

AYVOS Vaka Çalışması

SUNAN

Furkan ULU



Aşama 1- Model Eğitimi

Not : Yeterli GPU desteği olmadığından model eğitimi yapılamamıştır, onun yerine hazır model kullanılmıştır

a) YOLOv12 model eğitimi – Hiper-parametre ayarları ve nedenleri :

1. Epoch: 150-200- Kucuk-orta boyutlu veri setleri icin yeterlidir.

- Tray ve plate gibi objeler genelde birbirine benzeyen ve sık tekrarlanan nesneler. Bu yüzden erken epoch'larda model overfit etmeye meyilli.
- Başlangıçta 100 epoch yeterli olabilir ama validation mAP grafiğini izleyerek early stopping uygulanmalı.

2. imgsz = 640 / 960 :

- Tabağın sayısını tespit ettiğimiz için detaylar önemli. Modelin default imgsz değeri 640 fakat, Küçük resolution (640) küçük tabakları (Dondurma kalıpları) kaçırabilir. 1280'de ağır olabilir bu sebepten ötürü 960 uygun görülebilir.

3. batch = 8 veya 16 :

- Batch çok büyük olursa overfit riski artar, çok küçük olursa model öğrenemez. 8-16 arası genellikle en stabil sonuçları verir. GPU belleğine göre karar verilebilir.

4. learning-rate = 0.005 - 0.01 :

- Fazla yüksek olursa mAP'de ani iniş çıkışlar olabilir, düşük olursa öğrenme çok yavaşlar. 0.01 ile başlayıp CosineLR gibi scheduler ile düşürülür.

5. optimizer = SGD / Adam :

- overfitting'i baskılamak için drpoout, early_stopping ,weight_decay eklenebilir .

6. augmentation : hsv_h=0.015, scale=0.5, flipud=0.0, mosaic=1.0:

- Yansıma ve parlama gibi gerçek dünya sorunlarını simüle etmek için önemli. Ama fazla abartılırsa tabak sayısını bozar, o yüzden dengeli uygulanmalı.

7. conf_threshold = 0.5 - 0.7 :

- Özellikle inference sırasında doğru sayım için bu eşik önemli. Tabağa benzer ama olmayan şeyleri eleyebilmek için bir miktar yüksek tutulmalı.

b) YOLOv12 avantajları / dezavantajları:

Avantajlar:

- Gerçek zamanlı çalışabiliyor, çünkü dedektör + tracker yapısı hızlı.
- Yeni çıkan v12, önceki YOLO'lara göre daha az parametreyle daha yüksek accuracy sunuyor.
- Tek karede çalışması yeterli, videolardaki hareketli tepsi için frame-by-frame algılamaya uygun.
- GPU kullanımı verimli, RTX 3060-4060 sınıfı kartlarda çok rahat inference yapılabiliyor.

Dezavantajlar:

- Küçük tabakları kaçırabilir, özellikle motion blur varsa.
- Sabit ışık koşullarında overfit olabilir, çünkü parlak tabak ile zemini ayırt etmek zorlaşır.
- Tracking yoksa (tek frame ile karar veriliyorsa), bir frame'de olmayan tabak kayıtlardan düşebilir.
- Düşük donanımlarda yavaş kalabiliyor.

c) Alternatif 3 model önerisi ve nedeni:

1. EfficientDet
2. YOLOv8
3. Detectron2 (Faster R-CNN veya RetinaNet)

YOLOv8 neden alternatif olabilir?

- Aynı YOLO ailesinden geldiği için benzer eğitim pipeline'ına sahip.
- v12'ye göre daha stabil ve community desteği daha fazla.
- Hem edge cihazlara hem sunucuya uygun versiyonları var.
- Tray ve tabak gibi küçük ama düzenli objeleri YOLOv8 daha hassas layer yapısıyla daha iyi tanır.
- v12 kadar hızlı olmasa da, eğitim sırasında daha az data ile daha stabil sonuç verebilir.
- Ayrıca label formatları neredeyse birebir uyumlu, geçiş yapmak kolay.

d) Veri seti hazırlarken dikkat edilecek kurallar

- Dengeli sınıf dağılımı
 - Her sınıf yeterince örnek içermeli (Class başı 1000+ görsel) ve dengeli bir dağılım olmalı , yoksa model bir sınıfa saplanır.
- Doğru bounding box çizimi
 - Kutular obje ile tam oturmalı, çok boşluk veya fazla kesim olmamalı. Mümkün olduğunca bbox'lar farklı class'larla karışmamalı.
 - Bu çalışmada tabak ve tepsiler üst üste biniyor, bu yüzden etiketler pixel-level değilse kutular düzgün hizalanmalı.
- Çeşitlilik
 - Tespit edilen nesnenin hareket anında da tespit edilebilmesi için farklı açılarda ve farklı parlaklık düzeyleri simüle edilerek veri seti düzenlenmeli. Augmentasyon tek başına yeterli olmayabilir.
 - Bazı durumlarda arkaplan etiketsiz olarak verilerek nesnelerin arka plandan bağımsızlaştırılması yapılabilir. (Bazı case'lerde gerekebilir)
- Tutarlı etiketleme
 - Her obje aynı sınıfla ve benzer mantıkla etiketlenmeli. Nasıl başlandıysa o şekilde devam edilmeli.
- Bozuk veri olmamalı
 - Eksik label, hatalı image, corrupt dosya varsa eğitim sırasında hata çıkabilir.