

## Analisis Kualitas Lokalisasi Bounding Box Modifikasi YOLOv4-Tiny-3L pada Citra Penginderaan Jauh

### Bounding Box Localization Quality of Modified YOLOv4-Tiny-3L in Remote Sensing Images

Ahmad Fauzan<sup>1</sup>, Muhammad Januansyah Jayadi<sup>2</sup>, Andy Fadly<sup>3</sup>

<sup>1</sup>Department of Artificial Intelligence, Kristen Indonesia Paulus University, Makassar, 90241 Indonesia

<sup>2</sup>Department of Informatics Engineering, Kristen Indonesia Paulus University, Makassar, 90241 Indonesia

<sup>3</sup>Department of Digital Business, Muhammadiyah Education University of Sorong, Sorong, 98418 Indonesia

#### Article Info

##### Genesis Artikel:

Diterima, 13 April 2026

Direvisi, 29 Mei 2026

Disetujui, 13 Juni 2026

##### Kata Kunci:

Deteksi Objek

Penginderaan Jauh

Yolov4-Tiny

Kualitas Lokalisasi

Intersection Over Union

RSOD

#### ABSTRAK

Deteksi objek pada citra penginderaan jauh memerlukan evaluasi yang tidak berhenti pada AP50, karena aplikasi yang menuntut posisi objek akurat bergantung pada kualitas regresi bounding box pada ambang Intersection over Union (IoU) yang lebih ketat. Penelitian ini merupakan analisis sekunder atas data eksperimen YOLOv4-Tiny-3L dan YOLOv4-Tiny-Mod pada dataset Remote Sensing Object Detection (RSOD); tidak dilakukan pelatihan ulang model. Berbeda dari studi sebelumnya yang berfokus pada modifikasi arsitektur dan mAP agregat, kontribusi penelitian ini adalah karakterisasi degradasi lokalisasi melalui AP Decay, AP Slope sebagai indikator deskriptif rata-rata penurunan per interval, dan segmentasi tiga zona IoU. YOLOv4-Tiny-Mod menghasilkan AP lebih tinggi pada seluruh threshold, dengan peningkatan terbesar 13,83 poin persentase pada AP85; selisih rata-rata mencapai 6,29 poin pada zona moderate precision dan 5,66 poin pada strict localization. AP Decay endpoint kedua model hampir sama dan sedikit lebih besar pada model modifikasi (97,93 vs. 97,72), sehingga tidak digunakan sendiri sebagai bukti superioritas. Analisis per kelas juga menunjukkan peningkatan Recall Aircraft, sementara Recall Oiltank dan Playground tetap rendah dan tidak dapat diatribusikan secara kausal hanya dari data agregat. Temuan mengindikasikan bahwa CSPBlock dan kepala deteksi beresolusi 104×104 terutama memberi manfaat ketika tuntutan ketelitian lokalisasi meningkat.

#### ABSTRACT

Object detection in remote sensing imagery should not be evaluated only at AP50 because applications requiring accurate object positions depend on bounding-box regression quality at stricter Intersection over Union (IoU) thresholds. This study performs a secondary analysis of previously published YOLOv4-Tiny-3L and YOLOv4-Tiny-Mod experiments on the Remote Sensing Object Detection (RSOD) dataset; no model retraining was conducted. Unlike the earlier study, which focused on architectural modification and aggregate mAP, the present contribution characterizes localization degradation using AP Decay, AP Slope as a descriptive mean drop per threshold interval, and three IoU zones. YOLOv4-Tiny-Mod achieved higher AP at every evaluated threshold, with the largest gain of 13.83 percentage points at AP85; the mean advantage was 6.29 points in the moderate-precision zone and 5.66 points in strict localization. Endpoint AP Decay was nearly identical and slightly larger for the modified model (97.93 vs. 97.72), so it is not interpreted in isolation as evidence of superiority. Class-level results show a strong Recall improvement for Aircraft, whereas Oiltank and Playground retain very low Recall that cannot be causally explained from aggregate data alone. The results indicate that CSPBlock and the 104×104 detection head mainly benefit the model as localization requirements become stricter.

This is an open access article under the [CC BY-SA](#) license.



#### Penulis Korespondensi:

Ahmad Fauzan,

Department of Artificial Intelligence,

Kristen Indonesia Paulus University,

Email: ahmadfauzan.idcontact@gmail.com

## 1. PENDAHULUAN

Kemajuan teknologi penginderaan jauh dalam beberapa dekade terakhir telah mengubah cara manusia memahami dan memantau kondisi permukaan bumi. Citra yang dihasilkan dari platform satelit maupun wahana udara tanpa awak kini mampu merekam detail objek dengan resolusi spasial yang semakin tinggi, bahkan mencapai skala kurang dari satu meter per piksel. Kemampuan ini membuka peluang besar untuk berbagai keperluan strategis, mulai dari pemantauan pergerakan pesawat di kawasan bandara, pengawasan infrastruktur kritis, pengelolaan lalu lintas udara, hingga respons tanggap darurat di wilayah yang sulit dijangkau secara langsung [2]. Seiring meningkatnya volume data citra yang dihasilkan setiap harinya, analisis manual oleh operator manusia menjadi tidak lagi efisien dan rawan terhadap kesalahan persepsi, sehingga kebutuhan akan sistem deteksi objek otomatis berbasis kecerdasan buatan menjadi semakin mendesak.

Deteksi objek merupakan salah satu tugas fundamental dalam bidang visi komputer yang bertujuan mengidentifikasi keberadaan dan menentukan posisi objek tertentu dalam sebuah citra secara otomatis. Berbeda dengan klasifikasi citra yang hanya menghasilkan label kelas, deteksi objek menghasilkan keluaran berupa bounding box yang melingkupi setiap objek yang terdeteksi beserta label dan skor kepercayaannya [3]. Dalam konteks penginderaan jauh, tugas ini menghadapi sejumlah tantangan khas yang tidak dijumpai pada citra konvensional, antara lain variasi orientasi objek yang tidak terbatas, perbedaan skala objek yang sangat lebar dalam satu citra, kepadatan objek yang tinggi, serta latar belakang yang kompleks dan heterogen seperti runway, apron, dan area hijau di sekitar bandar udara [9].

Pendekatan berbasis pembelajaran mendalam, khususnya arsitektur jaringan saraf konvolusi, telah terbukti unggul dalam menangani kompleksitas tersebut dibandingkan metode konvensional berbasis fitur buatan manusia seperti Histogram of Oriented Gradients (HOG) dan Support Vector Machine (SVM). Di antara berbagai arsitektur yang ada, keluarga model YOLO (You Only Look Once) menjadi pilihan populer karena kemampuannya melakukan deteksi dalam satu tahap proses inferensi, sehingga menghasilkan keseimbangan yang baik antara kecepatan dan akurasi [3]. Salah satu varian yang dikembangkan khusus untuk perangkat dengan sumber daya komputasi terbatas adalah YOLOv4-Tiny-3L, yang menambahkan kepala deteksi ketiga pada resolusi  $52 \times 52$  di atas arsitektur YOLOv4-Tiny standar, sehingga meningkatkan kemampuan model dalam mengenali objek pada berbagai skala ukuran [4].

Meskipun YOLOv4-Tiny-3L menawarkan efisiensi yang baik, arsitektur ini masih memiliki keterbatasan dalam hal ekstraksi fitur untuk objek kecil dan penanganan latar belakang yang rumit. Untuk mengatasi hal tersebut, penelitian terdahulu telah melakukan modifikasi arsitektur YOLOv4-Tiny-3L dengan mengintegrasikan Cross-Stage Partial Block (CSPBlock) pada bagian backbone dan menambahkan kepala deteksi keempat pada resolusi  $104 \times 104$  [1]. CSPBlock bekerja dengan membagi aliran fitur menjadi dua jalur yang diproses secara terpisah kemudian digabungkan kembali, sehingga mengurangi redundansi informasi sekaligus memperkaya representasi fitur yang dihasilkan [5]. Hasil evaluasi pada dataset Remote Sensing Object Detection (RSOD) dan Aircraft Classification menunjukkan peningkatan nilai mean Average Precision (mAP) dibandingkan model dasar, dengan mAP50 meningkat dari 97,98 persen menjadi 98,41 persen pada dataset RSOD [1].

Namun demikian, evaluasi pada penelitian terdahulu masih bertumpu pada mAP agregat dan AP50 sebagai metrik utama perbandingan. Intersection over Union (IoU) mengukur rasio area irisan terhadap area gabungan antara bounding box prediksi dan ground truth; semakin tinggi threshold IoU, semakin kecil toleransi terhadap kesalahan pusat, lebar, dan tinggi kotak. Pada detektor YOLO, koordinat bounding box dipelajari sebagai masalah regresi, sehingga kualitas lokalisasi dipengaruhi tidak hanya oleh kemampuan klasifikasi objek, tetapi juga oleh representasi spasial feature map dan fungsi loss yang mengarahkan optimasi koordinat. Literatur bounding-box regression menunjukkan adanya ketidakselarasan antara loss berbasis jarak parameter kotak dan metrik IoU, yang mendorong pengembangan GIoU, DIoU, dan CIoU untuk memasukkan geometri overlap, jarak pusat, dan rasio aspek secara lebih langsung [18], [19].

Kegagalan model pada threshold IoU tinggi dapat terjadi walaupun AP50 sudah tinggi. Pada strict localization, pergeseran beberapa piksel atau sedikit kesalahan skala dapat menurunkan IoU secara tajam, terutama pada objek kecil yang hanya direpresentasikan oleh sedikit sel feature map. Downsampling yang agresif mengurangi detail spasial, sedangkan feature map beresolusi lebih tinggi mempertahankan informasi tepi dan posisi yang lebih halus. Di sisi optimasi, CSPNet membagi aliran fitur ke jalur parsial dan menggabungkannya kembali untuk mengurangi duplicate gradient information serta menjaga keragaman aliran gradien [17]. Oleh karena itu, kombinasi CSPBlock dan kepala deteksi  $104 \times 104$  secara teoritis berpotensi memperbaiki representasi spasial dan regresi koordinat, tetapi pengaruhnya harus diuji pada spektrum IoU, bukan disimpulkan dari AP50 saja [5], [9], [17].

Kesenjangan tersebut menjadi fokus artikel ini. Studi arsitektur sebelumnya [1] memperkenalkan modifikasi YOLOv4-Tiny-3L dan melaporkan hasil pelatihan serta mAP agregat. Artikel ini tidak mengklaim arsitektur, dataset, atau eksperimen pelatihan baru; seluruh nilai AP berasal dari eksperimen yang sama dan dinyatakan secara eksplisit sebagai data sekunder. Batas kontribusinya adalah analisis metodologis yang berbeda, yaitu memetakan bagaimana performa berubah ketika kriteria overlap diperketat, mengidentifikasi zona IoU tempat keunggulan model muncul, dan menafsirkan konsekuensinya bagi desain detektor. Pemisahan ini penting karena dua model dengan AP50 hampir sama dapat memiliki kualitas lokalisasi yang berbeda pada AP75–AP95, informasi yang tidak terlihat pada pelaporan AP50 tunggal dan relevan bagi komunitas machine learning yang mengevaluasi robustness lokalisasi [6], [18].

Berdasarkan batas kontribusi tersebut, penelitian ini menggunakan dua indikator deskriptif yang diusulkan khusus untuk analisis ini, yaitu AP Decay dan AP Slope, serta segmentasi tiga zona IoU: coarse detection (AP50–AP65), moderate precision (AP70–AP80), dan strict localization (AP85–AP95). Tujuan penelitian adalah: (1) mengkarakterisasi penurunan AP kedua model pada IoU 0,50–0,95; (2) mengukur penurunan endpoint dan rata-rata penurunan per interval secara transparan; (3)

membandingkan pola per kelas tanpa mengubah data eksperimen asli; dan (4) menjelaskan, dari perspektif arsitektur CNN, mengapa CSPBlock dan kepala deteksi  $104 \times 104$  lebih berpengaruh pada lokalisasi menengah hingga ketat. Dengan demikian, novelty artikel terletak pada kerangka interpretasi kualitas lokalisasi, bukan pada pengulangan klaim peningkatan akurasi dari publikasi sebelumnya.

## 2. METODE PENELITIAN

Penelitian ini menggunakan pendekatan analisis kuantitatif terhadap data sekunder yang bersumber dari hasil eksperimen yang telah dipublikasikan sebelumnya pada studi modifikasi arsitektur YOLOv4-Tiny-3L [1]. Tidak dilakukan pelatihan ulang model maupun pengambilan data baru dalam penelitian ini. Seluruh nilai Average Precision (AP) yang dianalisis merupakan hasil evaluasi model yang identik dengan yang telah dilaporkan pada referensi tersebut, namun diolah menggunakan kerangka analisis baru yang belum diterapkan pada studi sebelumnya.

### 2.1. Sumber Data dan Model yang Dianalisis

Data yang digunakan dalam penelitian ini bersumber dari hasil evaluasi dua model pada dataset Remote Sensing Object Detection (RSOD), yaitu model YOLOv4-Tiny-3L sebagai model dasar dan model YOLOv4-Tiny-Mod sebagai model hasil modifikasi arsitektur [1]. Dataset RSOD terdiri dari 976 citra penginderaan jauh beresolusi tinggi yang diperoleh dari platform Google Earth dan Sky Map, mencakup lebih dari 6.950 objek yang terbagi dalam empat kelas, yaitu Aircraft, Oiltank, Playground, dan Overpass [10]. Keempat kelas tersebut memiliki karakteristik geometris yang berbeda, sehingga menyediakan variasi yang memadai untuk analisis perbandingan kualitas lokalisasi antar kelas.

Data yang diekstrak dari hasil eksperimen terdahulu meliputi nilai Average Precision pada sepuluh ambang batas IoU, yaitu AP50, AP55, AP60, AP65, AP70, AP75, AP80, AP85, AP90, dan AP95, serta nilai mean Average Precision (mAP) sebagai rata-rata dari kesepuluh nilai tersebut. Selain data tingkat model secara keseluruhan, data evaluasi per kelas juga diekstrak, mencakup nilai Precision, Recall, dan F1-Score untuk setiap kelas objek pada kedua model. Data confusion matrix yang tersedia dari studi terdahulu turut digunakan sebagai bahan pendukung interpretasi pada tahap pembahasan.

Tabel 1. AP50-AP95 YOLOv4-Tiny-3L dan YOLOv4-Tiny-Mod

| Average Precision | Yolov4-Tiny-3L | Yolov4-Tiny-Mods |
|-------------------|----------------|------------------|
| AP50              | 97.98%         | 98.41%           |
| AP55              | 97.66%         | 97.93%           |
| AP60              | 95.57%         | 97.50%           |
| AP65              | 93.22%         | 96.05%           |
| AP70              | 90.32%         | 94.04%           |
| AP75              | 81.80%         | 86.76%           |
| AP80              | 63.75%         | 73.95%           |
| AP85              | 34.11%         | 47.94%           |
| AP90              | 12.37%         | 15.30%           |
| AP95              | 0.26%          | 0.48%            |
| mAP               | 66.70%         | 70.84%           |

### 2.2. Definisi Metrik Analisis Baru

Untuk mengukur pola penurunan kualitas lokalisasi secara ringkas, penelitian ini memperkenalkan dua indikator deskriptif yang tidak dimaksudkan sebagai metrik standar baru untuk benchmark, yaitu AP Decay dan AP Slope. Keduanya merupakan usulan analitis penulis untuk merangkum kurva AP50–AP95 dari eksperimen sekunder; rumus tersebut tidak diadaptasi dari metrik baku yang telah memiliki distribusi inferensial atau uji signifikansi tersendiri.

Metrik pertama, AP Decay, didefinisikan sebagai selisih absolut antara AP50 dan AP95 pada level model. Indikator ini menangkap kehilangan performa dari kriteria overlap paling longgar ke paling ketat:

$$\text{AP Decay} = \text{AP}_{50} - \text{AP}_{95}$$

AP Decay yang lebih besar menunjukkan penurunan endpoint yang lebih tajam. Namun indikator ini harus dibaca bersama level AP absolut, karena dua model dengan titik awal AP50 berbeda dapat memiliki decay serupa meskipun salah satunya tetap unggul pada seluruh threshold. Pada data penelitian ini, AP95 kedua model mendekati nol sehingga AP Decay juga mengalami efek endpoint saturation; karena itu, rata-rata zona IoU digunakan sebagai bukti deskriptif utama untuk membandingkan kualitas lokalisasi.

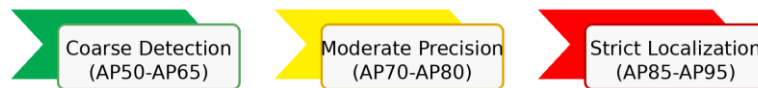
Metrik kedua, AP Slope, didefinisikan sebagai rata-rata penurunan endpoint per kenaikan satu langkah threshold 0,05. Istilah slope di sini bersifat deskriptif dan bukan koefisien slope hasil regresi linear:

$$\text{AP Slope} = (\text{AP}_{50} - \text{AP}_{95}) / 9$$

Pembagi 9 digunakan karena terdapat 10 threshold (0,50 sampai 0,95) yang membentuk 9 interval berjarak sama sebesar 0,05; hasilnya menyatakan rata-rata penurunan poin persentase per interval. Karena data yang tersedia berupa nilai AP agregat dari satu rangkaian eksperimen tanpa repeated random seeds atau prediksi per citra untuk bootstrap, AP Decay dan AP Slope tidak diberi interpretasi signifikansi statistik. Untuk analisis per kelas, AP95 tidak tersedia pada sumber data sekunder; oleh sebab itu penelitian tidak menghitung AP Decay per kelas. Sebagai gantinya digunakan AP50-to-mAP50:95 Localization Gap = AP50 – mAP50:95 sebagai indikator deskriptif terpisah, yang tidak ekuivalen dengan AP Decay.

### 2.3. Segmentasi Tiga Zona IoU

Untuk memudahkan interpretasi pola degradasi performa, sepuluh nilai threshold IoU dikelompokkan ke dalam tiga zona analisis berdasarkan tingkat ketaatan presisi yang dipersyaratkan.



Gambar 1. Diagram segmentasi tiga zona IoU

Zona pertama adalah Coarse Detection yang mencakup threshold AP50 hingga AP65. Pada zona ini, model dinilai berdasarkan kemampuan dasarnya dalam mengenali keberadaan objek tanpa tuntutan presisi posisi yang tinggi. Zona kedua adalah Moderate Precision yang mencakup threshold AP70 hingga AP80, merepresentasikan titik transisi di mana perbedaan kemampuan lokalisasi antar model mulai terlihat secara substansial. Zona ketiga adalah Strict Localization yang mencakup threshold AP85 hingga AP95, mengukur kemampuan model dalam menghasilkan bounding box yang benar-benar presisi dalam mencocokkan posisi dan dimensi objek sesungguhnya.

### 2.4. Prosedur Analisis

Analisis dilakukan melalui empat tahapan. Pertama, nilai AP50–AP95 ditabulasi ulang dan dihitung AP Decay serta AP Slope pada level model. Kedua, threshold dikelompokkan ke tiga zona IoU dan dihitung rata-rata AP per zona untuk membandingkan magnitude selisih kedua model. Ketiga, pada level kelas dihitung AP50-to-mAP50:95 Localization Gap karena AP95 per kelas tidak tersedia. Keempat, pola tersebut ditafsirkan bersama Precision, Recall, F1-Score, dan confusion matrix yang dilaporkan studi sumber. Tahap keempat bersifat explanatory-descriptive, bukan inferensi kausal; tanpa log pelatihan, distribusi sampel per kelas, prediksi per citra, dan eksperimen ablation, penyebab seperti class imbalance, resolusi objek, confidence threshold, atau Non-Maximum Suppression (NMS) hanya dapat diposisikan sebagai hipotesis yang perlu diuji lebih lanjut.

## 3. HASIL DAN PEMBAHASAN

Bagian ini menyajikan hasil analisis kuantitatif terhadap data evaluasi model YOLOv4-Tiny-3L dan YOLOv4-Tiny-Mod pada dataset RSOD. Penyajian hasil diorganisasikan berdasarkan empat tahapan analisis yang telah ditetapkan pada bagian Metode, meliputi tabulasi data AP lintas threshold, perhitungan AP Decay dan AP Slope, pengelompokan ke dalam tiga zona IoU, serta rekap data evaluasi per kelas objek.

### 3.1. Nilai Average Precision pada Seluruh Threshold IoU

Tabel 2 menyajikan nilai Average Precision kedua model pada sepuluh ambang batas IoU beserta AP Decay dan AP Slope yang dihitung berdasarkan Persamaan (1) dan Persamaan (2).

Tabel 2. Nilai AP pada Threshold serta AP Decay dan AP Slope

| Threshold | Yolov4-Tiny-3L | Yolov4-Tiny-Mods | Selisih |
|-----------|----------------|------------------|---------|
| AP50      | 97.98%         | 98.41%           | +0,43%  |
| AP55      | 97.66%         | 97.93%           | +0,27%  |
| AP60      | 95.57%         | 97.50%           | +1,93%  |
| AP65      | 93.22%         | 96.05%           | +2,83%  |
| AP70      | 90.32%         | 94.04%           | +3,72%  |
| AP75      | 81.80%         | 86.76%           | +4,96%  |
| AP80      | 63.75%         | 73.95%           | +10,20% |
| AP85      | 34.11%         | 47.94%           | +13,83% |
| AP90      | 12.37%         | 15.30%           | +2,93%  |
| AP95      | 0.26%          | 0.48%            | +0,22%  |
| mAP       | 66.70%         | 70.84%           | +4,14%  |
| AP Decay  | 97,72          | 97,93            | +0,21   |
| AP Slope  | 10,86          | 10,88            | +0,02   |

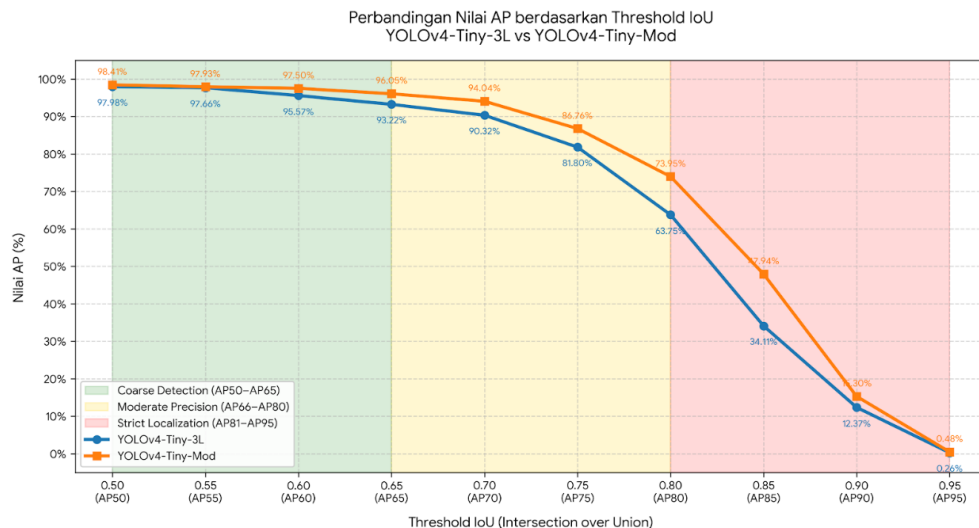
Berdasarkan Tabel 2, YOLOv4-Tiny-Mod menghasilkan AP lebih tinggi pada seluruh sepuluh threshold. Selisih terkecil terjadi pada AP95 (+0,22 poin persentase), sedangkan selisih terbesar pada AP85 (+13,83 poin). AP Decay endpoint justru sedikit lebih besar pada model modifikasi (97,93 vs. 97,72) dan AP Slope hampir identik (10,88 vs. 10,86). Temuan ini menegaskan keterbatasan indikator endpoint ketika AP95 kedua model sama-sama mendekati nol: superioritas model modifikasi tidak ditunjukkan oleh AP Decay yang lebih kecil, melainkan oleh level AP yang konsisten lebih tinggi dan pelebaran gap pada threshold menengah–ketat. Karena tidak tersedia repeated runs, seluruh selisih pada bagian ini diperlakukan sebagai magnitude deskriptif, bukan signifikansi statistik.

### 3.2. Analisis Berdasarkan Tiga Zona IoU

Tabel 3. Rata-rata Nilai AP per Zona IoU

| Zona IoU            | Threshold | Rata-rata 3L | Rata-rata Mod | Selisih |
|---------------------|-----------|--------------|---------------|---------|
| Coarse Detection    | AP50–AP65 | 96,11%       | 97,47%        | +1,37%  |
| Moderate Precision  | AP70–AP80 | 78,62%       | 84,92%        | +6,29%  |
| Strict Localization | AP85–AP95 | 15,58%       | 21,24%        | +5,66%  |

Berdasarkan Tabel 3, magnitude selisih kedua model berbeda antar zona. Pada Coarse Detection selisih rata-rata hanya 1,37 poin persentase; selisih kemudian melebar menjadi 6,29 poin pada Moderate Precision dan tetap tinggi sebesar 5,66 poin pada Strict Localization. Pola ini menunjukkan bahwa manfaat modifikasi arsitektur terutama muncul ketika kriteria overlap menjadi lebih ketat, bukan ketika tugas hanya menuntut deteksi kasar.



Gambar 2. Grafik garis perbandingan nilai AP kedua model

Gambar 2 memperlihatkan secara visual bahwa kurva kedua model bergerak berdekatan pada zona Coarse Detection, mulai terpisah pada zona Moderate Precision, dan menunjukkan perbedaan yang paling kontras di sekitar threshold AP85 sebelum kembali menyempit menjelang AP95.

### 3.3. Evaluasi Per Kelas Objek

Metrik evaluasi per kelas objek untuk kedua model, meliputi nilai Precision, Recall, F1-Score, mAP50, dan mAP50-95 sebagaimana diperoleh dari data eksperimen terdahulu [1].

Tabel 4. Metrik Evaluasi Per Kelas Objek pada Dataset RSOD

| Kelas      | Model           | Precision | Recall | F1-Score | mAP50  | mAP50-95 |
|------------|-----------------|-----------|--------|----------|--------|----------|
| All        | YOLOv4-Tiny-3L  | 0,91      | 0,96   | 0,93     | 97,98% | 66,70%   |
| All        | YOLOv4-Tiny-Mod | 0,94      | 0,94   | 0,94     | 98,41% | 70,84%   |
| Aircraft   | YOLOv4-Tiny-3L  | 0,903     | 0,940  | 0,921    | 95,65% | 58,54%   |
| Aircraft   | YOLOv4-Tiny-Mod | 0,933     | 0,998  | 0,965    | 96,33% | 62,03%   |
| Oiltank    | YOLOv4-Tiny-3L  | 0,937     | 0,283  | 0,428    | 99,15% | 72,50%   |
| Oiltank    | YOLOv4-Tiny-Mod | 0,955     | 0,283  | 0,428    | 99,30% | 75,64%   |
| Playground | YOLOv4-Tiny-3L  | 0,920     | 0,030  | 0,058    | 99,15% | 69,08%   |
| Playground | YOLOv4-Tiny-Mod | 0,958     | 0,030  | 0,058    | 99,59% | 74,84%   |

Berdasarkan Tabel 4, Aircraft menunjukkan kenaikan Recall dari 0,940 menjadi 0,998 dan F1-Score dari 0,921 menjadi 0,965. Sebaliknya, Recall Oiltank (0,283) dan Playground (0,030) tetap sangat rendah pada kedua model meskipun Precision dan mAP50 tinggi. Kombinasi Recall rendah dengan AP yang tinggi tidak harus kontradiktif karena Precision/Recall pada tabel dapat berasal dari satu operating threshold, sedangkan Average Precision mengintegrasikan precision–recall pada berbagai confidence threshold. Oleh sebab itu, data agregat ini belum cukup untuk menyimpulkan bahwa rendahnya Recall disebabkan class imbalance, NMS, atau resolusi; mekanisme tersebut dibahas sebagai hipotesis diagnostik pada Subbagian 3.7.

### 3.4. AP50-to-mAP50:95 Localization Gap Per Kelas Objek

Tabel 5. AP50-to-mAP50:95 Localization Gap Per Kelas Objek

| Kelas           | Localization Gap 3L<br>(AP50–mAP50-95) | Localization Gap Mod<br>(AP50–mAP50-95) | $\Delta$ Gap           |
|-----------------|----------------------------------------|-----------------------------------------|------------------------|
| Aircraft        | $95,65 - 58,54 = 37,11$                | $96,33 - 62,03 = 34,30$                 | -2,81 (Mod lebih baik) |
| Oiltank         | $99,15 - 72,50 = 26,65$                | $99,30 - 75,64 = 23,66$                 | -2,99 (Mod lebih baik) |
| Playground      | $99,15 - 69,08 = 30,07$                | $99,59 - 74,84 = 24,75$                 | -5,32 (Mod lebih baik) |
| All (rata-rata) | $97,98 - 66,70 = 31,28$                | $98,41 - 70,84 = 27,57$                 | -3,71 (Mod lebih baik) |

Tabel 5 tidak merepresentasikan AP Decay karena AP95 per kelas tidak tersedia. Tabel tersebut menunjukkan AP50-to-mAP50:95 Localization Gap, yaitu jarak antara kemampuan deteksi pada IoU 0,50 dan rata-rata performa pada rentang 0,50–0,95. Gap lebih kecil pada model modifikasi untuk Aircraft (-2,81 poin), Oiltank (-2,99), Playground (-5,32), dan agregat All (-3,71). Secara deskriptif, hal ini menunjukkan bahwa keuntungan AP50 pada model modifikasi lebih banyak dipertahankan ketika performa dirata-ratakan pada threshold yang lebih ketat, tetapi indikator ini tidak dapat menggantikan AP95 per kelas.

### 3.5. Interpretasi AP Decay dan AP Slope Keseluruhan Model

Hasil pada Tabel 2 memperlihatkan paradoks yang informatif: AP Decay endpoint YOLOv4-Tiny-Mod sedikit lebih besar, tetapi AP absolutnya lebih tinggi pada setiap threshold. Secara matematis, hal ini terjadi karena AP50 model modifikasi juga lebih tinggi sementara AP95 kedua model sama-sama hampir nol. Dengan demikian, AP Decay tidak boleh dibaca terpisah dari bentuk kurva. Zona Moderate Precision dan Strict Localization lebih diskriminatif untuk kasus ini karena mempertahankan informasi pada beberapa threshold, bukan hanya dua endpoint.

Dari perspektif arsitektur, kepala deteksi  $104 \times 104$  meningkatkan resolusi feature map yang digunakan untuk prediksi objek kecil. Untuk ukuran citra masukan yang sama, feature map yang lebih rapat memperkecil kuantisasi spasial per sel sehingga perubahan kecil pada pusat atau ukuran kotak dapat direpresentasikan lebih halus. Hal ini relevan pada threshold tinggi, ketika kesalahan koordinat kecil menghasilkan penurunan IoU yang besar. Prinsip multi-scale feature representation juga telah lama diketahui penting untuk localization karena lapisan beresolusi tinggi membawa detail posisi yang lebih kaya dibandingkan feature map yang telah mengalami downsampling berulang [9].

CSPBlock memberi mekanisme komplementer. CSPNet dirancang untuk mengurangi duplicate gradient information dengan memisahkan feature map ke dua jalur dan menggabungkannya kembali, sehingga variasi gradient flow dipertahankan tanpa menggandakan seluruh komputasi [17]. Pada konteks YOLOv4-Tiny-Mod, mekanisme ini secara teoritis dapat meningkatkan keragaman representasi yang sampai ke head regresi dan mengurangi hilangnya detail akibat transformasi berulang. Klaim ini merupakan penjelasan arsitektural yang konsisten dengan data, bukan bukti kausal tersendiri; verifikasiya memerlukan ablation yang memisahkan efek CSPBlock dari efek kepala  $104 \times 104$ . Dalam kerangka explainable analysis, pola gap yang membesar pada AP70–AP85 merupakan observasi yang dapat diuji pada ablation tersebut.

### 3.6. Analisis Pola Degradasi Berdasarkan Tiga Zona IoU

Pembagian threshold IoU ke dalam tiga zona memberikan interpretasi yang lebih stabil dibandingkan satu nilai endpoint. Pada Coarse Detection, kedua model sudah berada pada AP sangat tinggi sehingga ruang peningkatan kecil. Ketika threshold meningkat, fungsi evaluasi menjadi sensitif terhadap kesalahan pusat, skala, dan rasio aspek. Literatur GIoU serta DIOU/CIoU menunjukkan bahwa bounding-box regression membutuhkan objective yang selaras dengan geometri IoU; overlap saja tidak selalu memberikan gradient yang informatif, terutama ketika posisi pusat atau rasio aspek belum tepat [18], [19].

Pada Moderate Precision (AP70–AP80), selisih rata-rata melebar menjadi 6,29 poin persentase. Zona ini dapat dibaca sebagai titik ketika keuntungan resolusi spasial  $104 \times 104$  dan representasi backbone mulai terkonversi menjadi koordinat yang lebih tepat. Pada strict localization, keuntungan tetap ada tetapi kedua kurva turun tajam karena toleransi kesalahan menjadi sangat sempit. Untuk objek kecil, pergeseran satu-dua sel pada feature map atau sedikit kesalahan lebar/tinggi dapat menghasilkan penalti IoU yang jauh lebih besar daripada pada objek berukuran besar.

Pada Strict Localization (AP85–AP95), YOLOv4-Tiny-Mod masih unggul rata-rata 5,66 poin dan mencapai gap maksimum +13,83 poin pada AP85. Namun AP90 dan AP95 ikut turun tajam pada kedua model, menunjukkan bahwa penambahan head dan CSPBlock belum sepenuhnya menyelesaikan presisi koordinat ekstrem. Implikasi machine learning-nya

adalah pengembangan berikutnya tidak cukup hanya menambah kapasitas feature extractor; objective regresi yang lebih IoU-aware—misalnya GIoU, DIoU, atau CIoU—serta strategi box refinement perlu diuji secara terkontrol untuk melihat apakah kurva AP pada threshold  $\geq 0,85$  dapat dipertahankan [18], [19].

### 3.7. Interpretasi Per Kelas dan Analisis Anomali Recall

Analisis per kelas menunjukkan dua fenomena berbeda: kualitas lokalisasi pada deteksi yang berhasil dan kemampuan model menemukan seluruh objek. Localization Gap yang lebih kecil pada model modifikasi menunjukkan retensi performa yang lebih baik secara deskriptif, sedangkan Recall mengukur cakupan deteksi pada operating point tertentu. Karena data sekunder tidak menyediakan distribusi instance per kelas, histogram ukuran objek, confidence-score distribution, hasil sebelum/sesudah NMS, atau prediksi individual, penyebab rendahnya Recall tidak dapat ditetapkan secara kausal.

Aircraft merupakan kelas dengan perbaikan cakupan deteksi paling jelas: Recall naik dari 0,940 menjadi 0,998 dan F1-Score dari 0,921 menjadi 0,965. Secara arsitektural, objek pesawat pada citra penginderaan jauh sering berukuran kecil sehingga head  $104 \times 104$  memberi sampling spasial yang lebih rapat. Bentuk pesawat yang relatif terstruktur juga memungkinkan peningkatan representasi backbone diterjemahkan menjadi fitur yang lebih diskriminatif. Namun untuk memisahkan kontribusi resolusi head dan CSPBlock, diperlukan ablation head-only, CSP-only, dan kombinasi keduanya.

Playground memiliki Localization Gap yang berkurang paling besar (5,32 poin), tetapi Recall tetap 0,030. Pola ini menunjukkan bahwa peningkatan kualitas ranking/lokalisasi tidak otomatis meningkatkan cakupan deteksi pada confidence threshold yang digunakan. Beberapa hipotesis yang perlu diuji adalah ketidakseimbangan jumlah instance atau variasi intra-kelas, ukuran objek yang tidak cocok dengan anchor/scale assignment, confidence threshold yang terlalu konservatif, serta suppression oleh NMS ketika prediksi tumpang tindih. Class imbalance merupakan masalah yang dikenal pada dense one-stage detector dan dapat ditangani antara lain dengan loss yang memfokuskan pembelajaran pada hard examples, seperti Focal Loss [20]; namun penerapannya pada RSOD harus divalidasi melalui eksperimen baru, bukan diasumsikan dari tabel agregat.

Oiltank menunjukkan pola serupa: Localization Gap turun 2,99 poin tetapi Recall bertahan di 0,283. Bentuk melingkar yang relatif konsisten dapat membantu lokalisasi saat objek berhasil ditemukan, sementara background clutter, skala kecil, atau thresholding dapat tetap menghasilkan false negative. NMS juga dapat memengaruhi Recall pada objek yang berdekatan karena kandidat dengan overlap tinggi dapat dieliminasi; DIoU-NMS merupakan salah satu alternatif yang secara eksplisit mempertimbangkan jarak pusat [19]. Karena output per kandidat tidak tersedia dalam data sekunder, pengaruh NMS pada kelas ini tetap merupakan hipotesis yang perlu diuji dengan analisis sebelum dan sesudah suppression.

### 3.8. Implikasi terhadap Pengembangan Arsitektur dan Aplikasi

Temuan penelitian ini memberi implikasi arsitektural yang lebih luas daripada pemilihan model untuk satu domain. Kenaikan yang kecil pada Coarse Detection tetapi besar pada Moderate/Strict Localization menunjukkan bahwa peningkatan resolusi head dan pengelolaan aliran gradient melalui CSP lebih berkaitan dengan ketelitian spasial daripada sekadar keberhasilan klasifikasi objek. Bagi pengembangan YOLO/CNN ringan, evaluasi sebaiknya melaporkan kurva AP lintas IoU atau AP75/AP90 selain AP50 agar perbaikan pada feature extractor tidak tersembunyi oleh saturation di threshold rendah.

Pada aplikasi penginderaan jauh yang memerlukan posisi objek akurat, seperti estimasi posisi dan jarak antarobjek, zona Moderate Precision dan Strict Localization lebih relevan daripada AP50 semata. YOLOv4-Tiny-Mod memberikan keuntungan rata-rata 6,29 dan 5,66 poin pada kedua zona tersebut. Akan tetapi, hasil ini belum membuktikan robustness terhadap kondisi lingkungan baru karena analisis menggunakan dataset dan output eksperimen yang sama dengan studi sumber [1]. Validasi eksternal pada dataset lain diperlukan sebelum generalisasi operasional dilakukan.

Untuk tugas yang hanya membutuhkan keberadaan atau penghitungan kasar objek, selisih 1,37 poin pada Coarse Detection dapat membuat efisiensi komputasi menjadi pertimbangan yang lebih dominan. Sebaliknya, target strict localization memerlukan optimasi lebih lanjut pada objective regresi, scale assignment, dan box refinement. Eksperimen berikutnya sebaiknya membandingkan loss IoU-aware (GIoU/DIoU/CIoU), menilai head  $104 \times 104$  melalui ablation, dan melaporkan parameter/FLOPs serta latency agar trade-off akurasi lokalisasi dan biaya komputasi dapat dinilai secara lengkap [18], [19].

Khusus untuk Oiltank dan Playground, prioritas penelitian bukan hanya memperbaiki bounding box, tetapi juga meningkatkan Recall. Pengujian yang disarankan meliputi audit distribusi kelas dan ukuran objek, precision-recall curve per kelas, sensitivity analysis terhadap confidence threshold dan NMS IoU threshold, class-balanced sampling atau Focal Loss, serta error analysis berbasis false negative. Rangkaian ini akan membedakan apakah kegagalan berasal dari data, representasi fitur, score calibration, atau post-processing [19], [20].

## 4. KESIMPULAN

(1) Secara empiris, YOLOv4-Tiny-Mod menghasilkan AP lebih tinggi pada seluruh IoU 0,50–0,95, dengan keuntungan terbesar pada AP85 (+13,83 poin), rata-rata +6,29 poin pada Moderate Precision, dan +5,66 poin pada Strict Localization. AP Decay endpoint tidak lebih kecil pada model modifikasi (97,93 vs. 97,72), sehingga indikator tersebut tidak digunakan sendiri ketika AP95 kedua model telah mendekati nol. Pada level kelas, Aircraft menunjukkan perbaikan Recall paling jelas, sedangkan Oiltank dan Playground masih memiliki Recall rendah.

(2) Secara teoretis, pola tersebut mendukung hipotesis bahwa detection head  $104 \times 104$  mempertahankan detail spasial yang dibutuhkan untuk regresi koordinat, sedangkan CSPBlock menjaga keragaman aliran gradien. Manfaat arsitektur karena itu lebih terlihat pada lokalisasi menengah hingga ketat daripada pada AP50 yang sudah mengalami saturation. Penjelasan ini tetap bersifat mekanistik-deskriptif dan perlu diverifikasi melalui ablation terpisah untuk CSPBlock, high-resolution head, dan kombinasinya.

(3) Pengembangan berikutnya perlu memprioritaskan loss regresi IoU-aware (GIoU/DIoU/CIoU), box refinement dan konfigurasi NMS untuk strict localization, serta class-balanced sampling/Focal Loss dan confidence calibration untuk kelas dengan Recall rendah [18]–[20]. Karena penelitian ini menggunakan data sekunder agregat dari satu eksperimen tanpa repeated seeds dan external validation, temuan tidak diposisikan sebagai uji signifikansi statistik atau bukti kausal. Replikasi pada dataset lain dan repeated runs diperlukan untuk menguji generalisasi.

## REFERENSI

- [1] A. Fauzan dan N. N. R. A. Mokobombang, "Modification of YOLOv4-Tiny-3L Architecture to Improve the Accuracy of Airport Object Detection," dalam Proc. International Conference on Information Technology and Computer Science (ICTCS), 2025.
- [2] X. Cheng, B. Zhang, J. Chen, dan H. Ke, "Deep Learning-Based Object Detection Techniques for Remote Sensing Images: A Survey," *Remote Sensing*, vol. 14, no. 10, hal. 2385, Mei 2022. [Daring]. Tersedia: <https://www.mdpi.com/2072-4292/14/10/2385>
- [3] P. Jiang, D. Ergu, F. Liu, Y. Cai, dan B. Ma, "A Review of Yolo Algorithm Developments," *Procedia Computer Science*, vol. 199, hal. 1066–1073, 2022. doi: 10.1016/j.procs.2022.01.135
- [4] V. K. Sharma, P. Dhiman, dan R. K. Rout, "Improved Traffic Sign Recognition Algorithm Based on YOLOv4-Tiny," *Journal of Visual Communication and Image Representation*, vol. 91, hal. 103774, Mar. 2023. doi: 10.1016/j.jvcir.2023.103774
- [5] C.-Y. Wang, A. Bochkovskiy, dan H.-Y. M. Liao, "Scaled-YOLOv4: Scaling Cross Stage Partial Network," Feb. 2021, arXiv: arXiv:2011.08036. doi: 10.48550/arXiv.2011.08036
- [6] Z. Cai dan N. Vasconcelos, "Cascade R-CNN: High Quality Object Detection and Instance Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 5, hal. 1483–1498, Mei 2021. doi: 10.1109/TPAMI.2019.2956516
- [7] D. Cai, Z. Zhang, dan Z. Zhang, "Corner-Point and Foreground-Area IoU Loss: Better Localization of Small Objects in Bounding Box Regression," *Sensors*, vol. 23, no. 10, hal. 4961, Mei 2023. doi: 10.3390/s23104961
- [8] S. El Ghazouali, A. Gucciardi, F. Venturini, N. Venturi, M. Rueeggsegger, dan U. Michelucci, "FlightScope: An Experimental Comparative Review of Aircraft Detection Algorithms in Satellite Imagery," *Remote Sensing*, vol. 16, no. 24, hal. 4715, Jan. 2024. doi: 10.3390/rs16244715
- [9] J. Liu dan Z. Wang, "Small Object Detection Based on Deep Learning for Remote Sensing: A Comprehensive Review," *Remote Sensing*, vol. 15, no. 13, hal. 3265, Jun. 2023. [Daring]. Tersedia: <https://www.mdpi.com/2072-4292/15/13/3265>
- [10] RSIA-LIESMARS-WHU, "RSOD-Dataset," GitHub, Jan. 2025. [Daring]. Tersedia: <https://github.com/RSIA-LIESMARS-WHU/RSOD-Dataset>
- [11] Ying-Tung Hsiao, Jia-Shing Sheu, dan Hsu Ma, "Efficient Object Detection and Intelligent Information Display Using YOLOv4-Tiny," *Advances in Technology Innovation*, vol. 9, no. 1, hal. 42–49, Des. 2023. doi: 10.46604/aiti.2023.12682
- [12] N. T. Allo, Indrabayu, dan Z. Zainuddin, "A Novel Approach of Hybrid Bounding Box Regression Mechanism to Improve Convergency Rate and Accuracy," *International Journal of Intelligent Engineering and Systems*, vol. 17, no. 2, hal. 715–727, Apr. 2024. doi: 10.22266/ijies2024.0430.57
- [13] Y.-H. Liao dan J.-G. Juang, "Automatic Marine Debris Inspection," *Aerospace*, vol. 10, no. 1, hal. 84, Jan. 2023. doi: 10.3390/aerospace10010084
- [14] Z.-Q. Zhao, P. Zheng, S.-T. Xu, dan X. Wu, "Object Detection With Deep Learning: A Review," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 30, no. 11, hal. 3212–3232, Nov. 2019. doi: 10.1109/TNNLS.2018.2876865
- [15] M. L. Ali dan Z. Zhang, "The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection," *Computers*, vol. 13, no. 12, hal. 336, Des. 2024. doi: 10.3390/computers13120336
- [16] N. Jegham, C. Y. Koh, M. Abdelatti, dan A. Hendawi, "Evaluating the Evolution of YOLO (You Only Look Once) Models: A Comprehensive Benchmark Study of YOLO11 and Its Predecessors," Okt. 2024, arXiv: arXiv:2411.00201. doi: 10.48550/arXiv.2411.00201
- [17] C.-Y. Wang, H.-Y. M. Liao, Y.-H. Wu, P.-Y. Chen, J.-W. Hsieh, dan I.-H. Yeh, "CSPNet: A New Backbone That Can Enhance Learning Capability of CNN," dalam Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), 2020, hal. 390–391.
- [18] H. Rezatofighi, N. Tsoi, J. Gwak, A. Sadeghian, I. Reid, dan S. Savarese, "Generalized Intersection Over Union: A Metric and a Loss for Bounding Box Regression," dalam Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, hal. 658–666.
- [19] Z. Zheng, P. Wang, W. Liu, J. Li, R. Ye, dan D. Ren, "Distance-IoU Loss: Faster and Better Learning for Bounding Box Regression," *Proceedings of the AAAI Conference on Artificial Intelligence*, vol. 34, no. 07, hal. 12993–13000, 2020. doi: 10.1609/aaai.v34i07.6999.
- [20] T.-Y. Lin, P. Goyal, R. Girshick, K. He, dan P. Dollár, "Focal Loss for Dense Object Detection," dalam Proc. IEEE International Conference on Computer Vision (ICCV), 2017, hal. 2980–2988.