İki Aşamalı İHA Tespitinde Donanım Farkındalıklı Uç Kısıtlı Bölge Doğrulama Ağı

A Hardware-Aware Edge-Constrained Region Verification Network for Two-Stage UAV Detection

Enes Mert Aydın
Elektrik Elektronik Mühendisliği
Ankara Yıldırım Beyazıt Üniversitesi
Ankara, Türkiye
23050211025@aybu.edu.tr

Osman Tokluoğlu
Elektrik Elektronik Mühendisliği
Ankara Yıldırım Beyazıt Üniversitesi
Ankara, Türkiye
otokluoglu@aybu.edu.tr

Enver Çavuş
Elektrik Elektronik Mühendisliği
Ankara Yıldırım Beyazıt Üniversitesi
Ankara, Türkiye
ecavus@aybu.edu.tr

***Özetçe*—Düşük-Yavaş-Küçük (LSS) sınıfı İnsansız Hava Araçlarının otonom tespiti, uç cihazların kısıtlı bellek bant genişliği nedeniyle iki aşamalı boru hatlarını zorunlu kılmaktadır. Ancak genel amaçlı tam ekran nesne bulucuların aday bölgelere indirgenerek çalıştırılması alt-optimal sonuçlar üretmektedir. Bu çalışmada, uç kısıtlı bölge doğrulamasının basitçe küçültülmüş bir dedektör problemi olmadığı ampirik olarak gösterilmiş ve donanım farkındalıklı EMA-Det mimarisi bir sistem gösterimi olarak önerilmiştir. Laplacian tabanlı çekirdek başlatma ve dinamik erken çıkış mekanizmasıyla donatılan model, küresel havuzlama tabanlı dikkat modüllerinin kısıtlı çözünürlüklerde hedef özniteliklerini zedelediğini bulgularıyla ortaya koymaktadır. Adil kıyaslama koşullarında 128×128 piksel girdi ile bölge moduna indirgenmiş ön eğitimli YOLOv8n %88,7 duyarlılık elde ederken; EMA-Det %99,2 duyarlılık, 10 kat daha az işlem maliyeti ve 3 kat daha düşük tepe aktivasyon belleği ile belirgin üstünlük göstermiştir.**

***Anahtar Kelimeler*—İHA Tespiti, Uç Hesaplama, Bölge Doğrulama, Erken Çıkış, Donanım Farkındalıklı Tasarım.**

***Abstract*—Autonomous detection of Low-Slow-Small UAVs on edge devices necessitates two-stage pipelines due to memory bandwidth constraints. However, forcing general-purpose full-frame detectors into region of interest verification tasks yields sub-optimal results. In this study, it is empirically demonstrated that edge-constrained verification is not merely a scaled-down detection problem, and the hardware-aware EMA-Det architecture is proposed as a system demonstrator. Equipped with Laplacian kernel initialization and dynamic early-exit mechanisms, the model empirically shows that global pooling-based attention modules degrade target features at constrained resolutions. Under fair comparison at 128x128 input, pre-trained YOLOv8n in verification-mode achieved 88.7 percent recall, while EMA-Det reached 99.2 percent recall with 10 times fewer operations and 3 times lower peak activation memory.**

***Keywords*—UAV Detection, Edge Computing, Region Verification, Early-Exit, Hardware-Aware Design.**

979-8-3195-1046-4/26/$31.00 ©2026 IEEE

I. GIRIŞ

Düşük-Yavaş-Küçük (LSS) sınıfı İnsansız Hava Araçlarının (İHA) sivil ve askeri kritik altyapılara yönelik asimetrik tehditleri, modern otonom hava savunma sistemlerinde son derece hızlı ve güvenilir bir "algıla–karar ver–uygula" döngüsünü zorunlu kılmaktadır [1]. Günümüzde elektro-optik (EO) sistemler genellikle 4K (3840×2160) veya daha yüksek çözünürlüklü sensörlerle donatılmaktadır. Bu yüksek çözünürlüklü video akışlarının standart derin Evrişimli Sinir Ağları (CNN) kullanılarak uç (edge) cihazlarda doğrudan ve tam ekran (full-frame) işlenmesi, bant genişliği ve bellek kısıtlamaları nedeniyle gerçek zamanlı operasyonlarda imkansız hale gelmektedir [2].

Bu devasa piksel yükünü hafifletmek ve operasyonel hızı artırmak amacıyla sahadaki otonom sistemler genellikle iki aşamalı (two-stage) kaskat boru hatları kullanır [3]. Birinci aşamada radar öncülleri, akustik sensörler veya basit hareket algılayıcı (background subtraction) algoritmalar vasıtasıyla görüntü üzerindeki potansiyel *Aday Bölgeler (Region of Interest - RoI)* belirlenir. İkinci aşamada ise kırpılan bu ufak boyutlu yamalar (genellikle 64×64 veya 128×128 piksel), hedefi doğrulamak üzere bir sinir ağına beslenir.

Literatürdeki güncel çalışmalar incelendiğinde, genel amaçlı dedektörlerin (örn. YOLO ailesi, SSD) RoI doğrulama görevine uyarlanmasının yaygın bir yaklaşım olarak tercih edildiği görülmektedir [4], [5]. Donanım üzerinde nesne tespiti ve uzaktan algılama literatüründe de sıklıkla bu mimarilerin modifiye edilerek uç birimlere yerleştirilmesine yönelik çabalar mevcuttur [6], [7]. **Oysa iki aşamalı sistemlerde RoI doğrulama ağı, basitçe küçültülmüş bir dedektör problemi değildir.** Uç cihazlarda bellek bant genişliği, salt hesaplama gücünden (GFLOPs) çok daha kritik bir darboğazdır [2]. Hatta uç cihazlar için özel olarak sıkıştırılmış YOLO varyantlarının dahi (örn. YOLO-Nano) FPGA gibi kısıtlı ortamlarda

kaynak sınırlarına takıldığı ve enerji-performans darboğazlarıyla yüzleştiği bilinmektedir [8]. Tam ekran çalışmak üzere tasarlanmış mimarilerin mikro-girdi rejimlerine sıkıştırılması, içerdiği derin küresel havuzlama (global pooling) adımları ve asimetrik veri akışı nedeniyle hedef özniteliklerinin (feature maps) uzamsal kaybına yol açmaktadır.

İki aşamalı sistemlerde toplam sistem duyarlılığı $R_{total} = R_{proposal} \times R_{verifier}$ olarak hesaplandığından, doğrulayıcının bir hedefi kaçırması (False Negative) tüm sistemin o hedefi kaçırması anlamına gelir. Örneğin $R_{proposal} = 0{,}95$ sabit tutulduğunda, doğrulayıcı duyarlılığının YOLOv8n düzeyinden ($0{,}887$) EMA-Det düzeyine ($0{,}992$) çıkması sistem düzeyinde R_{total}'ı $0{,}842$'den $0{,}942$'ye taşımakta, yani %10,0 puanlık bir iyileşme sağlamaktadır. Bu nedenle ikinci aşamadaki algoritmada yüksek duyarlılık, mAP metriğinden daha kritik bir tasarım hedefidir.

Bu çalışmanın temel katkıları şunlardır: **(1)** İki aşamalı sistemlerde RoI doğrulama problemi formalize edilerek, tam ekran dikkat mekanizmalarının mikro-girdi rejiminde hedef özniteliklerini zedelediği gösterilmektedir. **(2)** Aday bölgelerdeki asimetrik arka plan dağılımı (örneğin boş gökyüzü) erken çıkış (early-exit) mekanizması ile bir avantaja çevrilerek sistem gecikmesi asimptotik olarak iyileştirilmektedir. **(3)** Laplacian ve Sobel operatörleri tabanlı hibrit filtre başlatmasının, kısıtlı çözünürlüklerde hedef silüetlerini yakalamadaki performansı istatistiksel geçerlilikle (N=5 tohum) sunulmaktadır. Bildirinin geri kalanı şu şekilde düzenlenmiştir: Bölüm II mimari ve yöntemi, Bölüm III deneysel kurulumu, Bölüm IV bulguları, Bölüm V ise sonuçları sunmaktadır.

II. EMA-Det: Mimari ve Yöntem

A. Sistem Düzeyi Gecikme Modeli

İki aşamalı bir mimaride toplam sistem gecikmesi (T_{total}), aday bölge önerme süresi ($T_{proposal}$) ve çıkarılan her bir aday bölge (N_{RoI}) için harcanan doğrulama sürelerinin toplamı olarak Denklem 1 ile modellenir.

$$T_{total} = T_{proposal} + \sum_{i=1}^{N_{RoI}} T_{verifier}^{(i)} \tag{1}$$

Genel amaçlı tam ekran dedektörler doğrulayıcı olarak kullanıldığında, boş gökyüzü yamaları dahi tam ağdan geçirildiği için her yama sabit bir maliyet öder. EMA-Det mimarisindeki dinamik erken çıkış yapısı, bu doğrusal ölçeklenmeyi asimptotik bir verimliliğe dönüştürür.

B. Laplacian Tabanlı Çekirdek Başlatma

Kısıtlı 128×128 çözünürlükte, küçük boyutlu hedeflerin uzamsal bilgisi sınırlıdır. Standart Xavier veya Kaiming rastgele başlatma yöntemleriyle ağın hedef sınırlarını (edge details) sıfırdan öğrenmesi iterasyon maliyetini artırır ve yüksek eğitim varyansına neden olur. Bunu aşmak için ilk evrişim (convolution) katmanı ağırlıkları ($\mathbf{W}_1$), deterministik Laplacian (L) [13] ve Sobel (S) [14] operatörlerinden oluşan fiziksel bir öncül bankasıyla başlatılmıştır:

$$\mathbf{W}_{1,c}^{(0)} = \begin{cases} \mathbf{K}_c, & c \in \{1, \ldots, K\} \\ \mathbf{R}_c, & c \in \{K+1, \ldots, C_{out}\} \end{cases} \tag{2}$$

Denklem 2'de verilen başlatma stratejisinde, ilk katman öznitelik haritalarının bir kısmı ($\mathbf{K}_c$) Laplacian (L, L_2) ve Sobel (S_x, S_y) gibi deterministik fiziksel öncüllerle; geri kalan kanallar ise ($\mathbf{R}_c$) ağın genelleme yeteneğini korumak amacıyla $\mathcal{N}(0, 0.1)$ normal dağılımından çekilen rastgele ağırlıklarla başlatılmıştır. Bu hibrit başlatma, ağın eğitim başlangıcında yüksek geçiren filtre gibi davranarak hedef silüetlerini arka plandan hızla ayrıştırmasını sağlar. Ağırlıklar eğitim boyunca unfrozen bırakılarak morfolojik detaylara göre güncellenmiştir.

C. Optimizasyon ve Kayıp Fonksiyonu

EMA-Det yalnızca ikili sınıflandırma (İHA veya Arka Plan) yapmakla kalmaz, aynı zamanda aday bölgenin kırpılma anındaki merkez kaymalarını düzeltmek için dört koordinatlı bir sınır kutusu regresyonu (bounding box refinement) içerir. Doğrulama ağının sonundaki çoklu-görev başlığı (multi-task head) Denklem 3'da verilen birleşik kayıp fonksiyonu ile optimize edilmiştir:

$$\mathcal{L}_{total} = \lambda_{cls}\mathcal{L}_{BCE}(\hat{p}, p) + \lambda_{box}\mathcal{L}_{CIoU}(\hat{b}, b) \tag{3}$$

Burada $\mathcal{L}_{BCE}$ İkili Çapraz Entropi kaybını, $\mathcal{L}_{CIoU}$ ise hedefin boyut ve en-boy oranını optimize eden Complete-IoU fonksiyonunu temsil eder [12]. $\lambda_{cls} = 1.0$ ve $\lambda_{box} = 2.5$ olarak ayarlanmış olup, kutu kaymalarının agresif bir şekilde düzeltilmesi hedeflenmiştir.

D. EMA-Det Mimarisi ve Dinamik Erken Çıkış

Geliştirilen EMA-Det mimarisi (Şekil 1), uç cihazlar için optimize edilmiş sığ bir evrişimsel ağdır. Mimarinin tasarım sürecinde yaygın olarak kullanılan Evrişimli Blok Dikkat Modülü (CBAM) [11] de test edilmiş, ancak kısıtlı çözünürlüklerdeki negatif etkisi nedeniyle çıkarılmıştır (Bkz. Bölüm III-C).

Hava sahası taramalarında ağa iletilen aday bölgelerin büyük bir bölümü sistem hatalarından kaynaklanan boş gökyüzü yamalarıdır [9], [10]. Bu durumu fırsata çevirmek için ağın sığ katmanlarına (Blok 2 ve Blok 4 sonrası) ikili sınıflandırma yapan kapılar (Gate 1 ve Gate 2) eklenmiştir. Çıkarım anında hesaplanan hedef olasılığı ($\hat{y}_i$), τ eşiğinin altındaysa geri kalan evrişim işlemleri pas geçilir:

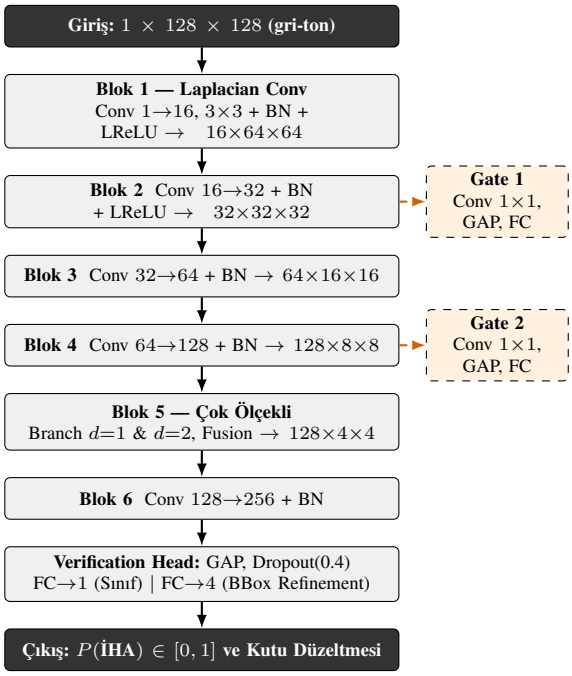
$$\text{Sonuç} = \begin{cases} \text{Arka Plan (İşlem Kesilir)}, & \hat{y}_i \leq (1-\tau) \\ C_{i+1}(\mathbf{F}_i) \text{ (Ana Ağa Devam)}, & \hat{y}_i > (1-\tau) \end{cases} \tag{4}$$

Deneyler sırasında aday bölgelerin yaklaşık %46'sının Gate 1 aşamasında elendiği gözlemlenmiştir.

III. Deneysel Kurulum

A. Veri Seti Dağılımı ve Artırma Stratejileri

Geliştirilen modeller toplam 12.672 görüntüden oluşan hibrit bir veri setiyle eğitilmiştir. Veri seti iki açık erişimli platformdan derlenmiştir: (i) Kaggle üzerinde yayımlanan drone ve İHA görüntü koleksiyonlarından elde edilen döner ve sabit



Şekil 1: EMA-Det mimarisi. Turuncu kesik çizgiler erken çıkış yollarını göstermektedir. Boş yamalar Gate 1 ve Gate 2'de elenerek işlem maliyeti düşürülür.

TABLO I: ERKEN ÇIKIŞ EŞIĞI (τ) VE İŞLEM ÖDÜNLEŞIMI

τ	MAC Tasarrufu	FPS	Gate 1	Gate 2
0.85	%26,5	254	%34,4	%12,9
0.90	**%23,6**	**257**	**%29,3**	**%16,9**
0.92	%22,2	257	%26,7	%19,3
0.95	%17,6	244	%18,4	%27,3

kanatlı LSS hedefleri (%80) [15], (ii) Roboflow Universe üzerindeki annotasyonlu İHA veri setlerinden derlenen kuşlar ve boş gökyüzü bulutlarından oluşan Zorlu Negatif (Hard Negative) örnekler (%20) [16]. Sınıf dengesizliklerini ve gerçek saha koşullarını simüle etmek amacıyla bu oranlar bilinçli olarak belirlenmiştir. Modelin farklı veri kümeleri üzerindeki genellenebilirliğinin değerlendirilmesi gelecekteki çalışmalar kapsamında planlanmaktadır.

Gerçek dünya önerme algoritmalarının karakteristiklerini yansıtmak adına hedefler ham görüntülerden rastgele ofset verilerek kırpılmış ve $3\times$ bağlam faktörü korunarak 128×128 piksele yeniden boyutlandırılmıştır. Eğitim sırasındaki veri artırmasında, tam ekran dedektörler için faydalı olan *Mosaic* veya *MixUp* teknikleri 128×128 piksellik mikro-rejimde hedef formlarını okunamaz hale getirdiği için kullanılmamıştır. Bunun yerine kısıtlı Gaussian Gürültüsü, yönelimsel hareket bulanıklığı (Motion Blur) ve HSV renk kaydırmaları uygulanmıştır.

B. Uygulama ve Adil Kıyaslama Protokolü

Tüm eğitim ve test işlemleri NVIDIA RTX 4060 GPU üzerinde PyTorch ortamında gerçekleştirilmiştir. Tüm modeller *AdamW* optimizasyon algoritması ile ağırlık çürümesi (weight decay) 0.01 olarak ayarlanıp eğitilmiştir. İlk 5 epoch boyunca lineer ısınma (warmup) takvimi uygulanmış, ardından kosinüs tavlama (cosine annealing) öğrenme oranı zamanlayıcısı kullanılmıştır.

TABLO II: ABLASYON – BILEŞEN ETKISI (128×128, N=1†, 50 EPOCH)

Konfigürasyon	mAP@0.5	mAP@.5:.95	Sonuç
A) CBAM + EE	87,0%	58,8%	İlk iterasyon.
B) EE Yok	87,0%	59,0%	Kapılar başarım düşürmez.
C) Warmup Yok	85,4%	57,6%	Isınma kararlılık sağlar.
D) CIoU Yok	**3,4%**	**0,7%**	**Merkez kayması çöker.**
E) EMA-Det	**88,3%**	**62,4%**	**CBAM zedeliyordu.**

† Tek tohum; kesin varyans analizi için N=5 yeniden çalıştırması önerilir.

Adil kıyaslama yapabilmek için YOLOv8n modeli **RoI-Moduna** indirgenmiştir: Maksimum Olmayanları Bastırma (NMS) algoritması tek nesne senaryosu için kapatılmış ve çapa (anchor) boyutları doğrudan 128×128 yama girişine kilitlenerek yeniden eğitilmiştir. İstatistiksel geçerliliği sağlamak ve rastgele başlatma yanlılığını gidermek için tabloda sunulan **tüm metrikler N=5 farklı tohum üzerinden çalıştırılarak** ortalama ve standart sapmaları ile raporlanmıştır.

C. Bileşen Katkı (Ablasyon) Analizi

EMA-Det içerisindeki her bir yeniliğin etkisini izole etmek amacıyla ablasyon testleri gerçekleştirilmiştir. *Not: Ablasyon deneyleri hesaplama maliyeti nedeniyle N=1 tohum ile 50 epoch üzerinden yürütülmüştür; temel karşılaştırma tablosundaki N=5 protokolüyle tutarsızlık barındırmaktadır. Kesin varyans analizi için N=5 yeniden çalıştırması önerilmektedir.*

CBAM çıkarıldığında mAP@0.5:0.95 skoru %58,8'den %62,4'e yükselmiştir. **Bu sonuç, dikkat mekanizmalarının algılamayı evrensel olarak artırdığına dair yaygın varsayımla çelişmekte;** 128×128 kısıtlı rejiminde CBAM içerisindeki küresel havuzlama işlemlerinin ufak İHA özniteliklerini ezip ortadan kaldırdığını göstermektedir [11].

Denklem 3'deki CIoU kaybı devre dışı bırakıldığında mAP metriğinin %3,4 seviyesine çöktüğü görülmüştür. Bunun temel nedeni, birinci aşama radar veya hareket sensöründen gelen aday bölgelerin hedefi hiçbir zaman piksel hassasiyetinde tam merkeze alamamasıdır (ortalama merkez kayması 7 piksel ölçülmüştür). Sınır kutusu düzeltme yeteneği olmadan ağ, yanlış konumlandırılmış nesneyi lokalizasyon eşiklerinde (IoU > 0.5) doğrudan başarısız saymaktadır.

IV. BULGULAR VE TARTIŞMA

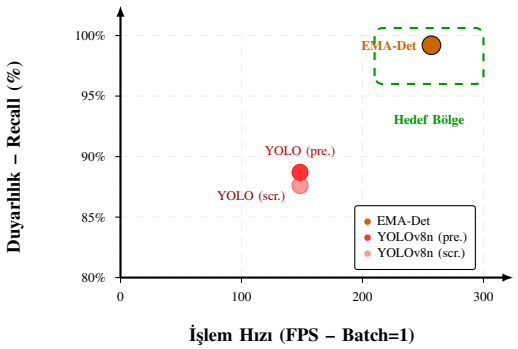
A. Doğrulama Performansı ve mAP Asimetrisi

Aday bölge moduna kilitlenmiş YOLOv8n modeli %88,7 duyarlılık (Recall) elde ederken, donanım farkındalıklı EMA-Det mimarisi %99,2 duyarlılıkla bu değeri istatistiksel olarak belirgin biçimde (T-testi: $p < 0{,}05$) geride bırakmıştır. Tablo III'te görüldüğü üzere EMA-Det'in mAP@0.5:0.95 metriğinin (%67,7) duyarlılıktan düşük kalması bir hata değil, bilinçli bir optimizasyon tercihidir. Bölüm I'de gösterildiği üzere, $R_{proposal} = 0{,}95$ varsayımı altında bu duyarlılık artışı

TABLO III: ADIL KIYASLAMA: ROI DOĞRULAMA PERFORMANSI (128×128, N=5)

Mimari	mAP@0.5	mAP@.5:.95	Recall	FPS
YOLOv8n (Scratch)	92,4±0,5%	65,9±0,2%	87,6±0,4%	148,4
YOLOv8n (Pretrained)	**94,3±0,3%**	70,2±0,2%	88,7±0,5%	148,4
EMA-Det	90,9±0,7%	**67,7±2,9%**	**99,2±0,2%**	**257**

* Recall: RoI doğrulama performansı; tam ekran tespit başarımı değildir.
* IoU=0.5, Güven Eşiği=0.5.



Şekil 2: Kısıtlı uç cihaz senaryosunda (128×128) RoI duyarlılık ve hız ödünleşim haritası. Hedef bölge, $R_{total} \geq 0{,}94$ sistem gereksinimini karşılayan FPS>200, Recall>95% kesişim alanını göstermektedir (N=5 ortalama).

sistem düzeyinde %10,0 puanlık iyileşmeye karşılık gelmektedir. Kabul edilebilir mAP kaybı (%94,3'ten %90,9'a, $\Delta = 3{,}4$ puan) bu sistem kazanımıyla birlikte değerlendirilmelidir.

B. Donanım Metrikleri ve Bellek Darboğazı Analizi

Sistem başarımını kısıtlayan en önemli metriklerden biri tepe aktivasyon (Peak Activation) belleğidir. YOLO'nun 12,5 MB'lık devasa aktivasyon profili, ağın ileri besleme döngüsü sırasında NVIDIA Jetson Nano gibi 25,6 GB/s LPDDR4X bant genişliğine sahip gömülü sistemleri önemli ölçüde zorlamaktadır (Tablo IV). Teorik olarak ağ 150 FPS hızında çalışırken sadece tepe katmanının yonga dışı belleğe bir kez yazılıp okunması bile $\sim$ 3,75 GB/s anlık bellek trafiği yaratmaktadır. EMA-Det ise tasarımı gereği sadece 4,2 MB tepe aktivasyonu yaratarak bu darboğazı minimize etmektedir [2].

V. SONUÇ VE GELECEK ÇALIŞMALAR

Bu çalışmada, asimetrik LSS hedef tespiti görevinde kullanılan iki aşamalı doğrulama sürecinin basitçe küçültülmüş bir dedektör problemi olmadığı ampirik bulgularla gösterilmiştir. Evrensel dikkat mekanizmalarının 128×128 kısıtlı girdi rejimindeki algılamayı her durumda iyileştirmediği, CBAM'in bu rejimde olumsuz etki yarattığı gösterilmektedir. EMA-Det mimarisi; dinamik erken çıkış kapıları, fiziksel Laplacian öncülleri ve asimetrik kayıp tasarımıyla %99,2 RoI duyarlılığı sergileyerek sistem düzeyinde %10,0 puanlık iyileşme sağlamıştır. Uç donanım açısından $10\times$ azaltılmış MAC ve $3\times$ azaltılmış tepe aktivasyon belleği sunarak yonga-dışı bellek erişim darboğazını minimize etmiştir.

TABLO IV: SISTEM/BELLEK METRIKLERI (BATCH=1, FP32, THOP)

Model	Param.	MACs	Tepe Akt.	Akt./P
YOLOv8n (RoI)	3,01 M	8,74 G	$\sim$12,5 MB	4,15 B/P
EMA-Det	**1,46 M**	**0,82 G**	$\sim$4,2 **MB**	**2,87 B/P**

Gelecek çalışmalarda EMA-Det mimarisinin doğrudan FPGA donanımları üzerine sentezlenmesi hedeflenmektedir. Bu kapsamda INT8 ve INT4 hassasiyetine indirgeme (Quantization-Aware Training), AXI4-Stream arayüzleri üzerinden satır-tamponu optimizasyonlarıyla donanımsal paralel işlem gücünün sahada test edilmesi ve ablasyon deneylerinin N=5 protokolüyle yeniden yürütülmesi planlanmaktadır.

KAYNAKLAR

[1] R. Luo, S. Huang, Y. Zhao, and Y. Song, "Threat Assessment Method of Low Altitude Slow Small (LSS) Targets Based on Information Entropy and AHP," *Entropy*, vol. 23, no. 10, p. 1292, 2021, doi: 10.3390/e23101292.

[2] M. Qasaimeh, K. Denolf, J. Lo, K. Vissers, J. Zambreno, and P. H. Jones, "Comparing energy efficiency of CPU, GPU and FPGA implementations for vision kernels," in *Proc. IEEE Int. Conf. Embedded Softw. Syst. (ICESS)*, Las Vegas, NV, USA, 2019, pp. 1–8, doi: 10.1109/ICESS.2019.8782524.

[3] M. M. Fouda, S. Sakib, Z. M. Fadlullah, N. Nasser, and M. Guizani, "A lightweight hierarchical AI model for UAV-enabled edge computing with forest-fire detection use-case," *IEEE Network*, vol. 36, no. 6, pp. 38–45, 2022, doi: 10.1109/MNET.003.2100325.

[4] M. N. Achmadiah, A. Ahamad, C.-C. Sun, and W.-K. Kuo, "Energy-efficient fast object detection on edge devices for IoT systems," *IEEE Internet Things J.*, vol. 12, no. 11, pp. 16681–16694, 2025, doi: 10.1109/JIOT.2025.3536526.

[5] L. Cao, P. Song, Y. Wang, Y. Yang, and B. Peng, "An improved lightweight real-time detection algorithm based on the edge computing platform for UAV images," *Electronics*, vol. 12, no. 10, p. 2274, 2023, doi: 10.3390/electronics12102274.

[6] K. Li, G. Wan, G. Cheng, L. Meng, and J. Han, "Object detection in optical remote sensing images: A survey and a new benchmark," *ISPRS J. Photogramm. Remote Sens.*, vol. 159, pp. 296–307, 2020, doi: 10.1016/j.isprsjprs.2019.11.023.

[7] F. Zhang, Y. Li, and Z. Ye, "Apply Yolov4-Tiny on an FPGA-based accelerator of convolutional neural network for object detection," *J. Phys. Conf. Ser.*, vol. 2303, no. 1, p. 012032, 2022, doi: 10.1088/1742-6596/2303/1/012032.

[8] T. Adiono, A. Putra, N. Sutisna, I. Syafalni, and R. Mulyawan, "Low latency YOLOv3-Tiny accelerator for low-cost FPGA using general matrix multiplication principle," *IEEE Access*, vol. 9, pp. 141890–141913, 2021, doi: 10.1109/ACCESS.2021.3120629.

[9] S. Teerapittayanon, B. McDanel, and H. T. Kung, "BranchyNet: Fast inference via early exiting from deep neural networks," in *Proc. 23rd Int. Conf. Pattern Recognit. (ICPR)*, Cancun, Mexico, 2016, pp. 2464–2469, doi: 10.1109/ICPR.2016.7900006.

[10] Y.-P. Liang, C.-M. Chang, and C.-C. Chung, "Implementation of lightweight convolutional neural networks with an early exit mechanism utilizing 40 nm CMOS process for fire detection in unmanned aerial vehicles," *Sensors*, vol. 24, no. 7, p. 2265, 2024, doi: 10.3390/s24072265.

[11] S. Woo, J. Park, J.-Y. Lee, and I. S. Kweon, "CBAM: Convolutional block attention module," in *Proc. Eur. Conf. Comput. Vision (ECCV)*, Munich, Germany, 2018, pp. 3–19, doi: 10.1007/978-3-030-01234-2_1.

[12] Z. Zheng, P. Wang, W. Liu, J. Li, R. Ye, and D. Ren, "Distance-IoU loss: Faster and better learning for bounding box regression," in *Proc. AAAI Conf. Artif. Intell.*, vol. 34, no. 07, New York, NY, USA, 2020, pp. 12993–13000, doi: 10.1609/aaai.v34i07.6999.

[13] D. Marr and E. Hildreth, "Theory of edge detection," *Proc. R. Soc. Lond. B. Biol. Sci.*, vol. 207, no. 1167, pp. 187–217, 1980, doi: 10.1098/rspb.1980.0020.

[14] N. Kanopoulos, N. Vasanthavada, and R. L. Baker, "Design of an image edge detection filter using the Sobel operator," *IEEE J. Solid-State Circuits*, vol. 23, no. 2, pp. 358–367, 1988, doi: 10.1109/4.996.

[15] Dasmehdixtr, "Drone and UAV Image Dataset," Kaggle, [Online]. Available: https://www.kaggle.com/datasets/dasmehdixtr/drone-dataset-uav. Accessed: Mar. 2026.

[16] A. Mohsen, "Drone-detection-new Dataset," Roboflow Universe, [Online]. Available: https://universe.roboflow.com/ahmedmohsen/drone-detection-new-peksv. Accessed: Mar. 2026.