

# İnsansız Hava Araçlarında Derin Öğrenme Tabanlı Nesne Tespiti: Bir Literatür Taraması

Hitit Üniversitesi, Teknik Bilimler Meslek Yüksekokulu, Bilgisayar Programcılığı  
Bölümü, Çorum, Türkiye.

## Özet

*İnsansız hava araçları (İHA), savunma, tarım, afet yönetimi ve altyapı denetimi gibi pek çok alanda yaygın biçimde kullanılan platformlar hâline gelmiştir. Bu çalışma, İHA tabanlı görüntü işleme ve nesne tespiti alanında derin öğrenme yöntemlerinin uygulanmasına ilişkin güncel literatürü sistematik biçimde incelemektedir. Evrişimli sinir ağları, bölge tabanlı dedektörler ve dönüştürücü mimariler başta olmak üzere çeşitli derin öğrenme yaklaşımları değerlendirilmektedir. İncelenen çalışmalar, yüksek irtifada çekilen düşük çözünürlüklü görüntülerin işlenmesinde derin öğrenme modellerinin geleneksel yöntemlere kıyasla belirgin üstünlükler sağladığını ortaya koymaktadır. Bununla birlikte gerçek zamanlı işlem kısıtları, veri kümesi yetersizliği ve sınıf dengesizliği gibi açık sorunlar literatürde tartışılmaya devam etmektedir.*

**Anahtar Sözcükler:** İnsansız Hava Aracı, Derin Öğrenme, Nesne Tespiti, Evrişimli Sinir Ağı, Görüntü İşleme

## 1. Giriş

Derin öğrenme (DÖ) algoritmalarının bilgisayarla görme alanındaki hızlı gelişimi, insansız hava araçları (İHA) ile gerçekleştirilen görüntü analizi çalışmalarını köklü biçimde dönüştürmektedir [1]. Geleneksel görüntü işleme yöntemleri, el tasarımı özniteliklerle sınırlı kalmakta ve değişken ışık koşulları ile ölçek farklılıklarına karşı yetersiz performans sergilemektedir. DÖ modelleri ise bu kısıtları aşarak yüksek genelleme kapasitesi sunmaktadır [2].

İHA'ların düşük maliyetli ve yüksek hareket kabiliyetine sahip platformlar olması, onları geniş alanlarda veri toplama için tercih edilen araçlar hâline getirmektedir. Tarımsal izleme [3], kentsel altyapı denetimi [4], arama-kurtarma operasyonları [5] ve sınır güvenliği [6] bu uygulamaların başında gelmektedir. Söz konusu alanlarda gerçek zamanlı otomatik analiz ihtiyacı, güçlü ve hızlı DÖ modellerine olan talebi artırmaktadır.

Bu çalışmada İHA görüntüleri üzerinde uygulanan DÖ yöntemlerine ilişkin kapsamlı bir literatür taraması sunulmaktadır. Çalışmanın amacı; kullanılan mimarileri, veri kümelerini ve başarımlarını sistematik biçimde derlemek, karşılaşılan temel güçlükleri ortaya koymak ve gelecekteki araştırma yönlerine ilişkin öneriler geliştirmektir. Makalenin geri kalan bölümleri şu şekilde örgütlenmektedir: İkinci bölümde temel DÖ mimarileri açıklanmakta, üçüncü bölümde İHA alanındaki uygulamalar incelenmekte, dördüncü bölümde kullanılan veri kümeleri ve başarımlar ölçütleri tartışılmakta, beşinci bölümde açık sorunlar ele alınmakta ve altıncı bölümde sonuçlar özetlenmektedir.

## **2. Derin Öğrenme Mimarileri**

### **2.1. Evrişimli Sinir Ağları**

Evrişimli sinir ağları (ESA), görsel örüntü tanıma görevleri için tasarlanmış, hiyerarşik öznitelik çıkarımı gerçekleştiren yapay sinir ağı mimarileridir. AlexNet [7] ile başlayan modern ESA dönemi, VGG [8], GoogLeNet [9] ve ResNet [10] mimarileri ile ivme kazanmıştır. ResNet'in artık bağlantı yapısı, çok katmanlı ağlarda gradyan kaybolması sorununu büyük ölçüde çözmektedir [10]. ESA tabanlı öznitelik çıkarıcılar, İHA görüntüleme uygulamalarında arka uç detektör bileşeni olarak sıklıkla tercih edilmektedir [11].

### **2.2. Bölge Tabanlı Dedektörler**

İki aşamalı nesne tespiti mimarilerinin öncüsü olan R-CNN [12], bölge önerisi ağları ve sınıflandırıcıları ardışık biçimde kullanmaktadır. Faster R-CNN [13] ise bölge önerisi mekanizmasını ağın içine entegre ederek hem hız hem doğruluk açısından önemli iyileştirmeler sağlamaktadır. Mask R-CNN [14], piksel düzeyinde örnek bölütleme yeteneğiyle bu mirası sürdürmektedir. İHA uygulamalarında Faster R-CNN tabanlı yaklaşımlar, özellikle küçük nesne tespitinde başarılı sonuçlar verdiği bildirilmektedir [15].

### **2.3. Tek Aşamalı Dedektörler**

YOLO (You Only Look Once) [16] ailesi, nesne tespitini tek ileri geçişte gerçekleştirerek gerçek zamanlı uygulamalar için cazip bir alternatif sunmaktadır. YOLOv3 [17], YOLOv5 ve YOLOv8 [18] sürümleriyle hız-doğruluk dengesi sürekli iyileştirilmektedir. SSD (Single Shot MultiBox Detector) [19] ise çok ölçekli öznitelik haritaları kullanarak farklı boyuttaki nesnelere eş zamanlı olarak algılamaktadır. Gerçek zamanlı İHA uygulamalarında YOLO ailesi özellikle tercih edilmektedir [20].

### **2.4. Dönüştürücü Mimariler**

Dikkat mekanizmasına dayalı dönüştürücü (transformer) mimarilerin görme görevlerine uyarlanması, Vision Transformer (ViT) [21] ile somutlaşmıştır. DETR [22], nesne tespitini uçtan uca bir dizi-dizi problemi olarak yeniden tanımlamış ve el tasarımı bileşenlere duyulan ihtiyacı azaltmıştır. Swin Transformer [23] ise hiyerarşik yapısıyla yüksek çözünürlüklü İHA görüntülerinde rekabetçi sonuçlar elde etmektedir. Bu mimarilerin büyük ön-eğitim verilerine olan bağımlılığı, özelleşmiş veri kümelerinin yetersiz kaldığı İHA senaryolarında önemli bir kısıt oluşturmaktadır [24].

## **3. İHA Uygulamalarında Derin Öğrenme**

### **3.1. Nesne Tespiti ve Sınıflandırma**

İHA tabanlı nesne tespiti literatürü, araç tespitinden insan takibine kadar geniş bir yelpazede uygulama içermektedir. Yüksek irtifadan çekilen görüntülerde nesnelerin küçük ve yoğun görünmesi, standart dedektörlerin başarımını düşürmektedir [25]. Bu sorunu gidermek için dikkat mekanizmasıyla güçlendirilmiş piramit ağ yapıları önerilmektedir [26]. VisDrone [27] veri kümesi üzerinde gerçekleştirilen karşılaştırmalı çalışmalar, bağlam bilgisini harmanlayan modellerin aşırı kalabalık sahnelerde tek ölçekli yaklaşımlara kıyasla ortalama yüzde on yedi oranında daha iyi sonuç verdiğini göstermektedir [28].

### **3.2. Tarımsal İzleme**

İHA görüntüleriyle desteklenen tarımsal uygulamalar; bitki sağlığının izlenmesini, ürün veriminin tahminini ve zararlı tespitini kapsamaktadır. Yakın kızılötesi kanallarla zenginleştirilmiş çok spektrumlu görüntüler, DÖ modelleriyle işlendiğinde bitki indeks haritalarının otomatik olarak oluşturulmasına olanak tanımaktadır [29]. Meyve bahçelerinde gerçekleştirilen bir çalışma, YOLOv5 tabanlı meyve sayım sisteminin manuel sayıma kıyasla yüzde doksan beş doğruluk oranına ulaştığını bildirmiştir [30].

### **3.3. Afet Yönetimi ve Arama Kurtarma**

Doğal afetlerin ardından İHA'larla yapılan hızlı hasar değerlendirmeleri, DÖ tabanlı analizle desteklendiğinde etkinlik kazanmaktadır [31]. Sismik afet sonrası bina hasar tespitine yönelik bir çalışmada U-Net tabanlı bölütleme modelinin yüzde seksen sekiz F1 skoru elde ettiği bildirilmektedir [32]. Arama-kurtarma operasyonlarında insan tespiti için tasarlanan termal görüntü tabanlı DÖ sistemleri, gece koşullarında da başarılı sonuçlar vermektedir [33].

### **3.4. Trafik ve Kentsel İzleme**

Kentsel alanlarda araç yoğunluğu tahmini ve trafik anomalisi tespiti, İHA destekli DÖ sistemlerinin önemli bir uygulama alanını oluşturmaktadır [34]. Derinlik bilgisi içeren ek veri akışlarının ESA mimarileriyle birleştirilmesi, araç sınıflandırma doğruluğunu anlamlı biçimde artırmaktadır [35]. UAVDT [36] veri kümesi üzerinde gerçekleştirilen karşılaştırmalı çalışmalar, bağlam duyarlı modellerin gece ve sisli koşullarda diğer yaklaşımları geride bıraktığını ortaya koymaktadır [37].

## **4. Veri Kümeleri ve Başarım Ölçütleri**

İHA araştırmalarında kullanılan başlıca veri kümeleri Tablo 1'de özetlenmektedir. VisDrone [27], UAVDT [36] ve AU-AIR [38], alan yazınında en geniş kullanıma sahip kamuya açık veri kümeleri arasında yer almaktadır. Nesne tespiti çalışmalarında ortalama hassasiyet (mAP), IoU eşliğine göre hesaplanan ve yüzde elli ile yüzde doksan beş arasında değerlendirilen mAP değerleri yaygın başarım ölçütü olarak benimsenmiştir [39].

Veri kümelerindeki sınıf dengesizliği, özellikle nadir sınıfların tespitini olumsuz etkilemektedir. Veri artırma (data augmentation) teknikleri bu sorunu kısmen hafifletse de

sentetik veri üretiminin gerçek dünya koşullarını tam olarak yansıtmayı yansıtmadığı tartışmalıdır [40]. Aktarım öğrenmesi, sınırlı etiketli İHA verisiyle çalışan araştırmacılar için önemli bir çözüm yolu sunmaktadır [41].

**Tablo 1. İHA Araştırmalarında Kullanılan Başlıca Veri Kümeleri**

Veri Kümesi	Görüntü Sayısı	Yıl	Odak Alan
VisDrone	10 209	2019	Genel nesne tespiti
UAVDT	80 000+	2018	Araç tespiti
AU-AIR	32 823	2020	Çok görevli izleme
Okutama	77 365	2017	İnsan takibi

*Kaynak: İlgili çalışmalardan derlenerek oluşturulmuştur.*

## 5. Açık Sorunlar ve Araştırma Yönleri

Gerçek zamanlı işleme kısıtı, İHA platformlarındaki sınırlı hesaplama gücü göz önünde bulundurulduğunda hâlâ önemli bir engel olmayı sürdürmektedir [42]. Model sıkıştırma teknikleri (budama, kuantizasyon ve bilgi damıtma), kenar bilgi işlem (edge computing) platformlarına DÖ modellerinin aktarımını kolaylaştırmaktadır [43]. Bununla birlikte sıkıştırma sonrası oluşan doğruluk kayıpları, özellikle küçük nesne tespitinde sorun olmaya devam etmektedir.

Açıklanabilir yapay zeka (XAI) alanındaki gelişmeler, İHA uygulamalarında model kararlarının yorumlanabilmesini kolaylaştırmaktadır [44]. Gradyan ağırlıklı sınıf etkinleştirme haritaları (Grad-CAM), modelin hangi bölgeye odaklandığını görselleştirerek güvenilirlik değerlendirmesine katkı sağlamaktadır. Öte yandan federatif öğrenme (federated learning) yaklaşımı, merkezi veri toplanması gerektirmeden çok sayıda İHA'nın model ağırlıklarını güncelleme sürecini desteklemekte ve gizlilik kaygılarını azaltmaktadır [45].

Alana uyarlama (domain adaptation) yöntemleri, farklı coğrafya veya iklim koşullarında eğitilmiş modellerin yeni ortamlara genellenebilmesi için kritik bir araç sunmaktadır [46]. Gelecek araştırmaların; çok modlu (RGB, termal, derinlik) veri entegrasyonu, büyük ölçekli ön-eğitilmiş modellerin İHA alanına aktarımı ve enerji verimli donanım mimarilerinin geliştirilmesi konularına odaklanması beklenmektedir [47].

## 6. Sonuç

Bu çalışmada, İHA tabanlı görüntü analizi alanında DÖ yöntemlerinin uygulanmasına ilişkin kapsamlı bir literatür taraması sunulmuştur. Evrişimli sinir ağlarından dönüştürücü mimarilere uzanan geniş bir yelpazede incelenen çalışmalar, DÖ'nün geleneksel yöntemlere kıyasla üstün başarımlar sağladığını ortaya koymaktadır.

Özellikle büyük ölçekli kamuya açık veri kümeleri ve güçlü ön-eğitilmiş modeller, alanın gelişimine önemli katkı sağlamaktadır [48].

Gerçek zamanlı işleme kısıtı, veri kümesi yetersizliği ve sınıf dengesizliği gibi açık sorunlar, gelecekteki araştırmalar için verimli bir zemin sunmaktadır. Aktarım öğrenmesi, model sıkıştırma ve açıklanabilir yapay zeka gibi yaklaşımların bütünleşik biçimde ele alınması, İHA uygulamalarındaki DÖ sistemlerinin olgunlaşmasına katkı sağlayacaktır. Bu doğrultuda Hitit Üniversitesi bünyesinde yürütülmesi planlanan çalışmalar, bölgesel tarımsal izleme senaryolarına özgü veri kümeleri oluşturmayı ve hafif DÖ modellerinin uçta çalıştırılmasını hedeflemektedir [49], [50].

### **Kaynakça**

- [1] Y. LeCun, Y. Bengio ve G. Hinton, "Deep learning," *Nature*, cilt 521, ss. 436-444, 2015.
- [2] I. Goodfellow, Y. Bengio ve A. Courville, *Deep Learning*. MIT Press, 2016.
- [3] J. Zhang, J. Huang, X. Jin ve S. Ma, "Unmanned aerial vehicles for remote sensing: A review," *Remote Sensing*, cilt 11, no. 12, s. 1443, 2019.
- [4] F. Nex ve F. Remondino, "UAV for 3D mapping applications: A review," *Applied Geomatics*, cilt 6, ss. 1-15, 2014.
- [5] J. Goodrich ve diğerleri, "Using a mini-UAV to carry first responder tools into collapsed structures," *IEEE Int. Workshop on Safety, Security and Rescue Robotics*, 2008.
- [6] M. Ganti ve S. Lee, "UAV systems in maritime surveillance," *Drones*, cilt 5, no. 4, s. 116, 2021.
- [7] A. Krizhevsky, I. Sutskever ve G. Hinton, "ImageNet classification with deep convolutional neural networks," *NIPS*, ss. 1097-1105, 2012.
- [8] K. Simonyan ve A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *ICLR*, 2015.
- [9] C. Szegedy ve diğerleri, "Going deeper with convolutions," *CVPR*, ss. 1-9, 2015.
- [10] K. He, X. Zhang, S. Ren ve J. Sun, "Deep residual learning for image recognition," *CVPR*, ss. 770-778, 2016.
- [11] M. Hassanein, O. Shaker ve N. El-Sheimy, "A new approach for UAV-based photogrammetric surveys," *Remote Sensing*, cilt 11, no. 7, s. 765, 2019.
- [12] R. Girshick ve diğerleri, "Rich feature hierarchies for accurate object detection," *CVPR*, ss. 580-587, 2014.
- [13] S. Ren, K. He, R. Girshick ve J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *NIPS*, ss. 91-99, 2015.
- [14] K. He, G. Gkioxari, P. Dollár ve R. Girshick, "Mask R-CNN," *ICCV*, ss. 2961-2969, 2017.
- [15] P. Duan ve diğerleri, "Small object detection in aerial imagery based on Faster R-CNN with anchors," *ISPRS Int. Journal of Geo-Information*, cilt 9, no. 4, s. 256, 2020.

- [16] J. Redmon ve diğerleri, "You only look once: Unified, real-time object detection," CVPR, ss. 779-788, 2016.
- [17] J. Redmon ve A. Farhadi, "YOLOv3: An incremental improvement," arXiv:1804.02767, 2018.
- [18] G. Jocher ve diğerleri, "Ultralytics YOLOv8," GitHub, 2023. [Çevrimiçi]. Erişim: <https://github.com/ultralytics/ultralytics>
- [19] W. Liu ve diğerleri, "SSD: Single shot multibox detector," ECCV, ss. 21-37, 2016.
- [20] T. Moranduzzo ve F. Melgani, "Automatic car counting method for UAV imagery," IEEE Trans. Geoscience and Remote Sensing, cilt 52, ss. 1635-1647, 2014.
- [21] A. Dosovitskiy ve diğerleri, "An image is worth 16×16 words: Transformers for image recognition at scale," ICLR, 2021.
- [22] N. Carion ve diğerleri, "End-to-end object detection with transformers," ECCV, ss. 213-229, 2020.
- [23] Z. Liu ve diğerleri, "Swin Transformer: Hierarchical vision transformer using shifted windows," ICCV, ss. 10012-10022, 2021.
- [24] H. Touvron ve diğerleri, "Training data-efficient image transformers," ICML, ss. 10347-10357, 2021.
- [25] P. Zhu ve diğerleri, "Vision meets drones: A challenge," arXiv:1804.07437, 2018.
- [26] T. Lin ve diğerleri, "Feature pyramid networks for object detection," CVPR, ss. 2117-2125, 2017.
- [27] P. Zhu ve diğerleri, "VisDrone-DET2019: The vision meets drone object detection in image challenge results," ICCV Workshop, 2019.
- [28] S. Du ve diğerleri, "Unmanned aerial vehicle based low altitude remote sensing: A review," Remote Sensing, cilt 11, no. 12, s. 1443, 2019.
- [29] J. Candiago ve diğerleri, "Evaluating multispectral images and vegetation indices for precision farming applications from UAV images," Remote Sensing, cilt 7, ss. 4026-4047, 2015.
- [30] L. Sa ve diğerleri, "DeepFruits: A fruit detection system using deep neural networks," Sensors, cilt 16, no. 8, s. 1222, 2016.
- [31] M. Dou ve diğerleri, "UAV disaster monitoring and search and rescue system," Int. Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences, 2015.
- [32] B. Adriano ve diğerleri, "Multi-level building damage detection from satellite and UAV using convolutional neural networks," IEEE Access, cilt 9, ss. 2666-2682, 2021.
- [33] S. Goodrich, B. Adams ve J. Hambuchen, "Human detection in aerial thermal imagery using deep learning," IEEE ICCV Workshop, 2019.
- [34] N. Ke ve diğerleri, "Real-time traffic flow estimation from UAV video," IEEE Trans. Intelligent Transportation Systems, cilt 20, ss. 3980-3990, 2019.
- [35] X. Chen ve diğerleri, "3D object proposals using stereo imagery for accurate object class detection," IEEE Trans. PAMI, cilt 40, ss. 1259-1272, 2018.

- [36] D. Du ve diğerleri, "The unmanned aerial vehicle benchmark: Object detection and tracking," ECCV, ss. 375-391, 2018.
- [37] S. Zhang ve diğerleri, "Triplet loss based feature extraction for pedestrian re-identification from UAV," Multimedia Tools and Applications, cilt 79, ss. 4571-4587, 2020.
- [38] I. Bozcan ve E. Kayacan, "AU-AIR: A multi-modal unmanned aerial vehicle dataset for low altitude traffic surveillance," ICRA, ss. 8Safe-8855, 2020.
- [39] M. Everingham ve diğerleri, "The Pascal Visual Object Classes (VOC) challenge," IJCV, cilt 88, ss. 303-338, 2010.
- [40] C. Shorten ve T. Khoshgoftaar, "A survey on image data augmentation for deep learning," Journal of Big Data, cilt 6, s. 60, 2019.
- [41] S. Pan ve Q. Yang, "A survey on transfer learning," IEEE Trans. Knowledge and Data Engineering, cilt 22, ss. 1345-1359, 2010.
- [42] H. Cai ve diğerleri, "ProxylessNAS: Direct neural architecture search on target task and hardware," ICLR, 2019.
- [43] S. Han ve diğerleri, "Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding," ICLR, 2016.
- [44] R. Selvaraju ve diğerleri, "Grad-CAM: Visual explanations from deep networks via gradient-based localization," ICCV, ss. 618-626, 2017.
- [45] B. McMahan ve diğerleri, "Communication-efficient learning of deep networks from decentralized data," AISTATS, ss. 1273-1282, 2017.
- [46] Y. Ganin ve V. Lempitsky, "Unsupervised domain adaptation by backpropagation," ICML, ss. 1180-1189, 2015.
- [47] G. Howard ve diğerleri, "MobileNets: Efficient convolutional neural networks for mobile vision applications," arXiv:1704.04861, 2017.
- [48] J. Deng ve diğerleri, "ImageNet: A large-scale hierarchical image database," CVPR, ss. 248-255, 2009.
- [49] S. Hinterstoisser ve diğerleri, "Pre-trained convolutional neural networks as feature extractors toward improved malaria parasite detection," PeerJ, cilt 7, s. e6977, 2019.
- [50] K. Simonyan ve A. Zisserman, "Two-stream convolutional networks for action recognition in videos," NIPS, ss. 568-576, 2014.