



Teknoloji Fakültesi

MARMARA ÜNİVERSİTESİ
TEKNOLOJİ FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

BİTİRME PROJESİ

Video Görüntüleri Üzerinden Cinsiyet Tespiti ve Kalabalık Analizi

PROJE YAZARI

Şule AKTÜRK

DANIŞMAN

Doç. Dr. Önder DEMİR

İL, TEZ YILI

İstanbul 2024



Teknoloji Fakültesi

**MARMARA UNIVERSITY
FACULTY OF TECHNOLOGY
DEPARTMENT OF COMPUTER ENGINEERING**

GRADUATION PROJECT

Video-based Gender Detection and Crowd Analysis

PROJECT AUTHOR

Şule AKTÜRK

ADVISOR

Doç. Dr. Önder DEMİR

CITY, THESIS YEAR

İstanbul 2024

MARMARA ÜNİVERSİTESİ
TEKNOLOJİ FAKÜLTESİ
BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ

Marmara Üniversitesi Teknoloji Fakültesi Bilgisayar Mühendisliği Öğrencisi
..... nın “.....” başlıklı bitirme projesi çalışması,
..../..../..... tarihinde sunulmuş ve jüri üyeleri tarafından başarılı bulunmuştur.

Jüri Üyeleri

Prof. Dr. Adı SOYADI (Danışman)

Marmara Üniversitesi (İMZA)

Doç. Dr. Adı SOYADI (Üye)

Marmara Üniversitesi (İMZA)

Dr. Öğr. Üyesi Adı SOYADI (Üye)

Marmara Üniversitesi (İMZA)

İÇİNDEKİLER

KISALTMALAR/ABBREVIATIONS.....	ii
ŞEKİL LİSTESİ.....	iii
TABLO LİSTESİ.....	iv
ÖZET	1
1. GİRİŞ	2
1.1 Literatür Taraması:.....	2
2. MATERYAL VE YÖNTEM.....	6
2.1 Algoritma Seçimi	6
2.2 Projenin Ana İşleyişi.....	6
2.3. Videolardan Fotoğraf Elde Etme	7
2.4 Fotoğrafları Etiketleme	8
2.5 YOLO Formatına Uygun Klasör Oluşturma.....	9
2.6 Eğitim Ortamını Oluşturma	10
2.7 YOLOv8 ile Eğitimi Gerçekleştirme	11
3. BULGULAR VE TARTIŞMA	13
3.1 PolygonZone Teknolojisiyle İnsan Sayma	16
4. SONUÇ	24
5. ARAŞTIRMANIN SINIRLAMALARI	25
6. İLERİYE YÖNELİK ÖNERİLER	26
KAYNAKÇA.....	27
ÖZGEÇMİŞ - RESUME.....	29

KISALTMALAR/ABBREVIATIONS

CPU: Central Process Unit

CCTV: Close Circuit TeleVision

YOLO: You Only Look Once

Faster R-CNN: Faster Region Based Convolutional Networks

MIT: Massachusetts Teknoloji Enstitüsü

SIMD: Single Instruction, Multiple Data

HOG: Histogram of Oriented Gradients

SVM: Support Vector Machine

MDP: Markov Decision Processes

SORT: Simple Online Real-Time Tracking

MOTA: Multiple Object Tracking Accuracy

MAP: Mean Average Precision

OpenCV: Open Source Computer Vision Library

IMDb: Internet Movie Database

SENResNet: Selective Ensemble Network based on ResNet

RNFT: Region-based Non-local Fully Convolutional Networks

R-CNN: Region-based Convolutional Neural Network

DNN: Deep Neural Network

GPU: Graphics Processing Unit

TPU: Tensor Processing Unit

ROI: Region of Interest

LOI: Line of Interest

ŞEKİL LİSTESİ

Şekil 2.1 Kalabalık Analizi İzlenecek Adımlar	7
Şekil 2.2 Veri Setinden Örnek Görüntü	8
Şekil 2.3 Etiketlenmiş Örnek Görüntü	8
Şekil 2.4 Örnek Etiket Dosya İçeriği	9
Şekil 2.5 Veri Seti Klasör Yapısı	10
Şekil 2.6 Data.yaml Dosya İçeriği	11
Şekil 2.7 GPU Bağlantı Kontrolü Kodu	11
Şekil 3.1 F1 Güven Eğriliği Grafiği	13
Şekil 3.2 Karışıklık Matrisi	14
Şekil 3.3 Modelin Eğitimi Sonucu Elde Edilen Grafikler	15
Şekil 3.4 Örnek Poligon Çizimi	18
Şekil 3.5 Görüntü Üzerinden Cinsiyet Tespiti ve Kişi Sayımı	19
Şekil 3.6 Video Üzerinden Cinsiyet Tespiti ve Kalabalık Analizi Sonuçları	22
Şekil 3.7 Etiketlemede Kadın-Erkek Dağılımları	23

TABLO LİSTESİ

Tablo 1: YOLOv8 mimarisi ile doğrulama veri setinden elde edilen performans sonuçları..16

ÖZET

Cinsiyete göre kalabalık analizi, çeşitli alanlarda önemli faydalar sağlamaktadır. Örneğin, alışveriş merkezleri ve halka açık etkinlikler gibi yerlerde güvenlik ve hizmetlerin daha iyi planlanmasını mümkün kılar. Ayrıca, bu tür analizler, pazarlama stratejilerinin belirlenmesinde, müşteri profillerinin çıkarılmasında ve işletme verimliliğinin artırılmasında kullanılabilir. Şirketler, demografik bilgilere dayanarak daha hedeflenmiş ve etkili kampanyalar düzenleyebilirler. Kamu alanlarında, hizmetlerin adil dağılımı ve ihtiyaçların doğru belirlenmesi açısından önemli veriler sağlar.

Bu çalışmada, YOLOv8 algoritması kullanılarak NTT Data şirketinin mola alanı görüntülerinden elde edilen videolar aracılığıyla 3000 fotoğraflık bir veri seti oluşturulmuştur. Bu veri seti, cinsiyet tespiti amacıyla erkek/kadın olarak etiketlenmiştir. Etiketleme işlemi sonrasında, Google Colab'ın bulut ortamı kullanılarak YOLOv8 algoritması ile model eğitilmiştir. Eğitilen model, Supervision'ın PolygonZone kütüphanesi kullanılarak video üzerinden kalabalık analizi gerçekleştirmiştir. Elde edilen sonuçlar incelendiğinde, cinsiyet tespitinde doğruluk oranının yaklaşık %86 olduğu belirlenmiştir.

ABSTRACT

Gender-based crowd analysis provides significant benefits in various fields. For instance, it enables better planning of security and services in places like shopping malls and public events. Additionally, such analyses can be utilized in determining marketing strategies, extracting customer profiles, and enhancing business efficiency. Companies can organize more targeted and effective campaigns based on demographic information. In public areas, it provides crucial data for the fair distribution of services and accurate determination of needs.

In this study, a dataset of 3000 images was created using videos of NTT Data company's break area, utilizing the YOLOv8 algorithm. This dataset was labeled as male/female for the purpose of gender detection. Following the labeling process, the model was trained using the YOLOv8 algorithm in the cloud environment of Google Colab. The trained model conducted crowd analysis on the video using the Supervision's PolygonZone library. Upon examining the obtained results, it was determined that the accuracy rate for gender detection was approximately 86%.

1. GİRİŞ

Günümüz dünyasında güvenlik ve insan hareketlerinin izlenmesi, şehir yönetimlerinden özel sektöre kadar geniş bir yelpazede kritik bir öneme sahiptir. Özellikle kalabalık alanların izlenmesi, insan yoğunluğunun tespit edilmesi ve cinsiyet ayrımının yapılması, pek çok endüstri ve uygulama alanında stratejik bilgiler sunarak verimlilik ve güvenlik açısından önemli katkılar sağlamaktadır. Bu çalışma, video görüntülerinden cinsiyet tespiti ve insan sayımı yapabilen bir yapay zekâ modelinin geliştirilmesini ele almaktadır.

Projede kullanılan veri seti, NTT Data firmasının mola alanına ait sabit bir kameradan elde edilen videolardan oluşmaktadır. Kamera, mobese benzeri bir güvenlik kamerası olup, alanın her yerini kapsamamakta ve kameranın açısında bazı kör noktalar oluşmaktadır. Bu nedenle alan içerisindeki bazı insanların görüntüleri net bir şekilde görünmemektedir. Ayrıca, çalışanların mola alanına giriş ve çıkışları merkezi bir kapıdan yapılmakta, bu da giriş-çıkış kontrolünü zorlaştırmaktadır. Bu kısıtlamalar altında doğru ve güvenilir bir kalabalık analizi yapabilmek amacıyla ileri düzey yapay zekâ algoritmalarının kullanılması gerekmektedir.

Çeşitli endüstrilerde, kalabalık analizi ve insan sayımı, operasyonel verimliliğin artırılması ve güvenlik protokollerinin güçlendirilmesi için kritik bilgiler sağlar. Örneğin, alışveriş merkezlerinde insan yoğunluğunun doğru bir şekilde tespit edilmesi, müşteri hizmetlerini optimize etmek ve güvenlik önlemlerini artırmak için önemlidir. Benzer şekilde, havaalanları gibi yoğun insan trafiğinin olduğu yerlerde, güvenlik kameralarından elde edilen verilerin doğru analizi, güvenlik ihlallerinin önlenmesine ve yolcu akışının daha verimli yönetilmesine katkıda bulunabilir. Stadyumlar ve büyük etkinlik alanlarında ise, kalabalık yönetimi ve acil durum planlarının uygulanmasında bu tür analizler hayati öneme sahiptir.

1.1 Literatür Taraması:

Kalabalık analizi ve insan sayımı üzerine yapılan çeşitli çalışmalar, bu teknolojinin farklı uygulama alanlarındaki etkinliğini ve potansiyel faydalarını ortaya koymaktadır. Örneğin, Li ve arkadaşları tarafından yapılan bir çalışmada [1], yoğun insan trafiğinin olduğu şehir meydanlarında kalabalık tespiti ve anomali algılama için derin öğrenme teknikleri kullanılmıştır. Bu çalışma, güvenlik kameralarından elde edilen görüntülerle anlık analizler yaparak potansiyel tehlikeleri erken aşamada tespit edebilmiştir.

Bir başka örnek olan Jefferson ve arkadaşlarının yaptığı [2] restoran video görüntüleri üzerinden hali hazırda sayma işlevi bulunan yapıya YOLOv2 ve InceptionV3 algoritmaları kullanılarak cinsiyet tespiti özelliği eklenmiş ve %83.86 doğruluk oranıyla başarıya ulaşılmış fakat gerek kamera açısından gerek de veri setinin çeşitliliğinin az olmasından kaynaklı bazı tespitlerin başarılı bir şekilde yapılamadığı görülmüştür.

Aman Ahmed ve Prateek Bansal tarafından gerçekleştirilen çalışmada [3], CCTV kameralarıyla kaydedilen videolar üzerinden insanların sayısını tespit etme ve bu kişilerin yaş ve cinsiyetini belirleme yeteneklerini geliştirmeyi amaçlamaktadır. MobilNet SSD modeli kullanılarak video çerçevelerinde insan tespiti yapılmakta, ardından yaş ve cinsiyet tahmin modülleri aracılığıyla analizler gerçekleştirilmektedir. Çalışmanın sonuçları, yüksek çözünürlüklü videolardan elde edilen görüntülerin daha doğru sonuçlar verdiğini göstermektedir. Ancak, düşük çözünürlüklü videolarda veya kişinin yüzünün net görünmediği durumlarda doğru sonuçlar elde edilememektedir.

Hussain ve arkadaşları [4], kalabalık sayımı, cinsiyet tanıma ve yaş tahmini görevlerini eşzamanlı olarak gerçekleştirebilen bir model geliştirmiştir. Bu modelin eğitiminde Faster R-CNN tabanlı bir tespit modeli kullanılmıştır. Bu model, özellikle büyük ölçekli CVML-OCC veri seti üzerinde eğitilerek, çeşitli veri setleri üzerinde (Mall, ShanghaiTech, WorldExpo10) test edilmiştir ve diğer yaklaşımlara kıyasla daha yüksek doğruluk oranları elde edilmiştir. Bu yenilikçi yaklaşım, kalabalık analizinde yaş ve cinsiyet kategorilerinin dahil edilmesiyle elde edilen gelişmiş doğruluk oranlarını göstermektedir.

Mudassar Raza ve arkadaşlarının yaptığı çalışmada [5], derin öğrenme yöntemlerinden istiflenmiş otomatik kodlayıcıların (stacked auto encoders) kullanıldığı yaya cinsiyetinin görünüme dayalı tanınmasını ele alıyor. Metodoloji olarak, bilgisayarlı görü tabanlı derin öğrenme yaklaşımı ve derin ayrışmalı sinir ağları aracılığıyla arka planı kaldırma işlemi kullanılıyor. Önerilen yaklaşım, PETA veri seti üzerinde eğitilerek test ediliyor ve MIT veri setiyle karşılaştırılıyor. Sonuçlar, ön, arka ve karışık görünüm için yüksek doğruluk oranları gösteriyor ve önerilen yöntemin mevcut çalışmalara göre daha sağlam olduğunu ortaya koyuyor. Ancak, yazarlar bu alanda daha fazla araştırma yapılması gerektiğini ve gelecekte cinsiyet tanıma çalışmalarının vücut parçalarına dayalı olarak genişletilebileceğini belirtiyorlar.

Vincenzo Carletti ve arkadaşlarının 2019 yılında yayınladıkları makalede [6], akıllı kameralar için etkili bir gerçek zamanlı cinsiyet tanıma sistemi oluşturulmuştur. Sistem, harici bir sunucuya ihtiyaç duymadan doğrudan akıllı kameranın içerisinde çalışabilen bir gömülü görüntü çözümü olarak tasarlanmıştır. Görüntü ediniminden cinsiyet sınıflandırmasına kadar olan tüm gerçek zamanlı bilgisayarlı görü işleme süreci, tek bir kare üzerinde gerçekleştirilmiştir. Bu süreç, yüz algılama, temsil ve takip gibi ara işlemleri içerir. Arka plan çıkarma işlemi, SIMD talimatlarına dayalı bir paralel uygulama ile yürütülür ve yüz algılama yalnızca önceden belirlenmiş hareketli pikseller üzerinde gerçekleştirilir. Yüz algılama süreci, akıllı kameralar için optimize edilmiş hızlı bir Viola-Jones algoritması kullanılarak hızlandırılmıştır. Cinsiyet sınıflandırması için ise, HOG feature extraction ve svm teknolojileri kullanılmıştır. Veri setinde aynı kişiye ait birden fazla yüzün kullanılması, önemli bir cinsiyet tanıma doğruluğu sağlamıştır. Bu sistem, akıllı kameraların tahtasında çalışabilen etkili bir gerçek zamanlı cinsiyet tanıma çözümü sunmaktadır.

2020 yılında yayınlanan makaleye göre [7], standart bir video kamera görüntüsü üzerinden bir tren istasyonunun görüntüleri kullanılarak insan tespiti yapıp, kişi sayım işlemi gerçekleştirilmiştir. Çalışmalarda PAMELA-UANDES veri seti kullanılmıştır. Kişi tanıma

aşamasında EspiNet, Faster-RCNN ve Yolov3 olmak üzere üç farklı derin öğrenme nesne algılayıcısı değerlendirilmiştir. Kişi izleme aşamasındaysa üç benchmark izleyici (Markov Decision Processes (MDP), SORT ve D-SORT) değerlendirilmiş ve yaklaşık %80 seviyesinde rekabetçi bir MOTA (Çoklu Nesne Takip Doğruluğu) elde edilmiştir. 322 video dizisi için insanlar sayılmış ve %95'in üzerinde bir F1 skoru elde edilmiştir.

Musa Peker ve arkadaşlarının [8] 2020 yılında yapmış oldukları insan sayma teknolojisinin 2 farklı donanım ve 4 farklı algoritmadaki performansları üzerine yapmış oldukları çalışmada Yolo'nun farklı 4 versiyonu olan; Yolov3, Yolov3-Tiny, Yolov4 ve Yolov4-Tiny algoritmalarının performanslarının karşılaştırılması sonucunda Yolov4-Tiny algoritmasının ideal olduğu sonucuna varılmıştır.

Mehmet Karahan ve arkadaşlarının 2022 yılında hazırladıkları makalede [9], yüz özelliklerinden yaş ve cinsiyet tespiti üzerine RetinaNet ve YoloV3 modelleri kullanılmıştır. Elde edilen sonuçlar MAP ve çıkarım süresi açısından karşılaştırılmıştır. Karşılaştırmalar sonucunda, YOLOv3 algoritmasının daha hızlı ve daha yüksek doğruluk oranıyla çalıştığı gözlemlenmiştir.

Julie Ann B. Susa ve arkadaşlarının 2022 yılında yazdıkları Derin Sinir Ağı Tabanlı Gözetim Tuvalet Kısıtlama Sistemi için Cinsiyet Tanımlaması adlı çalışmada [10], erkekleri ve kadınları ayırt eden bir tespit mekanizması sunulmuştur. Okullardaki her iki cinsiyeti de kategorize etmek için görüntü işleme kullanılarak cinsiyet tanımlama geliştirilmiştir. Cinsiyet tanımlama sisteminin uygulanmasında YOLOv3 teknolojisi kullanılmıştır. Çalışmanın sonuçlarında, kullanılan tespit modelinin %95,28 mAP değerine sahip olduğunu belirtmektedir. Bu anlamda kadın-erkek tespitinde YoloV3 algoritmasının başarılı olduğu görülmüştür.

Varnima E. K. ve arkadaşı C. Ramachandran ile birlikte yaptığı çalışmada [11], yüz görüntülerinden cinsiyet tahmini için kullanılabilecek gerçek zamanlı bir sistem oluşturmuşlardır. Kullandıkları teknik, You Only Look Once (YOLO) v3 nesne algılama algoritmasıdır. Eğitim için Darknet, test için ise Keras ve OpenCV kullanılmışlardır. Kullanılan veri seti, IMDb, Google Hint görüntüleri ve mobil kamera kullanılarak çekilen bazı Hint özel görüntülerinin bir kombinasyonudur. Test görüntüleri doğruluğu %84,69 olarak bulunmuştur. Önerilen çalışmanın amacı, sistemin izleme, güvenlik ile ilgili alanlar gibi durumlarda kullanılmasıdır.

Guangyong Zheng ve Yuming Xu'nun 2021 yılında yayınladıkları Derin öğrenmeye dayalı video dizilerinde verimli yüz algılama ve izleme çalışmasında [12], mevcut video yüz algılama ve izleme yöntemlerinin, ortam ışığı değişiklikleri, insan duruş değişiklikleri ve örtüşme gibi karşılaştığı mevcut zorlukları ele alarak kamuya açık yüz ve video veri setleri üzerinde yapılan kapsamlı deneysel çalışmalarda bulunmuşlardır. Bu çalışma sonucunda, önerilen SENResNet ve RNFT modellerinin doğruluk ve performans açısından son teknoloji karşılaştırma yöntemlerinden üstün olduğunu sonucuna ulaşmışlardır.

Aziza Srazhdinova ve arkadaşlarının 2020 yılında yayınladıkları makalede [13], nesnelerin gerçek zamanlı olarak tespit edilmesi ve izlenmesi görevlerini çözmek amacıyla, insan

tespitinde yolov3 algoritmasını ve insanların takip edilmesinde SORT (Simple Online Real-Time Tracking) algoritmasını kullanarak bir program oluşturmuşlar. Hız, performans ve doğruluk oranı açısından iyi bir sonuç veriyor olmasından kaynaklı bu algoritma tercih edilmiş.

Resul Daş ve arkadaşlarının 2020 yılında hazırladıkları makalede [14], durağan ve hareketli görüntülerde nesne tanıma ve takibiyle ilgili Tensorflow ve Faster R-CNN kütüphaneleri kullanılmıştır. Kullanılan kütüphaneler video, fotoğraf üzerinde test edilerek hem nesne tespiti ve takibi yapılmış hem de kullanılan teknolojilerin artı ve eksi yönleri üzerine sonuca varılmıştır.

Andrei de Souza Inacio ve arkadaşlarının yaptığı çalışmada [15], çeşitli DNN modelleri kullanılarak video üzerinden yaş ve cinsiyet tespiti yapabilen bir çalışma ortaya koymuşlardır. Bu çalışma sonucunda, yaş ve cinsiyet tespiti için kullanılmış 6 model (EfficientNet, InceptionV3, MobileNet, VGG16, SqueezeNet, SqueezeNet2) arasından en başarılı sonucu veren model EfficientNet olmuştur. Karşılaşılan zorlukların en büyük nedeni ise videodaki insanların görüntülerinin bazı noktalarda engellenmesinden kaynaklanmıştır. Başarılı etiketleme ve tespiti açısından zorlayıcı bir unsurdur. Benzer zorluklar kendi projemizde de yaşanmıştır.

Tüm bu çalışmalar, video görüntülerinden insan tespiti ve sayımı konusundaki araştırmaların önemini ve bu teknolojilerin gerçek dünya uygulamalarında nasıl kullanılabileceğini göstermektedir. Projemiz, bu literatür kapsamında, sabit kameralarla elde edilen video verilerinin analizi yoluyla güvenilir ve etkili bir kalabalık analizi yapmayı amaçlamaktadır. Elde edilen sonuçlar, özellikle güvenlik ve operasyonel verimliliğin artırılmasında önemli katkılar sağlayacaktır.

2. MATERYAL VE YÖNTEM

Yapay zekâ ve makine öğreniminde temelde iki öğrenme türü vardır: Denetimli öğrenme, denetimsiz öğrenme.

Denetimli Öğrenme (Supervised Learning) türünde model, giriş verilerine karşılık olarak gelen etiketleri yani doğru çıktı verileri ile eğitilir. Model, bu etiketleri öğrenerek yeni verileri sınıflandırır veya tahmin eder. Sınıflandırma, regresyon, görüntü tanıma gibi projelerde kullanılır.

Denetimsiz Öğrenme (Unsupervised Learning) türünde veri setinde herhangi bir etiket bulunmaz. Model, kendi kendine verinin yapısını anlamaya çalışır ve benzer özelliklere sahip verileri gruplandırır. Kümeleme, boyut indirgeme, veri görselleştirme, anomali tespiti gibi alanlarda kullanılır.

2.1 Algoritma Seçimi

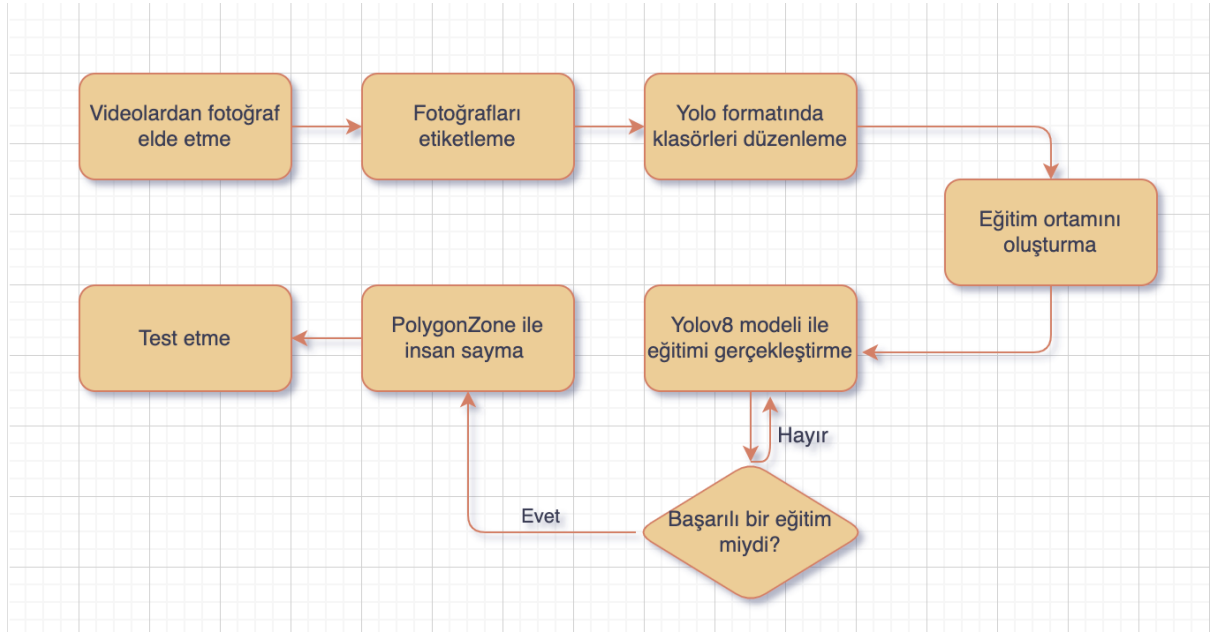
Video üzerinden cinsiyet tespitini yapabilmek için bir geliştirme algoritmasına ihtiyaç duyulmaktadır. Bu projede, cinsiyet tespiti için tercih edilen algoritma Yolov8 algoritması olmuştur.

Cinsiyet tespiti için YOLO (You Only Look Once) algoritmasının tercih edilmesinin temel nedeni, denetimli öğrenme yaklaşımının bu tür görsel sınıflandırma problemleri için oldukça uygun ve etkili olmasıdır.

Yolov8, Yolo (You Only Look Once) ailesinin sondan bir önceki versiyonudur. Yolov8, Ultralytics tarafından geliştirilmiştir. Yolov8, nesne algılama (object detection), görüntü sınıflandırma (image classification) gibi pek çok makine öğrenimi alanında kullanılabilen en yeni ve son teknoloji ürünlerinden biridir. Önceki sürümlerine ve benzeri teknolojilere kıyasla YOLOv8, optimize edilmiş hesaplama grafikleri ve hafif model tasarımı sayesinde daha hızlı tahmin yapabilir. Eğitim ve test aşamalarında yüksek doğruluk oranlarına ulaşır. YOLOv8, optimize edilmiş hesaplama grafikleri ve hafif model tasarımı sayesinde daha hızlı tahmin yapabilir. Önceden eğitilmiş modellerle transfer öğrenmesi yaparak eğitim süresini ve gereksinimlerini azaltır. Entegrasyon ve kullanım kolaylığı açısından sıklıkla tercih edilen Yolov8, hızlı nesne algılama, kullanıcının yaptığı pek çok işi otomatikleştirerek geliştirme aşamasında geliştiriciye büyük kolaylıklar sağlayan son teknoloji algoritmalarındandır. Ayrıca geniş ve detaylı dokümantasyon desteğiyle birlikte kullanıcı dostudur. Tüm bu özellikler ele alındığında, cinsiyet tespiti aşamasında Yolov8 algoritmasının kullanılması tercih edilmiştir.

2.2 Projenin Ana İşleyişi

Video üzerinden cinsiyete göre kalabalık analizi yapabilmek için izlenmesi gereken adımlar Şekil 2.1’de görülmektedir.



Şekil 2.1 Kalabalık Analizi İzlenecek Adımlar

2.3. Videolardan Fotoğraf Elde Etme

Projenin ilk adımında veri seti oluşturulur. Veri seti oluşturulma nedeni, cinsiyet tespiti işleminin gerçekleştirilebilmesi için eğitilecek yapay zekâ modelinin öğrenebilmesi için etiketlenmiş verilere ihtiyaç duymasıdır. Bu adım için Marmara Üniversitesi'nin anlaşmasıyla birlikte, NTT Data firmasının mola alanında bulunan güvenlik kamerasına ait veriler kullanılmıştır. 10 saatten fazla video görüntüsü bulunan bu videolar, VLC Media Player aracı kullanılarak ekran görüntüleri alınmıştır. Elde edilen ekran görüntüleri yaklaşık 6000 adettir fakat görüntüler üzerinde bulanıklık ve kalabalık açısından en uygun olanları seçilerek 3000 adet fotoğraf elde edilmiştir. Elde edilen fotoğraflar “img1.jpg” formatına çevrilmiştir. Hepsi aynı kameradan çekildiği için tüm görüntülerdeki kamera açısı aynıdır. Örnek görüntü Şekil 2.2'deki gibidir.



Şekil 2.2 Veri Setinden Örnek Görüntü

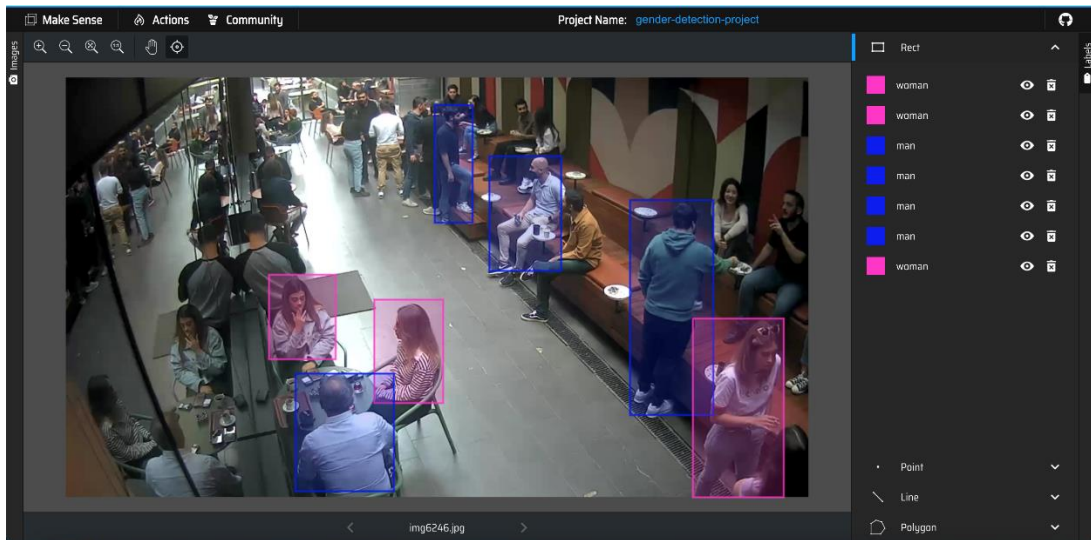
2.4 Fotoğrafları Etiketleme

Etiketleme işlemi, yapay zekâ ve makine öğrenimi modellerinin eğitim sürecinde kullanılan veri noktalarına anlamlı etiketler veya sınıflandırmalar atanmasıdır. Bu işlem, veri setindeki her bir örneğin neyi temsil ettiğinin belirlenmesi anlamına gelir.

Elde edilen 3000 adet fotoğraf, Makesense.ai sitesi kullanılarak etiketlenmiştir. Makesense etiketleme aracı; açık kaynak kodlu, ücretsiz, basit bir kullanıma sahip olup, herhangi bir uygulama indirmeye gerek kalmadan çevrimiçi olarak kullanılabildiğinden tercih edilmiştir.

Site üzerinde etiketleme işlemi için öncelikle fotoğraflar yüklenir, ardından görüntüler için kullanılacak etiketler belirlenir. Bu projede etiketleme işlemi dörtgen kullanılarak yapılmıştır fakat kullanılan teknolojiye göre farklı çokgenlerle de etiketlemeye site olanak sağlamaktadır.

Örnek etiketleme işlemi Şekil 2.3'teki gibidir.



Şekil 2.3 Etiketlenmiş Örnek Görüntü

Etiketleme işleminin ardından fotoğraflara ait etiketler elde edilir. Projede dörtgen formatında etiketleme yapıldığından etiketleme dosyasında 1 adet cinsiyete ait etiket verisi ve 4 adet koordinat verisi bulunur. Örnek dizilim aşağıdaki gibidir:

<class_id> <x_center> <y_center> <width> <height>

Burada class_id değeri, etiketin bulunduğu sınıfı ifade eder, örneğin burada “0” rakamı “erkek”, “1” rakamı ise “kadın” cinsiyetini temsil etmektedir.


x_center değeri, nesnenin merkezinin x eksenindeki konumunu ifade eder.

y_center değeri, nesnenin merkezinin y eksenindeki konumu ifade eder.

width değeri, nesnenin genişliğini ifade eder.

height değeri ise, nesnenin yüksekliğini ifade eder.

Örnek etiket dosyası ve içeriği Şekil 2.4’teki gibidir.



img6246.txt

1	0.461271	0.651978	0.092746	0.246403
1	0.318076	0.571942	0.089688	0.201439
0	0.375150	0.845324	0.132494	0.280576
0	0.815438	0.549460	0.112110	0.512590
0	0.521912	0.205935	0.050959	0.282374
0	0.619245	0.323741	0.096823	0.273381
1	0.905126	0.787770	0.122302	0.424460

Şekil 2.4 Örnek Etiket Dosya İçeriği

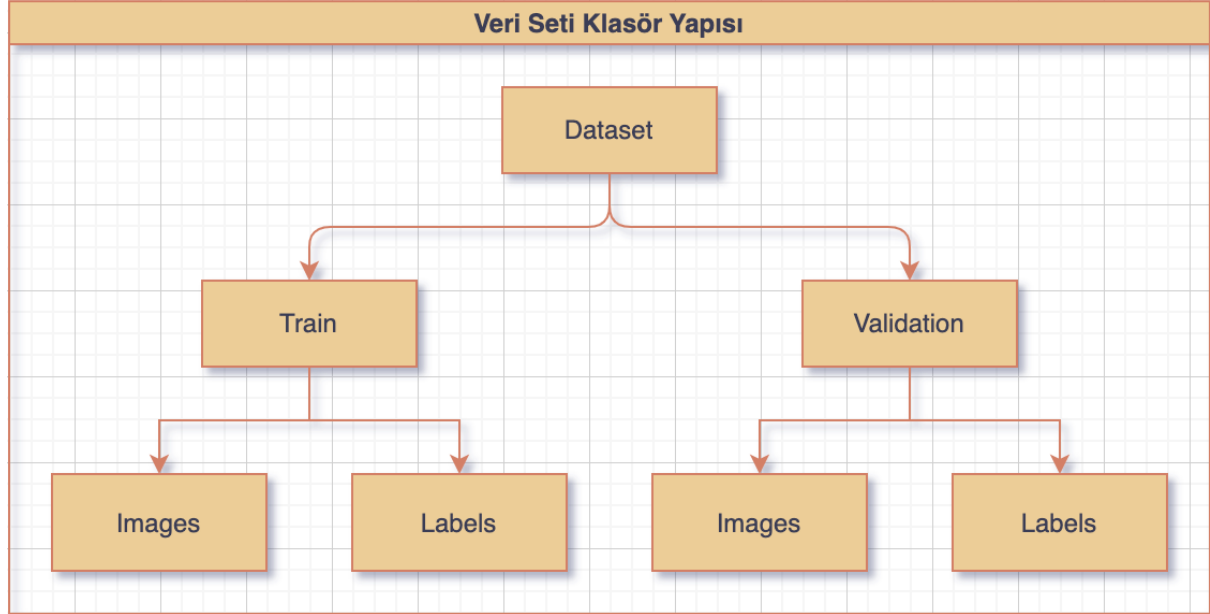
Txt formatında olan bu dosyaya göre img6246.jpg görüntüsünde belirtilen koordinatlarda 4 kadın ve 3 erkek kişi bulunmaktadır.

2.5 YOLO Formatına Uygun Klasör Oluşturma

3000 fotoğrafın etiketlenmesinin ardından Yolo algoritmasının klasör yapısı aşağıda verilmiştir. Bu klasör yapısına uygun şekilde 3000 adet görüntüyü eğitim(train) ve doğrulama(validation) veri setleri olmak üzere iki klasöre ayrılmıştır. Eğitim yani train veri seti, modelin öğrenme yapabilmesi için hazırlanan veri setidir. Model bu veri seti üzerinde tekrarlı olarak çalışarak verilerdeki desenleri ve ilişkileri anlamlandırmaya çalışır. Doğrulama yani validation veri seti ise, modelin performansını değerlendirmek için kullanılan veri setine denilir. Modelin eğitim veri setinde görmediği veriler üzerinde tahminde bulunarak test

edilmesiyle modelin genelleme yeteneği ölçülür. Bu ölçümler sayesinde modelin aşırı uyumu (overfitting) tespit edilmeye yardımcı olur ve modelin gelişimi kayıt altına alınır. Eğitim ve Doğrulama veri setlerinin ayrımı yapılmadan önce görüntüler, split algoritması kullanılarak karıştırılır, ardından %80 eğitim(train), %20 doğrulama(validation) olmak üzere aşağıdaki şemaya uygun olarak klasörler oluşturulmuştur.

Yolo formatına uygun veri setinin klasör yapısı Şekil 2.5’de verilmiştir.



Şekil 2.5 Veri Seti Klasör Yapısı

2.6 Eğitim Ortamını Oluşturma

Eğitim ortamı olarak Google Colab, kullanıcılarına belirli bir limite kadar ücretsiz GPU (Grafik İşlem Birimi) ve TPU (Tensor İşlem Birimi) sağlar. Yapay zekâ modelleri gibi yoğun hesaplamalı ve uzun süren işlemleri daha kolay ve kısa sürede tamamlayabilir. Google’ın diğer ürünleriyle etkileşimli bir şekilde çalışılmasına olanak sağlar. Ayrıca bulut tabanlı olduğundan kullanılan bilgisayarın performansından bağımsız bir şekilde çalışabilir. Bilgisayara herhangi bir kütüphane indirilmesine gerek kalmadan kütüphaneler kurulabilir, pek çok programlama dilini desteklemesinden kaynaklı bu projede tercih edilmiştir.

Ayrıca Google’ın diğer ürünleri ile entegrasyon kolaylığı bakımından Google Colab, geliştiricilere çok büyük fayda sağlamaktadır. Veri setini ve eğitim sonrası çıktıları ve büyük paketleri bilgisayarda ekstra yer kaplamadan bulut ortamı üzerinde depolanır. Bu anlamda yukarıda bahsedilen veri seti klasörü Google Drive’da saklanmıştır.

Eğitim öncesi son aşamada projede kullanılan sınıf isimleri ve dosyaların konum bilgileri data.yaml adlı bir dosyaya yazılmıştır. Dosyaların konumları Google Drive’daki konum bilgileridir. Data.yaml dosyası içeriği aşağıdaki Şekil 2.6’daki gibidir.

```

train: /mydrive/gender_detection/dataset/train/images
val: /mydrive/gender_detection/dataset/valid/images

nc: 2
names: ["man", "woman"]

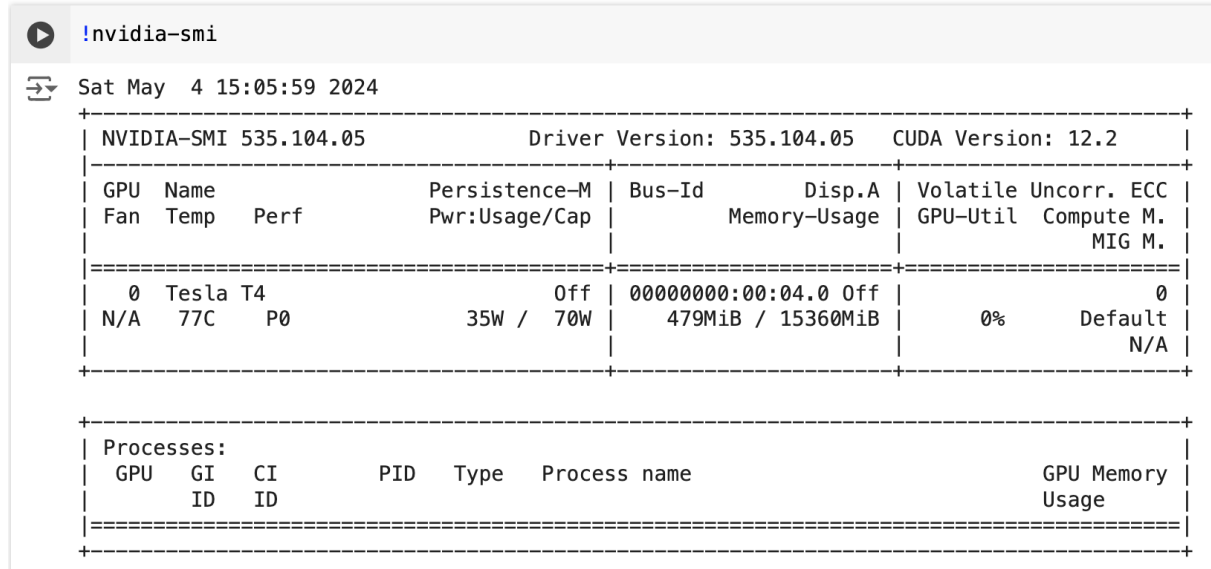
```

Şekil 2.6 Data.yaml Dosya İçeriği

Data.yaml dosyası, veri setinin içerik bilgilerini, sınıf sayısı ve isimlerini tutmak ve modele yol göstermesi için kullanılan dosyadır.

Google hesabı ile giriş yapıp Colab üzerinde yeni bir proje oluşturulduktan sonra yoğun hesaplamalar gerçekleştirileceğinden ilk adım olarak çalışma zamanı varsayılan CPU'dan T4 GPU olarak ayarlanır. T4 GPU, derin öğrenme modellerinin eğitimi ve çıkarımı sırasında CPU'ya göre çok daha hızlı işlem yapar. Bu, eğitim süresini büyük ölçüde kısaltır ve daha karmaşık modelleri daha kısa sürede eğitmenize olanak tanır. Ayrıca geniş bellek kapasitesi sayesinde bellek sınırlamaları nedeniyle model ve veri setlerini küçültme gerekliliğini ortadan kaldırır.

Eğitimin ilk adımı olarak, “! nvidia-smi” komutu girilerek bağlı olunan GPU bilgilerine erişilir. Örnek bağlantı aşağıdaki Şekil 2.7’deki gibidir.



NVIDIA-SMI 535.104.05 Driver Version: 535.104.05 CUDA Version: 12.2									
GPU	Name	Persistence-M	Bus-Id	Disp.A	Volatile Uncorr. ECC				
Fan	Temp	Perf	Pwr:Usage/Cap	Memory-Usage	GPU-Util	Compute M.			
						MIG M.			
0	Tesla T4	Off	00000000:00:04.0	Off	0%	0			
N/A	77C	P0	35W / 70W	479MiB / 15360MiB	0%	Default			
						N/A			

Processes:							
GPU	GI	CI	PID	Type	Process name	GPU Memory	
	ID	ID				Usage	

Şekil 2.7 GPU Bağlantı Kontrolü Kodu

2.7 YOLOv8 ile Eğitimi Gerçekleştirme

Eğitim ortamının ve veri setinin hazırlanmasının ardından eğitim işlemine başlanılır. İlk adım olarak Colab üzerinden veri setinin bulunduğu Google Drive hesabına erişim sağlanır. Bunun için aşağıdaki kod blokları sırasıyla çalıştırılır ve bağlantı izinleri verilir:

```
from google.colab import drive
```

```
drive.mount('/content/gdrive')
```

```
!ln -s /content/gdrive/My\ Drive/ /mydrive
```

Aşağıdaki komutla birlikte veri setinin bulunduğu proje dizinine gidilir:

```
%cd /mydrive/gender_detection/
```

Proje için gerekli kütüphaneler indirilir:

```
pip install ultralytics
```

```
from ultralytics import YOLO
```

```
model = YOLO()
```

```
model = YOLO("yolov8n.pt")
```

```
model.train(data="data.yaml", epochs=100)
```

Yukarıdaki kod bloğu, Ultralytics tarafından sağlanan YOLO modelini kullanarak bir nesne tespiti modeli eğitmek için kullanılır. Bu kod, YOLOv8n modelinin önceden eğitilmiş ağırlıklarını içeren yolov8n.pt dosyasını yükler. YOLOv8n, YOLO algoritmasının 8. versiyonunun "nano" (n) varyantıdır. "Nano" varyantı, daha küçük ve daha hızlı bir modeldir, genellikle daha az hesaplama gücüne sahip cihazlar veya daha hızlı çıkarım gereksinimleri için kullanılır.

Son satırda ise modeli, belirtilen dosyadaki (data.yaml) bilgiler üzerinde eğitir. data.yaml dosyası, eğitim veri setinin yolunu ve sınıf etiketleri gibi bilgileri içerir. Model, 100 epoch boyunca eğilecektir. Bir epoch, eğitim veri setinin bir tam geçişini ifade eder.

100 epoch'luk eğitim seçilmesinin sebebi, veri setinin büyük olmasıdır. Programın yeterince iyi öğrenmesi için epoch sayısının çok olması gerekir fakat elbette ki epoch sayısı çok fazla seçilirse aşırı uyum (overfitting) problemiyle karşılaşılabilir. Bunun için kontrollü bir şekilde işlemler sürdürülür.

3. BULGULAR VE TARTIŞMA

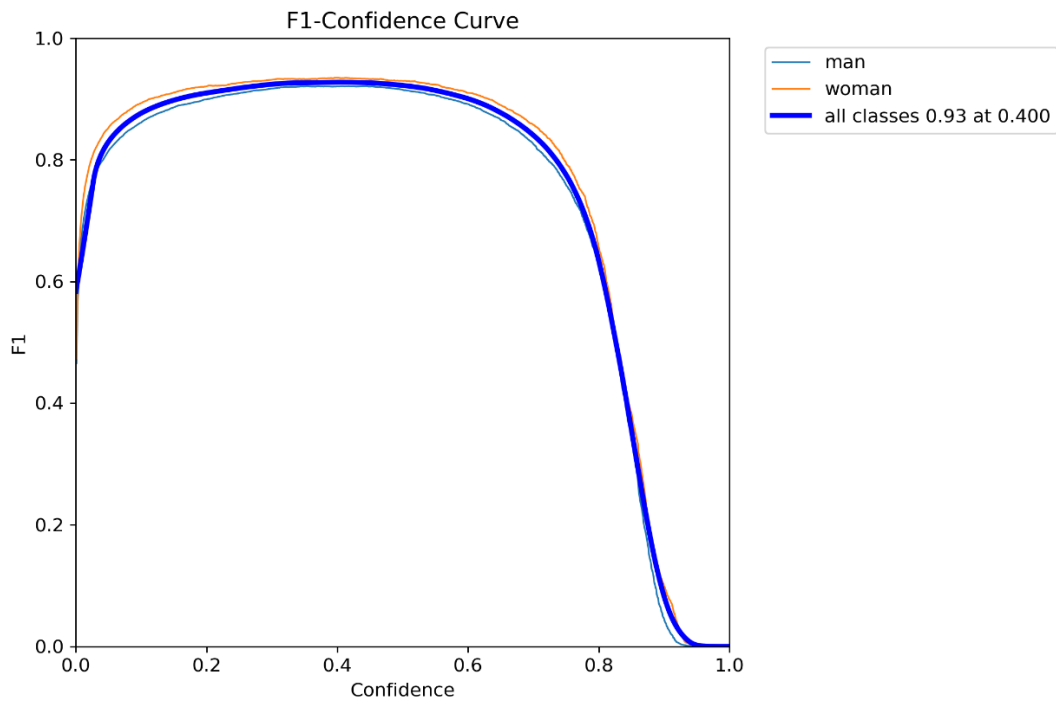
100 epoch'luk eğitim yaklaşık 8 saat sonunda sonlanmıştır. Eğitim sonunda aynı ortamda çekilmiş test videolarıyla program test edildiğinde programın başarı oranının yüksek olduğu görülmüştür.

Eğitim sonrasında bu programın performans ve doğruluğuyla ilgili detaylı bilgiye sahip olunabilen Karışıklık matrisi (confusion matrix) ve f1 skor değerleri incelendiğinde, programın doğruluk oranının %85.9 olduğu gözükmemektedir.

F1 skor, kesinlik (precision) ve duyarlılık (recall) değerlerinin harmonik ortalamasıdır ve modelin dengeleme kabiliyetini gösterir. Precision, modelin pozitif olarak tahmin ettiği örneklerin ne kadarının gerçekten pozitif olduğunu gösterir. Yani, modelin yaptığı pozitif tahminlerin doğruluk oranını ifade eder. Recall, gerçek pozitif örneklerin ne kadarının model tarafından doğru şekilde pozitif olarak tahmin edildiğini gösterir. Yani, pozitif örnekleri doğru şekilde bulma oranını ifade eder.

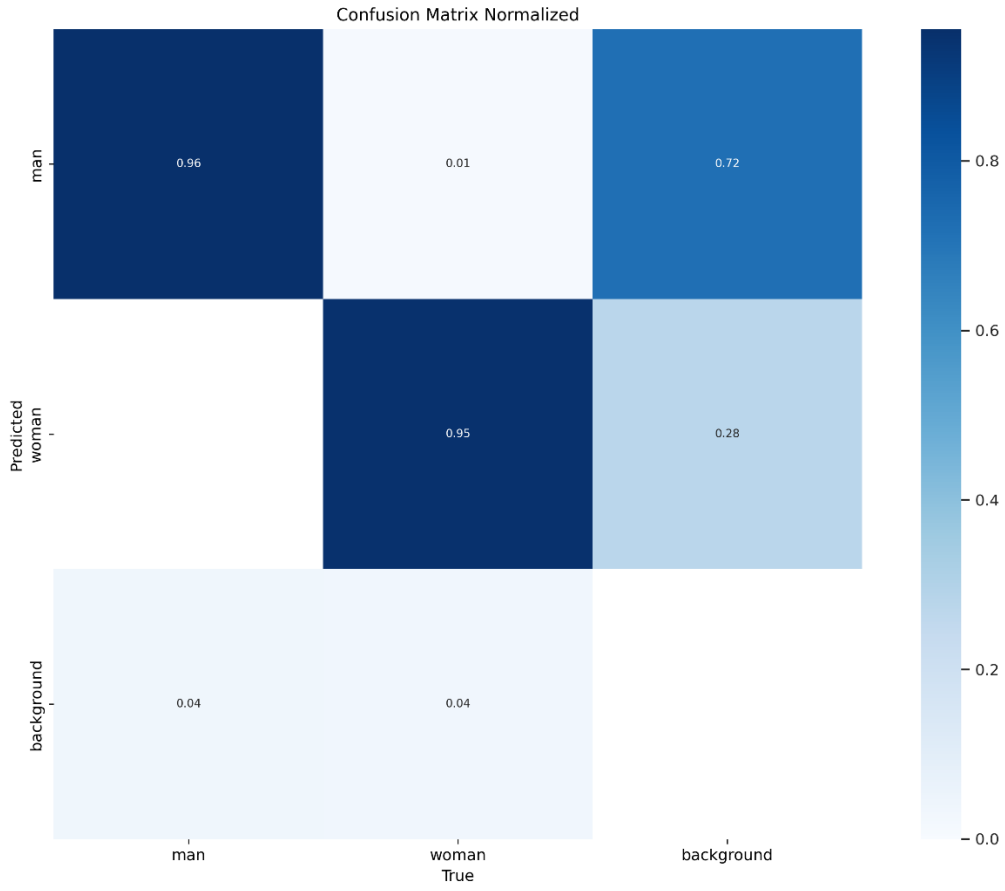
Şekil 3.1'de görünen F1 Güven eğrisi (F1-confidence curve), modelinizin farklı güven seviyelerinde F1 skorunu göstermektedir. Grafikteki kalın çizgi tüm sınıflar için ortalama F1 skorunu, ince çizgiler ise her sınıf için F1 skorunu göstermektedir.

Grafikte görülen optimal F1 skoru 0.93 ve bunun karşılık geldiği güven seviyesi 0.4 olarak belirtilmiştir yani, modelin tahminlerinin güven seviyesi en az %40 olduğunda en iyi F1 skoruna ulaşmaktadır. Bu güven seviyesinde, model hem doğru sınıflandırmalar yapma (precision) hem de mümkün olduğunca az hatalı negatif sonuç üretme (recall) konusunda en iyi dengeyi sağlamaktadır.



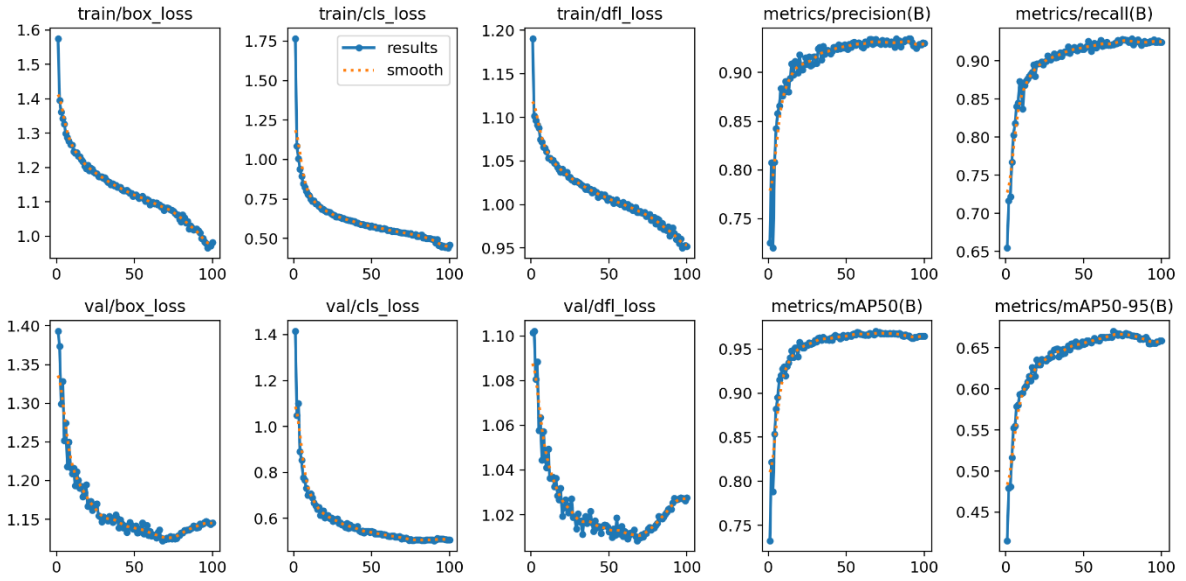
Şekil 3.1 F1 Güven Eğriliği Grafiği

İlgili eğitim sonucu elde edilen Karışıklık Matrisi (confusion matrix) görüntüsü Şekil 3.2’de verilmiştir:



Şekil 3.2 Karışıklık Matrisi

Şekil 3.3’te, modelin öğrenme sürecindeki eğitim ve doğrulama kayıplarıyla ilgili grafikler verilmiştir. Bu grafikler, modelin öğrenme ve doğrulama süreçleri hakkında genel yorum yapılabilmesine olanak tanır, aşırı uyum (overfitting) gibi ezbere öğrenme olup olmadığı üzerine yorum yapabilmemizi sağlar.



Şekil 3.3 Modelin Eğitimi Sonucu Elde Edilen Grafikler

Şekil 3.3'teki çıktılar incelenirse:

Kutu kaybı (Box loss): Nesne tespiti modellerinde, her bir nesnenin konumunu (genellikle sınırlayıcı kutular olarak adlandırılan dikdörtgen bölgelerin koordinatları) tahmin etmek önemlidir. Kutu kaybı, bu kutuların tahminlerinin gerçek kutularla ne kadar uyumlu olduğunu ölçer. Bu genellikle bir regresyon kaybı olarak formüle edilir ve öklid mesafesi veya benzeri bir ölçü kullanılarak hesaplanır.

Bu eğitim sonucunda, eğitim ve doğrulama kayıpları sürekli olarak azalmaktadır, bu da modelin öğrenme sürecinde iyileştiğini gösterir.

Sınıf kaybı (Class loss): Nesne tespiti modellerinde, her bir nesnenin hangi sınıfa ait olduğunu tahmin etmek de önemlidir. Sınıf kaybı, bu sınıf tahminlerinin gerçek etiketlerle ne kadar uyumlu olduğunu ölçer. Bu genellikle bir sınıflandırma kaybı olarak formüle edilir ve modelin nesneleri doğru sınıflara atamasını teşvik eder.

Bu eğitim sonucunda, eğitim ve doğrulama sınıf kayıpları da benzer şekilde azalmaktadır.

Özellik Seviye Mesafe kaybı (DFL loss): Genellikle yüz tanıma ve kişi takibi gibi görevlerde kullanılan bir kayıp fonksiyonudur. Bu kayıp fonksiyonu, bir kişinin yüzüne ait özellik vektörlerinin (genellikle bir yüz kodlayıcıdan elde edilir) birbirine ne kadar benzer olduğunu ölçer.

Bu eğitim sonucunda, değişken kayıplar da azalma eğilimindedir.

Kesinlik (Precision), Duyarlılık (Recall) ve mAP50: Bu metrikler zamanla artmakta, modelin doğruluğunun ve duyarlılığının iyileştiğini göstermektedir.

Genel olarak, model iyi bir performans sergilemektedir. Kesinlik, duyarlılık ve F1 skor gibi metrikler oldukça yüksektir ve bu da modelin hem doğruluğunun hem de duyarlılığının

yüksek olduğunu gösterir. Eğitim ve doğrulama kayıplarındaki sürekli azalma, modelin eğitim sürecinde iyileştiğini ve aşırı uyum (overfitting) gibi sorunların olmadığını göstermektedir.

Tablo 1’de, cinsiyete göre etiketleme sonucu eğitilen modelin Kesinlik (Precision), Duyarlılık (Recall), F1 Skor ve mAP50 değerlerine göre karşılaştırmalı sonuçları gösterilmiştir.

Tablo 1: YOLOv8 mimarisi ile doğrulama veri setinden elde edilen performans sonuçları

Sınıf	Kesinlik (Precision)	Duyarlılık (Recall)	F1 Skor	mAP50
Erkek (man)	0.883	0.956	0.918	0.965
Kadın (woman)	0.904	0.950	0.926	0.975
Ağırlıklı Ortalama (weighted average)	0.792	0.927	0.889	0.970

3.1 PolygonZone Teknolojisiyle İnsan Sayma

Sayma prensipleri genel olarak ikiye ayrılır. Bunlar, ROI ve LOI. Region of interest belirli bir alanı hedef seçer ve o alanda sayma işlemini yürütür. Line of interest ise belirli bir alan seçilmez, çizgi ile in-out benzeri yaklaşımla giren-çıkan olarak tüm alan değerlendirilir. Bu projede tercih edilen yaklaşım ROI ile Poligon (PolygoneZone) oluşturarak belirli bir alanda sayma işlemini gerçekleştirmektir.

Polygonezone metodu, Supervision kütüphanesi içinde, sıklıkla görüntü işleme uygulamalarında kullanılan bir fonksiyondur. Supervision kütüphanesi, genellikle bilgisayarla görme ve derin öğrenme alanlarında kullanılan Python kütüphanesidir. Polygonezone metodu, özellikle görüntü üzerinde belirli bölgeleri analiz etmek ve bu bölgelerdeki nesneleri sınıflandırmak veya tespit etmek amacıyla kullanılır.

Bu projede hedeflenen kalabalık analizi testlerinde belirli bir alandaki kalabalığı ölçmek istendiğinden, Supervision’ın PolygoneZone metodu kullanılmıştır.

PolygonZone’u projeye entegre ederken yine Google Colab üzerinde çalışma zamanı ayarlanır ve proje oluşturulur. Ardından gerekli kütüphaneler indirilir. Aşağıda projede ihtiyaç duyulan kütüphanelerin indirilme kodu gözükmemektedir:

```
!pip install ultralytics
import ultralytics
ultralytics.checks()
```

```
!python -m pip install 'git+https://github.com/facebookresearch/detectron2.git'
```

```
from IPython import display  
display.clear_output()
```

```
import detectron2  
print("detectron2:", detectron2.__version__)
```

```
!pip install supervision==0.2.0
```

```
from IPython import display  
display.clear_output()
```

```
import supervision as sv  
print("supervision", sv.__version__)
```

İlgili kütüphaneler kurulduktan sonra üzerinde test yapılacak video belirlenir ve sisteme yüklenir. Sisteme yüklendikten sonra videonun konum yolu aşağıdaki gibi verilir:

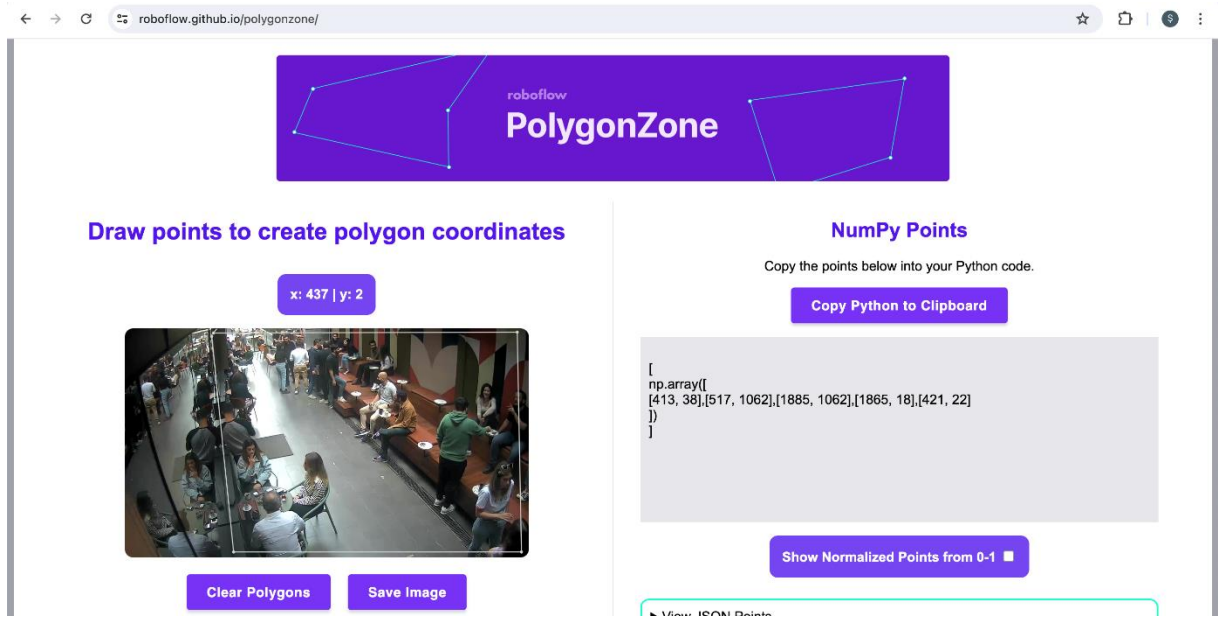
```
VIDEO_PATH = "/content/test-video-3.mp4"
```

Sonrasında, önceden eğitilmiş olan cinsiyet tespiti modelini yüklemek ve kullanıma hazır hale getirmek üzere aşağıdaki gibi modelin ağırlık dosyası olan “best.pt” dosyası yüklenir:

```
from ultralytics import YOLO  
  
model = YOLO('/content/best.pt')
```

Yüklenen ağırlık dosyası, modelin en iyi performans gösterdiği eğitim sonrası kaydedilen ağırlık dosyasıdır ve modelin yeni veriler üzerinde tahmin yapabilmesini sağlar.

Ağırlık dosyası başarılı bir şekilde yüklendikten sonra poligon çizimi ve cinsiyet tespiti aşaması öncesi, video üzerinde çizilecek poligonun koordinatlarının belirlenmesi gerekir. Bunun için Roboflow’un çizilecek poligonun koordinatlarını Python kodu olarak döndüğü <https://roboflow.github.io/polygonzone/> sitesinde test edilecek video üzerinde ilgili alanın çizimi yapılır ve koordinatlar alınır. Örnek işlem Şekil 3.4’te gösterildiği gibidir.



Şekil 3.4 Örnek Poligon Çizimi

Poligonun koordinatları da elde edildikten sonra artık yüklenen video üzerinde ilgili alan için poligon çizimi yapılır. Bu kısımda aşağıdaki kod bloğu kullanılır:

```
import numpy as np
import supervision as sv

np.bool_ = np.bool_

# initiate polygon zone
polygon = np.array([
    [390, 17], [500, 1080], [1900, 1080], [1900, 20], [390, 15]
])
video_info = sv.VideoInfo.from_video_path(MARKET_SQUARE_VIDEO_PATH)
zone = sv.PolygonZone(polygon=polygon, frame_resolution_wh=video_info.resolution_wh)

# initiate annotators
box_annotator = sv.BoxAnnotator(thickness=4, text_thickness=4, text_scale=2)
zone_annotator = sv.PolygonZoneAnnotator(zone=zone, color=sv.Color.white(),
    thickness=6, text_thickness=6, text_scale=4)

# extract video frame
generator = sv.get_video_frames_generator(MARKET_SQUARE_VIDEO_PATH)
iterator = iter(generator)
frame = next(iterator)
```

```

# detect
results = model(frame)[0]
detections = sv.Detections.from_yolov8(results)
detections = detections[(detections.class_id == 0) | (detections.class_id == 1)]
zone.trigger(detections=detections)

#color
color = sv.ColorPalette.from_hex(['#38b3e3', '#f06292'])

# annotate
box_annotator = sv.BoxAnnotator(thickness=4, text_thickness=4, text_scale=2, color=color)
labels = [f"{model.names[class_id]} {confidence:0.2f}" for _, confidence, class_id, _ in
detections]
frame = box_annotator.annotate(scene=frame, detections=detections, labels=labels)
frame = zone_annotator.annotate(scene=frame)

%matplotlib inline
sv.show_frame_in_notebook(frame, (16, 16))

```

Yukarıdaki kodla birlikte test videosunun ilk görüntüsü alınır ve burada best.pt modeline göre cinsiyet tespiti yapılır. Cinsiyet tespiti sonrası çerçeve içine alınan kişilerin cinsiyeti ve yüzde kaç doğruluk oranıyla bulunduğu bilgisi çerçevenin üstüne yazılır. Cinsiyete göre kadınsa pembe renkli bir çerçeve erkekse mavi renkli bir çerçeve ile renklendirilir. Ardından “PolygonZone” metoduyla birlikte belirlenen koordinat dizisine göre poligon oluşturulur. Tespiti yapılan kişilerin toplam sayısı poligonun ortasında sayıyla ifade edilir. Bu işlemler sonucunda elde edilen örnek görüntü Şekil 3.5’teki gibidir.



Şekil 3.5 Görüntü Üzerinden Cinsiyet Tespiti ve Kişi Sayımı

Fotoğraf üzerinde koordinatların ve tespitlerin kontrolünü yaptıktan sonra kalabalık analizinin video üzerinde yapılması için aşağıdaki kod bloğu çalıştırılır.

```
import numpy as np
import supervision as sv

# initiate polygon zone
polygon = np.array([
[390, 17],[500, 1080],[1900, 1080],[1900, 20],[390, 15]
])
video_info = sv.VideoInfo.from_video_path(MARKET_SQUARE_VIDEO_PATH)
zone = sv.PolygonZone(polygon=polygon, frame_resolution_wh=video_info.resolution_wh)

#Color
color = sv.ColorPalette.from_hex(['#38b3e3', '#f06292'])

# initiate annotators
box_annotator = sv.BoxAnnotator(thickness=4, text_thickness=4, text_scale=2, color=color)
zone_annotator = sv.PolygonZoneAnnotator(zone=zone, color=sv.Color.white(),
thickness=6, text_thickness=6, text_scale=4)

# Counter for female and male detections
female_count = 0
male_count = 0

def process_frame(frame: np.ndarray, _) -> np.ndarray:
    global female_count, male_count

    # Detect objects
    results = model(frame)[0]
    detections = sv.Detections.from_yolov8(results)
    detections = detections[(detections.class_id == 0) | (detections.class_id == 1)]
    zone.trigger(detections=detections)

    # Initialize counts
    female_count = 0
    male_count = 0

    # Dictionary to store the highest confidence class for each box
    highest_confidence_class = { }

    # Find the highest confidence class for each box
    for box, confidence, class_id, _ in detections:
```

```

# Convert box from numpy array to tuple to make it hashable
box_tuple = tuple(box)
if box_tuple not in highest_confidence_class or confidence >
highest_confidence_class[box_tuple][1]:
highest_confidence_class[box_tuple] = (class_id, confidence)

# Count female and male detections
for class_id, _ in highest_confidence_class.values():
if class_id == 1: # Female class
female_count += 1
elif class_id == 0: # Male class
male_count += 1

# Annotate the frame
labels = [f"{model.names[class_id]} {confidence:0.2f}" for _, confidence, class_id, _ in
detections]
frame = box_annotator.annotate(scene=frame, detections=detections, labels=labels)
frame = zone_annotator.annotate(scene=frame)

# Display female and male counts
total_count = female_count + male_count
cv2.putText(frame, f"Total Count: {total_count}", (20, 80),
cv2.FONT_HERSHEY_SIMPLEX, 1, (255, 255, 255), 2, cv2.LINE_AA)
cv2.putText(frame, f"Male Count: {male_count}", (20, 140),
cv2.FONT_HERSHEY_SIMPLEX, 1, (240, 98, 146), 2, cv2.LINE_AA)
cv2.putText(frame, f"Female Count: {female_count}", (20, 180),
cv2.FONT_HERSHEY_SIMPLEX, 1, (56, 179, 227), 2, cv2.LINE_AA)

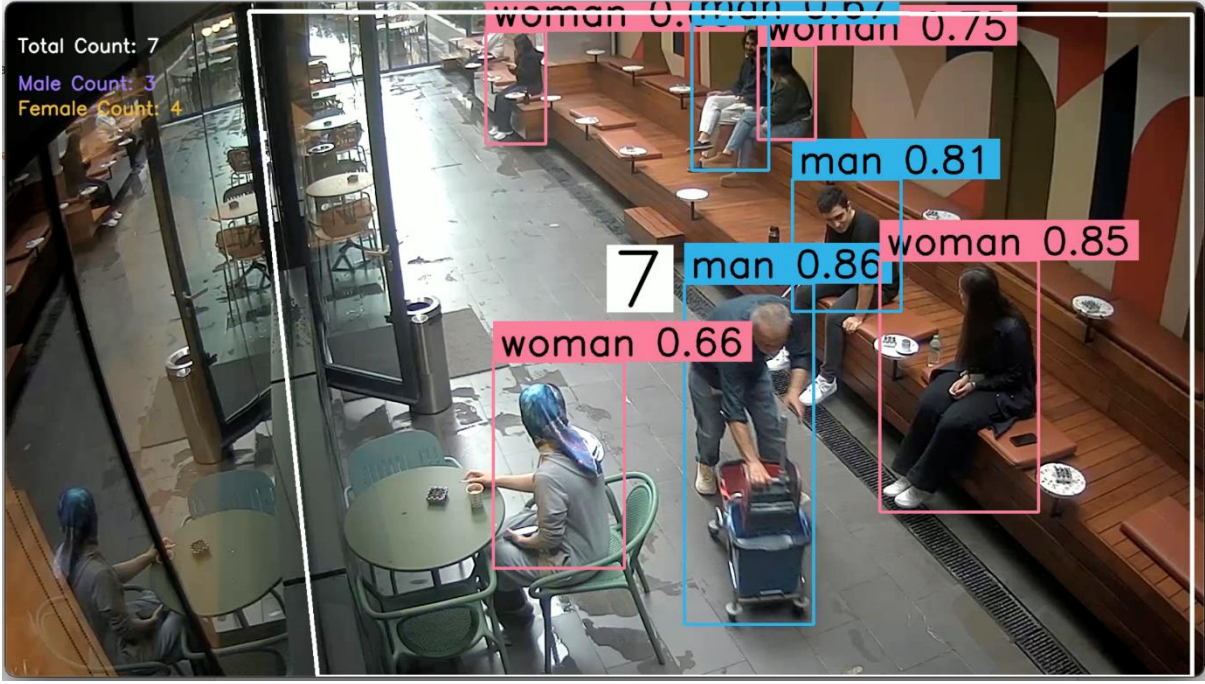
return frame

sv.process_video(source_path=MARKET_SQUARE_VIDEO_PATH,
target_path=f"{HOME}/test-video-4-colory.mp4", callback=process_frame)

from IPython import display
display.clear_output()

```

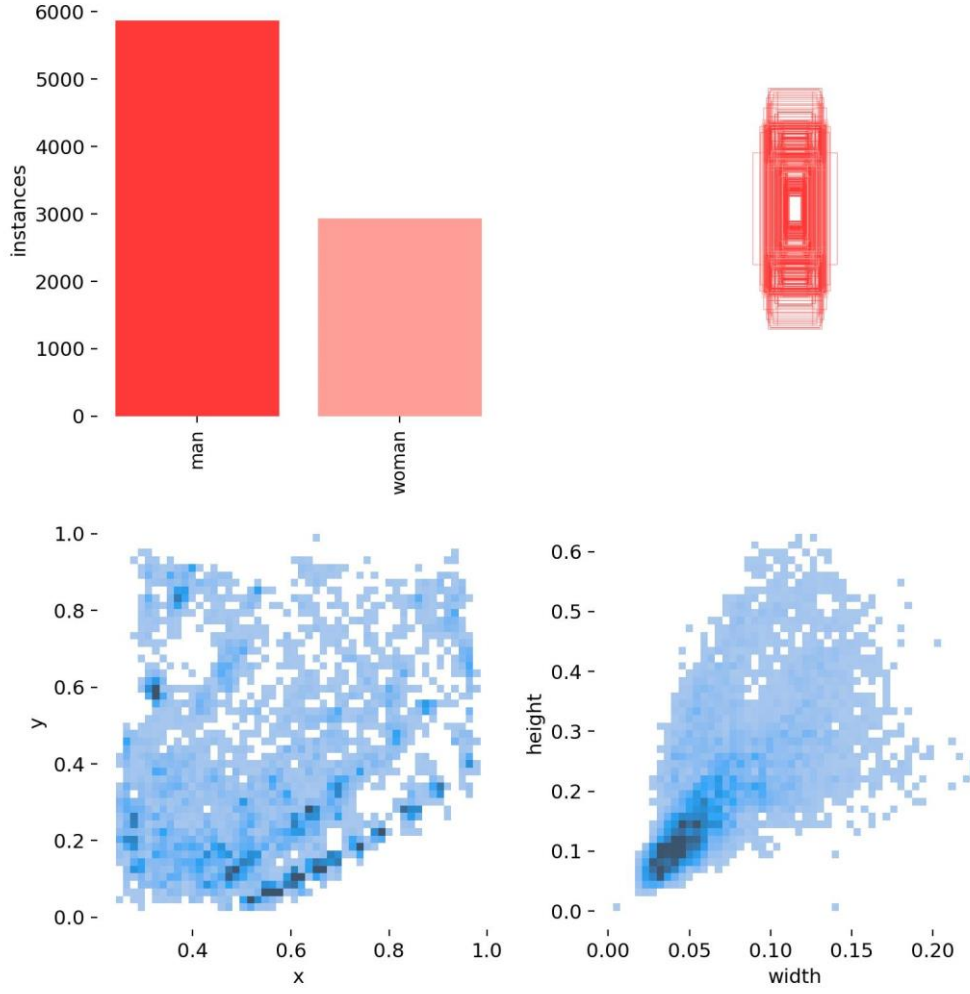
Yukarıdaki kod bloğu çalıştırıldıktan sonra video çerçevelere (frame) ayrılarak cinsiyet tespit ve sayım işlemleri yapılır. İşlem gerçekleştirilirken elde edilen kadın, erkek ve toplam kişi sayısı sol üst köşeye yazılmıştır.



Şekil 3.6 Video Üzerinden Cinsiyet Tespiti ve Kalabalık Analizi Sonuçları

Şekil 3.6’da görüldüğü üzere çizilen Poligon alanı içerisinde kalan kişiler, eğitilmiş model yardımıyla önce cinsiyet tespiti yapılarak etiketlenir ardından da kaç kişi olduğu bilgisi sayılarak sol üst köşede kalabalık oranı hakkında bilgi verilir.

Kamera açısının hep aynı kalmasından, veri setinde bulunulan ortamda insanların kalabalık olduğu durumlarda, cinsiyet etiketlemelerinde insanların tüm vücutları etiketlenemediği durumlar söz konusu olmuştur. Ya da kamera açısından kaynaklı olarak etiketlenen kişilerin görüntülerinin yeterince net olmadığı durumlarla karşılaşmıştır. Ayrıca Şekil 3.7’de görüldüğü gibi veri setinde etiketlenen kadın ile erkek oranlarının dengesizliği söz konusudur.



Şekil 3.7 Etiketlemede Kadın-Erkek Dağılımları

Yukarıda bahsedilen tüm bu sınırlayıcı etmenlerden kaynaklı olarak modelin bazı test videolarında yanlış cinsiyet tespiti ya da cinsiyet etiketinin birden fazla kez verilmesi gibi problemlerle karşılaşmıştır. Bundan sonraki çalışmalarda yukarıdaki etmenlere dikkat edilerek etiketleme ve geliştirme süreçlerinin ilerletilmesi, başarı oranının daha yüksek ve hata oranının daha az olacağı modeller elde edebilmek adına önemlidir.

4. SONUÇ

Bu çalışmada, YOLOv8 algoritması kullanılarak NTT Data şirketinin mola alanı görüntülerinin bulunduğu videolar kullanılarak 3000 fotoğraflık veri seti elde edilmiştir. Bu veri seti, cinsiyet tespiti yapılabilmesi adına erkek/kadın olmak üzere etiketlenmiştir. Etiketleme sonrası Google Colab'ın bulut ortamı kullanılarak YOLOv8 algoritması ile eğitilmiş ve Supervision'ın PolygonZone kütüphanesi kullanılarak video üzerinden kalabalık analizi yapılmıştır. Elde edilen sonuçlar incelendiğinde doğruluk oranı yaklaşık %86 oranında başarı sağlanmıştır. Ortam üzerinde kalabalıklık oranını cinsiyete göre ölçebilen bu proje, ileride daha da detaylandırılarak pek çok alanda fayda sağlayacaktır.

5. ARAŞTIRMANIN SINIRLAMALARI

1. **Yetersiz Veri Seti:** Veri setinin sınırlı olması, özellikle cinsiyet tespiti gibi hassas bir konuda modelin doğruluğunu etkileyebilir. Farklı ırklardan, yaş gruplarından ve vücut tiplerinden daha fazla örnek içeren bir veri seti elde etmek, modelin genelleme yeteneğini artırabilir.
2. **Kamera Kalitesi ve Açısı:** Video kayıtlarının kalitesiz olması veya kamera açılarının yetersiz olması, kişilerin tanıma sürecini zorlaştırabilir. Daha yüksek çözünürlüklü ve çeşitli açılardan çekilmiş görüntülerle çalışmak, modelin performansını artırabilir.
3. **Kalabalık Ortam ve Eksik Etiketleme:** Etiketleme sırasında ortamın kalabalık olması, bazı insanların tüm vücutlarının doğru şekilde etiketlenememesine neden olabilir. Bu durum, modelin yanlış sınıflandırma yapmasına yol açabilir. Daha fazla dikkatle ve titizlikle etiketleme yapılması veya otomatik etiketleme yöntemlerinin kullanılması, bu sorunu azaltabilir.
4. **Dengesiz Etiket Dağılımı:** Etiketleme sırasında erkeklerin kadınlara oranla daha fazla olması, modelin eğitiminde dengesizlik yaratabilir. Daha dengeli bir veri seti oluşturmak, modelin her iki sınıf üzerinde daha iyi performans göstermesini sağlayabilir.

6. İLERİYE YÖNELİK ÖNERİLER

1. **Veri Seti Zenginleştirme:** Farklı koşullarda çekilmiş, çeşitli insanlardan ve ortamlardan daha fazla veri eklemek, modelin genelleme yeteneğini artırabilir.
2. **Daha İyi Etiketleme Süreci:** Otomatik etiketleme araçları veya insan denetimli etiketleme süreçleri kullanarak daha doğru ve eksiksiz etiketleme yapmak, modelin performansını artırabilir.
3. **Model Geliştirme:** YOLOv8 algoritması üzerinde yapılan iyileştirmeler veya farklı bir nesne tespiti modeli kullanmak, daha doğru ve güvenilir sonuçlar elde etmeyi sağlayabilir.
4. **Çoklu Etiketleme Sorununu Çözme:** Modelin aynı kişiye birden fazla etiket verme sorununu çözmek için daha gelişmiş bir nesne tespiti yöntemi veya etiketleme düzeltme algoritmaları kullanılabilir.
5. **Performans Değerlendirmesi:** Elde edilen sonuçları daha detaylı incelemek ve modelin yanlış sınıflandırmalarını analiz etmek, gelecekteki çalışmalar için yol haritası çizebilir.

Bu öneriler, mevcut çalışmanın sınırlamalarını aşmak ve gelecekteki araştırmalara temel oluşturmak için kullanılabilir.

KAYNAKÇA

- [1] Liu, Z., Chen, Y., Chen, B., Zhu, L., Wu, D., & Shen, G. (2019). Crowd counting method based on convolutional neural network with global density feature. *IEEE Access*, 7, 88789-88798.
- [2] Keh, J. J., Cruz, M., Deticio, R., Tan, C. V., Jose, J. A., Dadios, E., & Fillone, A. (2020, December). Video-Based Gender Profiling on Challenging Camera Viewpoint for Restaurant Data Analytics. In *2020 IEEE 12th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment, and Management (HNICEM)* (pp. 1-5). IEEE.
- [3] Ahmed, A., Bansal, P., Khan, A., & Purohit, N. (2021, August). Crowd Detection and Analysis for Surveillance Videos using Deep Learning. In *2021 Second International Conference on Electronics and Sustainable Communication Systems (ICESC)* (pp. 1-7). IEEE.
- [4] Muzamal, J. H., Tariq, Z., & Khan, U. G. (2019, August). Crowd Counting with respect to Age and Gender by using Faster R-CNN based Detection. In *2019 International Conference on Applied and Engineering Mathematics (ICAEM)* (pp. 157-161). IEEE.
- [5] Raza, M., Sharif, M., Yasmin, M., Khan, M. A., Saba, T., & Fernandes, S. L. (2018). Appearance based pedestrians' gender recognition by employing stacked auto encoders in deep learning. *Future Generation Computer Systems*, 88, 28-39.
- [6] Carletti, V., Greco, A., Saggese, A., & Vento, M. (2020). An effective real time gender recognition system for smart cameras. *Journal of Ambient Intelligence and Humanized Computing*, 11(6), 2407-2419.
- [7] Velastin, S. A., Fernández, R., Espinosa, J. E., & Bay, A. (2020). Detecting, tracking and counting people getting on/off a metropolitan train using a standard video camera. *Sensors*, 20(21), 6251.
- [8] Peker, M., İnci, B., Musaoğlu, E., Çobanoğlu, H., Kocakır, N., & Karademir, Ö. İnsan Sayma Sistemi için Gömülü Cihazlarda Derin Öğrenme Mimarilerinin Performans Analizi.
- [9] Karahan, M., Lacinkaya, F., Erdonmez, K., Eminagaoglu, E. D., & Kasnakoglu, C. (2022). Age and gender classification from facial features and object detection with machine learning. *Journal of fuzzy extension and applications*, 3(3), 219-230.
- [10] Susa, J. A. B., Doculan, J. A. D., Merin, J. V., Zarate, J. M., Tria, M. L., Evangelista, R. S., & Reyes, M. C. (2022, December). Deep Neural Network-Based Gender Identification for Surveillance Restroom Restriction System. In *2022 International Conference on Emerging Technologies in Electronics, Computing and Communication (ICETECC)* (pp. 1-6). IEEE.
- [11] Varnima, E. K., & Ramachandran, C. (2020, June). Real-time Gender Identification from Face Images using you only look once (yolo). In *2020 4th International Conference on Trends in Electronics and Informatics (ICOEI)*(48184) (pp. 1074-1077). IEEE.

- [12] Zheng, G., & Xu, Y. (2021). Efficient face detection and tracking in video sequences based on deep learning. *Information Sciences*, 568, 265-285.
- [13] Srazhdinova, A., & Anvarov, S. (2020). Detection and tracking people in real-time with yolo object detector. *Challenges of Science*, 10(2020.010).
- [14] Daş, R., Polat, B., & Tuna, G. (2019). Derin öğrenme ile resim ve videolarda nesnelerin tanınması ve takibi. *Fırat Üniversitesi Mühendislik Bilimleri Dergisi*, 31(2), 571-581.
- [15] de Souza Inácio, A., Ramos, R. H., & Lopes, H. S. Deep Learning for People Counting in Videos by Age and Gender.

ÖZGEÇMİŞ

ŞULE AKTÜRK – JUNIOR FRONTEND DEVELOPER

Hello! My name is Şule. I am currently in the 4rd year of Computer Engineering at Marmara University. I am also a Junior Frontend Developer at Doğuş Teknoloji.

Education

Bachelor's Degree in Computer Engineering, Marmara University | 2020 - 2024

Professional Experience

Doğuş Teknoloji | September 2022 - Present

1.NetworkDry Laundry Firm Website Development

- Actively involved in Frontend development using Next.js and TypeScript to build a responsive and dynamic website.
- Optimized Google Tag Manager, Google Analytics, and SEO for enhanced performance and user insights.
- Implemented API interactions via Postman and established OAuth connections for Facebook and Google integrations.
- Focused on visually appealing designs with Tailwind CSS and integrated Firebase for a scalable architecture.

2.D-ONE Dashboard Web Development

- Utilized JavaScript, React, and AntDesign to create a user-friendly and intuitive dashboard.
- Designed and refined interfaces using Agile methods, managing API connections with Swagger.
- Employed React-Chart.js for data visualization and ensured the seamless transition from design to coding.

3.BugHunty Cyber Security Website

- Played an active role in developing the BugHunty Cyber Security Platform using Agile methodologies and Scrum.
- Developed new features, fixed bugs, and ensured web pages balanced functionality and aesthetics.

- Used React.js, TypeScript, and CSS for user-friendly web pages, and collaborated with cross-functional teams for timely project completion.
- Created high-quality mockups and prototypes, providing top-notch graphics and visual elements.

Technical Skills

- Languages: C, Java, Python, JavaScript, TypeScript, HTML, CSS, SQLite, PostgreSQL
- Frameworks & Tools: React, Next.js, Tailwind CSS, Ant Design, Swagger, Postman, Docker, Figma, Storybook
- Technologies: OAuth Connections, CI/CD, Azure DevOps
- Expertise: Frontend Development, Web Design, Wireframe & Design, Agile Methodologies, Project Management Tools
- Soft Skills: Strong Communication, Problem-Solving, Team Collaboration, Time Management, Adaptability, Attention to Detail, Creative Thinking, Leadership, Initiative and Proactiveness

In my role at Doğuş Teknoloji, I have honed my technical skills and demonstrated my ability to contribute effectively to various projects, combining creativity with technical expertise to deliver high-quality results.

Phone: +90 538 595 15 02

City: Istanbul / Turkey

Email: sule.aktrk1973@gmail.com

Linkedin: <https://www.linkedin.com/in/suleakturk/>

Github: <https://github.com/suleakturk>