadır. YOLO algoritması fotoğraf ve video formatlarında çalışabildiği gibi ağ üzerinden anlık olarak gelen yayın formatları üzerinde de yürütülebilmektedir. Algoritma, görüntülerden oluşan video ya da yayın dosyasındaki bütün görüntüler üzerinde aynı işlemi yaptığından dosyanın bütünlüğü, videonun eksiksiz olma zorunluluğu söz konusu değildir. Bulut bilişim sisteminde yürütülen algoritma, topun yatay ve dikey koordinatlarını, doğruluk oranını, kaç tane top bulunduğunu ve anlık saniye başına kare sayısını anlık çıktı olarak üretmektedir. Alınan bu çıktılar sürekli olarak yine bulut bilişim sistemi üzerinde bulunan bir dosyaya yazılmaktadır. Bu dosya sürekli olarak robot kolun bağlı bulunduğu ana bilgisayara gönderilmektedir. Bilgisayara anlık olarak gelen çıktı ile topun koordinatları bilinmekte, eğer hata sonucu ya da bilinçli olarak yapılan bir olay sonucu görüntüde birden fazla top bulunduğu takdirde daha önceden takibi yapılan topun hatasız takip edilmesine devam

edilmesi için ayrı bir algoritma yürütülmektedir. Görüntüdeki topun konumu kesin olarak kararlaştırıldıktan sonra ise robot kolun motorlarının beslenebilmesi için gereken açı hesaplama işlemi gerçekleştirilir. Robot kolun motorlarının hareketini sağlayacak beslemenin yapılabilmesi için gereken açı değerleri hesaplandıktan sonra motorlara açı değerleri PWM sinyallerine çevrilerek iletilir ve robot kolun hareketi sağlanır. İşleyiş Şekil 3'te net olarak görülebilmektedir.



Şekil 3. Sistem mimarisinin iş akış diagramı

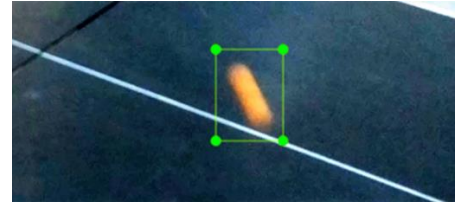
2.3. Teknik Yaklaşım ve Uygulama Detayları

2.3.1. Veri Setinin Hazırlanması

Bütün yapay sinir ağı algoritmalarında olduğu gibi YOLO algoritmasının da eğitilmesi gerekmiştir. Çalışmamız sırasında ilk başta veri seti olarak çeşitli sporlarda kullanılan farklı topların bulunduğu Google Open Images Dataset V4 isimli veri setinin kullanılması düşünülmüş, fakat diğer topların pinpon topuna benzememesi nedeniyle bundan vazgeçilmiştir. (Open Images Dataset, 2020)

Telefon kameraları ve dijital kameralar kullanılarak pinpon topunun durağan halinin fotoğrafları çekilmiştir. Pinpon topunun farklı renklerde mevcut olabilmesi nedeniyle beyaz, sarı ve turuncu renkteki pinpon topları kullanılmıştır. Masa tenisi maçlarında pinpon topu görüntülere her zaman dairesel biçimde yansımamaktadır. Hızına bağlı olarak dairesel şekli eliptik şekle dönüşebilmekte ve bulanıklaşabilmektedir (motion blur). Masa tenisi masasına çarptığı anda görüntü yakalanır ise V şeklide bir biçim elde edilebilmektedir. Bu nedenle telefon kameraları kullanılarak bütün bu olası durumları kapsayan birçok masa tenisi maçı videosu çekilmiştir.

Bu videolar daha sonra bilgisayar ortamında çerçevelere ayrılmıştır. İnternet üzerinde bulunan profesyonel masa tenisi maç videoları da çerçevelere bölünmüş ve veri setine eklenmiştir. Böylece pinpon topunun video içerisinde oluşturabileceği bütün biçimlerin veri setine eklendiğinden emin olunmuştur. Farklı çözünürlüklerde videolar üzerinde veri seti oluşturulmuştur. Bu sayede kamera kaynaklı sorunların önüne geçilmeye çalışılmıştır. Pinpon topunun içerisinde bulunmadığı, negatif görüntüler de veri setine eklenmiştir. Negatif görüntüler sıradan manzara, şehir ya da farklı nesnelerin görüntülerinden ve masa tenisi maçı sırasında topun görüntüde olmadığı durumlardan oluşturulmuştur. Böylece algoritmaya topun bulunduğu görüntüler ve topun bulunmadığı görüntüler aynı veri seti içerisinde verilmiş, algoritmanın doğruluk payının artırılması amaçlanmıştır.



Şekil 4. Etiketleme işleminden bir örnek

Çekilen fotoğraflardan, yakalanan videolardan ve internette edinilen videolarla birlikte yaklaşık 21,000 farklı pinpon topu görüntüsü ve negatif görüntü elde edilmiştir. Oluşturulan veri seti YOLO formatına göre pinpon topunun içerisinde olduğu kutular oluşturulacak LabelImg programının yardımıyla Şekil 4'te bulunan örnek etiketleme işlemi gibi etiketlenmiştir. YOLO formatı görüntü üzerinde bulunan nesnenin sınıfı, nesnenin orta noktasının piksel cinsinden X ve Y koordinatları ve nesnenin içinde bulunduğu kutunun piksel cinsinden genişlik ve yüksekliğinden oluşan, her satırda görüntüde bulunan diğer nesnelerin bahsedilen bilgilerinin yer aldığı bir formattır.

2.3.2. Algoritmanın Eğitilmesi

Hazırlanan veri seti görüntülerden ve etiket dosyalarından oluşmaktadır. Veri setinin algoritmaya girdi olarak verilebilmesi için dosyaların bulunduğu konumları, nesne sınıfı sayısını, nesnelerin adını, öğrenim ve test görüntülerinin ayrı konumlarını içeren bir yazı dosyası hazırlanmıştır. YOLO algoritması öğrenim süresi boyunca sürekli kendini yinelemektedir. Her yinelemenin sonunda öğrenimini test görüntüleriyle test edip hata oranı ortaya koymaktadır. Bu nedenle oluşturulan veri setinin %80'i öğrenim için, %20'si ise test işlemleri için ayrılmıştır. Test görüntülerinin ise %75'i pinpon topunun içerisinde bulunduğu pozitif görüntülerden, %25'i ise pinpon topunun bulunmadığı negatif görüntülerden elde edilmiştir.

YOLO algoritması hazırlanan veri setini daha hızlı öğrenmek ve çalıştırılma aşamasında daha hızlı çıktı vermek için grafik işlemci birimi kullanılabilmesini sağlamaktadır. Grafik işlemci biriminin işlem gücüne dayalı olarak öğrenme ve uygulama hızı değişmektedir. Bu nedenle bu çalışmada Google COLAB adındaki bulut bilişim sistemi üzerinde yer alan Nvidia TESLA P-100 adlı grafik işlemci birimi kullanılmıştır.

Algoritmanın daha az evrişimli katman kullanan versiyonu YOLO-tiny sürümü kullanılmıştır. Algoritmanın normal versiyonu 74 evrişimli katman kullanırken, YOLO-tiny sürümü ile birlikte 15 evrişimli katman kullanılmıştır (Bochkovskiy, 2020). Sadece tek bir sınıfa ait nesne tespit edileceği ve bu

nesnenin diğer nesnelere kıyasla daha basit bir yapıda olması nedeniyle YOLO-tiny sürümünün gerçek zaman üzerinde daha hızlı sonuç vereceği düşünülmüştür.

21,000 farklı görüntü içeren veri seti Nvidia TESLA P-100 grafik işlemci biriminde yaklaşık 7 saatte işlenmiş ve algoritma tarafından öğrenilmiştir. Öğrenme sırasında algoritmanın her yineleme sonucunda ortaya koyduğu hata oranı eğitim süresi boyunca giderek düşmüş, 6 saat 30 dakika sonunda hata payı %0.17 seviyesine indirgenmiştir. 30 dakika boyunca algoritma bu seviyenin altına inememiş, bu nedenle öğrenim durdurulmuştur.

2.3.3. Algoritmanın Uygulanması

Algoritma çalıştırılırken kameradan gelen anlık görüntüde arkaplanda bulunan ya da kisten algoritmanın hata vermesini sağlamak için konulmuş pinpon topuna benzeyen ya da kısa zaman aralığı içinde doğrusal olmayan hareket yaparken yakalanan ve hızlı hareket eden pinpon topunun oluşturduğu biçimlere benzeyen nesnelere pinpon topu olarak algılayabilmektedir. Bu sorunun önüne geçebilmek için her görüntü işlendikten sonra algoritma tarafından sağlanan doğruluk oranı, %25 ve üzerinde olacak şekilde ayarlanmış, sadece bu oranın üzerindeki çıktılar alınmıştır.

Masa tenisi masasından anlık olarak kayıt alacak olan kameradan gelen görüntülerle bir yayın oluşturulmuştur. Bunun sebebi pinpon topunun tespiti için kullanılan algoritmanın bir bulut bilişim sistemi üzerinde çalışacak olmasıdır. Oluşturulan yayın sayesinde algoritma ağ üzerinden anlık görüntüleri yakalayabilecek ve işleyebilecektir. Kameradan gelen görüntüler algoritmanın da üzerinde çalıştığı bulut bilişim sistemi üzerinde çalışan bir sunucuya gönderilmektedir. Sunucu gelen görüntüleri bir dosya üzerinden algoritmaya girdi olarak vermektedir.

Algoritma görüntü üzerinde tespit ettiği topun ya da topların merkez koordinatlarını X ve Y şeklinde piksel cinsinden çıktı olarak vermektedir. Çıktı, kameranın bağlı bulunduğu ve robot kolun sürüleceği bilgisayar üzerinde çalıştırılan farklı bir sunucuya yayın yapılmaktadır. Böylece kamera ile algoritma arasında iletişim sağlanmıştır.

Algoritma görüntüde bulunan bütün pinpon toplarını yakalamakta ve koordinatlarını çıktı olarak vermektedir. Bir görüntüde aynı anda birden fazla top bulunması ya da hata sonucu başka bir nesnenin top olarak algılanması halinde, koordinatları ve benzerlik oranları robot kolun sürüleceği bilgisayara iletilecektir. Bu nedenle takip edilen pinpon topunun konumu yanlış algılanabilecek, görüntüdeki diğer top ya da topları takip etmeye başlayabilecektir. Bu sorunun önüne geçilebilmesi için Şekil 5 üzerinde de görülebileceği gibi eğer birden fazla top varsa bir önceki görüntüde bulunan topun konumuna en yakın top seçilmektedir. Böylece zaten takip edilmekte olan top takip edilmeye devam etmekte, diğer topları ya da hata sonucu top olarak tespit edilen nesnelere gözardı edilmektedir.

2.3.3. Motor Açılarının Hesaplanması ve Robot Kolun Hareket Ettirilmesi

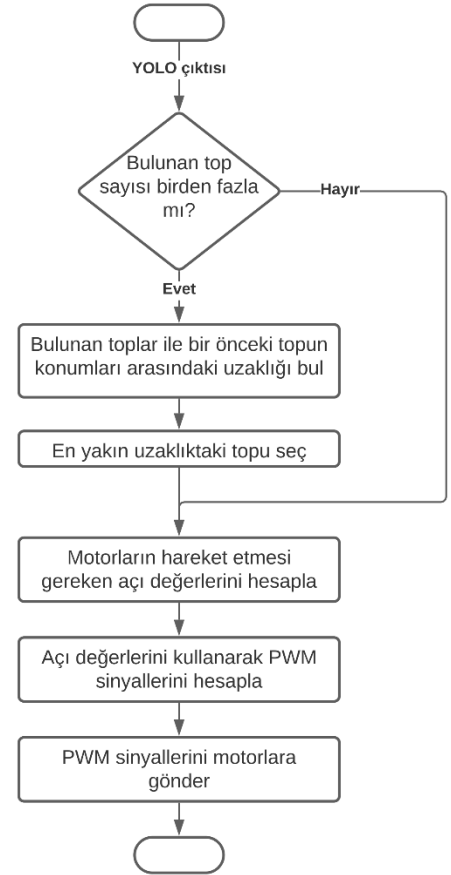
Algoritma tarafından sunucu üzerinden sağlanan ve bir sonraki görüntüde eğer birden fazla top varsa hangi topun takip edileceği kesinleştirildikten sonra, topun görüntü üzerindeki yatay ve dikey koordinatlarına bakılarak robot üzerinde bulunan

motorlara gerekli açı girdileri hesaplanmaktadır. Açı değerleri, kameranın bulunduğu noktaya, dolayısıyla görüntünün kaydedildiği görüş açısına ve kameranın robota göre göreceli konumuna bağlı olarak değişmektedir. Denklem I ve Denklem II'de bu hesaplamaların nasıl yapıldığı görülebilir. Bu çalışmada, kamera ve robot kolun konumları sabittir. Kamera robot kolun üzerinde, robot kol ile dikey ekseninde aynı hizada bulunmaktadır. Kameranın çözünürlüğü 1366x768 piksel boyutlarındadır. Algoritma tarafından sağlanan koordinatlar bu sınırlar içerisinde çıktı olarak gelmektedir. Kullanılan kameranın yatay görüş açısı 62,2° ve dikey görüş açısı 48,8°'dir. Bu nedenle robotun yatay ve dikey eksenlerdeki hareket alanı görüş açılarına göre 180°nin altına indirilmiş, tutarlılık sağlanmıştır.

$$a = (180 - A)/L \quad I$$

$$M = (K * a) + A/2 \quad II$$

A değeri kameranın yatay ya da dikey eksenindeki görüş açısını, L değeri kamera çözünürlüğünün yatay ya da dikey eksenin piksel cinsinden uzunluğunu, K değeri topun yatay ya da dikey eksenindeki piksel cinsinden konumunu gösterir. Ortaya çıkan M sonucu yatay ya da dikey eksen hareketi için kullanılacak motora gönderilecek PWM sinyali oluşturmak için gereken açıyı vermiştir.



Şekil 5. YOLO çıktısından motorların hareket ettirilmesine kadar olan iş akış diagramı

Robot kol üzerinde bulunan motorlar için gerekli açı bilgisi hesaplandıktan sonra, motorlara belirlenen açığa yönelmesi için gereken PWM sinyallerinin oluşturulması gerekmektedir. Robot kolda kullanılan motorlar aynı tipte olup, ikisinin de PWM sinyal aralığı 600µs-2400µs aralığındadır. 180° içerisinde verilmesi gereken her açı değeri;

yatay eksen için,

$$S = 2400 - M * 10 \quad \text{III}$$

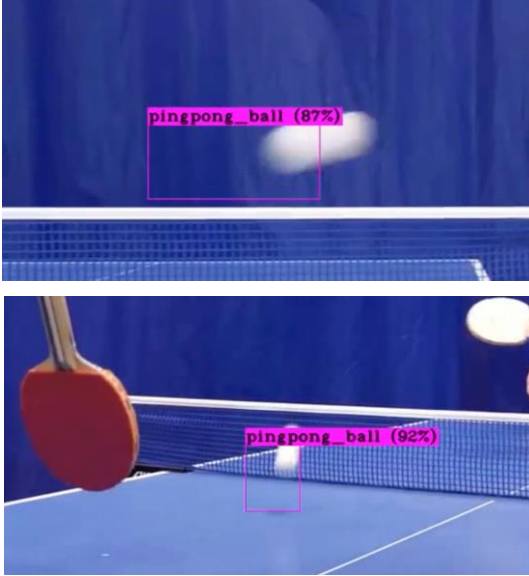
dikey eksen için,

$$S = 600 + M * 10 \quad \text{IV}$$

denklemleriyle PWM sinyaline dönüştürülerek bulunmuştur. PWM sinyalleri anlık olarak hesaplanmakta ve yine anlık olarak robot kola gönderilmekte ve robot kolun hareketi sağlanmaktadır.

3. Araştırma Sonuçları ve Tartışma

Algoritmaya anlık olarak verilen görüntünün uygulama üzerinden çıktısı video dosyalarına topun merkez konumunu ve topun içinde bulunduğu kutunun kenarları belirli olacak biçimde kaydedilmiştir. Bu çıktı videolarından elde edilen sonuca göre %99'un üzerinde bir başarı oranı sergilenmiş, top gerek durağan halinde gerek yüksek hızlarda aldığı eliptik biçimde, gerekse masaya çarptığı anda oluşturduğu V şeklinde yakalanmıştır. Robot kol ile algoritma arasındaki iletişim başarıyla sağlanmış, robot kolun topu büyük oranda başarıyla izleyebildiği görülmüştür.



Şekil 6. Video kaydı girdisiyle çalıştırılan algoritmadan alınan çıktudan iki görüntü. "pingpong_ball" nesnenin sınıfını, parantez içerisinde belirtilen yüzde ise bu nesneye olan benzerlik oranını göstermektedir.

Bazı durumlarda ise, ortam ışığına bağlı olarak top üzerinde yansıma olabilmekte, ya da top çok hızlı giderken görüntüde oldukça bulanık görüldüğünden arkaplanda bulunan duvar rengi ya da farklı nesnelere ile rengi karışabilmekte ve insan gözünün dahi topu bulmakta zorluk çektiği görüntüler meydana gelebilmektedir. Bu gibi durumlarda algoritma topu bulamamış, fakat top tekrar daha belirgin bir hal aldığı anda çok kısa sürede topu yakalamış ve robot kol top takibine devam edebilmiştir. Gerçek hayat koşullarında özellikle masa tenisi maçlarına bakılırsa topun daha iyi görülebilmesi için kameranın masaya belirli bir açıyla yukarıdan konumlandırıldığı, zemin renginin ise topun seyirciler tarafından daha kolay ayırt edilebilmesi için daha farklı bir renk seçildiği görülebilmektedir. Bu nedenle algoritma masa tenisi maçları üzerinde uygulandığında yüksek doğruluk oranları sergilemektedir.



Şekil 7. Takip edilen pinpon topunun izlediği yol görüntüsü

YOLO algoritması sonucu üzerinden gelen görüntüler dışında halihazırda bulut bilişim sistemine yüklenmiş masa tenisi maçı videoları üzerinde denendiğinde, saniyede 24 kare hızda olan videoları eşzamanlı olarak işleyebilmiş, ortalamada saniye başına 42 kare hızına ulaşılmıştır. Gerçek durum senaryoları için 24 kare ve 30 kare çekim yapan kameralar için oldukça yeterli olduğu kanaatine varılmıştır.

Robot kol masa tenisi masasının yan kenarında tam ortada durmaktadır. Kamera ise birden fazla kamera kullanılmadığı ve robotun topu yatay ve dikey eksende izlenmesi hedeflendiği için robot kol ile dikey eksen hizasında aynı konumda ve robot kolun hemen üzerinde yerleştirilmiştir. Fakat, robot kolun dikey eksenindeki hareketini sağlayan motorunda bağlı bulunan işaretçi ile kamera arasında yaklaşık 5cm yükseklik farkı oluşmuştur. Denklem I ve Denklem II yatay eksen için iyi sonuç vermiştir. Ancak robot kol dikey ekseninde Denklem I ve Denklem II ile belirlenen hareket alanının sınır açalarına yaklaştıkça, doğrusal olarak artan bir sapma oranıyla topun yerini doğru olarak göstermemeye başlamaktadır. Bu sorun Denklem I ve Denklem II üzerinde kamera ile robot kol arasındaki uzaklık farkı kullanılarak düzeltilmeye çalışılmış, büyük ölçüde başarı yakalanmıştır. Tek kamera yerine iki kamera kullanılabileceği, topun 3 boyutlu düzlemde tam konumunun çıkarılabileceği ve bu sayede hatanın düzeltilebileceği düşünülmektedir.

4. Sonuç

Yapılan çalışmada oluşturulan veri setiyle YOLO algoritması başarılı bir şekilde eğitilmiş, algoritmadan gerçek zamanlı dönüş alınabilmiş, topun yatay ve dikey eksenindeki merkez konumu büyük oranda doğru tespit edilmiş, robot kol işaretçisini büyük oranda topa doğru yöneltebilmiştir. YOLO algoritmasının doğrusal olmayan ve yüksek hızlarda hareket eden küçük nesnelere gerçek zamanlı uygulanması başarıyla gerçekleştirilmiştir.

Literatürde masa tenisi topunun takibi için Kalman filtresi (G. Chen, 2013) (P. R. Kamble, 2019), parçacık filtresi (Y. Huang, 2008) ve bulanık mantık algoritması (Hu Su, 2013) ya da basit Gaussian modellemesi (J. Liu, 2014) gibi farklı algoritmalar kullanılmıştır. Bu çalışmayla birlikte, bu tür literatürde daha önce kullanılan algoritmaların yanında, YOLO algoritmasının da pinpon topu gibi yüksek hızda doğrusal olmayan yönlerde hareket edebilen nesnelere tespiti için kullanılabileceği görülmüştür. Ayrıca, farklı nesnelere aynı anda eğitilebildikleri ve görüntü üzerinde eğitilen bütün nesnelere için tespit çalışması yapabildiği için birden fazla pinpon topu ya da birden fazla farklı yönlerde hareket eden ve farklı şekillerde nesnelere takibinin aynı görüntü üzerinde gerçekleştirilebileceği düşünülmektedir.

YOLO algoritmasının sadece tek bir görüntü üzerinde ya da önceden kaydedilmiş videolar üzerinde değil, aynı zamanda gerçek zamanlı video kayıtlarında da kullanılabileceği görülmüş, bu algoritmanın kullandığı işlem gücüne bağlı olarak çok hızlı bir şekilde canlı görüntülerden tespit yapabildiği görülmüştür.

Böylece birçok farklı gerçek senaryoda bu algoritmanın kullanılabileceği düşünülmektedir.

İlerleyen dönemlerde iki ya da daha fazla kamerayla topun 3 boyutlu hareket modelinin çıkarılmasıyla kesin konum bulunması ve robot kolun geliştirilip masa tenisi topuna karşılık verecek şekilde oyuncu konumuna getirilmesi planlanmaktadır.

Kaynakça

- J. Redmon, S. D. (2016). You Only Look Once: Unified, Real-Time Object Detection. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) (pp. 779-788). Las Vegas: IEEE.
- Y. Huang, J. L. (2008). A Method of Small Object Detection and Tracking Based on Particle Filters. ICPR 19th International Conference on Pattern Recognition (s. 1-4). Tampa, FL: IEEE.
- G. Chen, D. X. (2013). Visual Measurement of the Racket Trajectory in Spinning Ball Striking for Table Tennis Player. IEEE Transactions on Instrumentation and Measurement, 2901-2911.
- J. Liu, Z. F. (2014). Improved high-speed vision system for table tennis robot. 2014 IEEE International Conference on Mechatronics and Automation (pp. 652-657). Tianjin: IEEE.
- P. R. Kamble, A. G. (2019). Ball tracking in sports: a survey. Artificial Intelligence Review, vol. 52, no. 3, pp. 1-51.
- Bochkovskiy, A. (2020, 04 23). GitHub- AlexeyAB/darknet: YOLOv3. [github.com: https://github.com/AlexeyAB/darknet](https://github.com/AlexeyAB/darknet) sitesinden alınmıştır.
- Szeliski, R. (2011). Computer Vision: Algorithms and Applications (Texts in Computer Science) 2011th Edition. Springer.
- Aggarwal, C. C. (2018). Neural Networks and Deep Learning, A Textbook. Springer.
- Branko Ristic, S. A. (2003). Beyond the Kalman Filter: Particle Filters for Tracking Applications. Artech House.
- Corke, P. (2017). Robotics, Vision and Control: Fundamental Algorithms In MATLAB, Second Edition. Springer.
- Rafael C. Gonzalez, R. E. (2017). Digital Image Processing, 4th Edition. Pearson.
- Simon, D. (2006). Optimal State Estimation: Kalman, H Infinity, and Nonlinear Approaches 1st Edition. Wiley.
- Paul Scherz, S. M. (2016). Practical Electronics for Inventors, Fourth Edition. McGraw-Hill Education TAB.
- Hu Su, Z. F. (2013). Trajectory Prediction of Spinning Ball Based on Fuzzy Filtering and Local Modeling for Robotic Ping-Pong Player. IEEE Transactions on Instrumentation and Measurement, 2890-2900.
- Open Images Dataset. (2020, 01 29). Google OpenSource: <https://opensource.google/projects/open-images-dataset> sitesinden alınmıştır.