



KONYA TEKNİK ÜNİVERSİTESİ

Bilgisayar Mühendisliği

Bilişim Teknolojileri Uygulaması Dersi – Nesne Tespiti ve Takibi Uygulaması

Kamusal Alanlarda Terk Edilen Valizlerin Tespiti ve Takibi

Doç.Dr. Ömer Kaan Baykan

181220027

Şahin Emre ASLAN

İçindekiler

1	Şekiller Listesi	3
2	Tablolar Listesi	4
3	Özet	5
3.1	Projenin Amacı ve Önemi	5
3.2	Kaynak Araştırması	5
3.2.1	“Detecting Abandoned Luggage Items in a Public Space” (Kevin Smith, 2006)	5
3.2.2	“Localized Detection of Abandoned Luggage” (Chang Jing-Ying, 2010)	5
3.3	Dönem İçi Yapılan Çalışmaların Özeti	5
4	Proje Mevcut Durum Değerlendirmesi	6
4.1	Projede Ele Alınmış Durumlar	6
4.1.1	Durum 1 : İnsanın Bagajıyla Eşleştirilmesi.....	6
4.1.2	Durum 2 : İnsanın Bagajını Terk Etmesi	6
4.1.3	Durum 3 : Bagaj Sahibinin Yanına Başkalarının Gelmesi Durumu	6
4.2	Projenin Akış Şeması.....	7
5	Yazılım Mimarisi	8
5.1	CNN Tabanlı Derin Öğrenme Modelleri.....	8
5.2	Nesne Tespiti Model Seçimi.....	8
5.2.1	Nesne Tespiti Modelin Seçimi	9
5.3	Veri Setinin Hazırlanması ve Kaynaklar	9
5.3.1	Verilerin Etiketlenmesi	9
5.3.2	Verilerin Sayısı Ve Etiketler	10
5.3.3	Verilerin Toplandığı Kaynaklar	10
5.4	Nesne Tespiti Modelin Eğitilmesi	11
5.4.1	Google Colab.....	11
5.4.2	Modelin Eğitimi İçin HiperParametre Optimizasyonu	11
5.4.3	Eğitim Sonuçlarının İncelenmesi	11
5.5	Nesne Takibi.....	12
5.5.1	Derin Arama Takip Algoritması(Deep Sort).....	12
5.6	İnsan ve Eşyaların Eşleştirilmesi Durumu.....	13
6	Sonuçlar ve İnceleme	14
7	Kaynaklar	15
7.1	Kaynak Araştırması	15
7.2	Makale İçi Atıfta Bulunanlar	15
7.3	Veri Setlerinin Kaynakları.....	15

1 Şekiller Listesi

Şekil 1 Projenin Genel Akışı	7
Şekil 2. CNN Katmanları	8
Şekil 3 Temel Yolo Mimarisi	9
Şekil 4 YoloV5 Alt Modellerin Karşılaştırılması	9
Şekil 5 Pascal Voc Formatı Örneği	10
Şekil 6 Veri Setlerinin Kaynakları	10
Şekil 7 Model Performans Metrikleri.....	11
Şekil 8 Nesne Eşleştirme Fonksiyonunun Akış Şeması.....	13

2 Tablolar Listesi

Tablo 1 Veri Sayısı.....	10
Tablo 2 Sınıf - Örneklem Sayısı.....	10
Tablo 3 Eğitim Donanımı.....	11
Tablo 4 Hiper Parametreler	11

3 Özet

Projenin amacı, önemi, kaynak araştırılması gibi başlıklara yer verilmiştir.

3.1 Projenin Amacı ve Önemi

Tren istasyonları, otogarlar, şehir meydanları gibi yerlerde insanlar valizlerini terk edebiliyorlar. Bu durum şehirlerin güvenliği için bir tehdit oluşturabilir. Üstelik valizlerin kamera görüntüsünden insanlar tarafından takip edilmesi ise maliyetli bir süreçtir. Bu çalışma İP kameraların bulunduğu şehir meydanlarında, otogarlarda bilgisayarlı gözü teknolojilerini kullanarak takibi ve tespitini yapmayı önermektedir.

3.2 Kaynak Araştırması

3.2.1 “Detecting Abandoned Luggage Items in a Public Space” (Kevin Smith, 2006)

Bu çalışma eski bir çalışma olmasına rağmen yaklaşım benzerdir. Fakat bu çalışmada yalnızca insanın valizi veya çantayı bırakma durumu ele alınmış. İnsan çantayı bıraktıktan sonra etrafına iki daire çizilir. Bu çizilen dairelerin en dışta olanın dışına geçildikten sonra bir süre tutulur. Bu süre 30 saniyeyi geçtiğinde alarm başlatılarak tehdit unsuru olduğu kabul edilir.

3.2.2 “Localized Detection of Abandoned Luggage” (Chang Jing-Ying, 2010)

Geçmiş yıllarda yetersiz kalan donanım gücü nedeniyle baş omuz maskesi kullanılarak insan tespiti yapılmaya çalışılmış. Bu çalışmada valizlerin birden fazla terk edilmesi durumu için yetersiz kalmış. Fakat bu durumlar şuanda çalışmada ele alınmıştır.

3.3 Dönem İçi Yapılan Çalışmaların Özeti

1. Proje konusun belirlenmesi, literatür araştırması
2. Veri Setinin Oluşturulması, canlı yayınlardan uygun şartlar oluştuğunda verilerin toplanması
3. Gerçek zamanlı çıkarım yapabilecek hıza sahip uygun modelin seçimi
4. Model İçin Hiper Parametre Seçimi
5. Modelin Performans Değerlendirmesinin yapılması
6. Takip Algoritmalarının incelenerek çoklu takip yapabilen algoritmanın seçimi
7. Tespit modeliyle takip algoritmasının birleştirilmesi
8. Tespit edilen insan ve bagaj nesnesi arasında ki ilişkiyi kuran fonksiyonun yazılması

4 Proje Mevcut Durum Değerlendirmesi

Çalışmada yalnızca bilgisayarlı görü kısmı ele alınmıştır. İlerleyen aşamalarda geliştirilen sistem uygun uç cihaz seçilip, Nesnelerin Yapay Zekası(AIoT) projesine çevrilebilir. Projede ele alınmış durumlara [4.1'inci](#) başlıkta yer verilmiştir.

4.1 Projede Ele Alınmış Durumlar

4.1.1 Durum 1 : İnsanın Bagajıyla Eşleştirilmesi



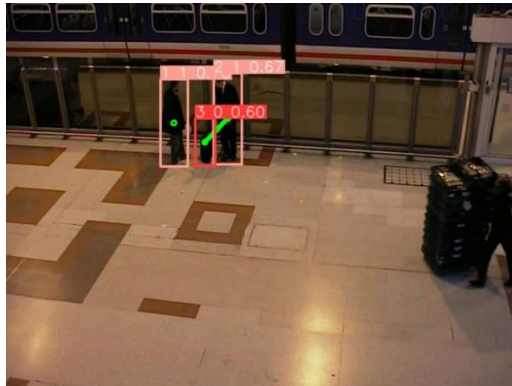
İnsan görüntüye girdiği andan itibaren valiziyle yeşil çizgiyle eşleşir, bu risk olmadığını ifade eder.

4.1.2 Durum 2 : İnsanın Bagajını Terk Etmesi



Valizinden ayrılmaya başladığı süreçte kırmızı çizgiyle işaretlenir, bu riskli durumu ifade eder.

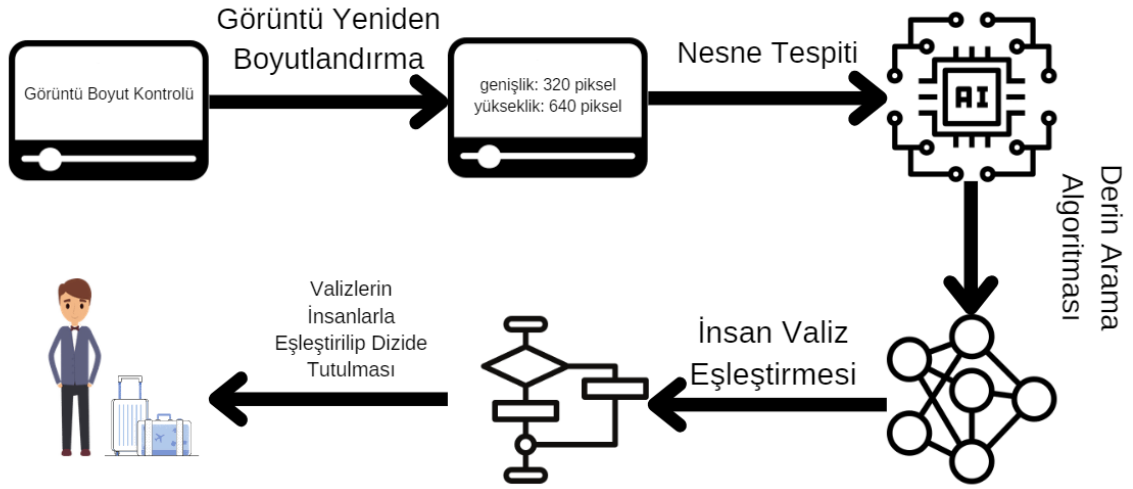
4.1.3 Durum 3 : Bagaj Sahibinin Yanına Başkalarının Gelmesi Durumu



Valiz yalnızca onu kamera görüntüsünün içine getiren kişiyle eşlenir. Böylelikle valizin kalabalık bir ortamda terk edildiği durumda başka insanlarla eşleştirilmemesi sağlanmış olur.

4.2 Projenin Akış Şeması

Projenin akış şemasına şekil 1’de yer verilmiştir.

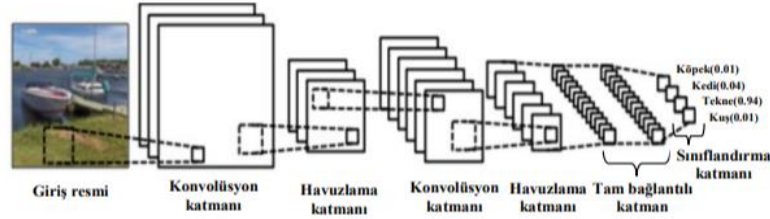


Şekil 1 Projenin Genel Akışı

5 Yazılım Mimarisi

5.1 CNN Tabanlı Derin Öğrenme Modelleri

CNN, görüntüyü çeşitli katmanlardan geçirerek özelliklerini tespit eder. Verilen görüntüleri birbirinden ayırt etmek (öğrenmek) için görüntü üzerindeki nesnelerin benzersiz özniteliklerini kullanır. CNN'ler art arda yerleştirilmiş birden fazla eğitilebilir bölümlerden oluşur. Ardından eğitici bir sınıflandırıcı ile devam edilir.



Şekil 2. CNN Katmanları

CNN'lerde giriş verileri alındıktan sonra Şekil 2'deki katmanlarda işlemler yapılarak eğitim süreci gerçekleştirilir. En sonunda doğru sonuç ile karşılaştırma yapmak için bir final çıktısı verir. Üretilen sonuç ile istenen sonucun farkı kadar bir hata oluşur. Bu hatanın bütün ağırlıklara aktarılması için geriye yayılım algoritması kullanılır. Her bir iterasyonda ağırlıkların güncellenmesi yapılarak hatanın azaltılması sağlanır.

[YOLO\[1\]](#), tek bir sinir ağı ile problemi uçtan uca optimize ederek doğrudan algılama performansına göre nesne tespiti yapmaktadır. Nesne tespitinde hız önemli bir kavramdır. YOLO'yu ön plana çıkaran özelliği de gerçek zamanlı nesne tespiti konusunda ki başarısıdır.

5.2 Nesne Tespiti Model Seçimi

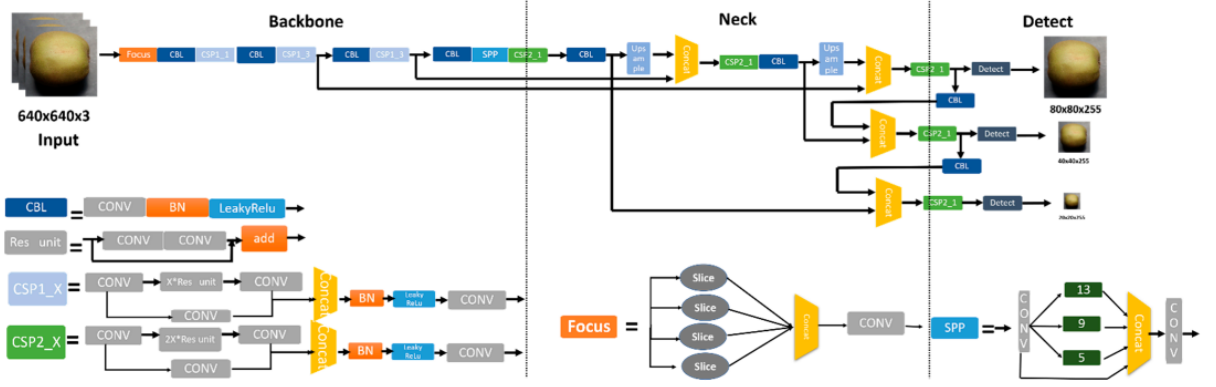
YOLO esas olarak üç ana bileşenden oluşur:

Omurga: Görüntü özelliklerini toplayan ve oluşturan evrişimli bir sinir ağı;

Boyun: Görüntü özelliklerini karşılaştıran ve birleştiren bir dizi ağ katmanı;

Baş: Görüntü özelliklerini tahmin edebilir, sınırlayıcı kutular oluşturabilir ve kategorileri tahmin edebilir. Güven, belirli koşullar altında sınıflandırmanın doğruluğunu gösterir.

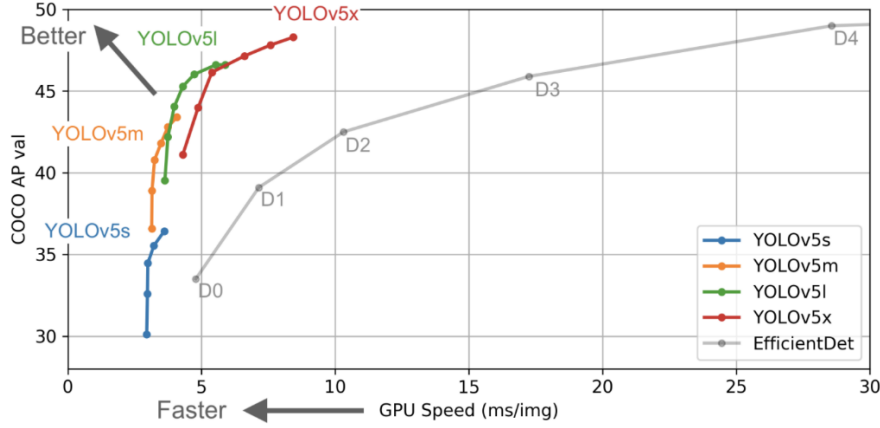
[YOLOv5\[2\]](#), YOLO serisinin yukarıda bahsedilen üç ana bileşenini kullanır. Şekil 3'te tek aşamalı nesne tespit dedektörünün genel şeması gösterilmiştir.



Şekil 3 Temel Yolo Mimarisi

Darknet kütüphanesi üzerine geliştirilen YOLOv4 modelinin Pytorch kütüphanesi üzerine yapılan yeni bir geliştirmesi ile YOLOv5 modeli üretilmiştir. YOLOv5 modelinin diğer YOLO sürümlerinden en önemli farkı geliştirildiği kütüphane olmuştur.

5.2.1 Nesne Tespiti Modelin Seçimi



Şekil 4 YoloV5 Alt Modellerin Karşılaştırılması

Nesne tespiti model seçimi projenin gereksinimlerine bağlı olarak seçilmiştir. Çalışmada tespit edilmesi beklenen nesnelerin çok küçük olmadığı göz önünde bulundurularak YOLOv5S modeli tercih edilmiştir. Bu tercih Şekil 4'teki hız ve doğruluk göz önünde bulundurularak yapılmıştır. Ayrıca Nvidia Jetson Nano[3] gibi uç cihazlarda YOLOv5S gerçek zamanlı çıkarım yapmak için yeterli donanımsal güce sahiptir.

5.3 Veri Setinin Hazırlanması ve Kaynaklar

Veriler tek tek internet üzerinden toplanmıştır. Probleme en uygun veri setini bulmak için böyle bir yöntemle başvurulmuştur.

5.3.1 Verilerin Etiketlenmesi

LabelIMG programı yardımıyla veriler PASCAL VOC[4] formatıyla .xml uzantılı olarak etiketlenmiştir. Örnek bir formata şekil 5'te yer verilmiştir.

```

<annotation>
  <folder>Resize_Data</folder>
  <filename>rotated2.jpg</filename>
  <path>C:\Users\Emre\Desktop\Bitirme Projem\VeriSeti\Resize_Data\Resize_Data\rotated2.jpg</path>
  <source>
    <database>Unknown</database>
  </source>
  <size>
    <width>224</width>
    <height>224</height>
    <depth>3</depth>
  </size>
  <segmented>0</segmented>
  <object>
    <name>hedef</name>
    <pose>Unspecified</pose>
    <truncated>1</truncated>
    <difficult>0</difficult>
    <bndbox>
      <xmin>40</xmin>
      <ymin>1</ymin>
      <xmax>195</xmax>
      <ymax>196</ymax>
    </bndbox>
  </object>
</annotation>

```

Şekil 5 Pascal Voc Formatı Örneği

5.3.2 Verilerin Sayısı Ve Etiketler

Tablo 1’de veri sayısı ve etiketlere yer verilmiştir.

Giriş Boyutları	Görsel Sayısı	Etiketler
(640,640,3)	2342 Adet	0, 1, 2, 3

Tablo 1 Veri Sayısı

Tablo 2’de yer alan 0 numaralı etiket valizleri, 1 numaralı etiket insanı, 2 numaralı etiket el çantalarını, 3 numaralı etiket ise sırt çantalarını temsil etmektedir.

Sınıf Kimliği:	0(Valiz)	1(İnsan)	2(El Çantası)	3(Sırt Çantası)
Örneklem Sayısı:	700	1480	180	72

Tablo 2 Sınıf - Örneklem Sayısı

5.3.3 Verilerin Toplandığı Kaynaklar

Veriler tren istasyonları, otobüs durakları, toplanma alanları gibi bir çok alan canlı izlenerek uygun şartlar gerçekleştiğinde toplanmıştır. Kaynaklara Şekil 6’da yer verilmiştir.

Herkese Açık Canlı Yayın Paylaşımları

La Plata, Missouri, USA | Virtual Railfan LIVE

Grosmont Station Webcam

Flagstaff, Arizona, USA | Virtual Railfan LIVE

Oluşturulmuş Veri Setlerinden Ayırdıklarım

Github - kevinlin311twABODA Repo

mircomanninoAbandoned-luggage-detection Repo

Pets - 2006 Dataset

Şekil 6 Veri Setlerinin Kaynakları

5.4 Nesne Tespiti Modelin Eğitilmesi

5.4.1 Google Colab

Colaboratory[5] öğrenciler, veri bilimciler, yazılım geliştiriciler için, tarayıcıda Pythonu yazmanızı ve çalıştırmanızı sağlar. Hiç yapılandırma gerektirmez. GPU'lara belli bir süre limiti ve bellek alanıyla ücretsiz erişim sağlar. Kolay paylaşım özelliği sayesinde toplu ve takım çalışması için idealdir. Eğitimin gerçekleştiği donanıma ait özellikler tablo 3'te yer almaktadır.

Donanım Bileşenleri	GPU	CPU	Ram
Donanım Özellikleri	Tesla T100 - CUDA	Intel(R) Xeon(R) CPU @ 2.20GHz	12GB

Tablo 3 Eğitim Donanımı

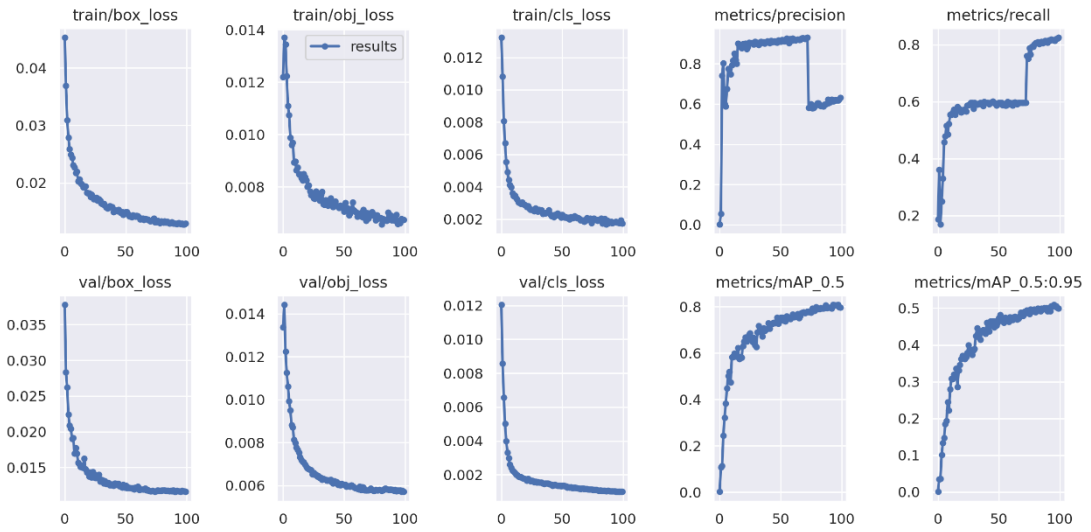
5.4.2 Modelin Eğitimi İçin HiperParametre Optimizasyonu

Learning Rate	Döngü Sayısı	Parti Boyutu
lr0: 0.01(SGD=1E-2, Adam=1E-3)	100	2

Tablo 4 Hiper Parametreler

Seçilen hiper parametrelere tablo 4'te yer verilmiştir. Hiper parametrelerin seçimi sezgisel olarak gerçekleştirilmiştir. En temel sebebi önceden eğitilmiş model kullanılmasıdır. COCO modeli içinde bir çok nesneye ait sınıfı bulundurmaktadır. Eğitim sürecinde ki ilk ağırlıklar rastgele olmak yerine tespit edilmesi beklenen 4 sınıfa daha yakın olacak şekilde başlatılır.

5.4.3 Eğitim Sonuçlarının İncelenmesi



Şekil 7 Model Performans Metrikleri

Eğitim sürecinde validasyon ve eğitim kümesi incelenir. Sonuçlar incelenirken daha doğru bir değerlendirme için validasyon kümesinin kayıp değerleri incelenir. Şekil 7’de ki tablo incelendiğinde görülecektir ki mAP değeri 0.95’tir.

TP → Doğru Sınıfın Doğru Tespit Edilmesi

FP → Doğru Sınıfın Yanlış Tespit Edilmesi

FN → Yanlış Sınıfın Doğru Olarak Tespit Edilmesi

TN → Yanlış Sınıfın Doğru Olarak Tespit Edilmesi

Hatırlama / Recall, doğru pozitif tahminlerin toplam pozitif tahmin sayısına oranıdır.

Doğru Sınıfın Doğru Tespit Edilmesi

Doğru Sınıfın Doğru T. Edilmesi + Doğru Sınıfın Yanlış Tespit Edilmesi

Kesinlik / Precision, doğru pozitif tahminlerin veri kümesindeki toplam pozitif örnek sayısına oranıdır.

Doğru Sınıfın Doğru Tespit Edilmesi

Doğru Sınıfın Doğru T. Edilmesi + Yanlış Sınıfın Doğru T. Edilmesi

Box_loss, c10U() değerlendirmesine bağlı olarak tahmin edilen sınırlayıcı kutuyla gerçek sınırlayıcı kutu arasında ki hatayı ifade eder.

Cls_loss, yolo görüntü içerisinde bir çok sınırlayıcı kutu oluşturur. Bu kutuların sınıflarının doğruluğu için bu parametre kullanılır.

5.5 Nesne Takibi

Nesneleri izlemek amacıyla kullanılan birçok farklı algoritma[6] vardır. Bazı algoritmalar yalnızca tek nesne takibini gerçekleştirirken, bazıları birden fazla nesnenin izlenmesi için kullanılmaktadır. Nesne izleme için kullanılan algoritmalar geniş gruplara veya kategorilere ayrılabilir. Bu çalışmada çoklu nesne takibi yapılacağından DeepSort takip algoritması tercih edilmiştir.

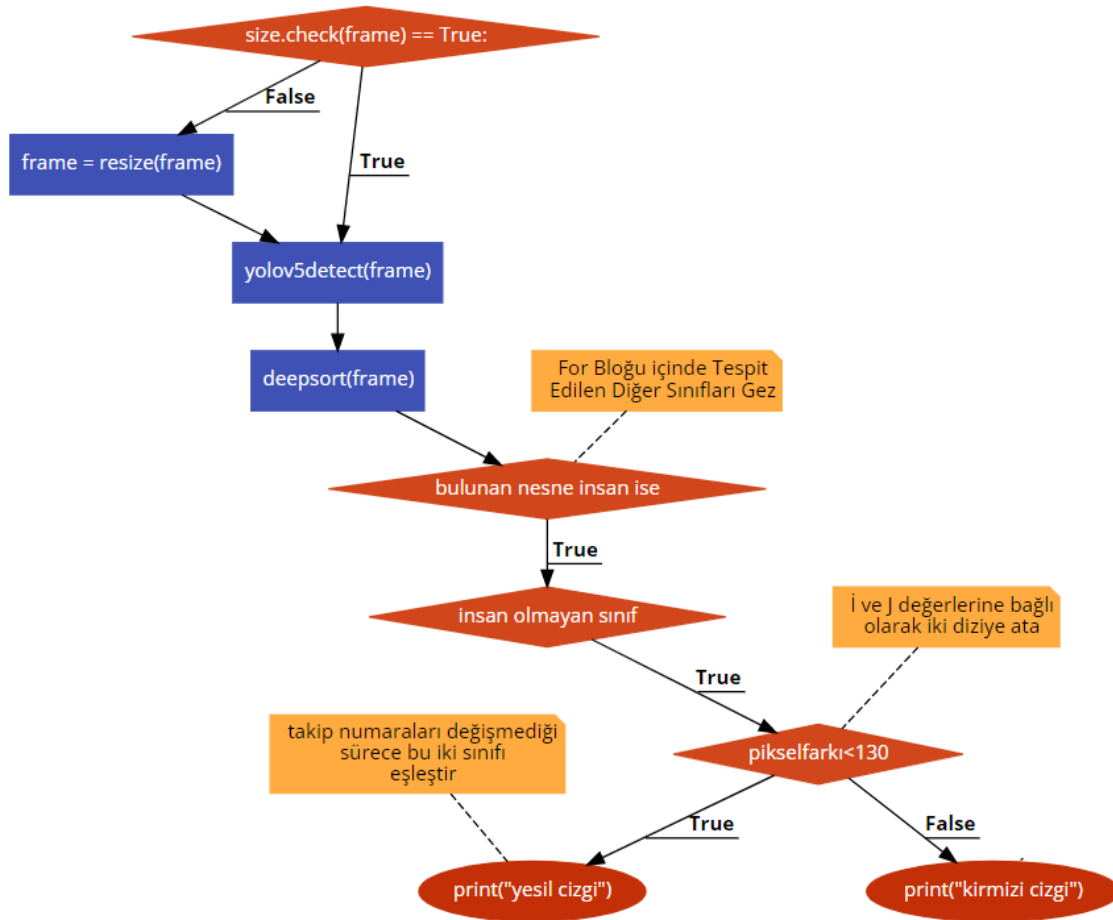
5.5.1 Derin Arama Takip Algoritması(Deep Sort)

Nesne takibi algoritmaları bu gibi çalışmalar için performansı olumlu etkiler. Derin Arama algoritmasında[7] kullanılmak üzere nesne sınıflandırılması amacıyla bir evrimsel sinir ağı tasarlanır. Yüksek bir başarı elde edilinceye kadar evrimsel sinir ağı eğitilir. CNN sayesinde sınıflandırılacak nesnenin en ayırt edici, nesneyi diğer nesnelerden ayıran en belirgin özelliği tespit edilmeye çalışılır. CNN yapısının son katmanında nesnelerin sınıflandırma işlemi yapılır. Bu sınıflandırma işlemi ise nesneyi temsil eden bir vektöre göre yapılır. Deep SORT algoritmasında ise tespit edilen her nesne sinir ağından geçirilerek bir vektör elde edilir ve bu vektörler kullanılarak iki nesnenin ilişkilendirilmesi yapılır.

5.6 İnsan ve Eşyaların Eşleştirilmesi Durumu

İnsan ve eşyaların tespiti YOLOv5S modeli tarafından yapılmaktadır. Tespit edilen nesnelerin takibi ise deepsort algoritması tarafından yapılır.

Test edilecek videolar sabit bir boyuta küçültülür. Bu küçültme işlemi valiz ve insan ilişkisinin eşik değerini kullanabilmek için yapılmaktadır. Gelen video 640x340 boyutlarına yaklaşacak şekilde ffmpeg yazılımı colab üzerinde çalıştırılarak videoya küçültme işlemi yapılır. Ardından valiz ve insanın tespiti yapılır. Tespit edilen insan ve valizlerin özneliklerine bağlı olarak takip numaraları atanır. Bu numaralar üzerinden aralarında ki mesafe piksel cinsinden eşiklenir. Bu eşik piksel değeri 130 olarak belirlenmiştir. Eğer ki tespit edilen insan sınıfının yakınında bir eşya tespit edildiyse ilişki kurulur. Tespit bitene kadar takip numaraları bir dizide tutulur. Kısaltmış akış şemasına şekil 8’de yer verilmiştir. Dizi içine alınan takip numaralı sınıflar başka bir sınıfla eşleştirilmezler. Aşağıda ki şemada buna yer verilmemiştir.



Şekil 8 Nesne Eşleştirme Fonksiyonunun Akış Şeması

6 Sonular ve İnceleme

Proje temel durumları gzetilerek yapıldığı iin bazı durumlarda yetersiz kalmıştır.

- Nesne Tespiti iin toplanan veri sayısı yetersiz gelmiştir. Model yeterli veriyle beslenememiştir.
- Oluşturulan veriseti balans bir veri seti, değildir. Bu yüzden model valizi, sırt antasına göre daha kolay tespit etmektedir. Bu durum hata matrisinde de görölmektedir.
- Yapılan model seçimi YOLOv5S, her ne kadar küçük bir modelde olsa mAP açısında yeterli değildir.
- Kamera görüş açısına giren insanlar yanyana ve çok kalabalık olduđu durumlarda tek kamerayla insan – valiz eşleştirmesi yapmak yeterli değildir. Aynı ortama farklı görüş açılardan kameralar yerleştirilerek görüntü eşleme yapılmalıdır. Bu sayede görüntüler düzeltilip üç boyutlu gerçek uzaklık hesaplanarak valiz – insan eşleştirilmesi yapılabilir.

7 Kaynaklar

Kaynaklara üç başlık altında yer verilmiştir.

7.1 Kaynak Araştırması

(Kevin Smith, 2006),

https://www.researchgate.net/publication/41387227_Detecting_Abandoned_Luggage_Items_in_a_Public_Space

(Chang Jing-Ying,

2010),https://www.researchgate.net/publication/45692814_Localized_Detection_of_Abandoned_Luggage

7.2 Makale İçi Atıfta Bulunanlar

[1] [https://www.cv-](https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf)

[foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf](https://www.cv-foundation.org/openaccess/content_cvpr_2016/papers/Redmon_You_Only_Look_CVPR_2016_paper.pdf)

[2] [GitHub - ultralytics/yolov5: YOLOv5 🚀 in PyTorch > ONNX > CoreML > TFLite](#)

[3] [Yapay Zeka ve Robotik İçin Jetson Nano Geliştirici Kiti | Nvidia](#)

[4] [The PASCAL Visual Object Classes Homepage \(ox.ac.uk\)](#)

[5] <https://colab.research.google.com/>

[6] <https://dergipark.org.tr/tr/download/article-file/1571016>

[7] <https://ieeexplore.ieee.org/document/7533003>

7.3 Veri Setlerinin Kaynakları

- <https://github.com/kevinlin311tw/ABODA>
- https://drive.google.com/drive/folders/1F3xAgdcmCZNEWpJY_HQkWcQneWvMg7xo
- <https://www.skylinewebcams.com/webcam/italia/lombardia/milano/milano-stazione.html>
- <https://www.youtube.com/watch?v=sKyqd7Y9TOk>
- <https://www.skylinewebcams.com/webcam/united-states/arizona/flagstaff/flagstaff-train-station.html>