



## Ceviz işleme hatları için YOLOv11 model temelli otomatik tespit ve sınıflandırma sistemleri

### Automatic detection and classification systems based on YOLOv11 model for walnut processing lines

Hakan Aktaş<sup>1,\*</sup> , Emrullah Polat<sup>2</sup> 

<sup>1,2</sup> Niğde Ömer Halisdemir Üniversitesi, Bilgisayar Mühendisliği Bölümü, 51240, Niğde Türkiye

#### Öz

Tarım ürünlerinin otomatik sınıflandırılması, kalite kontrol süreçlerinin hızlandırılması ve insan hatasının azaltılması açısından kritik öneme sahiptir. Bu çalışmada konveyör bant üzerinde bulunan çoklu ceviz bileşenlerinin gerçek zamanlı olarak tespiti ve sınıflandırılması amacıyla, YOLOv11 tabanlı bir derin öğrenme yöntemi önerilmiştir. Endüstriyel bir düzenek üzerinde görüntüler alınmış ve toplamda 1194 adet kabuk, 641 adet ceviz içi ve 458 adet zar görüntüsü etiketlenmiştir. Etiketlenen bu veriler ilk önce YOLOv11n modeli ile eğitilip test edilmiş olup; en yüksek recall değeri ceviz içi sınıfında 0.963 olarak ve en düşük recall değeri zar sınıfında 0.795 olarak hesaplanmıştır. Aynı etiketli veri seti daha büyük bir model olan YOLOv11L modeli ile eğitilip test edildiğinde en yüksek recall değeri ceviz içi sınıfında 0.977 ve en düşük recall değeri zar sınıfında 0.922 olarak hesaplanmıştır. Son olarak YOLO modellerinden farklı olarak RT-DETR nesne tespit algoritması alternatif bir model olarak kullanılmış ve sonuçları YOLOv11 modelleri ile kıyaslanmıştır. Her bir modelin kendine göre üstün yanları olup; elde edilen test sonuçlarına göre yüksek fps uygulamaları için YOLOv11n modeli, orta fps ve yüksek doğruluk için YOLOv11L modeli, orta - düşük fps değerlerinde kabuk ve ceviz içi sınıflarının yüksek doğrulukta tespiti için RT-DETR modeli önerilmektedir.

**Anahtar kelimeler:** Nesne tespiti, Nesne sınıflandırma, YOLOv11, RT-DETR, Ceviz tespiti ve sınıflandırma

#### 1 Giriş

Geleneksel görüntü işleme teknikleri, nesne sınıflandırma alanında uzun yıllardır kullanılmaktadır. Bu yöntemler genellikle kenar tespiti, renk eşikleme, morfolojik işlemler ve manuel özellik çıkarımı gibi adımları içermektedir. Örneğin, Histogram of Oriented Gradients (HOG) ve Scale-Invariant Feature Transform (SIFT) gibi öznitelik çıkarım yöntemleri, belirli nesne sınıflarının tanımlanmasında kullanılmıştır [1, 2]. Ancak, bu teknikler, özellikle karmaşık arka planlar, değişken aydınlatma koşulları ve benzer dokuya sahip nesnelerin ayırımında sınırlı başarı göstermektedir [3, 4]. Ayrıca, geleneksel yöntemlerin başarısı büyük ölçüde uzman bilgisine ve manuel ayarlamalara bağlıdır. Örneğin bir çalışmada, geleneksel

#### Abstract

Automatic classification of agricultural products is critical for accelerating quality control processes and reducing human error. In this study, a YOLOv11-based deep learning method is proposed for real-time detection and classification of multiple walnut components on a conveyor belt. Images were captured in an industrial setup, and a total of 1194 shell, 641 kernel, and 458 membrane images were labeled. These labeled data were first trained and tested with the YOLOv11n model; the highest recall value was calculated as 0.963 in the walnut class and the lowest recall value was calculated as 0.795 in the membrane class. When the same labeled dataset was trained and tested with the larger YOLOv11L model, the highest recall value was 0.977 for the walnut kernel class, and the lowest recall value was 0.922 for the membrane class. Finally, unlike the YOLO models, the RT-DETR object detection algorithm was employed as an alternative model, and its results were compared with those of the YOLOv11 models. Each model has its own advantages; based on the test results, the YOLOv11n model is recommended for high fps applications, the YOLOv11L model for medium fps with high accuracy, and the RT-DETR model is recommended for high accuracy detection of shell and walnut kernel classes at medium - low fps values.

**Keywords:** Object detection, Object classification, YOLOv11, RT-DETR, Walnut detection and classification

yöntemlerle yapılan nesne sınıflandırma işlemlerinin doğruluk oranlarının, derin öğrenme tabanlı yöntemlere kıyasla daha düşük olduğu rapor edilmiştir [5]. Bu sınırlamalar, özellikle gerçek zamanlı uygulamalarda ve büyük veri setlerinde daha belirgin hale gelmektedir [6]. Başka bir çalışmada, fıstık sınıflandırması için renk (RGB) tabanlı diskriminant analizi (DA) yöntemi kullanılarak, küçük kabuklu fıstıklarda %85 doğruluk gösterirken, renk ve eğim (slope) bilgisini birlikte kullanan k-NN algoritması %96 doğruluk oranına ulaşarak uzamsal bilginin sınıflandırma başarısını nasıl artırdığını ortaya koymuştur [7]. Benzer şekilde, ARM tabanlı Raspberry Pi kullanılarak geliştirilen bir çalışmada, renk, boyut ve şekil parametrelerine dayalı optik sınıflandırma algoritması

\* Sorumlu yazar / Corresponding author, e-posta / e-mail: haktas@ohu.edu.tr (H. Aktaş)

Geliş / Received: 23.06.2025 Kabul / Accepted: 02.10.2025 Yayımlanma / Published: 15.10.2025

doi: 10.28948/ngumuh.1723155

geliştirilerek kusurlu yer fıstığı sınıflandırılması üzerine çalışılmıştır. Çalışma sonucunda sınıflandırma doğruluğunun ve sistemin gerçek zamanlı çalışma hızının endüstriyel uygulamalar için uygun olmadığı gözlemlenmiştir [8]. Bu çalışmalarda da gösterildiği gibi geleneksel görüntü işleme yöntemleri, belirli ve kontrollü ortamlarda etkili olabilirken, dinamik ve karmaşık sahnelerde sınırlı performans sergilemektedir. Bu nedenle, daha esnek ve genelleme yeteneği yüksek olan derin öğrenme tabanlı yaklaşımlara yönelim artmaktadır [9].

Derin öğrenme mimarilerinden birisi olan ve sınıflandırmada kullanılan evrişimli sinir ağları (CNNs), görüntü sınıflandırma problemlerinde insan uzmanlığını gerektirmeyen otomatik özellik çıkarımı sayesinde büyük bir devrim yaratmıştır. AlexNet'in 2012 ImageNet yarışmasındaki başarısı [10], derin öğrenmenin görüntü tanımadaki üstünlüğünü ortaya koymuş ve sonrasında VGGNet [11], GoogLeNet/Inception [12], ResNet [13] ve EfficientNet [14] gibi farklı mimariler geliştirilmiştir. CNN'lerin başarısı, katman derinliğiyle birlikte çok karmaşık görsel görüntüleri öğrenme yeteneğinden kaynaklanmaktadır. Tarımsal ürünlerin sınıflandırmasında da CNN'ler yaygın olarak kullanılmakta olup; açık/kapalı antep fıstığı gibi zorlu sınıflandırma problemlerinde yüksek doğruluk oranları elde edilmektedir [15]. Ceviz parçalarının sınıflandırılması için geliştirilen CNN tabanlı bir çalışmada, ceviz parçalarını sınıflandırmak için ceviz içi, ceviz kabuğu ve ayrılmamış ceviz olarak 3 farklı sınıf kullanılmış ve %99.66 ceviz içi test doğruluğu ile endüstriyel uygulamalara pratik bir çözüm sunulmuştur [16].

Evrişimli sinir ağı mimarileri ile sınıflandırma yapılmak istenildiğinde tek bir sınıfa ait görüntünün mimariye verilmesi gerekmektedir. Ancak konveyör bant üzeri sınıflandırma işlemleri yapılırken farklı sınıflara ait nesnelere karışık halde gelmektedir. Böyle bir görünümün CNN mimarileri ile sınıflandırılması oldukça zordur. Aynı görüntü karesinde karışık halde bulunan nesnelere sınıflandırılması için YOLO (You Only Look Once) mimarileri daha uygundur. YOLO mimarisi, nesne tespiti ve sınıflandırma işlemlerini tek bir aşamada gerçekleştiren, gerçek zamanlı uygulamalar için optimize edilmiş bir derin öğrenme modelidir [17]. İlk olarak Redmon vd., tarafından önerilen YOLOv1, görüntüyü sabit bir ızgaraya bölerek her hücre için sınıf ve sınırlayıcı kutu tahminleri üretmektedir [18]. Bunu izleyen YOLOv2 [19] ve YOLOv3 [20] sürümleri, doğruluk ve hız açısından önemli iyileştirmeler sunmuştur. YOLOv4 [21] ve YOLOv5 [22] ile birlikte model kompakt hale getirilerek mobil ve gömülü sistemlerde de kullanım imkânı doğmuştur. Özellikle YOLOv8, segmentasyon ve çoklu nesne takibi gibi ileri düzey görevlerde de kullanılabilir hale getirilmiştir [23]. Tarımsal uygulamalarda YOLO mimarileri; elma, portakal, ceviz ve mısır gibi ürünlerin kalite kontrolü ve sayımı gibi görevlerde yüksek doğrulukta sonuçlar sağlamaktadır [24, 25]. Mısır tohumlarının çeşitlerini ve üzerindeki kusurları tespit etmek amacıyla yapılan bir çalışmada YOLOv7 tabanlı bir mimari geliştirilmiştir. Geliştirilen mimari ile orijinal YOLOv7 mimarisinden %2.59 oranında daha yüksek doğruluk elde edilmiştir [26].

Tarım alanında yapılan çalışmalardan bir tanesi de cevizlerin sınıflandırması ve ayıklamasıdır. Cevizlerin işlenmesi sürecinde otomatik kalite kontrol sistemlerinin geliştirilmesine yönelik çalışmalar, özellikle son yıllarda yoğunlaşmıştır. CNN ve YOLO tabanlı yaklaşımlar, hem ceviz içi ve ceviz kabuğu hem de ceviz zarı ayrımı gibi kritik görevlerde başarıyla uygulanmıştır. Kırılmış cevizlerin kabuk ve iç tespitini yapmak için gerçekleştirilen bir çalışmada YOLOX mimarisi kullanılmıştır [27]. Bu çalışmada elde edilen sonuçlar, tek aşamalı çoklu kutu algılama algoritması (Single Shot MultiBox Detector - SSD) ve YOLOv3 mimarilerine kıyasla yüksek doğruluğa ulaşmıştır. İnsansız hava araçları kullanılarak elde edilen ceviz ağacı veri seti başka bir çalışmada kullanılmış ve bu çalışmada ceviz ağaçlarının görüntüleri insansız hava aracı ile elde edilmiş ve ağaçlardaki cevizler etiketlenmiştir [28]. Elde edilen veriler geliştirilen w-YOLO derin öğrenme mimarisi ile kullanılmış ve cevizler tespit edilmiştir. Çalışmada Ortalama Doğruluk (mean Average Precision 0.5) değeri %97 olarak gözlemlenmiştir. Ceviz içlerinin kalitesini tespit edebilmek amacıyla yapılan başka bir çalışmada, X-ray ışın görüntüleri ile elde edilen farklı türde cevizler YOLOv5n tabanlı mimari kullanılarak sınıflandırılmıştır [29]. Yapılan çalışmada cevizler, sağlam, hafif buruşuk, orta derece buruşuk ve boş kabuk olarak sınıflara ayrılmıştır. Çalışma sonucunda farklı kalitelere sahip ceviz içleri %94.8 sınıflandırma doğruluğu ile tespit edilmiştir.

Bu çalışma, ceviz işleme hattında konveyör bant üzerinde bulunan ceviz içi, kabuk ve zar gibi benzer görsel özelliklere sahip bileşenlerin aynı kare içinde gerçek zamanlı olarak sınıflandırılmasına odaklanarak literatüre bu alanda katkı sağlanması amaçlanmaktadır. Mevcut çalışmalarda genellikle ceviz tespiti ya da kalite sınıflandırması tekil nesnelere üzerinde gerçekleştirilirken, bu çalışma çoklu nesne sahnelerinde sınıflandırmayı hedeflemektedir. Geleneksel yöntemlerin sınırlılıklarını aşarak YOLO mimarisinin çoklu nesne ayrımı yeteneğini kullanması, endüstriyel süreçlere doğrudan uygulanabilir bir çözüm sunmaktadır. Böylece, insan gücüne dayalı ayıklama işlemlerine alternatif, optimize edilmiş ve düşük gecikmeli bir yapay görme sistemi önerilmekte; bu yönüyle hem akademik hem de pratik açıdan özgün katkı sağlanmaktadır.

## 2 Materyal ve metod

### 2.1 Deney düzeneği ve donanım

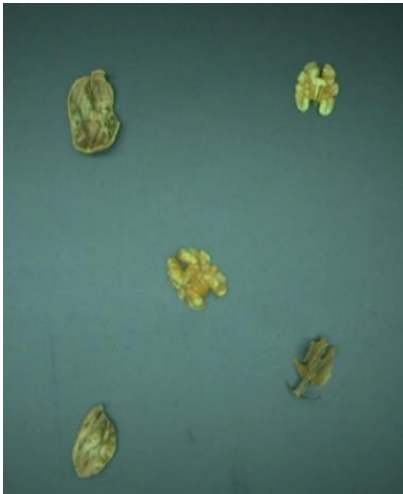
Bu çalışmada ceviz içi, ceviz kabuğu ve ceviz zarlarını yüksek doğrulukta sınıflandırmak ve eğitilen modelin endüstride kullanılabilir olması için Şekil 1'deki gibi endüstriyel bir test düzeneği kullanılmıştır. Şekil 1 ile verilen deney düzeneği konveyör bant, endüstriyel usb renkli kamera (3Mp, 2048 px × 1536 px, 55fps, global shutter, color camera) , 6 adet 12V şerit ledden oluşmaktadır. Aydınlatma olarak beyaz renk seçilmiş olup; arka fonun yansımaması için konveyör bandın rengi siyah olarak seçilmiştir. Modelin eğitimi ve testi için Intel i5 10400F, 32 GB RAM ve RTX 2060 özelliklerine sahip bir masaüstü bilgisayar, Python(v.3.6) programlama dili, YOLOv11 ve RT-DETR modelleri kullanılmıştır.



Şekil 1. Deney düzeneği

## 2.2 Veri seti oluşturma ve etiketleme

Veri seti oluşturmak ve etiketlemek için ceviz içi, ceviz kabuğu ve ceviz zarları konveyör bant üzerine yerleştirilmiş ve kameranın fps ayarı konveyör bant hızına senkron olacak şekilde ayarlanmıştır. Bahsedilen deney düzeneğinde alınan örnek veri seti görüntüsü Şekil 2 ile gösterilmiştir.



Şekil 2. Veri seti örnek görseli

Şekil 2 ile gösterildiği gibi toplamda 583 görüntü çekilmiştir. Bu çekilen görüntüler sırasıyla %70, %15 ve %15 olmak üzere eğitim test ve doğrulama verisi olarak ayarlanmıştır. Oluşturan veri setine ait görüntü sayıları Tablo 1 ile verilmiştir.

Tablo 1. Veri seti eğitim, doğrulama ve test görüntü sayıları

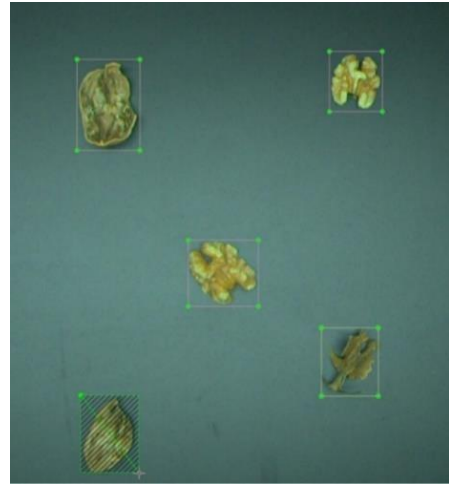
Eğitim (%70)	Test (%15)	Doğrulama (%15)	Toplam
407	88	88	583

Şekil 2 ile verilen görüntüleri etiketlemek için açık kaynak kodlu labeling programı kullanılmış ve Şekil 3 ile

verildiği gibi etiketleme işlemleri yapılmıştır. Tablo 1 ile verilen görüntü sayılarında toplamda 583 görüntü bulunup bu görüntülerdeki 1633 (ceviz içi, ceviz kabuğu ve ceviz zarı) nesne tek tek etiketlenmiştir. Bu etiketlerin eğitim, test ve doğrulama verilerine dağılımı ise Tablo 2 ile verilmiştir.

Tablo 2. Veri setinde bulunan etiketli nesne sayıları

Sınıflar	Eğitim	Test	Doğrulama	Toplam Nesne
Kabuk	883	155	156	1194
İç	448	107	86	641
Zar	302	79	77	458

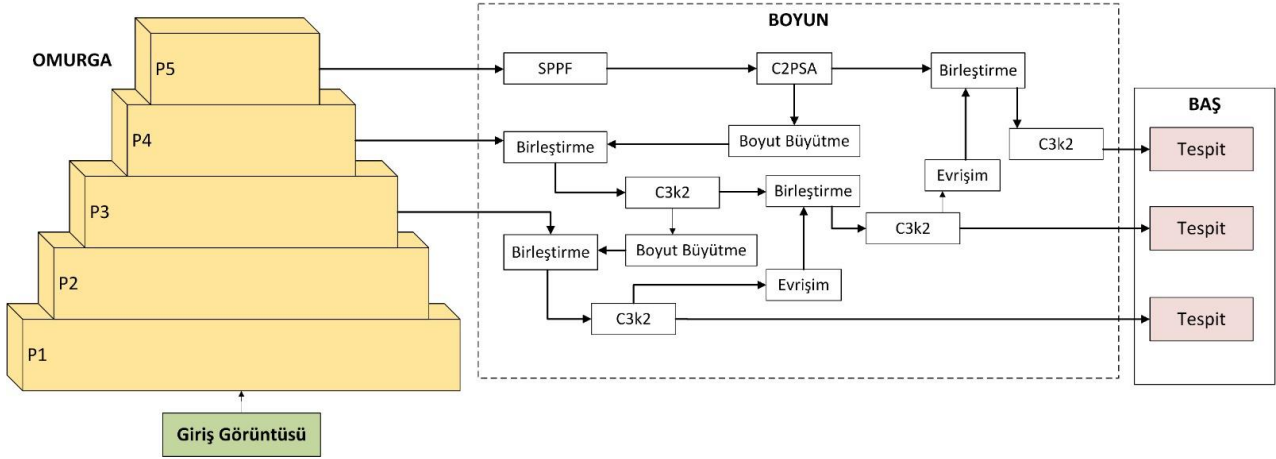


Şekil 3. Veri seti etiketleme örneği

## 2.3 YOLO mimarisi

YOLO (You Only Look Once), gerçek zamanlı nesne tespiti ve sınıflandırma işlemlerini tek bir evrişimli sinir ağı (CNN) üzerinden gerçekleştiren bir derin öğrenme mimarisidir. Bu mimari, geleneksel nesne tespiti yaklaşımlarından farklı olarak, giriş görüntüsünü sabit bir ızgaraya bölerek her hücrede nesne olup olmadığını ve varsa sınıf etiketini ve konum bilgisini doğrudan tahmin eder. Bu tek aşamalı yapı sayesinde hem tespit hem de sınıflandırma işlemleri eşzamanlı olarak yürütülür, bu da YOLO'yu yüksek hız ve düşük gecikme süresi gerektiren uygulamalarda oldukça avantajlı kılar. YOLO mimarileri, nesne tespiti görevlerinde hızlı geri besleme döngüsü, tek görüntü geçişinde çoklu nesne tespiti yapabilmeye ve çeşitli çözünürlüklerde çalışabilme özellikleri sayesinde özellikle mobil ve gömülü sistemlerde sıkça tercih edilmektedir.

Şekil 4 ile gösterilen YOLOv11 mimarisi [30] Ultralytics tarafından YOLOv8 temel alınarak geliştirilen güncel bir nesne tespit mimarisidir. Bu model, önceki versiyonlara kıyasla hem mimari hem de performans açısından önemli iyileştirmeler içermektedir. Temel yenilikler arasında C2f modülünün yerini alan C3k2 (Cross-Stage Partial with kernel size 2) bloğu bulunmaktadır. Bu yapı, daha verimli özellik çıkarımı sağlarken hesaplama maliyetlerini düşürmektedir. Modelin dikkat çeken bir diğer özelliği ise SPPF (Spatial Pyramid Pooling Fast) bloğu çıkışına entegre edilen C2PSA (Cross-Stage Partial with Position Self-Attention) modülüdür.



Şekil 4. YOLOv11 mimarisini

Bu modül, konumsal öz-dikkat mekanizması sayesinde özellikle karmaşık sahnelerde nesne tespit performansını artırmaktadır. Ayrıca, derinlemesine ayrılabilir evrişimlerin (DWConv) kullanımıyla model parametrelerinde optimizasyon sağlanmıştır. YOLOv11, ağ derinliği ve genişliğinde yapılan dengeli ölçeklendirmeler sayesinde farklı karmaşıklık seviyelerindeki tespit görevlerine uyum sağlayabilmektedir. Bu mimari yaklaşım, modelin hem yüksek doğruluk hem de gerçek zamanlı performans hedeflerini karşılamasını mümkün kılmaktadır. Özellikle nesne örtüşmelerinin yoğun olduğu senaryolarda gelişmiş tespit yetenekleri sunmaktadır [31].

#### 2.4 RT-DETR mimarisini

Gerçek Zamanlı Nesne Tespit Transformatörü (Real Time Detection Transformer, RT-DETR), yüksek doğruluk sağlayan ama gerçek zamanlı sistemlere uygulanması zor olan Transformatör tabanlı modellere hız ve verimlilik açısından çözüm getirmek için geliştirilmiştir. YOLO mimarilerinde bulunan gerçek zamanlı tespit yeteneklerine kıyasla, DETR yapıları sınırlı özellikler sunmaktadır. Buna karşın nesne tespit uygulamalarında güçlü performans vermektedir. 2023 yılında yapılan bir çalışmada önerilen RT-DETR [32] modeli, transformatör yapısını yeniden tasarlaması ile gerçek zamanlı çalışmayı mümkün kılmıştır.

RT-DETR mimarisini, hafifletilmiş bir encoder-decoder transformatör yapısı kullanmaktadır. Bu yapıya giriş görüntülerini beslemek için encoder yapısına bağlanan bir CNN yapısı bulunmaktadır. Giriş görüntüsü, CNN mimarisinden geçerek özellik haritası elde edilir. Elde edilen özellik haritası, transformer encoder yapısına beslenir. Encoder, multi-head attention yapılarını kullanarak görüntü içerisindeki nesnelerin ilişkilerini anlamlandırır. Encoder sonrasında gelen decoder yapısı bu ilişkileri ve çıkarılan özellikleri sorgular. Her bir nesne adayını tahmin eder. Mimari içerisinde kullanılan bu yapı sayesinde RT-DETR, nesne tespitinde doğruluk ve çıkarım hızı açısından dengeli bir performans ortaya sunmaktadır. Doğruluk kaybı yaşamadan gerçek zamanlı uygulamalarda kullanılabilen bir mimari olma özelliği taşımaktadır [33].

#### 2.5 Performans değerlendirme ölçütleri

Bu çalışmada nesne tespiti modelinin performansı üç temel ölçüt kullanılarak değerlendirilmiştir. Bunlar: Recall, mAP@50 ve mAP@50-95 değerleridir. Bu değerlerin açıklamaları şu şekildedir:

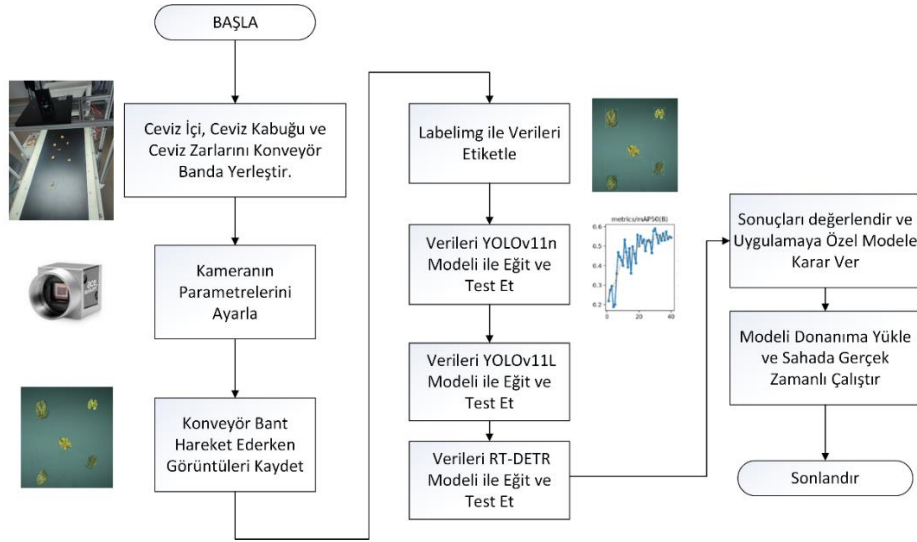
Recall (geri çağırma), modelin görüntüdeki gerçek nesnelere ne ölçüde doğru şekilde tespit edebildiğini gösterir ve doğru pozitif (True Positive) tespitlerin, doğru pozitif ve yanlış negatif (False Negative) tespitlerin toplamına oranı olarak hesaplanır. Yüksek recall değeri, modelin mevcut nesnelerin büyük çoğunluğunu başarıyla tanıdığını gösterir.

Ortalama Doğruluk (mean Average Precision) metriği, IoU eşiği 0.5 olarak alındığında hesaplanan mAP@50, modelin tespit doğruluğunu ölçmek için yaygın olarak kullanılan bir göstergedir; bu metrik, doğru tespit olarak değerlendirilen sonuçlara karşılık gelen precision-recall eğrisinin altındaki alanın hesaplanmasıyla elde edilir.

mAP@50-95 metriği, daha kapsamlı ve hassas bir değerlendirme sunar; bu metrik, IoU eşiği 0.5'ten 0.95'e kadar 0.05 artışlarla hesaplanan ortalama doğruluk (AP) değerlerinin ortalaması alınarak elde edilir. Böylece, modelin nesne konumlandırma başarısını farklı örtüşme koşulları altında daha ayrıntılı şekilde ölçmüş olur.

#### 2.6 Önerilen yöntem

Endüstride konveyör bant üzerindeki ceviz içi, kabuk ve zarların tespit ve sınıflandırılması için Şekil 5'teki gibi bir yöntem önerilmiştir. Birçok çalışma masa üstü veri seti ile yapılmakta olup; bu çalışmada konveyör bant üzerinde endüstriyel kamera ve aydınlatma altında görüntüler alınmış ve devamında bu görüntüdeki nesnelere tek tek etiketlenmiştir. Bu çalışmada etiketli verileri eğitip test etmek için üç farklı model kullanılmıştır. Bu üç model seçilirken yüksek fps değerlerinde işlem yapabilmek adına YOLOv11 mimarisinden YOLOv11n modeli, yine daha yüksek doğruluklar elde edebilmek adına YOLOv11L modeli seçilmiştir. Son olarak YOLO mimarilerinden farklı bir mimari ile nesne tespiti yapılması adına RT-DETR mimarisini kullanılmıştır. Bu mimariye ait parametre sayıları Tablo 3'te verilmiştir.



Şekil 5. Önerilen yöntem

Tüm bu modeller etiketli veriler ile tek tek denenmiş ve performans sonuçları karşılaştırılmıştır. Elde edilen bu karşılaştırma sonuçlarına göre uygulamaya özel bir modelin kullanıcı tarafından seçilmesi amaçlanmıştır.

Tablo 3. Kullanılan mimarilerin parametre sayıları

Mimari	Parametre Sayısı
YOLOv11n	2.6M
YOLOv11L	25.34M
RT-DETR	32.9M

### 3 Bulgular ve tartışma

Bu çalışmada Şekil 5 ile gösterilen önerilen yöntem kullanılarak YOLOv11n, YOLOv11L ve RT-DETR modelleri ile eğitim ve test işlemleri yapılmıştır. Modellerde kullanılan eğitim parametreleri Tablo 4 ile verilmiştir.

Tablo 4. Eğitim parametreleri

Parametre	Değer
Task	Detect
Mode	Train
Epochs	100
Batch size	16
Image size	640
Learning rate	0.01
Momentum	0.937
Weight Decay	0.0005
Data augmentation	Disabled

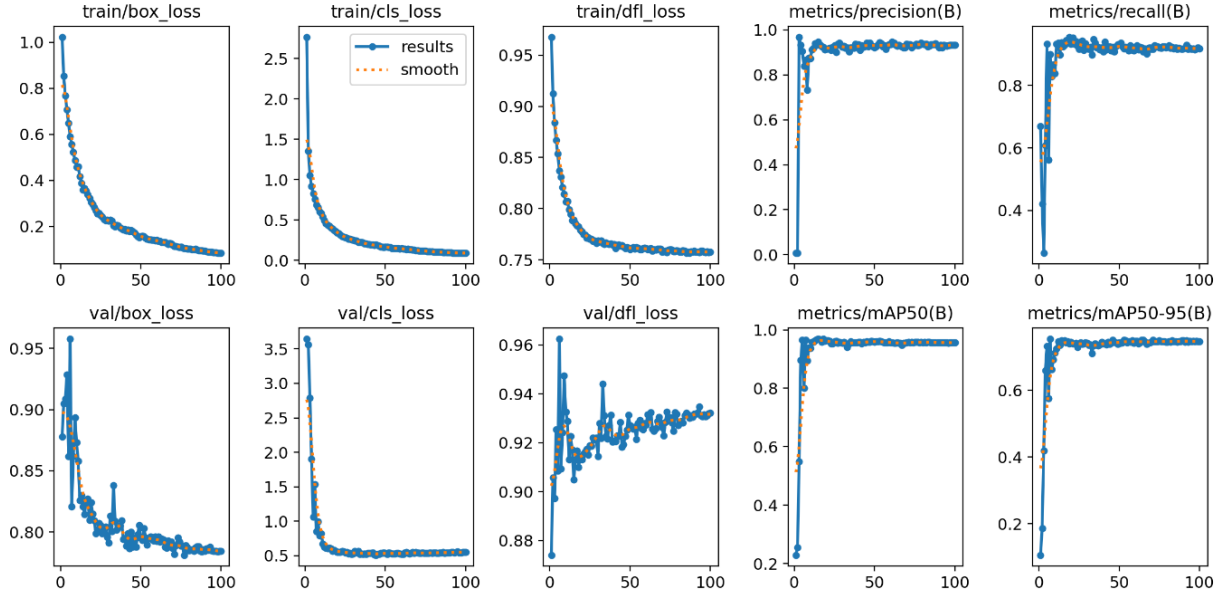
Bu parametreler kütüphanedeki varsayılan başlangıç parametreleri olup; sadece epochs değeri bu uygulamaya göre değiştirilmiştir. Eğitim grafiklerine bakıldığında loss değerleri düzgün olarak azaldığı için model eğitimlerinde overfitting görülmemiş ve daha düşük learning rate değeri kullanılmamıştır. Yine görüntüler kontrollü bir ortamda toplandığı için nesnelere net bir şekilde görülmekte bu yüzden image size 640 olarak uygun görülmüştür. Daha yüksek çözünürlük gerektiren durumlarda ise image size 960 olarak

seçilmesi daha uygun olacaktır. Ancak bu da eğitim süresini uzatacaktır. Tüm bu sebeplerden, bu parametrelerin bu çalışma için uygun olduğuna karar verilmiştir.

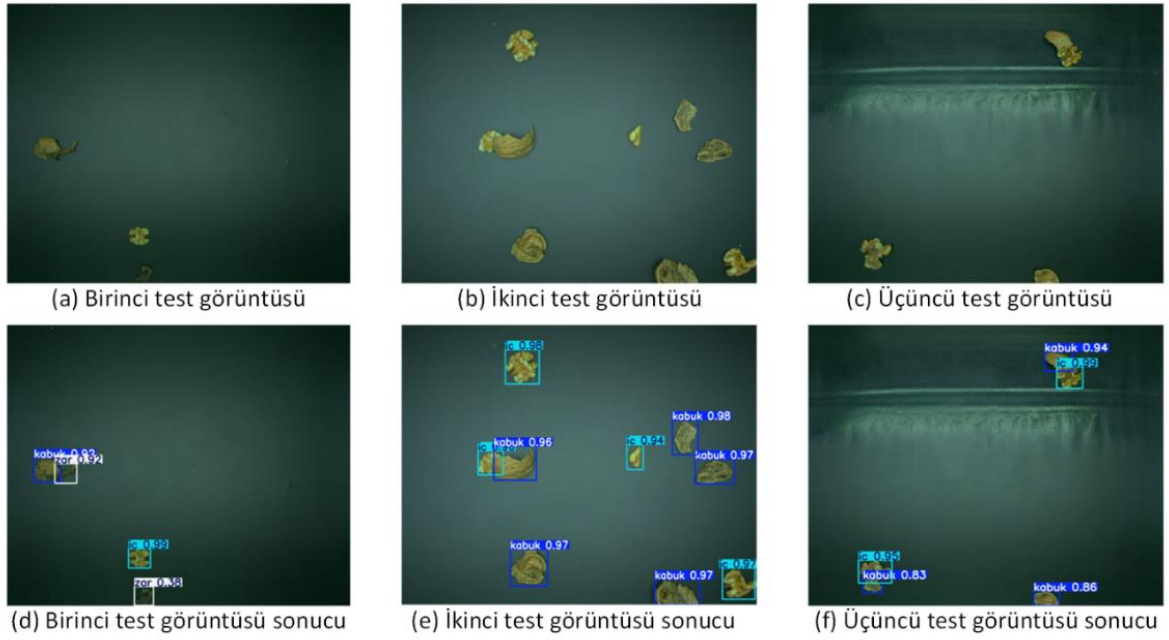
Tablo 2 ile sayıları verilen etiketli veriler YOLOv11n modeli kullanılarak Tablo 4 'deki parametrelere göre eğitilmiş ve yapılan eğitime ait grafikler Şekil 6 ile gösterilmiştir. Doğrulama ve eğitim veri setleri için her adımdaki kayıp ve doğruluk grafikleri Şekil 6 içerisindeki grafiklerde gözlemlenmektedir. YOLOv11n modeli ile eğitim ve test işlemleri tamamlandıktan sonra Şekil 5'teki önerilen yöntemle göre aynı veriler YOLOv11L ve RT-DETR modelleri ile tekrardan eğitilip test edilmiş ve tüm bu eğitimlere ait sonuçlar Tablo 5'te özet olarak verilmiştir. Tablodaki sonuçlara bakıldığında YOLOv11L ve RT-DETR modelleri daha büyük modeller olduğu için beklenildiği üzere bu iki modelde daha yüksek sonuçlar elde edilmiştir. Ceviz içi sınıfı genel olarak tüm modellerde iyi sonuçlar vermiş olup; en zor tespit edilen ise zar sınıfına ait görüntülerdir. YOLOv11n modeli diğer modellere göre düşük doğrulukta çalışmasına rağmen bazı görüntülerde tam doğrulukta çalıştığı gözükmemektedir. Bu tam doğrulukta çalışan görüntülere örnek Şekil 7'te verilmiştir. Şekil 7-d-e-f görüntülerine bakıldığında görüntülerdeki ceviz içi, kabuk ve zar sınıfları tam olarak tespit edilmiştir. YOLOv11L modelinin YOLOv11n modeline olan üstünlüğünü göstermek adına Şekil 8'de karşılaştırmalı görüntüler verilmiştir. Tablo 5'teki verilere göre YOLOv11L modeli tüm sınıflarda YOLOv11n modelinden daha iyi değerlere sahiptir. Örneğin recall değeri en düşük sınıf olan zar sınıfının tespitinde YOLOv11n modeli ile Şekil 8 c'de tek bir zar etiketi olması gerekirken üç tane zar etiketi çizilmiştir. Ancak Şekil 8 b'deki sonuçlara bakıldığında YOLOv11L modelinin zarı doğru bulup doğru etiketleme yaptığı açık ve net bir şekilde gözükmemektedir. Yine Şekil 8c'de ceviz içi olması gereken görüntü YOLOv11n modeli ile ceviz kabuğu olarak tespit edilmiştir. Ancak Şekil 8 b'deki sonuçlara bakıldığında YOLOv11L modelinin ceviz içini doğru bulup doğru etiketleme yaptığı gözükmemektedir.

**Tablo 5.** YOLOv11n, YOLOv11L ve RT-DETR modelleri ile her bir sınıf için elde edilen test sonuçları

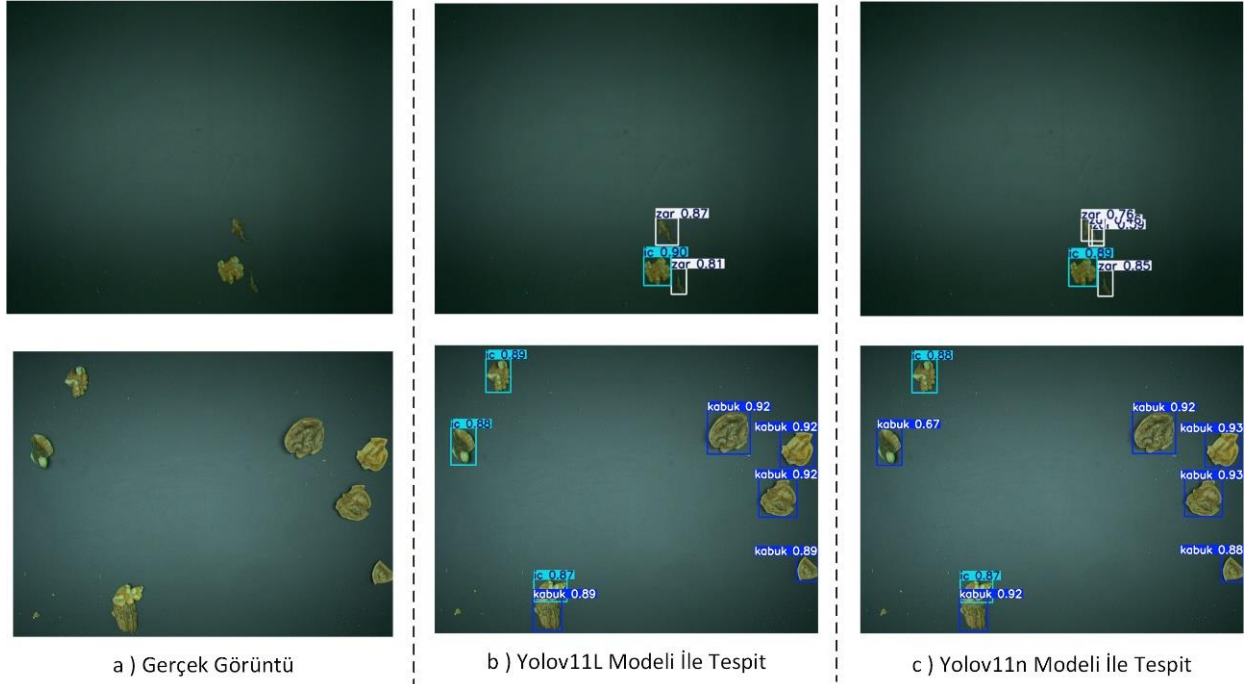
Sınıflar	YOLOv11n			YOLOv11L			RT-DETR		
	Recall	mAP50	mAP 50-95	Recall	mAP50	mAP 50-95	Recall	mAP50	mAP 50-95
Kabuk	0.923	0.940	0.739	0.949	0.979	0.857	0.987	0.993	0.867
Ceviz İçi	0.963	0.977	0.761	0.977	0.993	0.868	0.988	0.986	0.844
Zar	0.795	0.940	0.668	0.922	0.971	0.721	0.896	0.96	0.736



**Şekil 6.** YOLOv11n modeli ile gerçekleştirilen eğitim adımları



**Şekil 7.** YOLOv11n modeli ile test verisi kullanılarak elde edilen sonuçlar



Şekil 8. YOLOv11n ve YOLOv11L modellerinin test verileri ile kıyaslanması

Tüm bu modeller eğitildikten sonra tek bir görüntü için çıkarım süreleri hesaplanmış ve bu sürelerden de fps değerleri hesaplanmıştır. Hesaplanan değerler Tablo 6'daki gibidir.

Tablo 6. Çıkarım süreleri ve fps değerleri

Mimari	Çıkarım Süresi	Fps değeri
YOLOv11n	12.57 ms	79.58
YOLOv11L	28.82 ms	34.70
RT-DETR	44.33 ms	22.56

#### 4 Tartışma ve gelecek çalışmalar

Bu çalışmada önerilen yöntem ile YOLOv11n, YOLOv11L ve RT-DETR mimarileri kullanılarak ceviz içi, kabuk ve zar olmak üzere üç sınıfta nesne tespiti gerçekleştirilmiştir. Elde edilen sonuçlar, özellikle ceviz içi sınıfında yüksek başarı oranının tüm modellerde elde edildiği gözükmektedir. Ceviz içi için en yüksek recall parametresi RT-DETR modeli ile 0.988 olarak hesaplanmıştır. Yine en yüksek mAP50 ve en yüksek mAP50-95 değeri YOLOv11L modeli ile sırasıyla 0.993 ve 0.868 olarak hesaplanmıştır. En düşük değerlere sahip olan sınıf, ceviz zarı sınıfı olup bu sınıftaki en düşük recall değeri YOLOv11n modeli ile 0.795 olarak yine en yüksek recall değeri YOLOv11L modeli ile 0.922 olarak hesaplanmıştır. Gelecek çalışmalarda etiketli zar ve ceviz içi görüntülerinin sayısı en yüksek etiketli sınıf sayısı olan 1194'e kadar çıkarılması hedeflenmektedir. Bu sayede daha homojen bir veri seti elde edilip özellikle zar sınıfı için daha yüksek recall ve mAP50-95 değerlerinin hesaplanması hedeflenmektedir. Etiketli veriler ile eğitilen modeller yüksek tahmin oranları ile çalışmakta olup; Şekil 8'de görüldüğü üzere, temas eden nesnelerin dahi doğru şekilde sınırlayıcı kutular ile tespit edilmesi, modelin segmentasyon

açısından da güçlü olduğunu göstermektedir. Ancak deneysel çalışmalarda bir görüntüde en fazla yaklaşık 10 adet nesne bulunurken, gerçek saha uygulamalarında uygun kamera ve lens sistemleriyle tek bir karede 50'ye yakın nesne yer alabilmektedir. Bu gibi durumlarda YOLO mimarisinin tepkime süresi artabilir ve bu da gerçek zamanlı işleme için gerekli olan fps değerini olumsuz etkileyebilir. Dolayısıyla, bu tür yoğun sahnelerde çıkarım süresi kısa olan modellerin kullanılması daha doğru olacaktır. Bu açıdan bakıldığında YOLOv11n modelinin yüksek bant hızına sahip gerçek zamanlı sistemlere daha uygun olduğu görülmektedir. Ancak bant hızı daha yavaş olan uygulamalarda daha düşük fps değerlerinin kabul edilebilmesi halinde YOLOv11L ve RT-DETR modellerinin de kullanılmasının uygun olduğu görülmektedir. YOLOv11n modeli kullanıldığı halde daha yüksek fps değerlerine çıkılması gerekirse donanım optimizasyonu ve model hafifletme stratejileri büyük önem taşımaktadır. Gelecek çalışmalarda nesne sayısının daha yoğun olduğu ve konveyör bant hızının daha fazla olduğu görüntülerde yüksek fps değerlerinin elde edilmesi için optimizasyon çalışmalarının yapılması hedeflenmektedir. Ayrıca bu çalışmada yalnızca nesne tespiti yapılmış olup, nesne takibi (object tracking) gerçekleştirilmemiştir. Gerçek zamanlı uygulamalarda, özellikle konveyör bant gibi hareketli ortamlarda, nesneler birden fazla görüntüde tespit ediliyorsa bu tespit edilen nesnelerin takibi kritik bir adımdır. Bu nedenle, gelecekteki çalışmalarda Deep SORT gibi nesne takip algoritmalarının entegrasyonu ile sahada gerçek zamanlı nesne tespit ve takip sistemlerinin geliştirilmesi hedeflenmektedir.

#### Çıkar çatışması

Yazarlar çıkar çatışması olmadığını beyan etmektedir.

**Benzerlik oranı (iThenticate): %6**

#### Kaynaklar

- [1] N. Dalal ve B. Triggs, Histograms of oriented gradients for human detection. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), sayfa 886-893, San Diego, ABD, 20-25 Haziran 2005.
- [2] D. G. Lowe, Distinctive image features from scaleinvariant keypoints, International Journal Of Computer Vision, vol. 60, pp. 91-110, 2004.
- [3] N. O'Mahony, S. Campbell, A. Carvalho, S. Harapanahalli, G. V. Hernandez, L. Krpalkova, D. Riordan ve J. Walsh, Deep learning vs. traditional computer vision. Science and Information Conference, sayfa 128-144, 2019.
- [4] V. Pagire, M. Chavali ve A. Kale, A comprehensive review of object detection with traditional and deep learning methods. Signal Processing, 237, 110075, 2025. <https://doi.org/10.1016/j.sigpro.2025.110075>.
- [5] E. Karypidis, S. G. Mouslech, K. Skoulariki ve A. Gazis, Comparison analysis of traditional machine learning and deep learning techniques for data and image classification. arXiv preprint arXiv:2204.05983, 2022. <https://doi.org/10.37394/23206.2022.21.19>.
- [6] I. R. Ward, H. Laga ve M. Bennamoun, RGB-D image-based object detection: from traditional methods to deep learning techniques. in: RGB-D Image Analysis and Processing, Springer, sayfa 169-201, 2019.
- [7] R. P. Haff, T. C. Pearson ve N. Toyofuku, Sorting of in-shell pistachio nuts from kernels using color imaging. Applied Engineering in Agriculture, 26 (4), 633-638, 2010. <https://doi.org/10.13031/2013.32053>.
- [8] P. Vasishth, Desai ve A. Bavarva, Image processing method for embedded optical peanut sorting. International Journal of Image, Graphics & Signal Processing, 7 (12), 2015. <https://doi.org/10.5815/ijjgsp.2015.12.06>.
- [9] Z.-Q. Zhao, P. Zheng, S.-T. Xu ve X. Wu, Object detection with deep learning: A review. IEEE Transactions on Neural Networks and Learning Systems, 30 (11), 3212-3232, 2019. <https://doi.org/10.1109/TNNLS.2018.2876865>.
- [10] A. Krizhevsky, I. Sutskever ve G. E. Hinton, Imagenet classification with deep convolutional neural networks. Advances in Neural Information Processing Systems, 25, 2012. <https://doi.org/10.1145/3065386>.
- [11] K. Simonyan ve A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014. <https://doi.org/10.48550/arXiv.1409.1556>.
- [12] C. Szegedy, W. Liu, Y. Jia, P. Sermanet, S. Reed, D. Anguelov, D. Erhan, V. Vanhoucke ve A. Rabinovich, Going deeper with convolutions. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, sayfa 1-9, 2015.
- [13] K. He, X. Zhang, S. Ren ve J. Sun, Deep residual learning for image recognition. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, sayfa 770-778, 2016.
- [14] M. Tan ve Q. Le, EfficientNet: Rethinking model scaling for convolutional neural networks. International Conference on Machine Learning, sayfa 6105-6114, 2019.
- [15] H. Aktaş, T. Kızildeniz ve Z. Ünal, Classification of pistachios with deep learning and assessing the effect of various datasets on accuracy. Journal of Food Measurement and Characterization, 16 (3), 1983-1996, 2022. <https://doi.org/10.1007/s11694-022-01313-5>.
- [16] Z. Wu, K. Luo, C. Cao, G. Liu, E. Wang ve W. Li, Fast location and classification of small targets using region segmentation and a convolutional neural network. Computers and Electronics in Agriculture, 169, 105207, 2020. <https://doi.org/10.1016/j.compag.2019.105207>.
- [17] J. Redmon, S. Divvala, R. Girshick ve A. Farhadi, You only look once: Unified, real-time object detection. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, sayfa 779-788, 2016.
- [18] J. Redmon ve A. Farhadi, YOLOv3: An incremental improvement. arXiv preprint arXiv:1804.02767, 2018. <https://doi.org/10.48550/arXiv.1804.02767>.
- [19] A. Bochkovskiy, C.-Y. Wang ve H.-Y. M. Liao, YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934, 2020. <https://doi.org/10.48550/arXiv.2004.10934>.
- [20] G. Jocher, A. Stoken, J. Borovec, C. Liu, A. Hogan, L. Diaconu, J. Poznanski, L. Yu, P. Rai, R. Ferriday ve diğerleri, Ultralytics/YOLOv5: v3.0. Zenodo, 2020. <https://doi.org/10.5281/zenodo.3983579>.
- [21] C.-Y. Wang, A. Bochkovskiy ve H.-Y. M. Liao, YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, sayfa 7464-7475, 2023.
- [22] G. Jocher, A. Stoken, J. Borovec, C. Liu, A. Hogan, L. Diaconu, F. Ingham, J. Poznanski, J. Fang, L. Yu ve diğerleri, Ultralytics/YOLOv5: v3.1 – bug fixes and performance improvements. Zenodo, 2020. <https://doi.org/10.5281/zenodo.4154370>.
- [23] D. Guo, Z. Li, H. Shuai ve F. Zhou, Multi-Target vehicle tracking algorithm based on improved DeepSORT. Sensors, 24 (21), 7014, 2024. <https://doi.org/10.3390/s24217014>.
- [24] A. I. B. Parico ve T. Ahamed, Real time pear fruit detection and counting using YOLOv4 models and deep SORT, Sensors, 21 (14), 4803, 2021. <https://doi.org/10.3390/s21144803>.
- [25] J. Lei, W. Zheng, L. Zhang, W. Lv ve Y. Li, Detection of walnut internal quality via improved YOLOv5 and x-ray imaging. Journal of Food Process Engineering, 47 (10), e14742, 2024. <https://doi.org/10.1111/jfpe.14742>.

- [26] Y. Che, H. Bai, L. Sun, Y. Fang, X. Guo ve S. Yin, Real-Time detection of varieties and defects in moving corn seeds based on YOLO-SBWL. *Agriculture*, 15 (7), 685, 2025. <https://doi.org/10.3390/agriculture15070685>.
- [27] Y. Zhang, X. Wang, Y. Liu, Z. Li, H. Lan, Z. Zhang ve J. Ma, Machine Vision-Based Chinese Walnut shell--kernel recognition and separation. *Applied Sciences*, 13 (19), 10685, 2023. <https://doi.org/10.3390/app131910685>.
- [28] M. Wu, L. Yun, C. Xue, Z. Chen ve Y. Xia, Walnut recognition method for UAV remote sensing images. *Agriculture*, 14 (4), 646, 2024. <https://doi.org/10.3390/agriculture14040646>.
- [29] H. Zhang, X. Ning, H. Pu ve S. Ji, A novel approach for the non-destructive detection of shriveling degrees in walnuts using improved YOLOv5n based on X-ray images. *Postharvest Biology and Technology*, 214, 113007, 2024. <https://doi.org/10.1016/j.postharvbio.2024.113007>.
- [30] R. Khanam ve M. Hussain, YOLOv11: An overview of the key architectural enhancements. arXiv preprint arXiv:2410.17725, 2024. <https://doi.org/10.48550/arXiv.2410.17725>.
- [31] A. A. Lubis, A. Prasasta ve D. A. Sari, Towards efficient crowd counting and behavior analysis using YOLOv11. *International Journal of Technology and Modeling*, 4 (1), 35-47, 2025. <https://doi.org/10.63876/ijtm.v4i1.128>.
- [32] Y. Zhao, W. Lv, S. Xu, J. Wei, G. Wang, Q. Dang, Y. Liu ve J. Chen, DETRs beat YOLOs on real-time object detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, sayfa 16965-16974, 2024.
- [33] L. Bustamante ve J. C. Gutiérrez, Enhancing Real-Time Detection Transformer (RT-DETR) for Handgun Detection on Nvidia Jetson. *CLEI Electronic Journal*, 28 (3), 5-1, 2025.

