

**EGE ÜNİVERSİTESİ**

**MÜHENDİSLİK FAKÜLTESİ**

**BİLGİSAYAR MÜHENDİSLİĞİ BÖLÜMÜ**

**YAPAY ZEKÂ YÖNTEMLERİ (3+0)**

**2022-2023 BAHAR YARIYILI**

**PROJE-2 ARA RAPORU**

**TESLİM TARİHİ**

9/06/2023

**HAZIRLAYANLAR**

05200000096, Uğur Alp Balta

05200000761, Berat Duran

İçindekiler

[1) Problemin Tanımı 2](#_Toc137919084)

[2) Araştırma 2](#_Toc137919085)

[2.a) Linkler Ve Kaynaklar 2](#_Toc137919086)

[2.b) neler Öğrendik? 2](#_Toc137919087)

[3) Kullanılan Ortam ve Kütüphaneler 3](#_Toc137919088)

[4) Kullanılan Yöntem ve Özellikleri 4](#_Toc137919089)

[5) Deneyesel Çalışmalar 6](#_Toc137919090)

[5.a) Veriseti Nasıl Elde Edildi? 6](#_Toc137919091)

[5.b) Test: 6](#_Toc137919092)

[5.c) Tablolar, Grafikler, vb. 7](#_Toc137919093)

[6) Sonuç 10](#_Toc137919094)

[7) Ekler 11](#_Toc137919095)

[7.a)Başarım Yöntemleri 11](#_Toc137919096)

[7.b) Katkı 11](#_Toc137919097)

[7.c) Faydalanılan Hazır Kodlar 11](#_Toc137919098)

[7.d) Tanımlamalar 11](#_Toc137919099)

[8) Öz değerlendirme Tablosu 13](#_Toc137919100)

# 1) Problemin Tanımı

Kahve festivalinde belli bir anda içilen kahve sayısının drone yardımıyla bulunmasını amaçladığımız ödevimizde öncelikle soğuk kahve bulması konusunda YOLO v8 kütüphanesi kullanarak gerçekleştirmeyi amaçladık. Ardından bir dronu kahve içmekte olan kişilere yaklaşıp onları fotoğraflaması için programlamayı denedik. Projemiz tamamlanmış olup drone uçuşu özelliğinde eksiklikler mevcuttur.

# 2) Araştırma

## 2.a) Linkler Ve Kaynaklar

<https://www.youtube.com/watch?v=m9fH9OWn8YM&t=2443s>

<https://www.youtube.com/watch?v=gRAyOPjQ9_s&t=93s>

<https://docs.ultralytics.com/>

Boiler plate bazı kodlar ve bazı istenenlerin anlaşılması için ChatGpt kullanılmıştır.

## 2.b) neler Öğrendik?

YOLOv8, nesne algılama modeli için geniş ve çeşitli bir veri kümesi gerektirdiğini. Dataset, örnek görüntülerden ve her görüntüdeki nesnelerin etiketlerinden oluştuğunu öğrendik bunun dışında Dataset yapısı aşağıdaki bileşenleri içeriyor.

Görüntüler: Eğitim için kullanılacak gerçek dünya görüntüleri veya sentetik olarak oluşturulan görüntülerden oluşur.

Etiketler: Her görüntüdeki nesnelerin etiketleri. Etiketler, nesnenin sınıfını ve sınırlayıcı kutusunun (bounding box) koordinatlarını içerir.

**Train Parametreleri:**

YOLOv8 modelini eğitirken aşağıdaki parametrelerin ayarlanması önemlidir:

Batch Size: Eğitim için kullanılacak her bir veri yığınındaki görüntü sayısı. Daha büyük batch size'lar genellikle daha hızlı eğitim sağlar, ancak bellek sınırlamaları dikkate alınmalıdır.

Epochs: Eğitimde kullanılacak toplam geçit sayısı. Daha yüksek epoch sayısı, daha fazla öğrenme sağlayabilir, ancak aşırı öğrenmeye neden olabilir.

Learning Rate: Ağırlıkların güncellenme miktarını kontrol eden parametre. Uygun bir learning rate, modelin daha hızlı ve daha iyi bir şekilde öğrenmesini sağlar.

Optimizer: Ağırlıkların güncellenmesi için kullanılan optimizasyon algoritması. Yaygın olarak kullanılan optimizasyon algoritmaları arasında Adam, SGD (Stochastic Gradient Descent) ve RMSProp bulunur.

**Dataset Labeling:**

YOLOv8 için dataset oluştururken labeling süreci oldukça önemlidir. Labeling, her nesnenin sınıfını ve sınırlayıcı kutusunun koordinatlarını belirlemeyi içerir. YOLOv8, bounding boxların merkez koordinatlarını, genişliklerini, yüksekliklerini ve nesne sınıflarını tahmin eder. Labeling sürecinde aşağıdaki adımlar takip edilir:

Görüntüyü açın ve üzerindeki nesneleri tanıtın.

Her nesne için sınıf etiketini ve bounding box koordinatlarını belirleyin.

Bounding box koordinatlarını normalize edilir. Bunu labelImg kendisi yapabiliyor.

Labeling işleminde yapılan tercihler modelin eğiitm sürecinde etkilidir. Bizim örneğimizde buzlu kahve tanıma yapmayı amaçladık. Veri seti oluşturulurken yeterince veri yoktu ve verilerin geneli masa üstünde duran ve özellikle çevresinde dikkat dağıtıcı objeler olmayan fotoğrafları ve bu da modelin bu şekilde bir eğitim almasına yol açtı. Yeterince kahve içen insan veya kahve tutan insan fotoğrafı olmadığı için tanıma aşamasında kahve ile temasta bulunulduğunda tanıtım işlemi kötü sonuçlar vermeye başladı.

# 3) Kullanılan Ortam ve Kütüphaneler

Kullanılan ortam PyCharm Proffesional

classification için yolov8 kendimizin oluşturduğu ve label işlemi yaptığı set üzerinden eğitilmiştir. Görseller üzerinde bir eğitim olduğu için makine görmesi kullanıldığı söylenebilir. YOLO, torch ve cv2 projede kullanılan ana kütüphanelerdir. cv2 modelin performansını gözlemlemek ve sonuçları görsele dökmek için kullanılırken. torch gpu ile eğitimin hızlandırılması için kullanılmıştır. YOLO.train() eğitimin fonksiyonudur.

Torch - 2.0.1+cu117

YOLO -YOLOV8

cv2 - 4.7.0

## 

# 4) Kullanılan Yöntem ve Özellikleri

Projemiz Bilgisayarlı görü alanında olup. Özel nesne algılama kategorisi içindedir.

Projemizde yolov8l.pt modelini tercih ettik.

**Şema ve Modüller:**

metin, ekran görüntüsü, menü içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin, ekran görüntüsü, web sayfası, yazılım içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Parametreler:**

metin, ekran görüntüsü, yazı tipi, bilgi, enformasyon içeren bir resim

Açıklama otomatik olarak oluşturuldu

# 5) Deneyesel Çalışmalar

## 5.a) Veriseti Nasıl Elde Edildi?

Veristeni “google\_images\_download” kütüphanesni kullanarak googledan çektik ve hatalı fotoğrafları ayıkladıktan sonra labelImg yardımıyla fotoğrafları etiketledik. labelImg’in içersindeki YOLOv8’e formatındaki txt dosyasını kaydettik. Config.yaml dosyasını içerisine classımızı isimlendirdik ve train, validatiton, test setlerinin belirledik.

Verisetimiz farklı büyüklüklerdeki 100 fotoğraf üzerinden başlayıp indirilemeyen resimler sebebiyle yaklaşık 60 fotoğrafa kadar düşmüştür. Veriseti genel olarak etkileşimsiz kahve fotoğrafları üzerinden oluşup sadece soğuk kahveleri class olarak tanımlamıştır. Fotoğraflardaki çoklu kahveler hesaba katıldıuğında 100 küsüre yakın label içermektedir.

YOLO.train parametreleri

epoch, eğitilen epok sayısı

patience, değişim gözlemlenmeyen epokların sayısı

batch, batch başına düşen fotoğraf sayısı

imgsz, girdi fotoğraflarının

workers, veri yüklemesi başı çalışan işçi thread sayısı

lr0, anlık gelişim oranı

## 5.b) Test:

3 tür test yapılmıştır. Oluşturulan %29’luk bir test veriseti üzerinde denenmiştir.

Bazı grafikler gelişimin daha net görünmesi amacıyla eski hallerinde bırakılmıştır.

**Fotoğraf üzerine yapılan testler:**

**Ara Rapor:**

yemek, gıda, alkolsüz içki, fincan, kupa, kahve içeren bir resim

Açıklama otomatik olarak oluşturuldu

**Final:**

**yemek, gıda, fincan, kupa, alkolsüz içki, kahve içeren bir resim

Açıklama otomatik olarak oluşturuldu**

**Video üzerine yapılan testler:**

****

**Gerçek zamanlı yapılan testler:**



## 

metin, kolaj, kahve, yemek, gıda içeren bir resim

Açıklama otomatik olarak oluşturuldu

metin, diyagram, öykü gelişim çizgisi; kumpas; grafiğini çıkarma içeren bir resim

Açıklama otomatik olarak oluşturuldu

Görüldüğü üzere loss grafikleri, genel bir düşüş içerisindedir. Precision ve recall grafikleri 0.9’un üzerinde kalmıştır.

metin, ekran görüntüsü, dikdörtgen, diyagram içeren bir resim

Açıklama otomatik olarak oluşturuldu

Hataların bir kısmı labellingden kaynaklı olup aslında labeller’ın hatasıdır ve görüştürme işleme labeller’ın bu hatasını belli etmiştir.

diyagram, ekran görüntüsü, tasarım içeren bir resim

Açıklama otomatik olarak oluşturuldu

# 6) Sonuç

Proje başarılı bir şekilde istenilen objeyi tanımlayabilmektedir, istenilen işe göre datasetin güncellenmesi kanısına varılmıştır. Proje gerçek hayatta reklam amaçlarıyla kullanılabilmekle beraber objeler değiştirilip daha hayati işlerde de kullanılabilir. Örnek olarak orman yangını tespiti, PPE(Personal Protection Equipment) tanımlama gibi. Bu şekilde güvenlik sistemlerinin içerisine eklenilebilir.

Bu özel nesne algılama projesinin bizim için kazanımları arasında: nesneleri tanıma ve sınıflandırma becerilerimizi geliştirmesi ve derin öğrenme ve görüntü işleme alanındaki bilgilerimizi uygulamaya dökmeye sağlamamızı, Veri etiketleme süreci sayesinde etkili veri yönetimi ve ön işleme becerileri kazanmamızı, Modelin performansını iyileştirmek ve hata analizi yapmak için uğraşmamız dolayısıyla belirli optimizasyon becerileri kazanmamızı, Proje sırasında karşılaştığımız teknik zorlukları aşarak problem çözme yeteneklerimizi geliştirmemizi sağlamıştır.

# 7) Ekler

## 7.a)Başarım Yöntemleri,

model.train(data="config.yaml", epochs=50, batch = 16, imgsz=480, patience=0) bu satırdaki parametrelerle oynamak başarım oranını arttırmıştır.(örnek olarak eğitim sayısı). Tranformative learning yöntemi kulanılarak eğitilen veri üzeinden modelin tekrar tekrar eğitilmesi ve Batchlerin eğitim başına değişmesi gibi faktörler sonucu az sayıda verimiz olmasına rağmen. Çıplak gözle görülebildiği kadarıyla %15’ten fazla bir artış yakalamayı başardık.

## 7.b) Katkı

Bu çalışmaya benzer çalışmalar var olsa da kahve festivallerinde soğuk kahve tanımlama yönünde başka bir çalışma görmediğimizden projemiz bu niş alanda bir farklılık göstermektedir.

## 7.c) Faydalanılan Hazır Kodlar

Kütüphaneler haricinde herhangi bir hazır koddan faydalanılmamıştır. Ayrıca belirtildiği üzere bazı boiler plate kodlarda chatGPT’den faydalanılmıştır.

## 7.d) Tanımlamalar

#### 7.d.1) Kümeleme Algoritmaları için kullanılan performans ölçütlerini (clustering performance evaluation metrics) öğrenip iki tanesini kendi cümlelerinizle anlatınız.

İyi bir kümeleme algoritmasının kümesinin kümeler arası değişkenliği az olan (bir kümedeki veri noktaları birbirine benzer) ve kümeler arası değişkenliği fazla olan(kümeler diğer kümelerden farklı olması) kümelere sahip olması istenir.

Kümeleme için iki tür değerlendirme ölçütü vardır:

Dışsal Ölçütler: Bu ölçütler gerçek etiketlere ihtiyaç duyar ve pratikte bulunmayabilir.

İçsel Ölçütler: Bu ölçütler gerçek etiketlere ihtiyaç duymaz (tüm denetimsiz öğrenme sonuçlarına uygulanabilir)."

**Mutual Information (MI, NMI, AMI):** Küme atamaları arasındaki uyumu ölçen budışsal kümeleme ölçütüleri kümeleme sonuçlarının etiketlere olan benzerliklerini ölçerek performans değerlendirmesine yardımcı olurlar. Gerçek etiket değerlerine sahip olunmadığında kullanılmamalıdırlar.

3 türü bulunur:

Mutual Information (MI): MI, iki rastgele değişken arasındaki bağımlılığı ölçer. Kümeleme durumunda, MI, kümeleme sonuçlarındaki etiketlerle gerçek etiketler arasındaki ilişkiyi ölçmek için kullanılır. Daha yüksek MI değerleri, kümeleme sonuçlarının gerçek etiketlere daha yakın olduğunu gösterir. MI değeri, 0 ile maksimum değeri arasında değişir ve yüksek değerler daha iyi bir uyumu temsil eder.

Normalized Mutual Information (NMI): NMI, MI'nin normalleştirilmiş bir versiyonudur. NMI, kümeleme sonuçlarındaki etiketlerle gerçek etiketler arasındaki benzerliği ölçmek için kullanılır ve 0 ile 1 aralığında değerler alır. NMI, kümeleme sonuçlarındaki etiketlerin sınıf etiketleriyle uyumu gösterir. 1'e yakın bir NMI değeri, daha iyi bir uyumu temsil eder.

Adjusted Mutual Information (AMI): AMI, NMI'ye benzer şekilde kümeleme sonuçlarının gerçek etiketlerle olan uyumunu ölçer. Ancak AMI, kümeleme sonuçlarının rastgele performansla karşılaştırıldığında elde edilen uyumu dikkate alır. Rastgelelik düşünüldüğünde beklenen uyumu düzeltir. AMI yine 0 ile 1 aralığında değer alır ve 1'e yakın bir değer yine daha iyi bir uyumu temsil etmektedir.

**Silhouette Coefficient:** Kümeleme algoritmalarının performanslarını değerlendirmek için kullanılan bu yöntemde kümeleme sonuçlarının ne kadar homojen ve ayırt edilebilir olduğu ölçülür. Kullanmadan önce dikkat edilmesi gerekir ki; Eğer farklı tipteki algoritmaları karşılaştırıyorsanız yoğunluk temelli kümeleme algoritmaları için daha yüksek puan verdiğinden adil bir karşılaştırma olmayacaktır.

Silhouette Coefficient, her bir örnek için hesaplanır ve örneklerin kendi kümesindeki komşu örneklerle olan benzerliğini, ayrıca örneklerin komşu kümedeki örneklerle olan benzerliğini göz önünde bulundurur. Bir örneğin Silhouette Coefficient, şu formülle hesaplanır:

s(i) = (b(i) - a(i)) / max(a(i), b(i))

Burada, a(i) örneğin kendi kümesindeki diğer örneklerle olan benzerliğini temsil ederken, b(i) örneğin komşu kümedeki örneklerle olan benzerliği temsil eder. s(i) değeri -1 ile 1 arasında bir değer alır. Pozitif değerler, örneklerin doğru kümeye ait olduğunu, sıfır değeri ise örneklerin çakıştığını ve negatif değerler ise örneklerin yanlış kümelere ait olduğunu gösterir. Dolayısıyla, daha yüksek Silhouette Coefficient değerleri, daha iyi bir kümeleme yapısını yansıtır.

Sonuç olarak Silhouette Coefficient tüm örneklerin Silhouette Coefficient değerlerinin ortalaması alınarak hesaplanır.

**Kaynak:** <https://towardsdatascience.com/7-evaluation-metrics-for-clustering-algorithms-bdc537ff54d2>

#### 7.d.2) Derin Öğrenmede Transformer Model nedir? Tanımlayınız

Doğal dil işleme (NLP) ve diğer makine öğrenimi alanlarında kullanılan bir derin öğrenme modelidir. Bu model, özellikle dil işleme görevlerinde (metin çevirisi, metin sınıflandırma, konuşma tanıma vb.) büyük başarı elde etmiştir. Önceki NLP modellerinin aksine recursive veya convolutional yapılar yerine tamamen attention mekanizmalarına dayanır.

İki temel bileşenden oluşan model, encoder (kodlayıcı) ve decoder (çözücü) şeklinde ayırılır.

Encoder, giriş metnini bir dizi temsil vektörüne dönüştürürken, decoder, bu temsil vektörünü hedef metni oluşturmak için kullanır. Her iki bileşen de birden çok "transformer bloğu"ndan oluşur.

Her bir transformer bloğu ise yine iki alt katmana ayrılır ki bu katmanlar, multi-head attention mekanizması ve fully connected layer şeklinde ifade edilirler.

Multi-head attention mekanizması, girdi metninin içerdiği bilgilere dikkatini odaklayarak ilişkileri ve önemli özellikleri öğrenir. Fully connected layer’lar ise, dikkat mekanizmasının çıktılarını kullanarak sonuçları tahmin etmek veya metin oluşturmak için kullanılır.

Transformer modeli, büyük veri kümeleri üzerinde eğitilerek dil işleme görevlerinde yüksek performans gösterebilir. Ayrıca, paralel hesaplama yeteneklerini kullanarak GPU'lar üzerinde hızlı bir şekilde eğitilebilirler. Bu nedenle, büyük ölçekli dil işleme projelerinde sıkça tercih edilen bir modeldir. (örnek olarak; Google Translate/ OpenAI GPT)

**Kaynak:** <https://blogs.nvidia.com/blog/2022/03/25/what-is-a-transformer-model/>

# 8) Öz değerlendirme Tablosu

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
|  | **İstenen Madde** | **Var** | **Açıklama** | **Tahmini Not** |
| **1** | **Kapak Sayfası, Problemin Tanımı, Kullanılan Ortam, Yöntem ve Kütüphaneler, Araştırma (10)** |  | Seçmiş olduğumuz farklı problem dolasıyla tercih ettiğimiz ortam ve araştırmış olduğumuz kaynaklar açıklandı. | 10 |
| **2** | **Önerilen Yöntem (10)** |  | Bilgisayarlı alanındaki projemiz için yolov8 tercih etmiş olup bu model hakkındakişbazı detayları (şemavb.) ekledik. | 10 |
| **3** | **Deneysel Çalışmalar (10)** |  | Verisetini nasıl elde ettiğimize dair detaylı açıklamayla beraber farklı kaynaklar üzerindeki test aşamalarımız da listelendi. | 10 |
| **4** | **Proje Rapor Biçimi, Organizasyonu, Boyutu, Kalitesi, Kaynakça ve atıflar (10)** |  | Eski rapor templatini güncellediğimiz çalışmada kaynakçalar ilgili soruların içlerinde belirtildi. | 10 |
| **5** | **Sonuç (10)** |  | Elde ettiğimiz sonuç dolayısıyla projemizin işlevliliği ve edindiğimiz kazanımlar hakkında | 10 |
| **6** | **Ek 1: Başarım İyileştirme (10)** |  | Bazı parametreler ile oynayıp model üstündeki eğitim şeklini transformative learning geçirerek daha iyi sonuçlar elde etmeyi başardık. Fakat ara rapordan bu yana final dönemi dolayısyla fazla zamanımız bulunmadığından daha etkin başarı geliştirme yöntemlerine yönelemedik. Bu nedenle %20’lik bir puan eksiltmesine gidildi. | 8 |
| **7** | **Ek 2 (10)** |  | Projemizin bernzer projelere olan farklılıklarından kısaca bahsettik. | 10 |
| **8** | **Ek 3 (10)** |  | Herhangi bir hazır kod kullanmadığımızdan bu şekilde belirttik. Fakat istenen açıklama metodu bilinmediğinden %20lik bir puan eksiltmesine gidildi. | 8 |
| **9** | **Ek 4 (10): Her madde 5’er puan.** |  | Açıklamalar, ilgili kaynaklar vasıtasıyla detaylı şekilde yazıldı. | 10 |
| **10** | **Özdeğerlendirme Tablosu (10)** |  | Yapıldı, herhangi bir eksik saptanmadı | 10 |
| **100 üzerinden Toplam Not:** | | | | 96 |

**İş bölümü Detayları**

**1)-** Uğur Alp Balta-3.5 saat/ Beat Duran-3.5 saat

**2)** Uğur Alp Balta- 2.5 saat / Berat Duran-2.5 saat

**3)** Uğur Alp Balta- 4 saat / Berat Duran 4.5 saat

**4)** Uğur Alp Balta- 1.5 saat / Berat Duran- 1 saat

**5)** Uğur Alp Balta- 30 dakika / Berat Duran-30 dakika

**6)** Uğur Alp Balta- 30 dakika / Berat Duran-1 saat

**7)** Uğur Alp Balta- 5 dakika / Berat Duran-5 dakika (Benzer projeleri araştırma kısmı eklenmemiştir)

**8)** Uğur Alp Balta- 5 dakika / Berat Duran-5 dakika

**9)** Uğur Alp Balta- 1 saat

**10)** Uğur Alp Balta- 20 dakika