

Derin Öğrenme (FET312) / 2025-2026 Güz Dönemi

Yoğun Market Raf Ürünlerinin Nesne Tespiti

DeepDive5

Ad/Soyad	Öğrenci No
Gözde İçöz	23040301076
Sıla Özbeý	23040301110
Samet Güneş	23040101059
Muhammet Arslan	23040101028
Onat Tandoğdu	23040301087

GitHub/Repo Bağlantısı:

https://github.com/gozdeii/FET312_DeepDive5_DL_Project

İş/Bilimsel Soru:

Perakende sektöründe **yüksek yoğunluklu ve karmaşık raf görselleri** (cluttered scenes) içinde yer alan çok sayıda ürünü, potansiyel örtüşme (occlusion) durumlarında dahi yüksek doğrulukla tespit ederek stok durumunu, raf düzenini ve alanlarını otomatik olarak izlemek nasıl sağlanabilir?

Motivasyon/Amacı:

Projenin temel amacı, yoğun market raflarındaki ürünlerin **nesne algılama modelleriyle tespit ederek** tükenen veya yeniden stoklanması gereken ürünler ile yanlış sıraya yerleştirilmiş ürünlerin tespit etmektir. Bu sayede insan hatalarının düzeltilmesi ve raf düzeninin korunması hedeflenmektedir. Projemizde, bu zorlu görev için **farklı CNN mimarileri** hazırlanacak ve performansları değerlendirilecektir.

Görev Türü:

Nesne Tespiti (Object Detection).

Hedef Değişken(ler):

Projenin hedef değişkeni, nesne tespiti görevine özgü Hibrit bir problem olup; birincil olarak tespit edilen her bir ürünün konumsal bilgisini Sınır Kutusu'nun Merkez Koordinatları ve Boyutları ile tahmin eden Regresyon çıktıları gerektirir. Diğer yandan, ürünün 86 farklı kategoriye ait sınıfını ve nesneye ait olma olasılığını tahmin eden Çoklu Sınıflandırma çıktıları beklenir. Başarı kriterimiz, modelin doğrulukta sektör standardına ulaşması (özellikle mAP@0.5:0.95 değerinin %30 üzerine çıkması) ve uygulanan 5 farklı CNN mimarisinin arasında yüksek doğruluk ile gerçek zamanlı kullanım için gerekli olan Hız (FPS) arasında en uygun dengenin nicel olarak belirlenmesidir.

Başarı Kriterleri:

Eğitim hattının hatasız tamamlanması, [mAP@0.5](#) ve [mAP@0.5:0.95](#) değerlerinin sektör standartlarına (örn: %30 üzeri mAP) ulaşması ve Base Model aşamasında kullanılan farklı CNN mimarileri arasında en iyi performans/hız dengesinin belirlenmesi.

A. Önemli Noktalar ve Zaman Çizelgesi

Bu plan, 5 farklı mimarinin Base Model (Model 1) performansının hızlıca ölçülmesini ve ardından her üyenin kendi mimarisini Final Model (Model 2) için en yüksek performansa ulaşacaktır şekilde geliştirmesini öngörmektedir.

0. Hafta: Proje kapsamının belirlenmesi, SKU-110K veri setinin seçimi ve 5 farklı CNN mimarisinin ekip üyelerine atanması. GitHub deposu kurulumu ve dosya yapısının oluşturulması.

1. Hafta: Base Model Eğitimi ve Çıktıları (Model 1):

- Her grup üyesi, sorumlu olduğu ilk mimariyi (Model 1) kullanarak **tek epoch'luk (epochs=1)** minimal eğitimi tamamlar.
- Model 1 eğitim çıktıları (.pt ağırlıklar, loglar, grafikler) GitHub'daki /experiment_outputs klasörüne yüklenir.

2. Hafta: Vize Raporu (Ara Dönem) ve Model 2 Tanımlaması:

- 5 Base Modelin (Model 1) sonuçları toplanır ve karşılaştırımlı ilk analiz tamamlanır.
- **Ara Dönem Proje Raporu (Vize)** tamamlanır ve teslim edilir.
- Her üye, Model 1'in mimarisini temel alarak **Final Modeli (Model 2)** için uygulanacak iyileştirme ve optimizasyon stratejilerini belirler.

3-5. Hafta: Final Model Geliştirme ve Derin Eğitim (Model 2):

- (3 Hafta süresince) Tüm 5 üye, kendi **Final Modellerini (Model 2)** belirledikleri gelişmiş hiperparametreler ve veri büyütme teknikleriyle uzun süreli eğitime tabi tutar.
- Bu aşamada elde edilen en iyi model ağırlıkları ve performans metrikleri detaylı olarak kayıt altına alınır.

6. Hafta: Kapsamlı Performans Analizi:

- Tüm **10 farklı modelin** (5 Model 1 + 5 Model 2) final çıktıları (mAP, Loss, Hız) toplanır.
- Projenin temel hedeflerine ulaşıp ulaşılmadığı değerlendirilir ve **Kapsamlı Karşılaştırımlı Analiz Tabloları** hazırlanır.
- Rapor için gereken tüm grafikler ve görselleştirmeler oluşturulur.

7. Hafta: Proje Teslimi:

- Proje raporunun son hali (**Final Raporu**) tamamlanır ve teslim edilir.
- Proje sunum slaytları hazırlanır ve final sunum yapılır.

Üye Adı	Öğrenci No	Model 1 (Base Model - Minimal Test)
Gözde İçöz	23040301076	YOLOv5s
Sıla Özbeyp	23040301110	TinyGapNet
Samet Güneş	23040101059	TinyShelfNet
M. Arslan	23040101028	EfficientDetD0-Lite
Onat Tandoğdu	23040301087	RetinaNet

GitHub/Repo Bağlantısı: https://github.com/gozdeii/FET312_DeepDive5_DL_Project

İlgili Çalışmalar (Mini Literatür Taraması)

Projemiz, literatürdeki boşluğu, perakende alanındaki en zorlu veri kümesi (SKU-110K) üzerinde 5 farklı temel nesne tespiti mimarisinin performansını (YOLOv5s, YOLOv8s, SSD300, RetinaNet, TinyShelfNet) tek bir bütünlük çalışma kapsamında, standartlaştırılmış hiperparametreler ve veri büyütme teknikleriyle karşılaştırarak doldurmaktadır. Bu kapsamlı karşılaştırımlı analiz, her mimarinin yoğun nesne tespiti zorluklarına karşı gösterdiği Doğruluk (mAP) ve Hız (FPS) dengesini nicel olarak ortaya koyacaktır.

Temel Kaynakların Karşılaştırmalı Analizi

Referans No	Çalışmanın Kapsamı	Yöntem/Mimari	Veri Kümesi Ölçeği	Raporlanan Temel Metrikler
[1] Goldman et al. (2019)	Yüksek yoğunluklu perakende ortamlarında nesne tespiti zorluğunu ortaya koymak.	Farklı Base Modellerin Karşılaştırılması	SKU-110K	Yoğunluk Zorluğu
[2] Liu et al. (2016)	Nesne tespiti hızını artırmak için tek aşamalı mimari sunmak.	SSD (Single Shot MultiBox Detector)	PASCAL VOC, MS COCO	Hız (FPS) ve mAP dengesi
[3] Lin et al. (2017)	Yoğun nesne tespitinde sınıf dengesizliğini çözmek.	RetinaNet	MS COCO	Yüksek mAP başarısı

Teorik CNN Mimari Kıyaslaması (Hipotez)

Mimari	Tipi	Temel Odak Noktası	Ana Teknoloji	Beklenen Hız (FPS)	Yoğunluk (SKU-110K) Uygunluğu
YOLOv5s	Tek Aşamalı	Hız ve Denge	CSPNet	Çok Yüksek	Orta-Yüksek
TinyGAPNet	Tek Aşamalı	Doğruluk	Anchor-Free	Yüksek	Yüksek
EfficientDetD0-Lite	Tek Aşamalı	Verimlilik & Hafiflik	Bileşik Ölçekleme	Yüksek	Orta-Yüksek
RetinaNet	Tek Aşamalı	Sınıf Dengesizliği / Doğruluk	Focal Loss, FPN	Orta	Çok Yüksek
TinyShelfNet	İki Aşamalı / Hibrid	Raflar için Özelleşme	Bölgesel Öneri Mekanizması (RPN)	Orta-Düşük	Yüksek

Veri Kümesi Açıklaması (Dataset Description)

Alan	Detay
Adı	SKU-110K (Store Keeping Unit) Veri Seti
Kaynak	Bar-Ilan University, Tel Aviv University, Trax Retail
Bağlantı	https://github.com/eg4000/SKU110K_CVPR19

Lisans/Kullanım Hakları	Veri seti, çoğunlukla Akademik ve Ticari Olmayan kullanım için lisanslanmıştır. Verinin ticari ortamlardan toplanması nedeniyle ticari kullanım için özel sınırlamalar mevcuttur.
Açıklama	Veri kümesi, perakende raf görüntülerinden oluşur ve aşırı yoğun (densely packed) nesne tespiti senaryoları için tasarlanmıştır. Görseller, farklı coğrafyalardaki market raflarını yansıtmaktadır.

Şema: Veri Setindeki Değişkenler

Veri seti, hem görsel (girdi) hem de etiket (çıkıtı) değişkenlerinden oluşmaktadır:

- Girdi Değişkeni (Feature): Görüntü (Image).
 - Türü: RGB Piksel Matrisi (3 Kanal).
 - Birimi: Piksel (Çözünürlük değişkendir).
- Çıktı Değişkenleri (Target Variables): Nesne Tespiti görevinin iki ana bileşeni.
 - 1. Sınır Kutusu Konumu (Bounding Box): $b = (x, y, w, h)$
 - ♣ Türü: Regresyon (4 sürekli sayı).

Boyut, Ölçek ve Sınıf Dengesi

Alan	Değer
Toplam Görüntü Sayısı	11.762 görüntü.
Toplam Nesne Sayısı	Yaklaşık 1.2 Milyon etiketli nesne.
Nesne Yoğunluğu	Görsel başına ortalama 147.4 nesne.
Sınıf Sayısı	1 ürün sınıfı.
Veri Bölümlemesi	8.233 Eğitim (Train), 2.941 Test, 588 Doğrulama (Validation) görselleri.
Sınıf Dengesi (Oranı)	Veri seti tek bir “ürün” sınıfını içerir; bu nedenle sınıflar arasında oran farklılıklar veya sınıf dengesizliği (class imbalance) söz konusu değildir.

Etik, Gizlilik ve Önyargı

- Gizlilik ve Hassas Alanlar: Veri setinin içeriği market raflarındaki ürünlerdenoluştugu için, insan yüzü, kimlik bilgileri veya kişisel olarak tanımlanabilir herhangi bir veri içermemektedir. Bu nedenle, genel gizlilik ve kişisel veri koruma (KVKK) endişeleri bu proje için minimal düzeydedir.

Önyargı ve Sınırlamalar:

Ticari Sınırlama: Veri setinin ticari ortamlardan toplanması nedeniyle, projenin final çıktısının ticari amaçla kullanılması lisans kurallarına tabidir. Projemiz sadece akademik amaçlıdır.

Yoğunluk Önyargısı: Veri setinin yoğunluk üzerine kurulu olması, modelleri, diğer daha az yoğun (seyrek) sahnelerde gereğinden fazla nesne aramaya teşvik edebilecek bir önyargı yaratabilir.

Yöntemler ve Mimari

Bu proje, yüksek yoğunluklu SKU-110K veri seti üzerinde Nesne Tespiti (Object Detection) problemini çözmek için 5 farklı Konvolüsyonel Sinir Ağı (CNN) tabanlı Tek Aşamalı (Single-Stage) ve İki Aşamalı (Two-Stage) mimarı karşılaştırmalı olarak uygulamayı amaçlamaktadır.

Yöntemler ve Önerilen Çözüm

Önerilen çözüm, her grup üyesinin sorumlu olduğu mimarinin performansını, hızını ve kaynak verimliliğini adil bir zeminde test etmeye dayanır.

- Çözüm Yaklaşımı: Hibrit Karşılaştırmalı Nesne Tespiti.
- Uygulama Adımları:
 - Base Model Eğitimi (Model 1): Tüm 5 mimarinin minimal (1-epoch) eğitim ile başlangıç performanslarının ölçülmesi (Vize Raporu çıktısı).
 - Performans Analizi: Elde edilen 10 farklı modelin (5 Model 1 + 5 Model 2) sonuçlarının mAP@0.5, mAP@0.5:0.95, Hız (FPS) ve Model Büyüklüğü (Model Size) metrikleri üzerinden nicel olarak değerlendirilmesi.

Çözümünüzün Detaylandırılması: Ortak Süreçler

A. Veri Ön İşleme ve Hazırlama Adımları

- Veri Formatı Dönüşümü: Veri setinin COCO formatından (JSON) alınarak, her bir modelin gerektirdiği formata (YOLO için TXT etiket dosyaları veya diğerleri için Pascal VOC/XML) dönüştürülmesi.
- Görüntü Yeniden Boyutlandırma: Her modelin gerektirdiği giriş boyutuna

B. Tasarlanan Eğitim Süreçleri ve Hiperparametre Yaklaşımları

- Eğitim Süreci:
 - Model 1 (Base Model): 1 epoch (hızlı başlangıç testi).
- Hiperparametre Optimizasyonu:
 - Başlangıç: Her modelin orijinal makalelerindeki başlangıç öğrenme oranı (Learning Rate) ve ağırlık azaltma (Weight Decay) değerleri kullanılacaktır.
 - Optimizasyon Yaklaşımı: Model 2 geliştirilirken, Grid Search veya Rastgele Arama (Random Search) yaklaşımları kullanılarak öğrenme oranı, yavaşlatma (momentum) ve toplu iş boyutu (batch size) değerleri optimize edilecektir.

Her Grup Üyesinin Mimari Detayları

Model	Mimari Yapısı	Optimizer	Kullanılan Loss Fonksiyonu
YOLOv5s	CSPDarknet-eski tabanlı backbone + PANet neck + YOLO head	SGD	Clou Loss (box), BCE (obj + cls)

TinyGAPNet	4 katmanlı hafif CNN backbone + Global Average Pooling (GAP) tabanlı regresyon başlığı	AdamW	DFL (Distribution Focal Loss) + BCE + IoU Loss
TinyShelfNet	Raf görüntüleri için optimize edilmiş hafif CNN backbone + FPN benzeri neck + küçük anchor-free head	Adam	Smooth L1 veya IoU tabanlı loss (paper'a göre değişiyor)
EfficientDetD0-Lite	EfficientNet-Lite backbone + BiFPN neck + EfficientDet detection head	AdamW	Smooth L1 (box) + Focal Loss (cls)
RetinaNet	ResNet backbone + FPN neck + anchor-based one-stage head	Adam	Focal Loss (cls) + Smooth L1 (box)

Deney Tasarımı

1. Deney: Base Model Karşılaştırmalı Performans Kıyaslaması (Model 1)

Bu ilk deney, Vize Raporu'nun temelini oluşturur ve Model 1 (Base Model) kodlarının doğru çalıştığını kanıtlamayı amaçlar.

- Ana Amaç: Projede kullanılacak 5 farklı nesne tespiti mimarisinin (YOLOv5s, YOLOv8s, TinyShelfNet, EfficientDet, RetinaNet) yüksek yoğunluklu SKU-110K veri seti üzerindeki minimal başlangıç potansiyelini (1 epoch) hızlıca ölçmek.
- Train & Test Ayrımı ve Hedef Değer: Veri seti, orijinal dağılımı ile kullanılacaktır: 8.233 Eğitim ve 2.941 Test görselleri. Hedef değerler, her görüntüdeki nesnelere ait Sınır Kutusu Koordinatları olarak ayarlanmıştır.
- Validasyon: 588 Doğrulama (Validation) görseli, her modelin eğitim sonundaki genelleme eğilimini izlemek için kullanılacaktır.
- Hiper Parametre Optimizasyonu: Bu aşamada optimizasyon yapılmamıştır. Tüm modeller, literatürdeki standart başlangıç değerleri (default hyper-parameters) ile eğitilmiş ve minimal (1 epoch) sonuçları raporlanmıştır.
- Final Sonuç (Plan): 1. hafta sonunda 5 modelin de Base Model sonuçları (Kaybı ve mAP) elde edilecektir.

Değerlendirme Kriterleri ve Başarı Sıralaması

Projemizin başarısı, perakende otomasyonu için kritik olan iki temel metrik üzerinden değerlendirilecektir:

1. Doğruluk Metrikleri (Accuracy Metrics)

- mAP@0.5 (Mean Average Precision at IoU=50%): Modelin hem nesneyi doğru konumlandırma (lokalizasyon) hem de doğru sınıflandırma başarısını ölçen endüstri standartı temel metiktir. Yüksek yoğunluklu ortamlarda elde edilen sonuçlar literatürdeki diğer Base Model sonuçlarıyla kıyaslanacaktır.
- mAP@0.5:0.95 (COCO Metriği): Bu metrik, IoU eşiği %50'den başlayıp %95'e kadar %5'lük artışlarla hesaplanan ortalama hassasiyeti temsil eder. Bu, modelin hassas lokalizasyon yeteneğini gösteren en zorlu metrik olup, başarımızın nihai kriteri olacaktır.

2. Hız Metriği (Efficiency Metric)

- Hız (FPS - Frames Per Second): Modelin saniyede işleyebileceği kare sayısını ölçer. Gerçek zamanlı raf izleme uygulamaları için modelin düşük gecikme (low latency) ile çalışması kritik olduğundan, mAP ile birlikte değerlendirilen en önemli ticari metrik olacaktır.

Model	Epoch	mAP@0.5	mAP@0.5:0.95	FPS
YOLOv5s	1	0.0371	0.05	~65 FPS
	10	0.24	0.11	~65 FPS
TinyGAPNet	1	0.13	0.06	~70 FPS
	10	0.27	0.12	~70 FPS
TinyShelfNet	1	0.11	0.04	~60 FPS
	10	0.23	0.10	~60 FPS
EfficientDet-D0	1	0.09	0.04	~45 FPS
	10	0.21	0.09	~45 FPS
RetinaNet	1	0.17	0.04	~40 FPS
	10	0.28	0.10	~40 FPS

Kullanılan Araçlar ve Frameworkler

Alan	Detay	Açıklama
Derin Öğrenme Kütüphanesi	PyTorch 2.0.0+	Tüm Base ve Final Model eğitimleri için tercih edilen temel framework.
Veri İşleme ve Analiz	Python 3.10+, Pandas, NumPy, JupyterN	Veri setinin (SKU-110K) yüklenmesi, ön işlenmesi, analiz edilmesi ve etiket formatlarının dönüştürülmesi için kullanılır.
Model Mimarileri	Torchvision, Ultralytics	Model mimarilerinin (RetinaNet) ve özelleştirilmiş veri büyütme (Data Augmentation) tekniklerinin uygulanması için kullanılır.

	YOLOv5/v8, Albumentations	
Seed Kullanımı	42 (Global Seed)	Deneyleşen tekniklerin tekrarlanabilirliğini (reproducibility) sağlamak amacıyla tüm ekip üyeleri tarafından aynı global seed değeri (42) kullanılacaktır.

Geliştirilen Kodlar

Proje kapsamında geliştirilen tüm Base Model (Model 1) kod dosyaları, isimlendirme kurallarına uygun olarak aşağıdaki GitHub deposuna yüklenmiştir:

GitHub Repo Bağlantısı: https://github.com/gozdeii/FET312_DeepDive5_DL_Project

A. Bilimsel ve Literatür Kaynakları

No	Kaynak Adı	Açıklama
[1]	SKU-110K Veri Seti (Veri Kaynağı)	Goldman, E., et al. (2019). SKU-110K: Densely Packed Scenes for 110K Categories of Retail Items. <i>2019 IEEE/CVF International Conference on Computer Vision (ICCV)</i> .
[2]	EfficientDet Mimarisi(M. Arslan)	Tan, M., & Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection.*
[3]	RetinaNet Mimarisi (O. Tandoğdu)	Lin, T. Y., et al. (2017). Focal Loss for Dense Object Detection. <i>2017 IEEE International Conference on Computer Vision (ICCV)</i> .
[4]	TinyShelfNet Mimarisi (S. Güneş)	[Yerel veya Özel Mimari Referansı]. Bu mimari, perakende raf tespiti için özelleştirilmiş, hafifletilmiş bir CNN yapısını temsil eder.
[5]	Yolov5s (G. İçöz)	Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. (YOLO'nun temel mimarilerinden biri referans alınmıştır).
[6]	TinyGAPNet (S. Özbeyp)	Bu mimari, Global Average Pooling (GAP) tabanlı hafif bir CNN yaklaşımı olup, yoğun market raflarında tek sınır kutusu regresyonu görevleri için optimize edilmiş özel bir modeldir.

B. Kod ve Geliştirme Kaynakları

No	Kaynak Adı	Açıklama
[5]	PyTorch Derin Öğrenme Kütüphanesi	Paszke, A., et al. (2019). PyTorch: An Imperative Style, High-Performance Deep Learning Library. <i>Advances in neural information processing systems</i> .

[6]	Proje GitHub Deposu	DeepDive5 Grubu. (2025). Yüksek Yoğunluklu Nesne Tespiti Projesi Kodları. <i>GitHub Repository</i> . Bağlantı: https://github.com/gozdeii/FET312_DeepDive5_DL_Project
-----	------------------------	--