

Derin Öğrenme Yöntemleri ile Şüpheli Davranış Tespiti

Suspicious Behavior Detection With Deep Learning Methods

Duygu ÇALIŞKAN¹ , Önder DEMİR^{2*} 

1Marmara Üniversitesi, Fen Bilimleri Enstitüsü, Bilgisayar Mühendisliği, 34722, Göztepe Kampüsü, Kadıköy, İstanbul

2Marmara Üniversitesi, Teknoloji Fakültesi, Bilgisayar Mühendisliği, 34722, Göztepe Kampüsü, Kadıköy, İstanbul

*odemir@marmara.edu.tr

Öz

Hızla gelişen teknoloji ile birlikte askeri, güvenlik ve bilişim sektöründe bilgi güvenliği sorunları meydana gelmiş bu sorunların çözümü için donanımsal ve yazılımsal yöntemler geliştirilmiştir. Görüntü işleme teknikleri yardımıyla bu güvenlik sorununa görüntüden şüpheli davranış tespiti ile gerçek zamanlı bir çözüm getirilmeye çalışılmaktadır. Görüntüden nesne ve pozisyon tanımlama, davranış belirleme ve derin öğrenme üzerine yapılmış birçok çalışma mevcuttur. Derin öğrenme, insanın düşünce yapısını veri setlerindeki örüntüler ile deneyimleyerek bilgisayara öğreten, yapay zekâ kavramının içinde barındırdığı makine öğrenmesi konusunun özel bir halidir. Makine öğrenme algoritmaları, belirlenmiş bir denkleme dayanmaksızın doğrudan bilgi verilerinden hesaplama yöntemlerini kullanarak öğrenir ve modellenirler. Derin öğrenme, yüz tanıma, plaka tanıma, nesne algılama, insansız hava araçlarında nesne tespiti, hareket algılama, otonom sürüş teknolojileri ile şerit tespiti gibi birçok gelişmiş sürücü yardım teknolojisi gibi problemleri çözmek için kullanılır. YOLO(You Only Look Once) algoritması, konvolüsyonel sinir ağlarını (CNN) kullanarak nesne tespiti yapan bir derin öğrenme algoritmasıdır. YOLO, konvolüsyonel sinir ağları ve benzerleri algoritmalara göre daha yüksek performansa sahip bir algoritma olduğundan; çalışma kapsamında gerçek zamanlı tespitte en yakın performans elde edilmek istendiğinden, bu çalışmada YOLO algoritması ile çalışılmıştır. Gerçekleştirilen tez çalışması 3 adımdan oluşmaktadır. Birinci adımda 1116 etiketli görsel ile Marmara Üniversitesi Kriminal Davranış/Nesne Veri Seti(MÜKDN) oluşturulmuş, ikinci adımda derin öğrenme için kullanılan evrişimsel sinir ağı YOLOv4 modeli tasarımı ve konfigürasyonu yapılmıştır. Son olarak üçüncü adımda sistem eğitilmiştir. Evrişimsel sinir ağı ağırlıkları farklı sistemlerde kullanılmak üzere elde edilmiş ve proaktif model gerçekleştirilmiştir.

Anahtar Kelimeler: Derin Öğrenme, YOLO, CNN, MÜKDN

Abstract

With the rapidly developing technology, information security problem have occurred in the military, security and information sector, and hardware and software methods have been developed to solve these problems. With the help of image processing techniques, a real-time solution is tried to be brought to this security question by detecting suspicious behavior from the image. There are many studies on object and position identification, behavior determination and deep learning from the image. Machine learning algorithms learn and model using computational methods directly from information data without relying on a defined equation. It is used to solve problems such as deep learning, face recognition, license plate recognition, object detection, object detection in unmanned aerial vehicles, motion detection, autonomous driving technologies and many advanced driver assistance technologies such as lane detection. YOLO (You Only Look Once) algorithm is a deep learning algorithm that performs detects objects using convolutional neural networks (CNN). YOLO is an algorithm that can respond faster in performance than convolutional neural networks and their derivatives. For this reason, since it is desired to exhibit the closest performance to real-time detection with in the scope of the study, the YOLO algorithm will be used in this study. The thesis work carried out consists of 3 steps. In the first step, the Marmara University Criminal Behavior/Object Data Set (MÜKDN) was created with 1116 labeled images, in the second step, the YOLOv4 model of the convolutional neural network used for deep learning was designed and configured. Finally, in the third step, the system is trained. Convolutional neural network weights were obtained for use in different systems and a proactive model was implemented.

Keywords: Deep Learning, YOLO, CNN, MÜKDN

Sorumlu Yazar/Corresponding Author: Önder DEMİR Tel:, e-posta: odemir@marmara.edu.tr

Gönderilme/Submitted: 13.01.2022, Düzenleme/Revised:20.03.2022, Kabul/Accepted: 23.03.2022

I. GİRİŞ

İnsan davranışının ve çevresiyle etkileşiminin otomatik olarak anlaşılması, çeşitli alanlardaki potansiyel uygulamaları nedeniyle günümüzde ilgi çekici bir araştırma alanı olmuştur. Bu araştırma alanları insan davranışını çok yönlü (duygular, ilişkisel tutumlar, eylemler, yüz ifadeleri vb.) modellemeye çalışır.

Gerek uluslararası, gerek ulusal veya ticari özel sektör girişimlerinde bilgi insanlık tarihi kadar köklü bir geçmişe dayanan ve insanlığın var olduğu sürece de önemini korumaya; vazgeçilmez bir güç olmaya devam edecek, çok önemli bir olgu ve bir teknolojidir.

Yüz yıllardır bilgi sızıntılarının neden olduğu olgular genelde eylemler, davranışlar gerçekleşikten ve istenmeyen sonuçlarla karşılaşıldıktan sonra tespit edilebilmektedir. 21. yüzyılda teknolojinin de hayatımızın her alanına girmesi ile bu bilgi sızıntılarında önceki dönemlere göre kıyaslandığında teknoloji ile; yani akıllı cihazlar aracılığı ile gerçekleştirilmekte ve her geçen gün artış eğiliminde olduğu görünmektedir [1-2].

Yapay Zekâ, en genel ifade ile insana özgü bilişsel, öğrenme, muhakeme etme özelliklerinin yüz yıllardır makinelerle aktarma çalışmaları olarak ifade edilebilecek geniş bir çalışma alanıdır.

Derin öğrenme, insanın düşünce yapısını veri setlerindeki örüntüler ile deneyimleyerek bilgisayara öğreten, yapay zekâ kavramını içinde barındırdığı makine öğrenmesi konusunun birçok multidisipliner alan ile çalışılabilir özel bir halidir [3].

Yüz ifadelerimiz, hareket pozisyonumuz, nesnelere suç eğilimi hakkında beklenenden daha fazlasını ifşa edebilir. Bir kişinin davranışları hakkında bilgi sahibi olabilir ve insanların yüzlerine bakarak duygularını çıkarılabilir; ve insan davranışının spesifik hali şüpheli eğilimleri hakkında önemli çıkarımlar elde edilebilir.

Derin öğrenme modellerindeki çeşitli son gelişmeler, görüntüleri kullanarak anlamsal örüntü tanıma performansını büyük ölçüde artırdı. Bir bireyin duygusal durumu ve diğer belirli karakter özellikleri veya pozisyon özellikleri gibi çeşitli durum tahminleri, yüz görüntülerinden tahmin edilebilir. Bu motivasyonla, bu çalışmada, çeşitli derin öğrenme mimarilerinin öğrenme yeteneklerini kullanarak, yüz ve nesne görüntülerinden suç eğilimi veya (suç tahmini/tespiti) çıkarımı yapılması gerçekleştirilmiştir. Çalışma, suçluların, tam olarak görüntü hırsızlığı yaparken yakalanırsa, bilinmeyen bir kişi tarafından herhangi bir suç eğilimini belirlemek için kullanılabilir bir dizi yüz/nesne özelliği olduğunu ortaya koymuştur. Geliştirilen bu

çalışmada 1116 etiketli görsel ile Marmara Üniversitesi Kriminal Davranış/Nesne Veri Seti(MÜKDN) oluşturulmuş, derin öğrenme için kullanılan evrimsel sinir ağı YOLOV4 modeli tasarımı ve konfigürasyonu yapılmış ve son olarak gerçekleştirilen proaktif model ile sistem eğitilmiş ve diğer sistemlerde de kullanılmak üzere evrimsel sinir ağı ağırlıkları elde edilmiştir.

Görüntü işleme algoritmaları ve derin öğrenme yöntemleri birlikte kullanılarak birçok alanda uygulamalar geliştirilmiştir. Tıbbi görüntülerden hastalık teşhisinde, ilaç üretimine, kalite kontrol uygulamalarına, insansız hava araçları kullanımından, kalabalık analizine, dil işlemeye kadar daha bir çok konuda derin öğrenme ile görüntü işleme alanında çalışmalar yapılmıştır.

Halgur ve ark. 2021' de, BT ve X-ray gibi göğüs radyografilerini kullanarak koronavirüs vakalarını tespit etmede X-ışınları ve BT tarama görüntüleri üzerine basit bir evrim sinir ağı (CNN) ve modifiye edilmiş önceden eğitilmiş AlexNet modeli uygulanmıştır. Deneyle sonuçları, kullanılan modellerin önceden eğitilmiş ağı üzerinden %98'e kadar doğruluk ve modifiye CNN kullanılarak %94,1 doğruluk sağlayabildiğini göstermektedir[4].

Mehdi ve ark. 2020 yılında hastalık teşhisi ve rehabilitasyon süreci için, optimize edilmiş özellik çıkarma ve ardından doğrusal diskriminant analizi sınıflandırması (TLRN-LDA) modelini önermiş ResNet50 Derin öğrenme modeli ile transfer öğrenme konsepti kullanılarak geliştirilmiştir. 31 sınıfın zorlu standart karşılaştırmalı ImageCLEF-2012 veri kümesi üzerinde kapsamlı deneyler gerçekleştirmişlerdir. Geliştirilen yaklaşım, aynı veri kümesindeki en son teknoloji yaklaşımlarla karşılaştırıldığında %10'a kadar daha yüksek olan %87,91'lik gelişmiş ortalama sınıflandırma doğruluğu sağlar[5].

Jauedi ve arkadaşları, 2020 yılında İnsan eylemi tanıma için yeni bir hibrit derin öğrenme modeli ile insan eylemi tanıma alanında yeni bir yaklaşım önermişlerdir. Bu yaklaşım, video içeriğinin analizine ve çıkarma özelliklerine dayanmaktadır. Hareket özelliklerini, GMM ve KF yöntemleri kullanılarak insan hareketi takibi ile sunmaktadır. Bir diğer özellikleri, Kapılı Tekrarlı Birimli Tekrarlayan Sinir Ağları modeli kullanılarak video dizisindeki her bir karenin tüm görsel özelliklerine dayandırmalarıdır. Bu yaklaşımın temel avantajları, videonun her karesinde ve her anındaki tüm özelliklerin analizi ve çıkarılmasıdır. Bu çalışmanın deneysel sonuçları, yüksek sınıflandırma oranı elde etmek için yeni yaklaşımın güçlü performansını göstermektedir. Gerçekten de önerilen yöntem, insan etkinliği analizi gibi çeşitli uygulamalarda ve alanlarda kullanılabilir. Vücut hareketine dayalı tıbbi uygulama, el hareketlerini kullanan akıllı ara yüz ve

daha fazlası. Gelecekteki çalışmalar için amaç, video sınıflandırma süresini azaltmak ve UCF Sport, UCF101 ve daha zorlu veri kümelerinin sınıflandırma oranını iyileştirmek için yaklaşımlarının değerlendirilmesi olarak sunulmuştur[6].

Ashwin ve arkadaşlar 2020 yılında, hintli öğrencilerin yüzlerini, el hareketlerini ve vücut duruşlarını kullanan e-öğrenme ve sınıf ortamları için duygusal veri tabanı oluşturarak yaptıkları çalışmada öğrencinin yüz ifadelerini tanımayı ve bunları Ekman'ın temel duygularına göre sınıflandırmayı içermektedir. Duyuşsal durumlar, öğrencilerin yüz ifadeleri, el hareketleri ve vücut duruşları olmak üzere üç bileşenin tümü kullanılarak sınıflandırılır. İfadelerin sınıflandırılması evrişimli sinir ağları (CNN) kullanılarak gerçekleştirmişlerdir. Çalışma sonucunda öğrencilerin yüz ifadeleri, el hareketleri ve vücut duruşları kullanılarak hem e-öğrenme (tek görüntü çerçevesinde tek öğrenci) hem de sınıf ortamları (tek görüntü çerçevesinde birden çok öğrenci) için çok modlu duyuşsal veri tabanı başarıyla oluşturulmuştur. Nötr de dahil olmak üzere 11 farklı duygusal durum için altın standart çalışması kullanılarak açıklamalar yapıldı. Anlatıcılar, Cohen' in kabul ettiği duygulanım durumlarına karşı ayırım yaparken güvenilir bir şekilde hemfikirlerdir. $=0.48$. Her öğrencinin her modalitesinde nesne lokalizasyonu gerçekleştirilir ve sınırlayıcı kutu koordinatları duyuşsal durumla birlikte saklanır. Duyuşsal durum sınıflandırmasının tespiti ve sınıflandırılmasında %83 ve %76 doğruluk elde etmişlerdir. Mevcut çalışma görüntüden duygu analizi konusunda güncel bir çalışma niteliğindedir[7].

Jammalamadaka ve arkadaşları 2017 yılında, bir sorgu modalitesi olarak poz kullanarak görüntü ve video arama için yeni bir yaklaşımı başarıyla gösterdiler. Dans ve spor video verileri üzerinde poz tanımlayıcılarını elde etmek ve poz alımını gerçekleştirmek için derin pozlet yöntemi ve derin poz gömme yöntemi olmak üzere iki yol önerdiler. İlk yöntemde, poz uzayının 'poza duyarlı' derin pozletler kullanılarak ayrıştırılabileceğini gösterdiler. Bu derin pozlet dedektörleri, belirli bir pozda vücut parçalarının bir alt kümesini modeller. CNN'yi kullanmışlardır. Poz için bir özellik temsili oluşturmada temel yapıtaşları olarak kullanılmıştır. İkinci yöntemde, bir görüntünün daha düşük boyutlu, poza duyarlı bir Alana doğrudan nasıl eşlenebileceğine değinmişlerdir. Daha sonra, her iki yöntemi kullanarak pozlamanın, rakip poz alma yöntemleriyle eşit olduğunu ampirik olarak göstermişlerdir[8].

Dieogo ve arkadaşları 2019 yılında durağan görüntülerden insan poz tahmini için yeni bir regresyon yöntemi sundular. Yöntem parça tabanlı

algılama haritalarını dolaylı olarak öğrenmek için derin bir evrişimli ağa entegre edilebilen türevlenebilir bir işlem olan soft-argmax işlemine dayanmaktadır ve bu, regresyon yöntemlerinden elde edilen son teknoloji puanlar üzerinde önemli bir gelişme sağlamıştır. Gelecekteki bir çalışma olarak, 3B poz tahmini veya pozdan insan eylemi tanımayı tamamen ayırt edilebilir bir şekilde sağlamak için yaklaşımlarına başka yöntemler de eklenebilir olduğunu belirtmişlerdir[9].

Baccouuche ve arkadaşları 2011 yılında, KTH veri setini kullanarak önsel bir modelleme olmaksızın, yalnızca eğitim örneklerinden otomatik öğrenmeye dayanarak, 3D Evrişimsel ağının genişletilmesiyle mekansal-zamansal özellikleri otomatik öğrenen model geliştirdiler. İnsan eylemlerinin dizilerini sınıflandırmak için nöral tabanlı bir derin model kullandılar. İki adımlı şema olarak sunulan modelde, spatiotemporal özellikleri otomatik olarak öğrenir ve bunları tüm dizileri sınıflandırmak için kullanmışlardır[10].

İnsan eylemlerini tanımak için en popüler modern yöntemler arasında Laptev ve ark. [11] Dollar ve ark. [12] Chen ve ark. [13] Gao ve ark. Liu ve ark[14], hepsi manuel olarak tasarlanmış mekansal-zamansal ilgi noktaları etrafında hesaplanan mühendislik hareket ve doku tanımlayıcılarını kullanır.

Daş ve ark., 2019 yılındaki derin öğrenme yöntemleri ile hareketli nesne takibi çalışmalarında Google'ın açık kaynak kodlu olarak kullanıma sunduğu Tensorflow kütüphanesini kullanarak nesne takibi için Faster R-CNN modeli ele alınmışlardır. Bu kütüphane yardımıyla fotoğraflar, video görüntüler üzerinde nesne tanıma gerçekleştirmiş; kütüphanelerin güçlü ve zayıf yönlerini değerlendirmişlerdir [15].

Yılmaz ve ark., 2020 yılındaki çalışmalarında, Faster CNN(Faster Region Based Convolutional Networks) ağı kullanarak geliştirdikleri çalışmada derin öğrenme teknikleriyle beraber Faster R-CNN ağı kullanılmışlardır. 502 Adet resim verisinden oluşan veri seti oluşturulmuş. Eğitilen model resim, video ve kamera üzerinden alınan anlık görüntüler ile çalıştırılmış ve yüksek başarı oranı elde edilmiştir. Modelin real-time gıda üretim tesisinde kalite kontrol sürecinde kullanılabileceği belirtilmiştir [16].

II. MATERYAL VE YÖNTEM

Bu bölümde derin öğrenme ile şüpheli davranış tespitinde kullanılan materyaller ve izlenen yöntem anlatılmıştır.

2.1. Yapay Zekâ

Yapay zekâ, antik roma uygarlığına kadar dayanan insan gibi düşünebilme, insan bilincini oluşturmayı hedefleyen, muhakeme edebilme, ön

görüde bulunma gibi kompleks işlemlerin modellenerek makinelere aktarılması olarak karşımıza birçok farklı problemin çözümünde kullanılan önemli bir bilim alanıdır.

Yapay Zekâ, en genel ifade ile insana özgü bilişsel, öğrenme, muhakeme etme özelliklerinin yüz yıllardır makinelere aktarma çalışmaları olarak ifade edebilecek geniş bir çalışma alanıdır[17].

Ramon Lull 1302 yılında yayımlanan The Ultimate General Art adlı kitabında yapay zekâdan; kavramların kombinasyonlarından oluşan yeni bir bilgi türünden ilk kez bahsetmiştir [18].

Akabinde 2. Dünya savaşında Allen Turingin "Makineler Düşünebilir mi?" sorusu ve Turing testi ile yapay zekâ fikri ortaya atıldı[19].

1956 yılında ise Mc Carthy ve arkadaşları tarafından gerçekleştirilen konferansta Ramon Lull'un bahsettiği yeni bilgi türünün adı yapay zekâ olarak isimlendirilmiştir. Mc Carthy, yapay zekâ isminin babası olarak tarihte yerini aldı[20].

Yapay Zekâ gelişimi dikkate alındığında elbette ki karşımıza elektriksel ve donanımsal çalışmaların sınırlamaları çıkmaktadır. Bu donanımsal ilerlemeler ve algoritmalarındaki gelişmelerle araştırmacıların dönemsel duraksamaları tarih sürecinde gözlemlense de. 1970'ler ve 1980'lerden sonraki donanımsal gelişmeler ile birlikte farklı yeni modeller geliştirilmiştir ve makine öğrenmesi ile farklı bilgi işlem teknikleri ortaya çıkmıştır. Bu tekniklerden en önemlisi de insan sinir sistemi modelini örnek alan yapay sinir ağıdır.

Bilgisayar donanımlarındaki gelişmeler insan sinir ağını modellenebilmesini mümkün kılacak düzeyde işlemleri tamamlayabilen yeni yapay zekâ modelleri tasarlanmıştır. Bu çalışma keşfine derin öğrenme modeli denilmektedir. Bu model görüntü işleme, ses işleme, genom yapıları, öngörü sistemleri vb. hayal gücünüze kalmış sistemlere katkıda bulunmuştur.

2.1.1 Yapay Sinir Ağları

1943 yılında ilk yapay sinir ağı modeli nörolog Warren McCulloch ve matematikçi Walter Pitts tarafından tanımlanmıştır. Biyolojik bir nöron hücresi; dendrit, hücre gövdesi, aksonlar ve sinapslardan oluşmaktadır diye tanımlanmıştır [21].

Biyoloji hücresinden esinlendiğini belirttiğimiz sinir ağıımız, 1957'de Rosenblatt yapay sinir ağlarının temel yapı taşı olan algılayıcı (perceptron) tanımını yapmıştır [22].

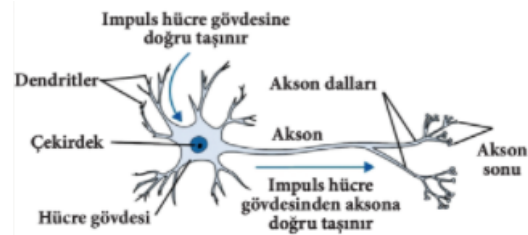
Şekil 1'de gösterilen McCulloch-Pitts modeline dayanan nöron m tane giriş parametresi(x_j) almaktadır ve alınan her giriş parametresi için bir yönlendirilmiş ağırlık parametresine(w_j) bağlanır. Sistemde giriş ve ağırlıklar doğrusal bir şekilde toplanır ve aktivasyon fonksiyonuna (ϕ) gönderilir. Aktivasyon fonksiyonunun çıkışı nöronun

çıkışını(y_k) üretir.. Bu işlem Denklem 1' deki gibi formülize edilebilir [21].

$$y_k = \phi(sk) = \phi(\sum w_k x_j m_j = 0) \quad (1)$$

İstenen çıktının üretilebilmesi için sistemin ağırlıkları dikkatle seçilerek nöronun eğitimi gerçekleştirilir. Birden fazla nöronun kullanılmasıyla oluşturulan sitemlere yapay sinir ağı denilmektedir [23].

Tez çalışmasında kullanılan yapay sinir sisteminde bu yapılar; işlemci elemanları, toplama fonksiyonları, transfer fonksiyonları, yapay nöron çıkışları ve ağırlıkları ifade etmektedir.

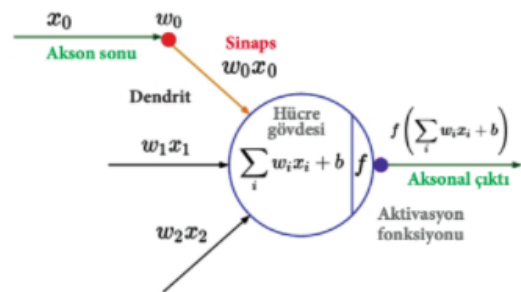


Şekil 1. Nöron Matematiksel Modeli[24]

2.1.2 Derin Öğrenme

Derin öğrenme, yapay sinir ağları algoritmalarını kullanarak çok katmanlı mimarilerde çok boyutlu veriler ile çalışma imkânı sağlayan, makine öğrenmesi alanının bir alt dalıdır. Derin öğrenme; yüz tanıma, plaka tanıma, nesne algılama, hareket algılama, şerit tespiti, yaya algılama, otomatik park etme ve gelişmiş araç sürüş yardımcısı, pozisyon tanımlama, duygu algılama, doğal dil işleme, ilaç keşfi gibi birçok karmaşık problemleri çözmek için kullanılmıştır ve bu çalışmalarda ciddi bir başarı oranı elde edilmiştir. Resim tanıma ve görüntü işleme için özellikle uygun bir alandır [3].

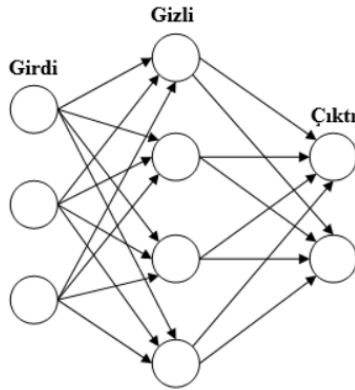
Şekil 1' deki biyolojik bir sinir hücresi referans alınarak Şekil 3 teki basit bir yapay sinir ağı modeli gösterilmiş ve bu yapay sinir ağının matematiksel modelinde Şekil 2' de gösterilmiştir.



Şekil 2. Nöron Biyolojik Modeli[24]

Yapay sinir ağı mimarisi bir sinir hücresi referans alınarak incelendiğinde giriş katmanı, gizli katman veya katmanlar ve çıkış katmanından oluşmaktadır.

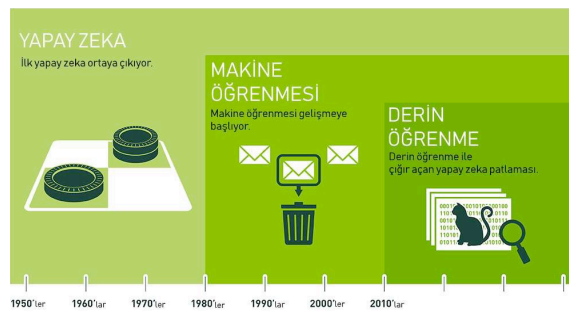
Derin öğrenme metotları danışmanlı öğrenme ve danışmansız öğrenme olarak iki ana gruba ayrılmaktadır [25]. Danışmansız öğrenmede veri setinde veri ile alakalı herhangi bir etiket mevcut değildir. Danışmalı da ise bunun aksine veri setinde gerekli etiketlemeler yapılarak veriler sisteme etkin bir şekilde verilerek öğrenilmesi sağlanmaktadır. Yapılan çalışmamızda veri setimiz öncelikle belirlediğimiz classlara uygun şekilde etiketlenip danışmalı öğrenme kullanılarak eğitim ve test işlemleri gerçekleştirilmiştir.



Şekil 3. YSA Yapısı[24]

2.1.2.1 Derin Öğrenme Mimarileri

Derin öğrenme, sisteme verilen bir veri seti ile sonuçları tahmin eden ve birden fazla katmandan oluşan makine öğrenmesi yönteminin özelleşmiş halidir. Derin öğrenme metotları danışmalı ve danışmansız öğrenme olarak iki gruba ayrılmaktadır. Derin öğrenme, makine öğrenmesinin; makine öğrenmesi ise yapay zekânın alt dalını oluşturmaktadır. Şekil 4 'de yapay zekâdan derin öğrenmeye gelinen sürecin kronolojisi görülmektedir.



Şekil 4. Yapay Zekâ Tarihsel Gelişim Süreci

2.1.2.2 Evrişimli Sinir Ağı

CNN, görüntü analizi için tasarlanmış derin bir sinir ağıdır. Geleneksel çok katmanlı algılayıcıların aksine, bir görüntüyü temel özelliklerine indirgemek için evrişimsel(convolution) ve havuzlama(pooling) adı verilen iki işlem kullanır. Klasik bir CNN yapısı

Şekil 5' de gösterilmiştir. Evrişim katmanı konvolüsyon (convolution), aktivasyon fonksiyonu (relu) ve havuzlama (pooling) basamaklarını içermektedir.

CNN çalışma prensibi bakımından yüksek boyutlu verilerden düşük boyutlu verilerin çıkarılması prensibine dayanmaktadır. Sınıflandırma katmanında düşük boyutlu öznelikler ile kategorilerin eşleşmesi sağlanır ve tam bağlı yapay sinir ağı mimarisine sahiptir [27].

R-CNN

R-CNN, R. Girshick ve ark tarafından 2013 yılında çoklu objeye sahip görsellerde kolay bir şekilde CNN çalıştıramadığı için R-CNN mimarisi önerilmiştir. R-CNN nesne tespitinde CNN tabanlı önerilen ilk model olduğundan daha sonra geliştirilen diğer modellere referans olmuştur. Fakat bazı eksiklik ve problemleri vardır. Bu problemlerden ilki, eğitimin ayrı aşamalardan oluşması, ikincisi ise her bölge teklifinde özellikler çıkarılarak kaydedilmekte ve bu işlem ayrıca DVM algoritması içinde uygulanmakta olduğundan çok fazla depolama alanına gereksinim duymakta bu da eğitimin maliyetli olmasına neden olmaktadır. Üçüncüsü ise grafik işlemcilerin kullanılmasına rağmen eğitimden sonra nesne tespiti için yaklaşık olarak bir dakikaya ihtiyaç duymasındır[27].

Fast R-CNN

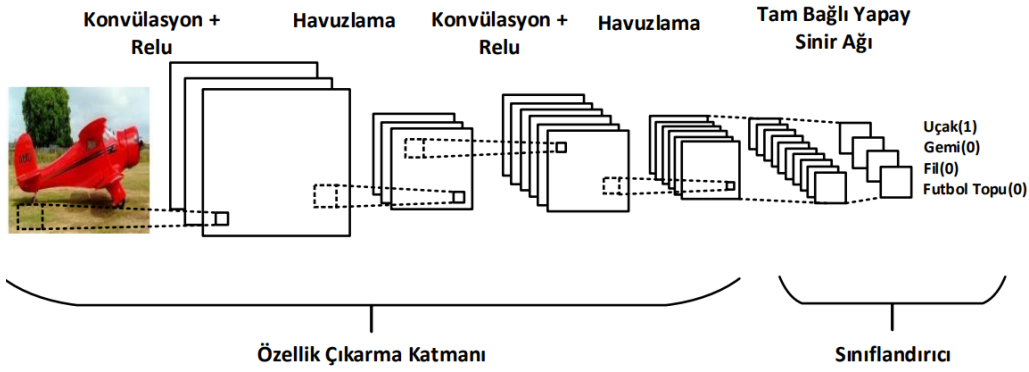
Fast R-CNN , Girshick ve ark. tarafından 2015 yılında R-CNN'in katmanlarındaki hesaplamaların her bölge teklifi için ayrı ayrı değil tüm resim için tek bir defa yapılması ile geliştirilerek sunulmuştur. Bu yolla eğitim süresi azalmış ve R-CNN'den daha iyi bir şekilde eğitilmesi sağlanmıştır [28-29].

Faster R-CNN

Faster R-CNN, Ren ve ark. tarafından Fast R-CNN algoritmasının bölge teklifleri üretme adımını geliştirilerek sunulmuştur. Faster R-CNN algoritmasının çalışma prensibi konvolüsyon katmanlarını hem nesne hem de bölge tekliflerini üretmekte ortak olarak kullanılmasına dayanmaktadır [29].

Faster R-CNN iki aşamada açıklayalım ilk aşama olarak Region Proposal Network, herhangi bir boyutta girdiyi alır ve obje skoruna göre dikdörtgen teklifi ortaya çıkarır bunu, evrişim katman tarafından oluşturulan öznelik haritası üzerinde küçük bir ağı kaydırarak yapar.

Fast R-CNN mimarisine sokulur ve bir sınıflandırıcı ile objenin sınıfı, regressor ile de bounding box'u tahmin edilir [30].

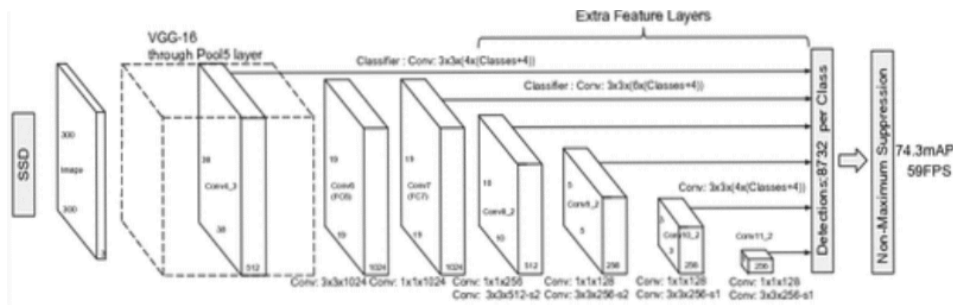


Şekil 5. Klasik Bir CNN Mimarisi[18]

Single Shot Multi Box Detector (SSD)

Gerçek zamanlı sistemler için Liu ve arkadaşları tarafından 2016 yılında öne sürülen Single Shot Multi Box Detector (SSD) modeli arka plan bilgisini kullanarak nesneyi algılamaktadır. SSD gerçek zamanlı olarak nesne algılama için tasarlanmıştır. SSD nesne algılama için iki aşamadan

oluşmaktadır; özellik haritalarını çıkarmak ve nesne algılamak için evrişimsel filtreleri kullanmak olarak. Doğruluk açısından teknolojinin başlangıcı olarak kabul edilen bu modelde düşük çözünürlükteki görüntülerin kullanılması ile daha hızlı R-CNN elde edilebilir[31-32]. Tüm süreç saniyede 7 defa çalışmaktadır. Single Shot Multi Box Detector (SSD) mimarisi Şekil 6' da gösterilmiştir[31-32].



Şekil 6. Single Shot Multi Box Detector (SSD)

You Only Look Once (YOLO)

YOLO nesne algılama ve nesne sınıflandırma için tek bir CNN ağı kullanan bir derin öğrenme algoritmasıdır. "You Only Look Once" (Sadece Bir Kez Baksın) anlamına gelmektedir. Joseph Redmon ve arkadaşları nesne algılamayı, doğrudan görüntü piksellerinden sınırlayıcı kutu koordinatlarına ve sınıf olasılıklarına kadar tek bir regresyon problemi olarak ele almışlardır. GoogLeNet modelinden esinlenmiştir. Tek bir evrişimli ağı ile aynı anda birden çok sınırlayıcı kutuyu ve bu kutular için sınıf olasılıklarını tahmin eder. Her sınıf için görüntüde tekrar tekrar işlem yapmak yerine algoritmanın adından da anlaşılacağı üzere görüntüye bir defa bakılır ve tüm sınıflar için olasılıkları ve sınırlayıcı kutu koordinatları oluşturulur.

Ağda 24 evrişim katmanı ve ardından 2 tam bağlantılı katman vardır. GoogLeNet tarafından kullanılan başlangıç modülleri yerine, Lin ve diğerlerine benzer şekilde 1×1 azaltma katmanlarını ve ardından 3×3 evrişim katmanlarını kullanmışlar, tam görüntü üzerinde eğitim gerçekleştirdiklerinden

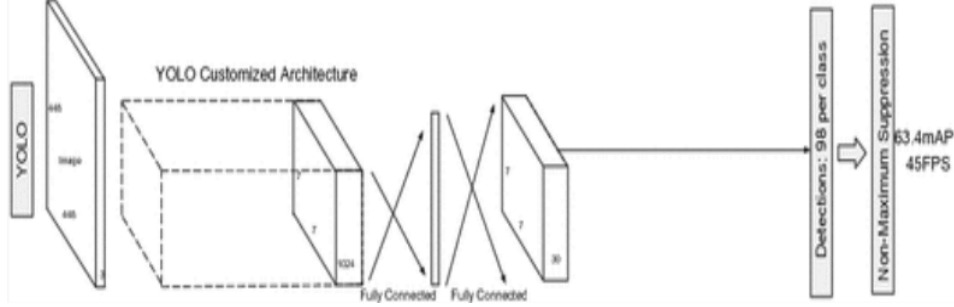
algılama performansını doğrudan optimize etmişlerdir. Bu sayede nesneyi ve nesnenin koordinat bilgisini öğrenmek oldukça hızlı bir şekilde gerçekleşmektedir. YOLO birleşik mimarisi sayesinde ileri derecede hızlıdır. YOLO 'nun ilk versiyonu saniyede 45 kare (fps) işleyebilecek yetenektedir. Fast YOLO olan versiyonu ise saniyede 155 kare işleyebilmektedir. Bu sayesinde gerçek zamanlı nesne tespitinde çok başarılı bir sistemdir[33].

YOLO mimarisi Şekil 7' de gösterilmiştir.

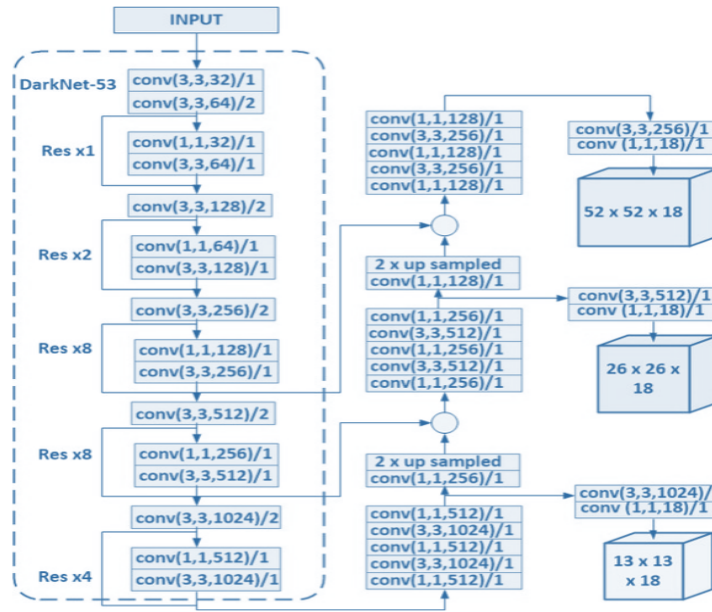
YOLO'nun YOLOv1, YOLOv2, YOLOv3 ve YOLOv4 versiyonları bulunmaktadır. Son kararlı versiyon olan YOLOv4, Alexey Bochkovskiy tarafından 2020'de yayınlandı. YOLOv4'ün Evrişimsel Sinir Ağı (CNN) doğruluğunu iyileştirdiği söylenen çok sayıda özellik var.

YOLOv4'ün, YOLOv3'ten farkı omurgasıdır. YOLOv4 CSPDarknet53 omurgasına sahiptir. YOLOv4 kafaları yolov3 ile aynıdır. Başlıklar tahmin bölümüdür ve iki türüdür. Biri Yoğun Tahmin (Tek aşamalı dedektör) diğeri ise Seyrek

Tahmin (iki aşamalı dedektör)[34]. Şekil 8’de YOLOv4’un mimarisi gösterilmiştir



Şekil 7.You Only Look Once(YOLO)



Şekil 8.YOLOv4 Mimarisi

2.2 Şüpheli Davranış

Uluslararası, ulusal ve şirketler özelinde bilgi çok büyük bir güç olup mevcut otoritelerce gizliliğine önem verilmektedir. Bilgi sızıntılarında dünyada bazen büyük bir yatırım hamlesine, bazen de ulusal veya uluslararası savaşa sürüklenme boyutuna varacak dünyada örnekleri bulunan bir olgudur. Bilgi sızıntılarındaki donanım ve yazılımsal alınan tedbirlere rağmen insan faktörü bu önlemlere rağmen sızıntıya yol açabilecek faaliyetlerde bulunmaktadır. Derin öğrenme algoritması ile cep telefonu veya küçük kamera aracılığı ile görüntü alırken bu olayı kayıt edip sisteme şüpheli hareket diye rapor edecek bir sistem üzerinde çalışılmıştır.

Davranışları inceleyen ve bunu güvenlik sistemlerine entegre eden market hırsızlığını raporlayabilecek bir uygulama VaakEye güvenlik sistemleri tarafından çalışılmış ve hızlı hareket etme, kamerayı tarama ve nesneyi üzerindeki çanta, cep vb. yerlere sıkıştırma üzerinde derin algoritma ile

çalışmışlardır. Ve bu hareketler tespit edildiğinde müşteriyi olası şüpheli olarak kayıt etmektedir. VaakEye 'in bu algoritmasının görselleri Şekil 9’ de gösterilmiştir[35].



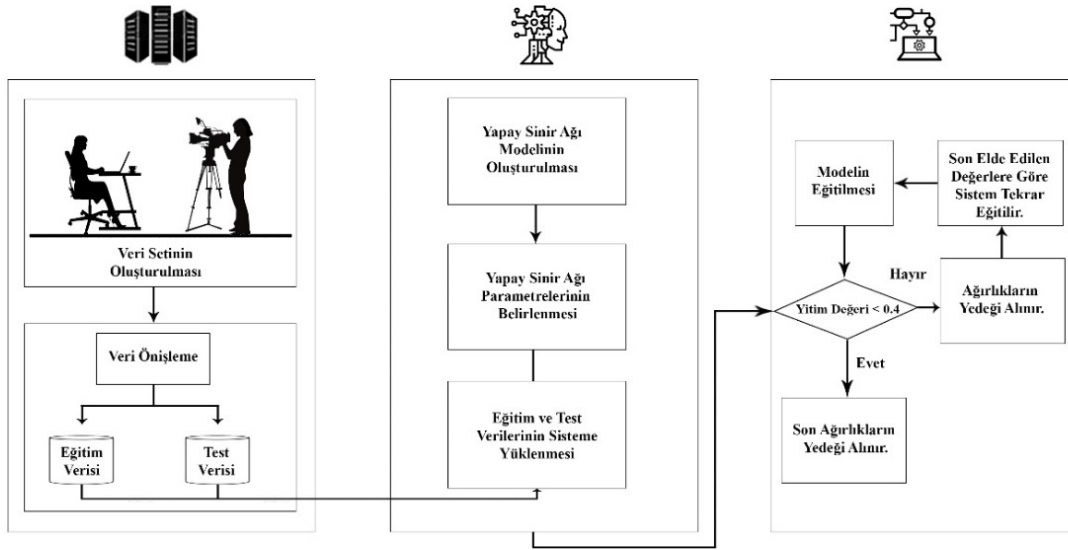
Şekil 9.VaakEye Hırsızlık Algoritma Çıktısı[35]

2.3 Geliştirilen Sistem

Gerçekleştirilen tez çalışması Şekil 10’ daki sistem mimarisinde görüldüğü gibi temel olarak 3 adımdan oluşmaktadır.

İlk olarak çalışmada kendi veri setimiz oluşturulmuştur. İç mekanlar da (Office, homeoffice, derslik, laboratuvar vb.) çekilen görüntü ve video kaydı gibi görsel kaynaklar belli ön işleme evresinden geçirilerek etiketli bir veri seti elde edilmiştir. Yeterli verinin elde edilebilmesi için Opencv ile veri çoğullama işlemi uygulanmıştır. Veri çoğullama işleminden sonra kendini tekrar eden verileri yapay sinir ağı modelinde overfitting'e neden olmaması için temizliyoruz. Ve son olarak etiketleme işlemini yapmak için python programlama (pilow kütüphanesi), makesense vb. programlar ile LabelTool program kullanıyoruz. Ve bu alanda yeni ve büyük veri seti olan Marmara

Üniversitesi Kriminal Davranış/Nesne (MÜKDN) oluşturulmuştur. Ön işleme tamamlandıktan sonra etiketli MÜKDN veri setindeki örnekler, %70 eğitim ve %30 test verisi olarak ikiye ayrılmıştır. İkinci adım olarak derin öğrenme için kullanılan evrişimsel sinir ağı YOLOv4 modeli tasarımı ve konfigürasyonu yapılmıştır. Veri seti üzerinde YOLO algoritması kullanılarak, kamera açısından bağımsız şüpheli görüntü alma eylemi, şüpheli davranış yüzeyleri tespiti gerçekleştirilmiştir. Son adımda ise sistemin eğitilerek, diğer sistemlerde kullanılmak üzere evrişimsel sinir ağının ağırlıkları elde edilmiştir.



Şekil 10. Sistem Mimarisi

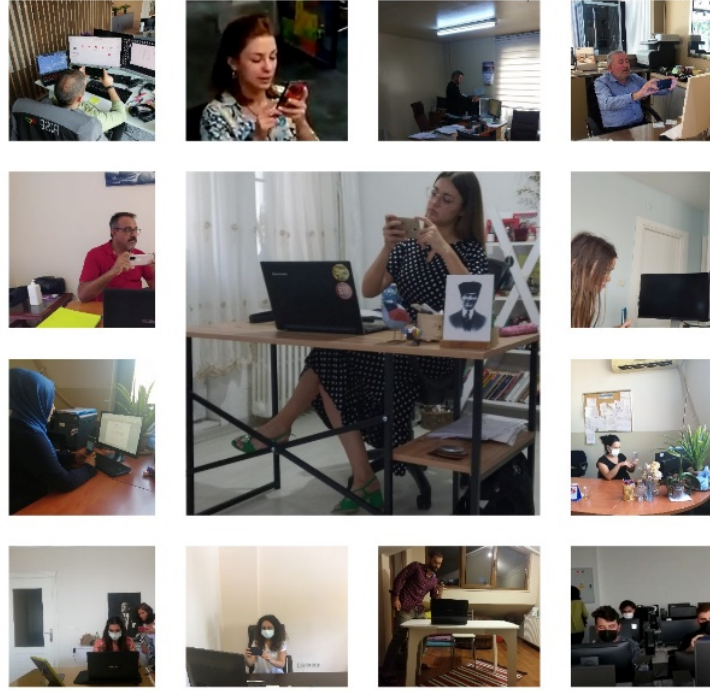
2.3.1 Kullanılan Veri Seti

Yapılan çalışmada araştırmacıların video çekimlerimle elde edilmiş MÜKDN veri seti kullanılmıştır. Veri hazırlama ve ön işleme evresinde, iç meknlarda çekilen görüntü ve video gibi görsel kaynaklar belirli ön işleme işlemlerinden geçirilerek derin öğrenmede kullanılmak üzere etiketli MÜKDN veri seti haline getirilmiştir.

Görüntüler İstanbul, Eskişehir ve Hatay'ın çeşitli semt ve bölgelerinde bulunan office, homeoffice, Marmara Üniversitesi Teknoloji Fakültesi

Bilgisayar laboratuvarlarında katılımcıların gönüllü olarak verilen görüntülerinden elde edilmiştir. Şekil 11'de veri setindeki görüntülerden örnekler gösterilmiştir.

Görüntüler temin edilirken Samsung HMX-QF30 kamerası ve Sony HDR-PJ410 kamerası, bilimüm cep telefonu kameraları kullanılmıştır. Görüntü temini sırasında gönüllülük esasına dayalı çekimler yapıldığından herhangi bir kişinin özel hayatı ihlal edilmemiştir.



Şekil 11. Office, Homeoffice'lerden Elde Edilen Veri Setinden Örnekler

2.3.1.1 Veri Çoğullama

Elde edilen görüntüleri sisteme uygun bir şekilde kullanmak için hem de YOLO küçük nesnelere iyi performans göstermediğinden YOLO algoritmasına uygun formatta kullanabilmek adına görüntülerde yeniden boyutlandırma işlemi yapılmıştır. Görüntüler 1024x768 boyutuna getirilmiştir.

Konvolüsyonel sinir ağları ile danışmanlı öğrenme çalışmalarında veri setinde birbirini tekrar etmeyen çok fazla örneğe gereksinim duyulmaktadır. Bunun nedeni örnek sayısı ile uygulama performansının doğru orantılı olarak artmasıdır. Uygulama performansını arttırmak için ham veri open cv veri artırma işlemlerinden çoğullama, aynalama, beyazlatma vb. yöntemlerini kullanılarak artırılmıştır.

2.3.1.2 Veri Temizleme

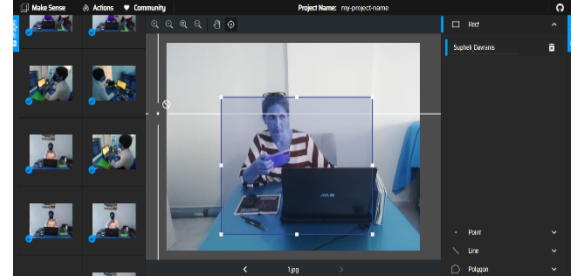
Elde edilen tüm video ve fotoğraflardan belirlenen senaryoya uygun olan görüntüler seçilip temizlenmiştir. Senaryoda karşılaşmayacağımız ve veri de overfitting'e neden olacak tekrarlar temizlenmiştir.

2.3.1.3 Veri Etiketleme

Etiketleme aşaması için veri setimize 1 den başlayarak veri setinin sonuna kadar sıralı isimlendirme işlemi uygulanmıştır.

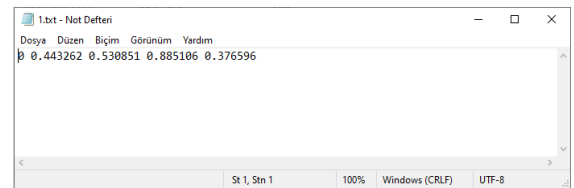
YOLO algoritmasına uygun şekilde etiketleme yapabilmek için kullanılacak alternatif programlar mevcuttur. Bunlara örnek olarak Supervise.ly, Hasty.ai, CVAT, Darwin, Heartex, Scalabel, Segments.ai, Make-Sense, LinkedAI,

LabelImg, RectLabel, LabelBox, LabelMe ve DataTurks ve makesense.ai vb. programlar verilebilir. Bu çalışma kapsamında makesense.ai ve labelBox ile tüm veri seti görüntüleri belirlenen senaryoya uygun şekilde Şekil 12' de görüldüğü gibi etiketlenmiştir.



Şekil 12. Veri Etiketleme İşlemi

Her bir görüntü için tespit edilen koordinatlar ve görüntüde kaç adet nesne olduğu bilgisi görüntü adı ile aynı isimde oluşturulan bir text dosyasına yazılarak kayıt edilmektedir. Şekil 13' te görüldüğü gibi her bir görüntü için etiketli .txt uzantılı dosya oluşturulmuştur.



Şekil 13. Etiketli Veri Örneği

2.3.1.1 Veri Setinin Normalize Edilmesi

Etiketlenen görüntülerin uzunluk(dH) ve genişliği(dW), şüpheli davranış hareketinin hangi koordinatta olduğu bilgisini veren (X₀,X₁,Y₀,Y₁) noktaları ve ŞDH etiket değeri görüntü verisi ile aynı isimdeki *.txt uzantılı(text) dosyası içerisine yazılmaktadır. Veri setindeki tüm örnekler normalizasyon işlemi uygulanarak etiketli MÜKDN veri seti elde edilmiştir. Bu işlem sistemde kullanılacak YOLO mimarisine uygun hale getirmek için yapılmıştır.

(X₀,X₁,Y₀,Y₁) koordinatları normalizasyon işlemi ile 0 ve 1 aralığına indirgenmiştir. Bu işlem ile görüntü içerisindeki etiketli ŞDH'nin, merkez noktası koordinatları (X,Y), yüksekliği(H) ve genişliği(W) bilgileri elde edilmektedir. Yapılan normalizasyon işleminin formülü Denklem 2 'de verilmiştir.

$$\begin{aligned} X &= \frac{X_1 + X_0}{2} \times \frac{1}{dW} \\ Y &= \frac{Y_1 + Y_0}{2} \times \frac{1}{dH} \\ W &= (X_1 - X_0) \times \frac{1}{dW} \\ H &= (Y_1 - Y_0) \times \frac{1}{dH} \end{aligned} \quad (2)$$

X : Etiketli verinin orta noktasının X koordinatı.

Y : Etiketli verinin orta noktasının Y koordinatı.

W : Etiketli verinin genişliği

H : Etiketli verinin yüksekliği

Normalize edilmiş ve etiketlenmiş görüntü örneği Şekil 14' de gösterilmiştir.



Şekil 14. Verinin Normalize Edilmesi ve Etiketlenmesi

2.3.1.4 Test ve Eğitim Sınıflarının Belirlenmesi

Ön işleme tamamlandıktan sonra etiketli MÜKDN veri setindeki örnekler, %70 eğitim ve %30 test verisi olarak ikiye ayrılmıştır.

2.1.2 Colabratory ve Darknet

Sistem mimarisini YOLOv4 modeline göre derin öğrenme yöntemleri ile eğitmek için darknet ağı ve colabratory ide 'si tercih edilmiştir. Bu bölümde darknet ve Colabratory kurulumu ve konfigürasyonları hakkında bilgi verilecektir.

2.1.2.1 Darknet

2.1.2.2 Colabratory

III.BULGULAR VE TARTIŞMA

Bu çalışmada bir yapay sinir ağı çatısı/framework) olan darknet ve gerçek zamanlı nesne tespit aracı olan YOLO, kriminal/şüpheli görüntüler üzerinde şüpheli görüntü tespiti yapmak(izinsiz görüntü alma) amacıyla kullanılmıştır. Veri seti olarak Marmara Üniversitesi Kriminal Davranış/Nesne (MÜKDN) veriseti oluşturulmuş ve kullanılmıştır. Veri seti YOLO formatına uygun bir şekilde makesense.ai'de etiketlenmiştir.

YOLO ile derin öğrenme yapmadan önce darknet 'in okuyacağı sinir ağı konfigürasyon dosyasında bazı değişiklikler yapılarak, sinir ağının küçük nesnelere de tespit edebilir hale gelmesi sağlanmış ve eğitim işleminin süresi için optimizasyonlar yapılmıştır.

YOLO 'nun eğitilmesi colabratory framework' ünde GPU üzerinde gerçekleştirilmiştir. Eğitim yaklaşık olarak 12 saat sürmüştür.

3.1 Kullanılan Metrik Yöntemler

Derin öğrenme sistemlerinde modelin eğitim süreci sonunda elde ettiği başarı ver performansı belli ölçütlere göre değerlendirilmektedir. Bu bölümde çalışmalarda kullanılan modelin değerlendirilmesi için kesinlik (precision), duyarlılık (recall) , F1 skor ve mAP sonuçlarına değinilmiştir

3.1.1 Kesinlik (Precision), Duyarlılık (Recall) , F1 skor ve mAP

Derin öğrenmede öğrenme başarısı olasılık modeli olan kesinlik (precision) değeri üzerinden hesaplanmaktadır. Bu skala için bazı temel kavramlar vardır. Bu kavramlardan dördü gerçekleşen öğrenmeyi test verisi üzerinde uygulandığında elde edilen tahmini sınıflandırmak için kullanılmaktadır. Bu kavramlar doğru pozitif(TP), yanlış pozitif(FP), doğru negatif(TN) ve yanlış negatif(FN) değerleridir.

Bu değerler, kullanıcı tarafından şüpheli davranış olarak etiketlenmiş verinin etiket değeri ve yapıla eğitim sonucundaki ağırlık değerlerine göre sistem tarafından yapılan tahmin sonucunun karşılaştırması ile elde edilmektedir. Şüpheli davranış olarak etiketlenmiş verinin pozitif olduğu varsayılırsa, tahmin sonucu eğer pozitif ise TP, negatif ise FN değeri oluşmaktadır. Yine aynı şekilde şüpheli davranış olarak etiketlenmiş verinin negatif olduğu

varsayılırsa, tahmin sonucu eğer pozitif ise FP, negatif ise TN değeri oluşmaktadır. Karmaşıklık matrisini Tablo 1’de görmek mümkündür.

Tablo 1. Karmaşıklık Matrisi

| | Etiketlenmiş Durum | | |
|---------------------|--------------------|---------------------|---------------------|
| | Toplam Durum | Etiket Pozitif | Etiket Negatif |
| Tahmin Edilen Durum | Tahmin Pozitif | Doğru Pozitif (TP) | Yanlış Pozitif (TP) |
| | Tahmin Negatif | Yanlış Negatif (FN) | Doğru Negatif (TN) |

Karmaşıklık matrisinin terimleri kullanılarak her bir örnek için Precision(hassasiyet) ve Recall(geri çağırma) değerleri hesaplanır. Kesinlik(Precision) değerinin formülünü Denklem 3’de görmek mümkündür.

$$P = \frac{\text{Doğru Pozitifler}}{\text{Toplam Pozitif Sonuçlar}} = \frac{TP}{TP+FP} \quad (3)$$

Duyarlılık(Recall) değerinin formülünü Denklem 4’de görmek mümkündür.

$$R = \frac{\text{Doğru Pozitifler}}{\text{Tespit edilen şüpheli davranış sayısı}} = \frac{TP}{TP+FN} \quad (4)$$

Test verisindeki her bir örnek için bir test çalıştırılır ve elde edilen sonuca göre recall(r) ve precision(p) değerleri yüzdelik olarak hesaplanır. Bu

değerlere göre bir r-p grafiği oluşur. r-p grafiği altında kalan alan Average Precision(Ortalama hassasiyet - AP) olarak hesaplanmaktadır. Average Precision formülünü Denklem 5’ de görmek mümkündür.

$$AP = \int_0^1 p(r) dr \quad (5)$$

Bu formül pratikte sonlu bir toplam olarak Denklem 6’ daki gibi kullanılır.

$$AP = \sum_{k=1}^n P(k) \cdot \Delta r(k) \quad (6)$$

n: nesne sayısı

P(k) : Her bir k değeri için precision

$\Delta r(k)$: Her bir k değeri için recall değerindeki değişim.

Her bir sınıf için AP hesabı yapıldıktan sonra tüm sınıflar için Mean Average Precision(mAP) değeri Denklem 7’ deki gibi hesaplanır.

$$mAP = \frac{\sum \text{E sınıf APi}}{\text{Toplam Sınıf Sayısı}} \quad (7)$$

Tablo 2’ de YOLOv4’ e ait F1-Skor, mAP, IoU ve tespit zamanı değerleri gösterilmiştir. Tablo 3’te YOLOv4’ün tek sınıf ile eğitim başarısı mAP değerleri iterasyon sayısına göre gösterilmiştir. Şekil

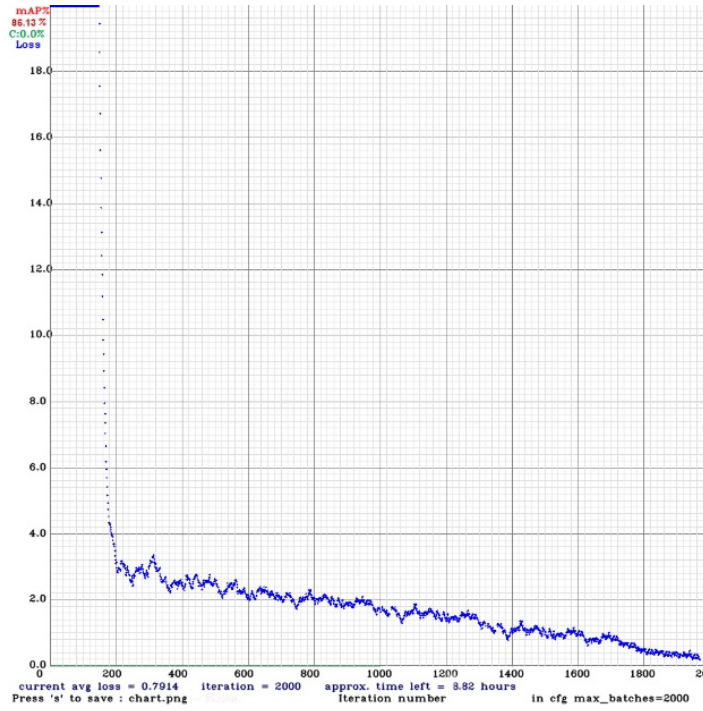
17’ de şüpheli davranış tespit modelimizin YOLOV4 algoritması kullanılarak 2000 iterasyonda tamamlanan eğitiminin çizdirilen mAP değerleri grafiği gösterilmiştir.

Tablo 3. YOLO ‘nun Tek Sınıf İle Öğretildiğinde Başarısı

| Tekrar(Iteration) | mAP |
|-------------------|-------|
| 500 | 29,68 |
| 1000 | 78,31 |
| 1500 | 82,86 |
| 2000 | 86,13 |

Tablo 2. YOLOv4’ün Metrik Hesaplama Değerleri

| Model Adı | F1 Skor | mAP | Iou | Tespit Zamanı | Güven Skoru Eşik Değeri |
|-----------|---------|------|------|---------------|-------------------------|
| YOLOv4 | 0,721 | 0,50 | 0,79 | 0,232 | 0,25 |



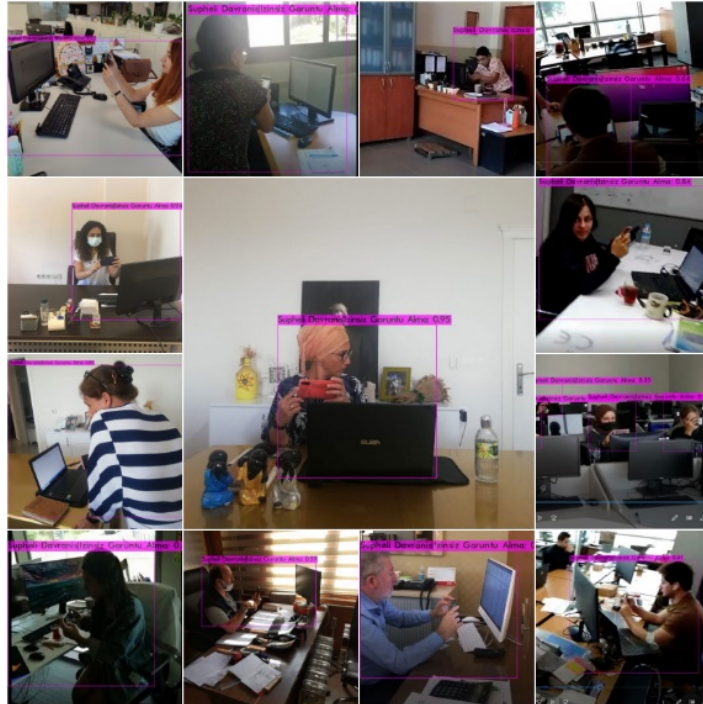
Şekil 17.Şüpheli Davranış Eğitim Grafiği

Şekil 18'de sistemin şüpheli davranış olarak doğru tahmin ettiği örnek görüntüler gösterilmiştir.

Örneklerde şüphelilerin gizli bir şekilde görüntü alma eylemi görülmektedir.

Şekil 19'da modelin başarısız olduğu tespit edemediği şüpheli görüntü örnekleri

gösterilmiştir. Görseller incelendiğinde şüpheliler cep telefonlarını iki avuçlarının içine sakladıklarında veya iki ellerini havaya kaldırırken başka bir cisim tuttuklarında model izinsiz görüntü alınmasını tespit edememiştir.



Şekil 19.Şüpheli Davranış | İzinsiz Görüntü Alınması Tespit Edilmiş Görüntüler



Şekil 19.Şüpheli Davranış | İzinsiz Görüntü Alınması Tespit Edilememiş Görüntüler

3.2 Sonuç ve Değerlendirmeler

Bu çalışmada gerçek zamanlı bir nesne tespit aracı olan YOLO Kriminal/şüpheli davranış görüntülerinde şüpheli davranış tespiti amacıyla kullanılmıştır ve başarılı sonuçlar elde edilmiştir. YOLO ile yapılan çalışmalar çoğunlukla günlük hayatta gözle görülebilen nesnelere makineye öğretilip tespit etmesi amacıyla kullanılmaktadır. Bu alandaki yeni veri seti olan Marmara Üniversitesi Kriminal Davranış/Nesne (MÜKDN) veri seti üzerinde yapılan bu çalışma sayesinde YOLO'nun davranış tespitinde kullanılabilirliği desteklenmiştir. Bir nesne tespit aracı olan YOLO ile şüpheli davranış tespitinde %70'in üzerinde başarı elde edilmiştir.

Derin öğrenmede başarıyı etkileyen en önemli faktörlerden biri elde edilmesi oldukça maliyetli olan eğitim verisinin çok olmasıdır. Bu çalışmada veri seti %70 eğitim - %30 test şeklinde ayarlanarak eğitim gerçekleştirildiğinde, mAP değeri 0,5'in altında çıkmıştır. %86,66'in üzerindeki başarıya ulaşılmıştır. Daha büyük bir veri seti ile başarının daha da yüksek çıkması beklenmektedir.

GPU gücünün makine öğrenmesi için fazlasıyla yeterli olduğu günümüzde en önemli ihtiyaç veridir. Makine ne kadar çok veri ile eğitirse, insan davranışının makinelerce taklit edilebilmesi o kadar iyi sağlanabilir. Günümüzde teknoloji kullanımının (sosyal medya vb.) artması sayesinde çok fazla veri biriktirmektedir. Etik kurallarına da uyararak, kişisel veri olmayacak verilerden veya kişisel verileri barındırmayacak şekilde toplanacak verilerle veri setleri oluşturulması ve bu verilerin gelecekte

yapılacak bilimsel çalışmalara da fayda sağlaması için paylaşılması önemlidir.

Bu çalışmada kullanılan Marmara Üniversitesi Kriminal Davranış/Nesne (MÜKDN) veri seti oluşturulmuş olup, YOLO formatında etiket yapısı kazandırılmıştır, bu veriler paylaşılacaktır.

Bu çalışmada kullanılan *.cfg dosyasındaki değerler optimum değerler değildir. Ancak mevcut hali ile şüpheli davranış tespiti yapılabilmektedir. *.cfg dosyasında yapılabilecek başka ayarlar ve YOLO katman sayısının artırılması gibi değişiklikler ile YOLO'nun başarısını artırabilir.

YOLO tek sınıf ile eğitilmiş olup, sınıf sayısının eğitimdeki başarıya etki ettiği ve tek sınıflı eğitimde yüksek başarıya daha hızlı ulaşıldığı görülmüştür.

Bu çalışmada sadece izinsiz görüntü alma şüpheli davranış tespiti yapılmıştır. YOLO ile başka davranış türleri (hırsızlık, şiddet vb.) üzerinde de çalışılabilir ve farklı başarı oranlarına ulaşılabilir.

REFERANSLAR

- [1] Radu, R. G. (2012). The monopoly of violence in the cyber space: Challenges of cyber security. In Power in the 21st Century (pp. 137-150). Springer, Berlin, Heidelberg.
- [2] Bıçakcı, S. (2014). NATO'nun gelişen tehdit algısı: 21. yüzyılda siber güvenlik. Uluslararası İlişkiler Dergisi, 10(40), 100-130.
- [3] Lecun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition.

- Proceedings of the IEEE, 86(11), 2278–2324. <https://doi.org/10.1109/5.726791>
- [4] Maghdid, H., Asaad, A. T., Ghafoor, K. Z. G., Sadiq, A. S., Mirjalili, S., & Khan, M. K. K. Diagnosing COVID-19 pneumonia from x-ray and CT images using deep learning and transfer learning algorithms, 26. <https://doi.org/10.1117/12.2588672>
- [5] Hassan, M., Ali, S., Alquhayz, H., & Safdar, K. (2020). Developing intelligent medical image modality classification system using deep transfer learning and LDA. *Scientific Reports*, 10(1), 12868. <https://doi.org/10.1038/s41598-020-69813-2>
- [6] Jaouedi, N., Boujnah, N., & Bouhlel, M. S. (2020). A new hybrid deep learning model for human action recognition. *Journal of King Saud University - Computer and Information Sciences*, 32(4), 447–453. <https://doi.org/10.1016/j.jksuci.2019.09.004>
- [7] T.S., A., & Guddeti, R. M. R. (2020). Affective database for e-learning and classroom environments using Indian students' faces, hand gestures and body postures. *Future Generation Computer Systems*, 108, 334–348. <https://doi.org/10.1016/j.future.2020.02.075>
- [8] Jammalamadaka, N., Zisserman, A., & C.V., J. (2017). Human pose search using deep networks. *Image and Vision Computing*, 59, 31–43. <https://doi.org/10.1016/j.imavis.2016.12.002>
- [9] Luvizon, D. C., Tabia, H., & Picard, D. (2019). Human pose regression by combining indirect part detection and contextual information. *Computers & Graphics*, 85, 15–22. <https://doi.org/10.1016/j.cag.2019.09.002>
- [10] Moez Baccouche, Franck Mamalet, Christian Wolf, Christophe Garcia, and Atilla Baskurt. LNCS 7065 - Sequential Deep Learning for Human Action Recognition.
- [11] Laptev, I., Marszalek, M., Schmid, C., & Rozenfeld, B. Learning realistic human actions from movies, 1–8. <https://doi.org/10.1109/CVPR.2008.4587756>
- [12] Dollar, P., Rabaud, V., Cottrell, G., & Belongie, S. Behavior Recognition via Sparse Spatio-Temporal Features, 65–72. <https://doi.org/10.1109/VSPETS.2005.1570899>
- [13] Ming-Yu Chen and Alexander Hauptmann. MoSIFT: Recognizing Human Actions in Surveillance Videos.
- [14] Liu, J., & Shah, M. Learning human actions via information maximization, 1–8. <https://doi.org/10.1109/CVPR.2008.4587723>
- [15] Daş, R., Polat, B., & Tuna, G. (2019). Derin Öğrenme ile Resim ve Videolarda Nesnelerin Tanınması ve Takibi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 571–581. <https://doi.org/10.35234/fumbd.608778>
- [16] Yılmaz, O., Aydın, H., & Çetinkaya, A. (2020). Faster R-CNN Evrimsel Sinir Ağı Üzerinde Geliştirilen Modelin Derin Öğrenme Yöntemleri ile Doğruluk Tahmini ve Analizi: Nesne Tespiti Uygulaması. *Avrupa Bilim ve Teknoloji Dergisi*, (20), 783–795.
- [17] Wikipedia (Ed.). (2021). Yapay zekâ. (2021, September 13). https://tr.wikipedia.org/w/index.php?title=Yapay_zekâ&oldid=25937799
- [18] Bonner, A. (2007). The art and logic of Ramon Lull: A user's guide. *Studien und Texte zur Geistesgeschichte des Mittelalters: Bd. 95*. Brill.
- [19] Turing, A. M. (2009). Computing machinery and intelligence. In *Parsing the turing test* (pp. 23-65). Springer, Dordrecht..
- [20] John McCarthy, Marvin L. Minsky, Nathaniel Rochester, & Claude E. Shannon. (1955). A Proposal for the Dartmouth Summer Research Project on Artificial Intelligence (27 Number 4(2006)). *AI Magazine Volume 27 Number 4 (2006) (© AAAI)*.
- [21] McCulloch, W. (1943). A logical calculus of the ideas immanent in nervous activity, 5.
- [22] Rosenblatt, F. (1957). *The perceptron, a perceiving and recognizing automaton Project Para*. Cornell Aeronautical Laboratory.
- [23] Christopher, M. B. (2016). *Pattern Recognition And Machine Learning*: Springer-Verlag New York
- [24] Karpathy, A. (2018). Stanford university cs231n: Convolutional neural networks for visual recognition. URL: <http://cs231n.stanford.edu/syllabus.html>.

- [25] Yapay Sinir Ağları ~ Papatya Bilim Yayınevi. (2020, October 24). <http://www.papatyabilim.com.tr/yapaySinirAglari.htm>
- [26] Fırıldak, K., & Talu, M. F. (2019). Evrimsel Sinir Ağlarında Kullanılan Transfer Öğrenme Yaklaşımlarının İncelenmesi. *Computer Science*, 4(2), 88-95.
- [27] Girshick, R., Donahue, J., Darrell, T., Malik, J., & Merzan, E. (2014). R-CNN for Object Detection. In *IEEE Conference*.
- [28] Girshick, R., Donahue, J., Darrell, T., & Malik, J. (Eds.). Rich feature hierarchies for accurate object detection and semantic segmentation.
- [29] Gkioxari, G., Girshick, R., & Malik, J. (Eds.). Contextual Action Recognition with R*CNN.
- [30] Ren, S., He, K., Girshick, R., & Sun, J. (2017). Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 39(6), 1137–1149. <https://doi.org/10.1109/TPAMI.2016.2577031>
- [31] Leibe, B., Matas, J., Sebe, N., & Welling, M. (Eds.) (2016a). *Computer Vision – ECCV 2016*. : Vol. 9905. Springer International Publishing.
- [32] Leibe, B., Matas, J., Sebe, N., & Welling, M. (2016b). *Computer Vision – ECCV 2016*, 9905. <https://doi.org/10.1007/978-3-319-46448-0>
- [33] Joseph Redmon, Santosh Divvala, Ross Girshick, & Ali Farhadi. You Only Look Once: Unified, Real-Time Object Detection.
- [34] Papers with Code - YOLOv4 Explained. (2021, September 13). <https://paperswithcode.com/method/yolo4>
- [35] bigumigu.com -Hırsızlığı % 81 Doğrulukla Önceden Tespit Edebilen Algoritma.(2021, Sebtember17).<https://bigumigu.com/haber/vaakeye-hirsizligi-81-dogrulukla-onceden-tespit-edebilen-algoritma/>
- [36] pjreddie.com, "Darknet: Open Source Neural Networks in C", (2021, September 13). <https://pjreddie.com/darknet/>
- [37] github.com," YOLOv4 / Scaled-YOLOv4 / YOLO - Neural Networks for Object Detection (Windows and Linux version of Darknet) ",(2021, September 13). <https://github.com/AlexeyAB/Darknet>