

# Derin Öğrenme (FET312) / 2025-2026 Güz Dönemi

## Yoğun Market Raf Ürünlerinin Nesne Tespiti

### DeepDive5

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

GitHub/Repo Bağlantısı:

[https://github.com/gozdeii/FET312\\_DeepDive5\\_DL\\_Project](https://github.com/gozdeii/FET312_DeepDive5_DL_Project)

Sunum Bağlantısı:

<https://www.youtube.com/watch?v=Nm397tAwJUs>

Üye Adı	Model 1	Model 2
Gözde İçöz	YOLOv5s	Faster R-CNN
Sıla Özbeý	TinyGapNet	MobileNetV3
Samet Güneş	TinyShelfNet	TinyShelfNet++
M. Arslan	EfficientDetNet-B0	EfficientDetD0-Lite
Onat Tandoğdu	RetinaNet (ResNet-18)	RetinaNet (ResNet-34)

### 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	<a href="https://github.com/eg4000/SKU110K_CVPR19">https://github.com/eg4000/SKU110K_CVPR19</a>
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.
Train-test split	Yaklaşık 75%-25%, veri setinin önceden belirlenmiş oranıdır.

## **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 yaratır.

## **Yöntemler ve Mimari**

Bu proje, yüksek yoğunluklu SKU-110K veri seti üzerinde Nesne Tespiti (Object Detection) problemini çözmek için 10 farklı Konvolüsyonel Sinir Ağı (CNN) tabanlı Tek Aşamalı (Single-Stage) ve İki Aşamalı (Two-Stage) mimariyi karşılaştırmalı olarak uygulamayı amaçlamaktadır.

Çalışmalarda modellerin kendi özellikleriyle öğrenmesine izin verilmiştir. Modellerin eğitimde girdileri için görüntüler yeniden boyutlandırılmış ve normalize edilmiştir ancak veri artırma (data augmentation) teknikleri uygulanmamıştır.

## **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)
Faster R-CNN	Cross Entropy Loss (cls) + Smooth L1 Loss (bbox)	SGD	Cross Entropy Loss (cls) + Smooth L1 Loss (bbox)
TinyGAPNet	4 katmanlı hafif CNN backbone + Global Average Pooling (GAP) tabanlı regresyon başlığı	AdamW	DFL (Distribution Focal Loss) + BCE + IoU Loss
MobileNetV3	Mobil cihazlar için optimize edilmiş hafif CNN backbone + basit regresyon	Adam	Smooth L1 Loss (bbox)

TinyShelfNet	Raf görüntüleri için optimize edilmiş <b>hafif CNN backbone + FPN</b> benzeri neck + küçük anchor-free head	Adam	Smooth L1 veya IoU tabanlı loss (paper'a göre değişiyor)
TinyShelfNet++	Hafif CNN backbone + FPN tabanlı neck + anchor-free detection head	Adam	IoU tabanlı Loss (box) + BCE
EfficientDetD0-Lite	EfficientNet-Lite backbone + BiFPN neck + EfficientDet detection head	AdamW	Smooth L1 (box) + Focal Loss (cls)
EfficientNet-B0	EfficientNet-Lite backbone + bounding box regresyon modeli	Adam	Smooth L1 (box)
RetinaNet	ResNet 18 backbone + FPN neck + anchor-based one-stage head	Adam	Focal Loss (cls) + Smooth L1 (box)
RetinaNet	Resnet 34 backbone + FPN neck + anchor-based one-stage head	Adam	Focal Loss (cls) + Smooth L1 (box)

## Performans Sıralaması

Model	mAP@0.5	Precision	Recall	F1
YOLOv5s	0.24	0.40	0.47	0.43
Faster R-CNN	<b>0.36</b>	<b>0.60</b>	0.60	0.58
TinyGAPNet	0.27	0.38	0.45	0.41
MobileNetV3	0.22	0.25	0.39	0.37
TinyShelfNet	0.23	0.39	0.41	0.40
TinyShelfNet++	0.30	0.49	0.52	0.50
EfficientNet-B0	0.25	0.41	0.46	0.43
EfficientDet-D0	0.34	0.57	<b>0.63</b>	<b>0.61</b>
RetinaNet 18	0.28	0.46	0.51	0.48
RetinaNet 34	0.32	0.52	0.56	0.54

## Performans Analizi

Bu çalışmada tek bir “en iyi” modelden bahsetmek yerine, problemin gereksinimlerine göre en uygun modelin değiştiği görülmüştür. Hız ve düşük hesaplama maliyetinin öncelikli olduğu senaryolarda **Yolov5** ve **TinyGapNet**, gerçek zamanlı çalışmaya uygun yapıları sayesinde öne çıkmaktadır. Doğruluğun kritik olduğu ve çıkışım süresinin ikinci planda kaldığı durumlarda ise **Faster R-CNN** ve **EfficientDet-D0** daha yüksek tespit hassasılıeti ve kararlı tahminler sunmuştur.

## 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)	Deneyleşenlerin tekrarlanabilirliğini (reproducibility) sağlamak amacıyla tüm ekip üyeleri tarafından aynı global seed değeri (42) kullanılacaktır.

## A. Bilimsel ve Literatür Kaynakları

No	Kaynak Adı	Açıklama
[1]	<b>SKU-110K Veri Seti (Veri Kaynağı)</b>	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]	<b>EfficientDet Mimarisi(M. Arslan)</b>	<b>Tan, M., &amp; Le, Q. V. (2020). Efficientdet: Scalable and efficient object detection.*</b>
[3]	<b>RetinaNet Mimarisi (O. Tandoğdu)</b>	Lin, T. Y., et al. (2017). Focal Loss for Dense Object Detection. <i>2017 IEEE International Conference on Computer Vision (ICCV)</i> .
[4]	<b>TinyShelfNet Mimarisi (S. Güneş)</b>	[Yerel veya Özel Mimari Referansı]. <i>Bu mimari, perakende raf tespiti için özelleştirilmiş, hafifletilmiş bir CNN yapısını temsil eder.</i>
[5]	<b>Yolov5s (G. İçöz)</b>	Redmon, J., & Farhadi, A. (2018). YOLOv3: An Incremental Improvement. (YOLO'nun temel mimarilerinden biri referans alınmıştır).
[6]	<b>TinyGAPNet (S. Özbeyp)</b>	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: <a href="https://github.com/gozdeii/FET312_DeepDive5_DL_Project">https://github.com/gozdeii/FET312_DeepDive5_DL_Project</a>