

İnşaat Sahalarında Güvenlik Kaskı Algılama için YOLO Algoritmalarının Performans Analizi

*Makale Bilgisi / Article Info

Alındı/Received: 08.10.2025

Kabul/Accepted: 18.02.2026

Yayımlandı/Published: 03.06.2026

Performance Analysis of YOLO Algorithms for Safety Helmet Detection in Construction Sites

Mahmut SAĞLIYAN^{1*} , Ahmet ÖZCAN² ¹ Kapadokya Üniversitesi, Bilgisayar ve Bilişim Teknolojileri Fak., Yazılım Geliştirme Bölümü, Nevşehir, Türkiye² Kapadokya Üniversitesi, Bilgisayar ve Bilişim Teknolojileri Fak., Yapay Zeka ve Makine Öğrenmesi Bölümü, Nevşehir, Türkiye

© 2026 The Authors | Creative Commons Attribution-NonCommercial 4.0 (CC BY-NC) International License

Öz

İnşaat sektörü, iş kazalarının en fazla yaşandığı alanlardan biridir. Bu nedenle, işçilerin sağlığını korumak ve kazaları önlemek amacıyla alınacak güvenlik önlemleri, özellikle güvenlik kaskılarının kullanımı, büyük önem taşımaktadır. Derin öğrenme yöntemleri kaskıları tespit edebilse de düşük ışık koşulları, yanlış kullanım ve güvenlik kaskılarının tasarımlarındaki farklılıklar gibi zorluklar performansı etkileyebilmektedir. Bu zorlukların üstesinden gelebilmek için YOLO (You Only Look Once) gibi hızla gelişen derin öğrenme tabanlı nesne tespit algoritmalarının performanslarını karşılaştırmak gereklidir. Bu çalışma, YOLO algoritmasının çeşitli sürümlerinin büyük veri kümeleri üzerindeki performans analizini sunmaktadır. İnşaat sahalarında güvenlik kaskılarının kullanımını izlemek ve iş kazalarını önlemeye yardımcı olmak amacıyla derin öğrenme tabanlı bir nesne tespit sistemi geliştirilmiştir. Bu amaçla, Hard Hat Workers ve BAC_HIEN_CONSTRUCTION_SAFETY_2024 veri setlerinde YOLOv5s, YOLOv8s, YOLOv11s ve YOLOv12s sürümleri değerlendirilmiştir. Deneysel sonuçlar, her bir modeli ortalama doğruluk, kesinlik, duyarlılık ve F1 skoru açısından, ayrıca işleme süresi ve gerçek zamanlı tespit yeteneği bakımından incelemiştir. Bulgular, güvenlik ihlallerinin tespiti için en uygun YOLO sürümünün belirlenmesine dair değerli bilgiler sunmakta ve iş güvenliği sistemlerinde daha etkili yapay zeka tabanlı nesne tespit çözümlerinin geliştirilmesine katkı sağlamaktadır.

Anahtar Kelimeler: Derin öğrenme; YOLO; Nesne tespiti; İş sağlığı ve güvenliği

Abstract

The construction industry consistently reports a high incidence of work-related injuries. Consequently, effective safety interventions—particularly the enforcement of protective helmet use—are essential for mitigating risk. Although contemporary deep learning methods can automatically detect helmet compliance, their performance may deteriorate under adverse conditions (e.g., low illumination), noncompliant wear (e.g., loosely fastened or tilted helmets), and variability in helmet appearance. To address these challenges, this study systematically benchmarks YOLO (You Only Look Once) object detection models. We develop a deep learning-based monitoring system for construction sites and evaluate four YOLO variants—YOLOv5s, YOLOv8s, YOLOv11s, and YOLOv12s—on the Hard Hat Workers and BAC_HIEN_CONSTRUCTION_SAFETY_2024 datasets. Models are compared using mean average precision (mAP), precision, recall, and F1-score, together with inference time and real-time detection capability. The results provide empirical guidance on the trade-off between accuracy and speed across YOLO versions, highlighting the models best suited for reliable helmet detection under realistic site conditions. These findings inform the selection and deployment of robust, AI-driven compliance monitoring solutions and contribute to the design of safer, more effective occupational safety systems.

Keywords: Deep learning; YOLO; Object detection; Occupational health and safety

1. Giriş

Yapay zekâ teknolojilerinin son yıllarda gösterdiği baş döndürücü gelişim, çeşitli alanlarda insan yeteneklerini tamamlayan ya da aşan sistemlerin ortaya çıkmasına neden olmuştur. Bu gelişimin temel yapı taşlarından biri olan makine öğrenmesi, bilgisayarların açık şekilde programlanmaksızın veriler üzerinden öğrenmesine

imkân tanımaktadır (Manorat vd. 2025). Özellikle derin öğrenme yöntemleri, verilerdeki karmaşık ve doğrusal olmayan desenleri öğrenme yetenekleri sayesinde, görsel içeriklerdeki nesneleri yüksek doğrulukla tespit edebilir ve

konumlarını hassas biçimde belirleyebilir (Dogra vd. 2024). Derin öğrenmenin gelişmesiyle birlikte bilgisayarla görü alanında da büyük ilerlemeler yaşanmış; bu alan, görüntülerin analiz edilerek otomatik biçimde yorumlanmasını sağlamaktadır (Huang vd. 2023).

Bilgisayarla görü sistemleri günümüzde endüstriden sağlığa, güvenlikten otonom araçlara kadar birçok uygulama alanında kullanılmaktadır. Bu sistemlerin en önemli alt alanlarından biri olan nesne tanıma, görüntülerde yer alan nesnelerin konumlarının ve türlerinin tespit edilmesini sağlar. Özellikle hızlı ve doğru

sonuçlar üreten algoritmalar sayesinde, gerçek zamanlı uygulamalar mümkün hale gelmiştir (Gheorghe vd. 2024, Mostafa vd. 2024). Bu amaçla geliştirilen en başarılı algoritmalarından biri olan YOLO (You Only Look Once), tek aşamalı ve tüm görüntüyü tek bir ileri besleme işlemiyle değerlendiren yapısıyla dikkat çekmiş; hem işlem hızı hem de doğruluk açısından öne çıkmıştır. Redmon vd. (2016) tarafından sunulan YOLO algoritması, o günden bu yana çeşitli sürümlerle geliştirilmiş ve nesne tespiti konusunda endüstri standardı hâline gelmiştir. YOLO algoritması, tarım alanında Badgujar vd. (2024), enerji alanında Vourkos vd. (2025), sağlık alanında Baldini vd. (2025) gibi çalışmalarla gösterildiği üzere birçok uygulama alanında kullanılmaktadır.

Bu gelişmelerin iş sağlığı ve güvenliği alanındaki uygulamaları da dikkat çekicidir. Özellikle inşaat sektöründe, çalışanların güvenliğini sağlamak amacıyla KKD (kişisel koruyucu donanımların) kullanımı büyük önem taşımaktadır. Dünya genelinde inşaat sektöründe meydana gelen kazaların önemli bir kısmı, güvenlik ekipmanlarının yetersiz veya hatalı kullanımından kaynaklanmaktadır (Vukicevic vd. 2022). Bu bağlamda, çalışanların özellikle güvenlik kaskı kullanıp kullanmadığının otomatik olarak tespiti, iş kazalarının önlenmesinde kritik rol oynamaktadır. İnsan gözlemine dayalı denetim süreçlerinin hem zaman alıcı hem de hataya açık olması, bu alanda otomatik denetim sistemlerinin geliştirilmesini gerekli kılmaktadır (Jalonen vd. 2023).

Bu çalışmanın temel amacı, işçilerin güvenlik kaskı takıp takmadığını gerçek zamanlı olarak tespit edebilen bir sistem kurmak ve bunun için YOLO mimarisinin farklı sürümlerinden YOLOv5s, YOLOv8s, YOLOv11s ve YOLOv12s modellerini kullanarak bu modellerin başarımını karşılaştırmaktır. Bu sayede hem akademik hem de endüstriyel uygulamalar için en uygun modelin belirlenmesi ve iş sağlığı alanında yapay zekâ tabanlı karar destek sistemlerinin yaygınlaştırılması hedeflenmektedir.

2. Literatür Özeti

Güvenlik kaskı tespiti üzerine yapılan çalışmaların çoğu, YOLO mimarisi gibi gerçek zamanlı nesne algılama algoritmalarının gücünü kullanmaktadır. Özellikle inşaat, otomotiv ve iş güvenliği gibi kritik alanlarda, kişisel koruyucu donanımların kullanımını tespit etmek amacıyla bu teknolojilerin kullanımına yönelik birçok çalışma yapılmıştır. Gheorghe vd. (2024), YOLO algoritmasının otomotiv sektöründeki nesne tespiti uygulamalarını incelemiş ve gerçek zamanlı performansın önemini vurgulamıştır. Bu çalışma, YOLO'nun farklı senaryolardaki başarımını ortaya koyması açısından temel bir referans

niteliği taşımaktadır. Bununla birlikte, Zhou vd. (2021) tarafından geliştirilen AT-YOLO mimarisi ile güvenlik kasklarının çeşitli sahnelerde yüksek doğrulukla tespit edilebildiği gösterilmiş, bu da algoritmanın esnekliğini ortaya koymuştur.

Literatürde özellikle inşaat sektörüne odaklanan birçok çalışmada, güvenlik kaskı kullanımının otomatik olarak tespitine yönelik çeşitli yöntemler geliştirilmiştir. Zhang vd. (2025), baş bölgesi eşleştirmesi ve durum izleme teknikleri ile kask kullanımını tespit etmeye yönelik bir sistem önermiştir. Bu çalışmada davranışsal izleme ile nesne tespiti bir arada değerlendirilmiş ve gerçek zamanlı performans testleri yapılmıştır. Lee vd. (2023), CNN tabanlı iki aşamalı kask tespit yöntemlerini karşılaştırarak, farklı mimarilerin doğruluk oranlarını değerlendirmiştir. YOLO mimarisinin farklı versiyonlarının yüzey algılama performanslarına yönelik bir diğer çalışmada ise, doğruluk ve işlem süresi gibi metrikler karşılaştırılmıştır (Raj & Prabadevi, 2025).

Kask tespitine yönelik algoritmalar yalnızca inşaat sahalarıyla sınırlı kalmayıp, trafik güvenliği gibi farklı alanlarda da kullanılmaktadır. Said vd. (2024), trafik yönetim sistemlerinde AI tabanlı kask ihlali tespit sistemlerini geliştirmiş ve bu yöntem gerçek zamanlı sistemlerde başarılı sonuçlar vermiştir. Benzer şekilde, Kwak ve Kim (2023), hem güvenlik kaskı hem de maske kullanımını tespit edebilen sistemler geliştirerek kimlik tanıma özelliklerini de entegre etmişlerdir. Bu, çok bileşenli kişisel koruyucu donanım denetiminde önemli bir adımdır. Jamtsho vd. (2021), plaka tespiti ile entegre bir kask algılama sistemini önererek, kask takmayan motosiklet sürücülerini belirlemeyi amaçlamıştır.

Geliştirilen sistemlerin bir diğer boyutu ise hava görüntüleri üzerinden kask ve diğer güvenlik ekipmanlarının tespit edilmesidir. Liu vd. (2024), drone görüntüleri üzerinden yüksekte çalışan işçilerin emniyet kemeri kullanıp kullanmadığını YOLO algoritması ile analiz etmiş ve bu alanda yeni bir uygulama alanı açmıştır. Lin vd. (2023), YOLOv5s modeli ile motor sürücülerinin kask kullanımını belirlemeye yönelik bir sistem önermiştir. Chen vd. (2022) ise havadan alınan görüntülerde güvenlik kaskı tespiti için iyileştirilmiş YOLOv4 tabanlı bir sistem geliştirmiştir. Çiftçi vd. (2024), YOLOv5n ve YOLOv8n modellerini kullanarak 2 farklı veri seti üzerinde inşaat sahalarında güvenlik ekipmanı tespiti gerçekleştirmiş ve bu iki modelin performansını karşılaştırmıştır.

Ayrıca, çalışanların kimliklerini belirleyen ve kişisel koruyucu ekipman kullanımını gerçek zamanlı olarak tespit eden entegre sistemlerin de geliştirildiği görülmektedir. Örneğin, Son ve Kim (2023), inşaat

makineleriyle çalışan personelin güvenliğini artırmak amacıyla kişi takibi ve kask denetimini birleştiren sistemler tasarlanmıştır. Benzer şekilde, Nazlı vd. (2024), YOLOv5 modeli kullanarak kişisel koruyucu donanım kullanımını değerlendiren gerçek zamanlı sistem önermiştir. Han vd. (2021), Single Shot Multibox Detector tabanlı yeni bir nesne algılama algoritması sunarak, çok ölçekli algılama yapısının karmaşık ortamlarda bile etkili olduğunu göstermiştir. Li vd. (2024) ise YOLOv4 mimarisini geliştirerek YOLO-PL modelini önermiş ve bu modelin kask tespiti performansını artırdığını belirtmiştir.

Yukarıda özetlenen literatür, güvenlik kaskı tespiti alanında YOLO mimarisinin güçlü bir temel sunduğunu ortaya koymaktadır. Mevcut çalışmalar, farklı senaryolarda (trafik, inşaat, havadan çekilmiş görüntüler) ve farklı donanım konfigürasyonlarında çeşitli YOLO modellerinin başarılarını gözler önüne sermiştir. Ancak bu çalışmaların çoğu tek bir model üzerinde yoğunlaşmakta ya da farklı YOLO versiyonların karşılaştırmasına yeterince yer vermemektedir. Bu noktada, bizim çalışmamız YOLO mimarisinin dört farklı versiyonu olan YOLOv5s, YOLOv8s, YOLOv11s ve YOLOv12s modellerini hem doğruluk hem de işlem süresi bakımından karşılaştırarak, güvenlik kaskı tespiti açısından en uygun modeli belirlemeyi amaçlamaktadır. Bu yönüyle çalışmamız, literatürdeki mevcut teknik yaklaşımları genişletmekte ve uygulamaya dönük katkılar sunmaktadır.

3. Materyal ve Metot

3.1 Veri Seti

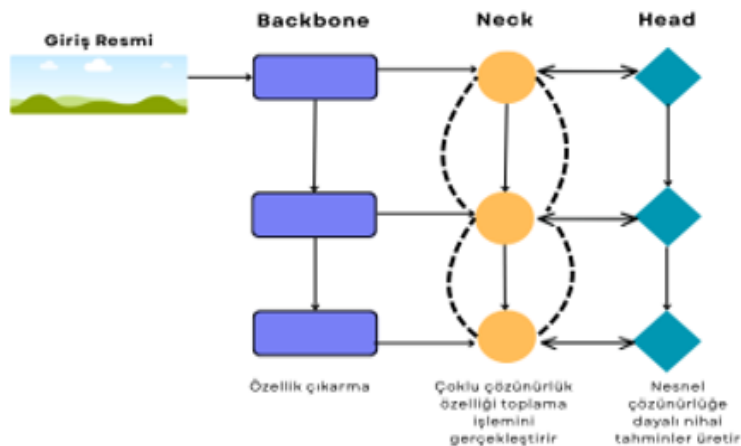
Model eğitiminde kullanılan veri setleri, makine öğrenme projelerinde sıklıkla kullanılan platformlardan birisi olan Roboflow üzerinden elde edilmiştir. Hard Hat Workers veri seti iş yeri ortamlarında kask takması gereken çalışanların nesne algılama veri setidir (Northeastern University - China, 2022). Açıklamalar ayrıca, bir bireyin kask takmadan mevcut olabileceği durumlar için sadece

"helmet" ve "head" sınıflarını içerir. Veri ön işleme ve veri artırma işlemleri uygulanmış olan bu veri setinde eğitim seti için 14748, Doğrulama seti için 1413 ve test seti için de 706 görüntü olmak üzere veri seti toplamda 16867 görüntüden oluşmaktadır. Diğer bir veri seti olan BAC_HIEN_CONSTRUCTION_SAFETY_2024 ise "helmet", "no helmet" ve "person" olmak üzere 3 farklı sınıf içerir. Eğitim seti için 15566, doğrulama ve test veri setinde 1945'er adet görüntü kullanılmıştır. Toplamda 19456 görüntüden oluşan bu veri setinde veri ön işleme ya da veri artırma işlemi uygulanmamıştır (Constructionsafety, 2024).

3.2 Kask Tespitinde Kullanılan Yöntemler

Bu çalışmada, görüntü tabanlı nesne tespiti görevini gerçekleştirmek amacıyla evrişimli sinir ağları temelli bir yaklaşım olan YOLO algoritması tercih edilmiştir. CNN mimarileri, görüntü verilerinden düşük ve yüksek seviyeli öznitelikleri çıkarma ve bu öznitelikleri katmanlar boyunca öğrenbilme yetenekleri sayesinde bilgisayarla görü alanında geniş bir uygulama yelpazesinde yaygın şekilde kullanılmaktadır. YOLO algoritması, CNN tabanlı mimari yapısı aracılığıyla giriş görüntüsünü tek bir ileri besleme işlemiyle analiz ederek eşzamanlı olarak hem sınıflandırma hem de nesne konumlandırma tahminlerini gerçekleştirebilmektedir. Bu bütünleşik yapı, modelin hem hızlı çalışmasını hem de gerçek zamanlı uygulamalara uygunluğunu mümkün kılmaktadır.

Çalışma kapsamında, YOLO algoritmasının farklı sürümleri olan YOLOv5s, YOLOv8s, YOLOv11s ve YOLOv12s modelleri üzerinde karşılaştırmalı analizler gerçekleştirilmiştir. Her bir model, mimari bileşenler bakımından farklı tasarım yaklaşımlarına sahiptir. YOLO mimarisi Şekil 1'de belirtildiği üzere genel olarak üç temel bileşenden oluşmaktadır: Backbone (Öznitelik Çıkarıcı), Neck (Ara Katman) ve Head (Tahmin Katmanı).



Şekil 1 YOLO ana bileşenleri

3.2.1 Backbone (Öznitelik Çıkarıcı)

Backbone katmanı, giriş görüntüsünden düşük ve yüksek seviyeli öznitelikleri çıkarmak amacıyla yapılandırılmıştır:

YOLOv5s sürümünde, backbone olarak CSPDarknet mimarisi kullanılmaktadır. Bu yapı, CSP (Cross Stage Partial) blokları sayesinde bilgi akışını dengelemekte, modelin parametre sayısını azaltırken doğruluk performansını artırmayı amaçlamaktadır.

YOLOv8s sürümünde, backbone olarak daha optimize ve modüler bir yapı olan C2f (Concatenate-to-Fuse) tabanlı mimari tercih edilmektedir. Bu yapı, önceki sürümlere kıyasla daha az parametre ile daha yüksek doğruluk sağlamaktadır.

YOLOv11s ve YOLOv12s sürümleri, henüz literatürde tam standartlaşmamış deneysel veya özel geliştirilmiş modeller olup, backbone katmanlarında sıklıkla Transformer tabanlı mimariler veya hibrid CNN-Transformer blokları kullanılmaktadır. Özellikle YOLOv12s modelinde, CNN yapıları ile dikkat mekanizmalarının (attention mechanisms) entegrasyonunu içeren gelişmiş backbone mimarileri uygulanmaktadır.

3.2.2 Neck (Ara Katman)

Neck katmanı, farklı çözünürlük seviyelerinde çıkarılan öznitelikleri birleştirerek hem küçük hem de büyük nesnelerin daha etkin biçimde algılanmasını sağlamaktadır:

YOLOv5s modelinde, neck bileşeni olarak PANet (Path Aggregation Network) yapısı kullanılmaktadır. Bu yapı, alt ve üst katmanlardan gelen bilgileri birleştirerek dengeli bir özellik dağılımı elde edilmesine olanak tanımaktadır.

YOLOv8s modelinde, PANet yerine daha verimli ve optimize edilmiş olan BiFPN (Bidirectional Feature Pyramid Network) benzeri bir yapı kullanılmaktadır. Bu yapı, çift yönlü bilgi akışını optimize ederek farklı seviyelerdeki özniteliklerin daha etkin şekilde harmanlanmasını sağlamaktadır.

YOLOv11s ve YOLOv12s sürümlerinde, neck katmanında attention mekanizmaları ve Transformer Fusion Block gibi ileri düzey birleşim stratejileri uygulanmaktadır. Bu sayede farklı seviyelerde çıkarılan öznitelikler daha anlamlı ve detaylı şekilde entegre edilmektedir.

3.2.3 Head (Tahmin Katmanı)

Head katmanı, öznitelik haritaları üzerinden sınıf ve konum tahminlerini (bounding box koordinatları) gerçekleştiren son aşamayı oluşturmaktadır:

YOLOv5s ve YOLOv8s modellerinde, sırasıyla anchor tabanlı (YOLOv5s) ve anchor-free (YOLOv8s) tahmin

mekanizmaları kullanılmaktadır. Anchor-free yapı, hiperparametre gereksinimini azaltarak daha esnek ve hızlı öğrenme sağlamaktadır.

YOLOv11s ve YOLOv12s sürümleri ağırlıklı olarak anchor-free head yapılarını benimsemekte olup, bunun yanında Dense Prediction Head ve Dynamic Head gibi gelişmiş tahmin katmanları da kullanılmaktadır. Bu gelişmiş yapıların amacı, özellikle karmaşık sahnelerdeki küçük veya örtüşen nesnelerin daha hassas ve güvenilir şekilde tespit edilebilmesidir.

Kullanılan yöntemin ana bileşenlerine ait diyagram şekil 1 de verilmiştir.

3.3 Eğitim Süreci

Model eğitimleri Jupyter Notebook hizmeti sağlayan Google Colab da T4 GPU kullanılarak gerçekleştirilmiştir. Eğitimler önceden eğitilmiş ağırlıklar (transfer learning) kullanılarak gerçekleştirilmiştir. Tüm YOLO versiyonları için eğitimlerde kullanılan parametreler çizelge 1' de verilmiştir.

Çizelge 1 Model Eğitiminde Kullanılan Parametreler

Parametre	Açıklama	Değer
size (pixels)	Giriş görüntü boyutu	640
batch_size	Her iterasyonda kullanılan görüntü sayısı	16
epoch	Eğitim tekrar sayısı (tüm veri seti üzerinden)	20
lr0	Başlangıç öğrenme oranı	0.01
lrf	Son öğrenme oranı çarpanı (final LR factor)	0.01
momentum	Momentum değeri	0.937
weight_decay	Ağırlık çürüme katsayısı (regularizasyon)	0.0005
warmup_epochs	Isınma (warmup) epoch sayısı	3.0
warmup_momentum	Isınma süresince başlangıç momentum değeri	0.8
warmup_bias_lr	Isınma süresince bias katmanları için öğrenme oranı	0.1

3.4 Değerlendirme Ölçütleri

Model performansını değerlendirmek amacıyla; ortalama doğruluk (mean Average Precision), kesinlik (precision), duyarlılık (recall), F1-skoru (F1-score) ve FPS (Frame Per Second) gibi yaygın başarımlar ölçütlerinden faydalanılmıştır. Kesinlik, modelin pozitif olarak sınıflandırdığı örnekler içerisindeki doğru sınıflandırılmış örneklerin oranını ifade eder ve doğru pozitif (TP) sayısının toplam pozitif tahmin sayısına (TP + FP) bölünmesiyle hesaplanır. Yüksek kesinlik değeri, modelin tahmin ettiği nesnelerin büyük bir çoğunluğunun gerçekten doğru olduğunu göstermektedir.

$$Precision = \frac{TP}{TP+FP} \quad (1)$$

Duyarlılık ise modelin, tüm gerçek nesneler içerisinde kaç tanesini doğru şekilde tespit ettiğini gösterir ve doğru pozitif (TP) sayısının, gerçek pozitiflerin toplamına (TP + FN) oranlanmasıyla elde edilir. Bu metrik, modelin kaç nesneyi tespit edebildiğini ölçmektedir.

$$Recall = \frac{TP}{TP+FN} \quad (2)$$

mAP (mean Average Precision), her bir sınıfa ait kesinlik-duyarlılık eğrisi altında kalan alanın (AP) ortalaması alınarak hesaplanır ve modelin genel tespit performansını özetleyen önemli bir göstergedir. IoU (Intersection over Union), nesne tespiti ve görüntü segmentasyonu gibi bilgisayarlı görü uygulamalarında kullanılan bir başarımlı ölçütüdür. Bir modelin tahmin ettiği kutunun gerçek kutu ile ne kadar örtüştüğünü ölçer. Özellikle mAP@0.5 (IoU eşik 0.5) ve mAP@0.5:0.95 (IoU eşik aralığı 0.5'ten 0.95'e kadar 0.05 artışlarla) değerleri, nesne tespit modellerinin doğruluğunu farklı hassasiyet seviyelerinde değerlendirmek için kullanılmaktadır.

$$mAP@50 = \frac{1}{N} \sum_{i=1}^N AP_i^{(IoU=0.5)} \quad (3)$$

$$mAP@50:95 = \frac{1}{10} \sum_{k=0}^9 \frac{1}{N} \sum_{i=1}^N AP_i^{(IoU=0.5+0.05k)} \quad (4)$$

F1-skor, kesinlik ve duyarlılık değerlerinin harmonik ortalaması alınarak elde edilir ve bu iki ölçüt arasında denge sağlayan bir metrik olarak kullanılır. Özellikle veri dengesizliği durumlarında modelin başarımını objektif biçimde yansıtmaya açısından önem taşımaktadır. Bu metriklerin birlikte kullanılması, modellerin farklı yönlerden değerlendirilmesine olanak sağlamış ve YOLO serisi modellerin doğruluk, hassasiyet ve genelleme yeteneklerinin kapsamlı bir biçimde analiz edilmesine zemin hazırlamıştır.

Çizelge 2 YOLO modellerinin Hard Hat Workers veri setindeki performans değerlendirilmesi

Model	Kesinlik	Duyarlılık	F1 Skoru	mAP 50	mAP 50-95	Eğitim Süresi	İnf (ms)	Pre (ms)	Post (ms)	FPS (İnf)	FPS (Total)
YOLOv5s	0,949	0,938	0,944	0,971	0,666	98 dk	7,4	0,5	1,2	135	110
YOLOv8s	0,943	0,934	0,938	0,967	0,667	94 dk	8,6	0,4	1,1	116	99
YOLOv11s	0,947	0,933	0,940	0,970	0,677	104 dk	8,8	0,4	1,1	113	97
YOLOv12s	0,948	0,938	0,943	0,971	0,680	140 dk	12,6	0,5	1	79	71

Çizelge 3: YOLO modellerinin BAC_HIEN_CONSTRUCTION_SAFETY_2024 veri setindeki performans değerlendirilmesi

Model	Kesinlik	Duyarlılık	F1 Skoru	mAP 50	mAP 50-95	Eğitim Süresi	İnf (ms)	Pre (ms)	Post (ms)	FPS (inf)	FPS (Total)
YOLOv5s	0,824	0,674	0,742	0,720	0,450	106 dk	7,6	0,3	1,3	131	109
YOLOv8s	0,829	0,679	0,747	0,732	0,467	104 dk	8,9	0,4	1,1	112	96
YOLOv11s	0,819	0,671	0,737	0,738	0,468	109 dk	9,3	0,3	1,1	107	93
YOLOv12s	0,827	0,682	0,748	0,743	0,475	151 dk	13,5	0,4	1	74	67

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (5)$$

Son olarak FPS, bir modelin veya sistemin saniyede kaç görüntü işleyebildiğini gösteren performans ölçütüdür. Nesne tespiti gibi bilgisayarla görme görevlerinde, FPS değeri modelin gerçek zamanlı çalışmaya uygun olup olmadığını belirlemede kullanılan bir metriktir. Modellerin FPS (inference) değerlerini hesaplamak için kullanılan formül denklem 6' da belirtilmiştir. FPS (total) hesaplamada kullanılan formül ise denklem 7' de verilmiştir.

$$FPS = \frac{1000}{\text{time inference}} \quad (6)$$

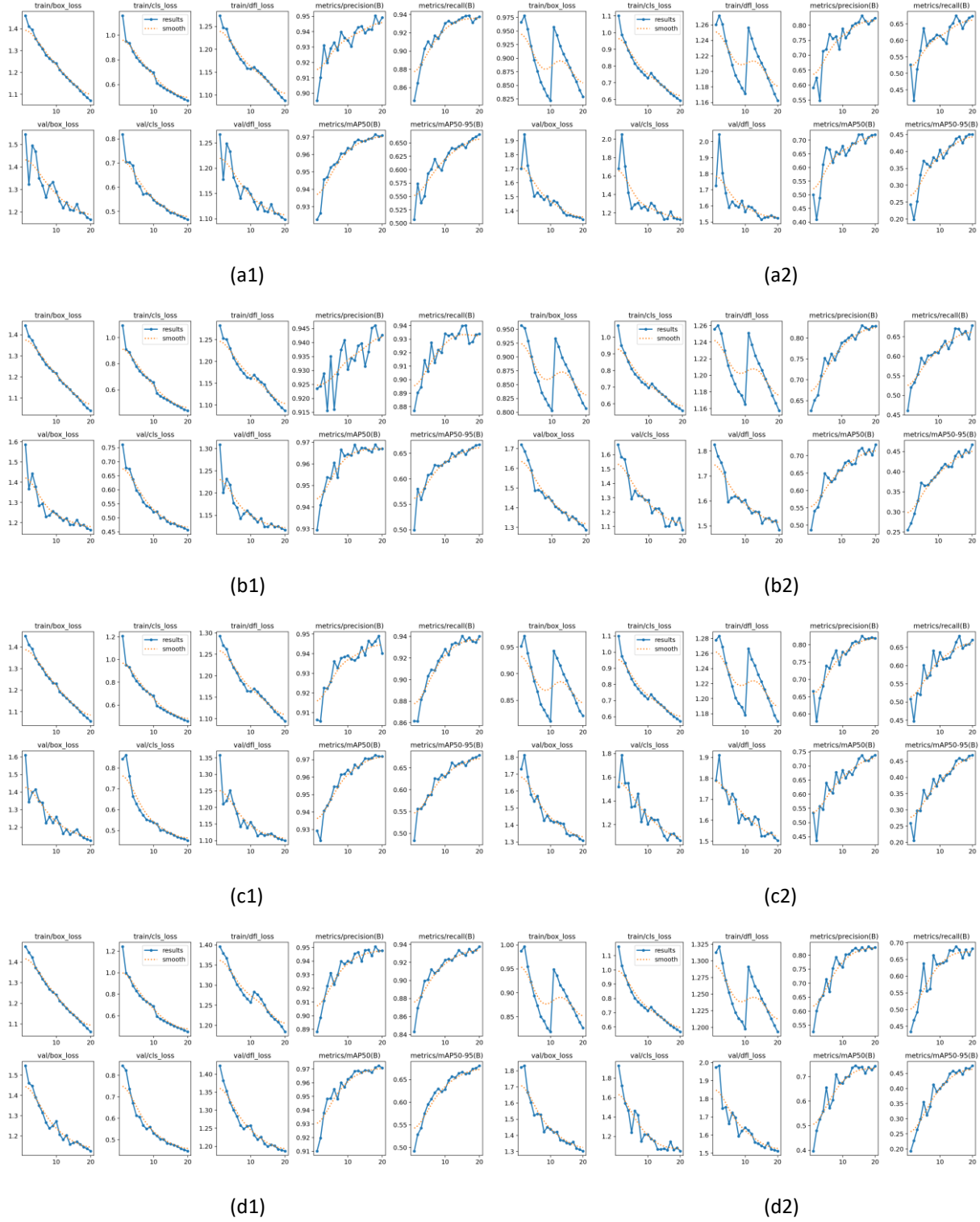
$$FPS = \frac{1000}{\text{preprocess} + \text{inference} + \text{postprocess}} \quad (7)$$

Time inference tek bir görüntü için ortalama çıkarım süresi, preprocess, model tahmini öncesinde görüntünün yeniden boyutlandırılması ve normalize edilmesi gibi hazırlık işlemlerinin süresi, postprocess ise model çıktılarının NMS (Non-Maximum Suppression) yöntemi ve eşikleme gibi adımlarla işlenerek son tahminlerin üretilmesi için gerekli süredir. "inf", "pre" ve "post" sırasıyla inference, preprocess ve postprocess süreleridir. 1000 katsayısı ise 1 saniyenin ms cinsinden karşılığıdır.

4. Araştırma Sonuçları ve Tartışma

4.1 Performans Değerlendirmesi

Kask tespitini gerçekleştirmek amacıyla gerçekleştirilen model eğitimimizde epoch sayısı 20 olarak belirlenmiştir. Birinci ve İkinci veri setlerinde sırasıyla 16867 ve 19456 adet görüntü ve her görüntüde çok sayıda kask nesnesinin bulunması nedeniyle her iki veri seti için de 20 epoch gibi düşük sayılabilecek bir değer ile eğitimler gerçekleştirilmiştir. Eğitim sonucu elde edilen değerler çizelge 2 ve çizelge 3 de verilmiştir.



Şekil 2 YOLO modellerinin iki farklı veri seti üzerindeki performans sonuçları. Alt görseller sırasıyla; (a1)–(a2) YOLOv5s, (b1)–(b2) YOLOv8s, (c1)–(c2) YOLOv11s ve (d1)–(d2) YOLOv12s modellerinin Hard Hat Workers ve BAC_HIEN_CONSTRUCTION_SAFETY_2024 veri setlerindeki sonuçlarını göstermektedir.

Şekil 2’de, dört farklı YOLO modelinin iki veri seti üzerindeki performans karşılaştırması sunulmuştur. Burada “1” numaralı görseller Hard Hat Workers veri setini, “2” numaralı görseller ise BAC_HIEN_CONSTRUCTION_SAFETY_2024 veri setini temsil etmektedir.

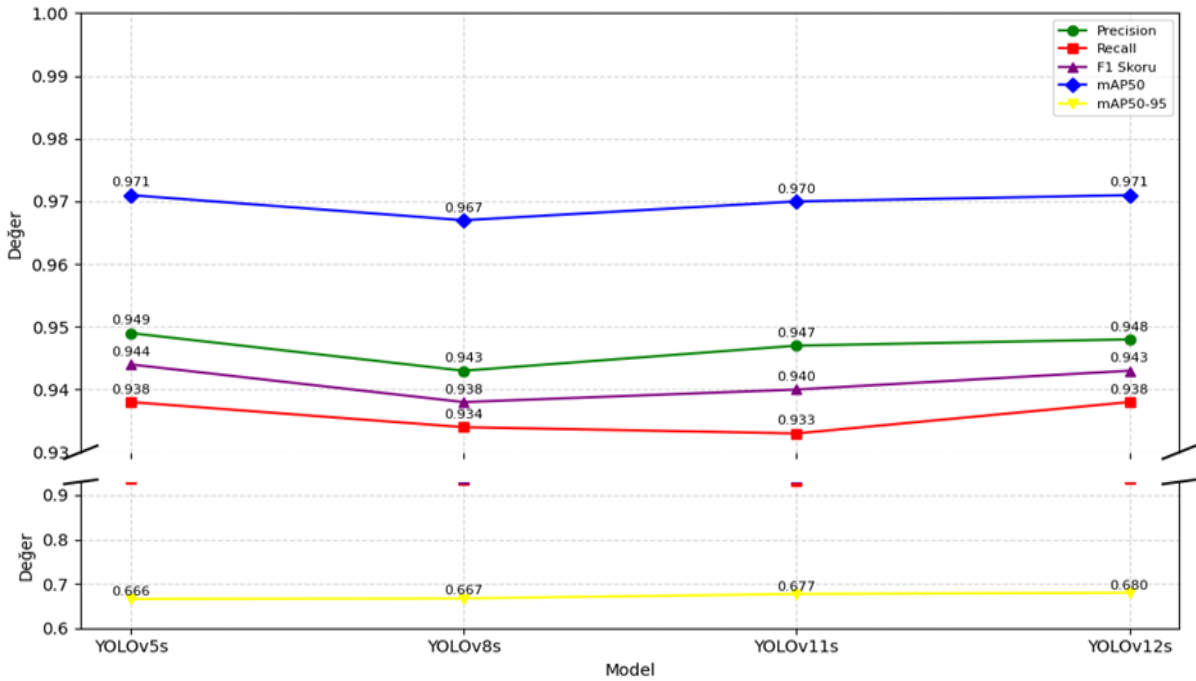
YOLO modellerinin sonuçları, Hard Hat Workers veri setinde BAC_HIEN_CONSTRUCTION_SAFETY_2024 veri setine kıyasla daha yüksek ve daha tutarlı bir performans sergilediğini göstermektedir. BAC_HIEN veri setinde kayıpların daha dalgalı seyretmesi ve doğruluk metriklerinin düşmesi, veri setinin daha karmaşık ve

öğrenmesi zor bir yapıya sahip olduğunu ortaya koymaktadır.

Çizelge 2'deki sonuçlara baktığımızda gerçekleştirilen eğitim sonuçlarına göre, tüm modeller başarılı performans göstermiştir. YOLOv5s modeli 0,949 kesinlik, 0,938 duyarlılık ve 0,944 F1 skoru elde ederek genel olarak en yüksek doğruluk oranını sağlamıştır. Benzer şekilde, YOLOv12s modeli 0,948 kesinlik, 0,938 duyarlılık ve 0,943 F1 skoru ile oldukça yakın bir performans göstermiştir. mAP50 değerleri incelendiğinde, YOLOv5s ve YOLOv12s modelleri 0,971 oranında en yüksek mAP50

başarısını yakalamıştır. mAP50-95 değerlerinde ise YOLOv12s modeli 0,680 ile en yüksek değere ulaşmıştır.

Eğitim süresi açısından değerlendirildiğinde, YOLOv8s modeli 94 dakika ile en kısa sürede eğitimi tamamlamıştır; buna karşın YOLOv12s modelinin eğitim süresi 140 dakika olmuştur. Bu sonuçlar, özellikle YOLOv12s modelinin doğruluk ve genelleme kapasitesini artırırken eğitim süresinde belirgin bir artış meydana getirdiğini göstermektedir. Modellerin bu metriklerdeki dağılımı ve aralarındaki farklılıkların daha net görülebilmesi adına elde edilen performans ölçütleri Şekil 3' te görselleştirilmiştir.



Şekil 3 YOLO modellerinin Hard Hat Workers veri seti üzerinde performans karşılaştırması

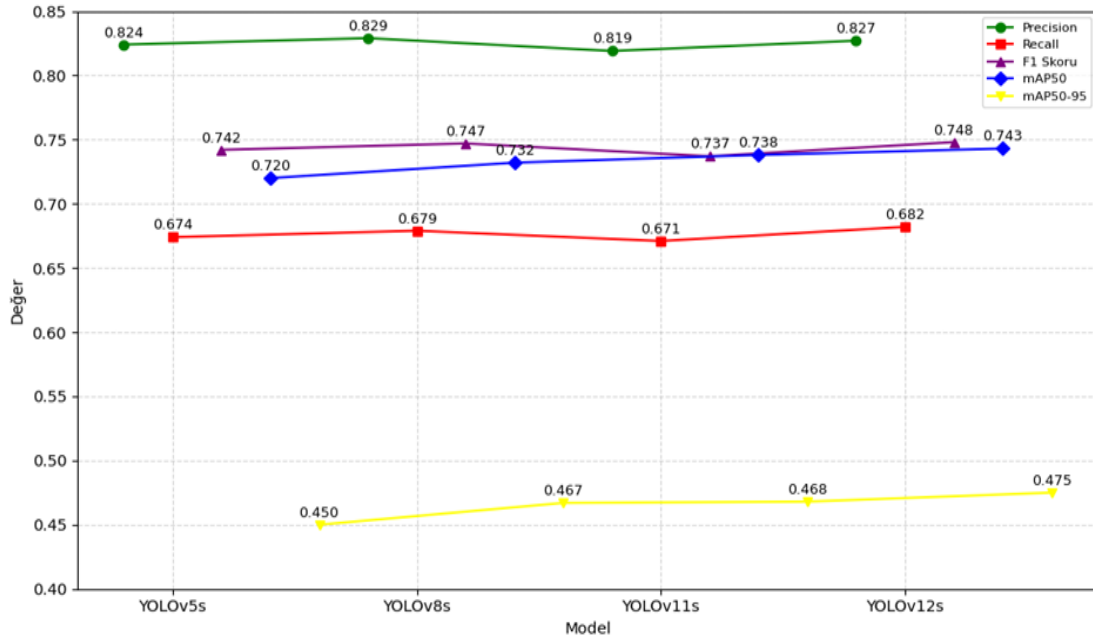
Bu bulgular, farklı YOLO modelleri arasında performans-maliyet dengesi açısından önemli farklılıklar olduğunu ortaya koymaktadır. Özellikle YOLOv12s modeli, mAP50-95 metriğinde elde ettiği sonuçla daha karmaşık ve zorlayıcı nesne tespit senaryoları için daha uygun bir alternatif olarak öne çıkmaktadır. Bununla birlikte, bu modelin eğitim süresi anlamında daha fazla kaynak ve zaman gerektirmesi, uygulama alanına göre model seçiminde dikkatli olunması gerektiğini göstermektedir. YOLOv5s ve YOLOv8s modelleri ise daha kısa eğitim süreleriyle makul doğruluk seviyeleri sunduğundan, kaynak kısıtı olan uygulamalarda tercih edilebilir seçeneklerdir. Çizelge 3' te verilen BAC_HIEN_CONSTRUCTION_SAFETY_2024 veri seti üzerinde yapılan değerlendirmelerde ise genel başarı oranlarının, Hard Hat Workers veri setine kıyasla daha düşük seviyelerde seyrettiği görülmektedir. Bu doğrultuda, kesinlik değerleri 0,819 ile 0,829 arasında

değişirken, duyarlılık değerleri 0,671 ile 0,682 arasında kalmıştır. F1 skorları açısından en iyi performans 0,748 ile YOLOv12s modelinde elde edilmiştir. mAP50 ve mAP50-95 değerlerinde de benzer bir eğilim gözlemlenmiş olup, mAP50 bakımından en yüksek değer 0,743 ile yine YOLOv12s modelinde elde edilmiştir. Eğitim süreleri ise diğer veri setine göre genel olarak biraz daha uzun sürmüştür ve YOLOv12s modeli için 151 dakika olmuştur.

Şekil 4'te verilen grafiğe baktığımızda, genel başarı düzeylerinin Hard Hat Workers veri setine göre daha düşük kaldığı gözlemlenmiştir. Bununla birlikte, tüm metriklerde tutarlı bir biçimde en yüksek değerleri elde eden YOLOv12s modeli, özellikle zorlu veri senaryolarında daha güçlü bir genelleme yeteneği sunduğu görülmüştür. Kesinlik, duyarlılık ve F1 skoru değerleri sırasıyla 0,827, 0,682 ve 0,748 olan bu model, aynı zamanda mAP50 ve mAP50-95 metriklerinde de diğer modellere üstünlük sağlamıştır. Ancak bu başarının, daha uzun süren eğitim

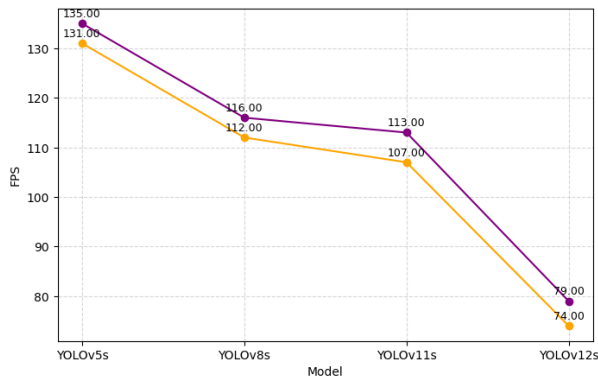
süreciyle birlikte geldiği, örneğin eğitimin 151 dakika sürmesiyle belirginleşmiştir. Buna karşın, YOLOv5s ve YOLOv8s gibi modeller daha kısa sürede eğitilmelerine rağmen kabul edilebilir doğruluk düzeylerini koruyarak zaman/verimlilik açısından avantaj sağlamaktadır. Elde

edilen bulgular, model seçiminde yalnızca doğruluk odaklı değil; aynı zamanda eğitim maliyeti, uygulama senaryosu ve veri kümesinin yapısal özelliklerinin birlikte değerlendirilmesi gerektiğini ortaya koymaktadır.



Şekil 4 YOLO modellerinin BAC_HIEN_CONSTRUCTION_SAFETY_2024 veri seti üzerinde performans karşılaştırması

Çalışmamızda FPS (inference) modellerin hesaplama verimliliğini ve mimari hızını doğrudan yansıttığı için temel karşılaştırma metriği olarak kullanılmıştır. FPS (total), preprocess, inference ve postprocess aşamalarının toplamını ifade eden uçtan uca gecikme süresini temsil etmektedir. Bu değer, gerçek zamanlı uygulamalardaki toplam işlem süresini göstermesi açısından bilgilendirici olmakla birlikte; veri akışındaki ek işlemleri de kapsadığından model mimarilerinin doğrudan karşılaştırılmasında kullanılmamıştır ancak gerçek zamanlı sistemlerdeki uçtan uca gecikmenin daha bütüncül bir şekilde değerlendirilebilmesi amacıyla çizelgelerde bilgilendirici ek bir metrik olarak sunulmuştur. Modellere ait FPS grafiği şekil 5 'te verilmiştir.



Şekil 5 YOLO modellerinin Hard Hat Workers ve BAC_HIEN_CONSTRUCTION_SAFETY_2024 veri setlerindeki FPS süreleri

Inference süreleri dikkate alındığında Hard Hat Workers veri setinde, YOLOv5s modeli 7,4 ms/görüntü ile en düşük çıkarım süresine ulaşmış ve bu değer yaklaşık 135 FPS'ye karşılık gelmiştir. Bu sonuç, YOLOv5s'in gerçek zamanlı nesne tespiti uygulamaları için yüksek düzeyde uygunluk taşıdığını ortaya koymaktadır.

YOLOv8s ve YOLOv11s modelleri sırasıyla 8,6 ms (116 FPS) ve 8,8 ms (113 FPS) ile birbirine yakın performans sergilemiş, bu modeller de yüksek FPS değerleri sayesinde gerçek zamanlı uygulamalar açısından tatmin edici bir hız sunmuştur. Buna karşın, YOLOv12s modeli 12,6 ms/görüntü çıkarım süresiyle 79 FPS seviyesinde kalmış ve önceki sürümlere kıyasla daha düşük bir hız sergilemiştir.

BAC_HIEN_CONSTRUCTION_SAFETY_2024 veri setinde elde edilen sonuçlara göre, YOLOv5s modeli 7,6 ms/görüntü çıkarım süresi ile en yüksek hız değerine ulaşmış ve saniyede ortalama 131 görüntü işleyebilme kapasitesine sahip olmuştur. Bu performans, modelin gerçek zamanlı nesne tespiti uygulamaları için oldukça elverişli olduğunu göstermektedir.

YOLOv8s ve YOLOv11s modelleri sırasıyla 8,9 ms (112 FPS) ve 9,3 ms (107 FPS) değerleri ile birbirine yakın performans sergilemiştir. Bu değerler, her iki modelin de hız açısından gerçek zamanlı uygulamalara uygun seviyede olduğunu ortaya koymaktadır. Bununla birlikte, YOLOv12s modeli 13,5 ms çıkarım süresi ile 74 FPS hızında

çalışmış ve önceki sürümlere kıyasla daha düşük bir hız göstermiştir. Her iki veri seti aynı model, aynı çözünürlük ve aynı donanımda test edilmesine rağmen FPS değerlerinin farklı çıkmasının temel nedeni, veri setlerindeki nesne yoğunluğunun ve sınıf sayısının farklı olmasıdır. BAC_HIEN veri seti, “person”, “helmet” ve “no helmet” olmak üzere daha fazla nesne içermekte ve görüntü başına üretilen tahmin kutusu sayısı Hard Hat Workers veri setine göre belirgin şekilde daha yüksektir. Bu durum FPS değerinin düşmesine neden olmuştur.



Şekil 6: Hard Hat Workers veri seti üzerinde gerçekleştirilen nesne tespit sonuçları



Şekil 7: BAC_HIEN_CONSTRUCTION_SAFETY_2024 veri seti üzerinde gerçekleştirilen nesne tespit sonuçları

5. Sonuç

Bu çalışma, YOLO tabanlı nesne tespiti algoritmalarının iş sağlığı güvenliği alanında kask tespiti amacıyla kullanılan farklı veri setlerinde gösterdiği başarıyı çok yönlü bir şekilde inceleyerek, model tercihinin veri yapısı, etiket kalitesi, model boyutu ve eğitim süresi parametreleriyle ilişkilendirilmesi gerektiğini ortaya koymuştur. YOLOv12s, en iyi F1 skoru ve mAP metriklerine ulaşmasına rağmen, artan eğitim süresiyle dikkat çekerken, YOLOv5s ve YOLOv8s diğer modellere göre yüksek FPS değeri, hızlı eğitim süresi ve kabul edilebilir performansı ile iyi bir

Şekil 6 ve şekil 7, çalışma kapsamında kullanılan test görüntüsünde YOLO tabanlı modellerin güvenlik kaskı tespit performanslarını görsel olarak ortaya koymaktadır. Görselde YOLOv12s modelinin nesneleri yüksek doğrulukla tespit ettiği gözlenmektedir. Diğer modellerde benzer doğrulukta sonuçlar vermiş olduğundan dolayı tek modelin tahmin görüntüsü sunulmuştur.

alternatif sunmuştur. Her iki veri seti üzerindeki performans karşılaştırmaları, model başarısının yalnızca mimari yapıya değil; aynı zamanda veri kümesinin yapısına, örnek çeşitliliğine ve öğrenme sürecinin parametrelerine bağlı olarak değişkenlik gösterebildiğini ortaya koymuştur. Bu durum, model seçiminin tek bir doğruluk metriğine göre değil; hedef uygulamanın gerekleri doğrultusunda bütüncül bir değerlendirmeye yapılması gerektiğini göstermektedir. Verilerin zorluk derecesinin arttığı ikinci veri setinde tüm modellerde görülen başarımların düşüşü, model geliştirmelerinin yanı sıra

veri kalitesinin artırılması gerektiğini vurgulamaktadır. Sonuç olarak, YOLO mimarileri arasındaki tercihlerin yalnızca metrik değerlerden ibaret olmayıp, kullanılacak veri setinin özellikleri, işlem süresi, donanım imkanları ve hedeflenen başarımla kriterleriyle bir bütün halinde değerlendirilmesi gerektiğini ortaya koymuştur.

Etik Standartlar Bildirgesi

Yazarlar tüm etik standartlara uyduklarını beyan ederler.

Yazarlık Katkı Beyanı

Yazar 1: Fikir Sahibi, Kaynaklar, Araştırma, Metodoloji, Deney, Görselleştirme, Yazma – orijinal taslak

Yazar 2: Araştırma, Kaynaklar, Doğrulama, Deney, Metodoloji, Yazma

Çıkar Çatışması Beyanı

Yazarların bu makalenin içeriğiyle ilgili olarak beyan edecekleri hiçbir çıkar çatışması yoktur.

Verilerin Kullanılabilirliği

Bu çalışma sırasında oluşturulan veya analiz edilen tüm veriler, yayınlanan bu makaleye dahil edilmiştir.

6. Kaynakça

- Badgujar, C. M., Poulouse, A., & Gan, H. (2024). Agricultural object detection with You Only Look Once (YOLO) Algorithm: A bibliometric and systematic literature review. *Computers and Electronics in Agriculture*, 223, 109090. <https://doi.org/10.1016/j.compag.2024.109090>
- Baldini, C., Migliorelli, L., Berardini, D., Azam, M. A., Sampieri, C., Ioppi, A., ... & Mattos, L. S. (2025). Improving real-time detection of laryngeal lesions in endoscopic images using a decoupled super-resolution enhanced YOLO. *Computer Methods and Programs in Biomedicine*, 260, 108539. <https://doi.org/10.1016/j.cmpb.2024.108539>
- Chen, W., Liu, M., Zhou, X., Pan, J., & Tan, H. (2022). Safety Helmet Wearing Detection in Aerial Images Using Improved YOLOv4. *Computers, Materials & Continua*, 72(2). <https://doi.org/10.32604/cmc.2022.026664>
- Çiftçi, M., Türkdamar, M. U., & Öztürk, C. (2024). Leveraging YOLO Models for Safety Equipment Detection on Construction Sites. *Journal of Computing Theories and Applications*, 1(4), 492-506. <https://doi.org/10.62411/jcta.10453>
- Dogra, A. K., Sharma, V., & Sohal, H. (2024). A survey of deep learning techniques for detecting and recognizing objects in complex environments. *Computer Science Review*, 54, 100686. <https://doi.org/10.1016/j.cosrev.2024.100686>
- Gheorghe, C., Duguleana, M., Boboc, R. G., & Postelnicu, C. C. (2024). Analyzing Real-Time Object Detection with YOLO Algorithm in Automotive Applications: A Review. *CMES-Computer Modeling in Engineering & Sciences*, 141(3). <https://doi.org/10.32604/cmes.2024.054735>
- Han, G., Zhu, M., Zhao, X., & Gao, H. (2021). Method based on the cross-layer attention mechanism and multiscale perception for safety helmet-wearing detection. *Computers and Electrical Engineering*, 95, 107458. <https://doi.org/10.1016/j.compeleceng.2021.107458>
- Huang, Y., Qian, Y., Wei, H., Lu, Y., Ling, B., & Qin, Y. (2023). A survey of deep learning-based object detection methods in crop counting. *Computers and Electronics in Agriculture*, 215, 108425. <https://doi.org/10.1016/j.compag.2023.108425>
- Jalonen, T., Al-Sa'd, M., Mellanen, R., Kiranyaz, S., & Gabbouj, M. (2023). Real-time damage detection in fiber lifting ropes using convolutional neural networks. *arXiv preprint arXiv:2302.11947*. <https://doi.org/10.1109/JSEN.2024.3521118>
- Jamtsho, Y., Riyamongkol, P., & Waranusast, R. (2021). Real-time license plate detection for non-helmeted motorcyclist using YOLO. *Ict Express*, 7(1), 104-109. <https://doi.org/10.1016/j.ict.2020.07.008>
- Kwak, N., & Kim, D. (2023). Detection of worker's safety helmet and mask and identification of worker using deeplearning. *Computers, Materials and Continua*, 75(1), 1671-1686. <https://doi.org/10.32604/cmc.2023.035762>
- Lee, J. Y., Choi, W. S., & Choi, S. H. (2023). Verification and performance comparison of CNN-based algorithms for two-step helmet-wearing detection. *Expert Systems with Applications*, 225, 120096. <https://doi.org/10.1016/j.eswa.2023.120096>
- Li, H., Wu, D., Zhang, W., & Xiao, C. (2024). YOLO-PL: Helmet wearing detection algorithm based on improved YOLOv4. *Digital Signal Processing*, 144, 104283. <https://doi.org/10.1016/j.dsp.2023.104283>
- Lin, H., Jiang, F., Jiang, Y., Luo, H., Yao, J., & Liu, J. (2023). A Model for Helmet-Wearing Detection of Non-Motor Drivers Based on YOLOv5s. *Computers, Materials & Continua*, 75(3). <https://doi.org/10.32604/cmc.2023.036893>
- Liu, Y., Zhang, J., Shi, L., Huang, M., Lin, L., Zhu, L., ... & Zhang, C. (2024). Detection method of the seat belt for workers at height based on UAV image and YOLO algorithm. *Array*, 22, 100340. <https://doi.org/10.1016/j.array.2024.100340>
- Manorat, P., Tuarob, S., & Pongpaichet, S. (2025). Artificial intelligence in computer programming education: A systematic literature review. *Computers and Education: Artificial Intelligence*, 100403.

<https://doi.org/10.1016/j.caeai.2025.100403>

Mostafa, S. A., Ravi, S., Zebari, D. A., Zebari, N. A., Mohammed, M. A., Nedoma, J., ... & Ding, W. (2024). A YOLO-based deep learning model for Real-Time face mask detection via drone surveillance in public spaces. *Information Sciences*, 676, 120865. <https://doi.org/10.1016/j.ins.2024.120865>

Nazli, N. A. N. M., Sabri, N., Aminuddin, R., Ibrahim, S., Yusof, S., & Nasir, S. D. N. M. (2024). A real-time system for detecting personal protective equipment compliance using deep learning model YOLOv5. *Procedia Computer Science*, 245, 647-656. <https://doi.org/10.1016/j.procs.2024.10.291>

Raj, G. D., & Prabadevi, B. (2025). ENHANCING SURFACE DETECTION: A COMPREHENSIVE ANALYSIS OF VARIOUS YOLO MODELS. *Heliyon*, 11(3), e42433. <https://doi.org/10.1016/j.heliyon.2025.e42433>

Redmon, J., Divvala, S., Girshick, R., & Farhadi, A. (2016). You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 779-788). <https://doi.org/10.48550/arXiv.1506.02640>

Said, Y., Alassaf, Y., Ghodhbani, R., Alsariera, Y. A., Saidani, T., Rhaïem, O. B., ... & Hleili, M. (2024). AI-Based Helmet Violation Detection for Traffic Management System. *CMES-Computer Modeling in Engineering & Sciences*, 141(1), 733-749. <https://doi.org/10.32604/cmcs.2024.052369>

Son, H., & Kim, C. (2021). Integrated worker detection and tracking for the safe operation of construction machinery. *Automation in Construction*, 126, 103670. <https://doi.org/10.1016/j.autcon.2021.103670>

Vourkos, E. G., Agathokleous, R. A., Panayides, A. S., Kalogirou, S. A., & Christoforou, E. G. (2025). Improving photovoltaics hotspot and dirt accumulation detection with computer vision: A systematic YOLO model comparison based on field aerial images. *Sustainable Energy Technologies and Assessments*, 78, 104355. <https://doi.org/10.1016/j.seta.2025.104355>

Vukicevic, A. M., Djapan, M., Isailovic, V., Milasinovic, D., Savkovic, M., & Milosevic, P. (2022). Generic compliance of industrial PPE by using deep learning techniques. *Safety science*, 148, 105646. <https://doi.org/10.1016/j.ssci.2021.105646>

Zhang, Y., Huang, S., Qin, J., Li, X., Zhang, Z., Fan, Q., & Tan, Q. (2025). Detection of helmet use among construction workers via helmet-head region matching and state tracking. *Automation in Construction*, 171, 105987. <https://doi.org/10.1016/j.autcon.2025.105987>

Zhou, Q., Qin, J., Xiang, X., & Tan, Y. (2021). Algorithm of Helmet Wearing Detection Based on AT-YOLO Deep Mode. *Computers, Materials & Continua*, 69(1), 159-174. <https://doi.org/10.32604/cmc.2021.017480>

İnternet kaynakları

Bac Hien Construction Safety 2024 Veri Seti, https://universe.roboflow.com/constructionsafety/bac_hien_construction_safety_2024 , (10.02.2026)

Hard Hat Workers Veri Seti, <https://universe.roboflow.com/joseph-nelson/hard-hat-workers>, (10.02.2026)