

# Evaluasi Dampak Keterbatasan Anotasi terhadap Kinerja YOLOv11 dalam Deteksi Multi-Objek Buah

Nugroho Fadillah Yudha Putra  
Program Studi Informatika  
Universitas Islam Indonesia  
Yogyakarta, Indonesia  
18523156@students.uii.ac.id

Arrie Kurniawardhani, S.Si., M.Kom.  
Program Studi Informatika  
Universitas Islam Indonesia  
Yogyakarta, Indonesia  
arrie.kurniawardhani@uii.ac.id

*Abstract—Proses anotasi manual pada dataset pelatihan seringkali terbatas oleh waktu dan sumber daya, sehingga menghasilkan data yang tidak sepenuhnya dianotasi. Penelitian ini bertujuan untuk mengevaluasi seberapa besar penurunan performa yang terjadi pada model deteksi objek YOLOv11 ketika dilatih menggunakan anotasi terbatas, dibandingkan dengan model beranotasi lengkap. Lima kelas buah seperti apel, jeruk, naga, durian, dan alpukat digunakan dalam eksperimen menggunakan dataset buatan sendiri, dengan distribusi 70% pelatihan, 20% validasi, dan 10% pengujian. Model dilatih di Google Colab menggunakan GPU Tesla T4, dan performa dievaluasi menggunakan precision, recall, mAP@0.5, dan mAP@0.5:0.95. Hasil menunjukkan penurunan signifikan pada model dengan anotasi terbatas, seperti turunnya mAP@0.5 hingga 33,84%. Analisis confusion matrix mengungkapkan bahwa kelas buah naga dan jeruk mengalami penurunan akurasi paling tajam, diduga akibat warna mencolok namun kontras rendah terhadap latar, serta kompleksitas bentuk. Studi ini menyoroti pentingnya kelengkapan anotasi dalam mendukung akurasi deteksi, sekaligus memberikan wawasan tentang batas toleransi performa model dalam kondisi pelabelan yang tidak ideal.*

**Keywords**—YOLOv11, Deteksi Objek, Anotasi Terbatas, Evaluasi Model, Multi-objek Buah.

## I. PENDAHULUAN

Deteksi objek dalam citra digital menjadi tantangan mendasar dalam *computer vision*, khususnya pada sektor pertanian dan sistem distribusi pangan, yang sangat bergantung pada keakuratan identifikasi objek di berbagai kondisi lingkungan. Kualitas dan kelengkapan anotasi dalam dataset pelatihan sangat mempengaruhi kinerja sistem deteksi objek secara keseluruhan [1]. Salah satu permasalahan yang umum terjadi adalah keterbatasan waktu dan sumber daya dalam melakukan proses anotasi secara menyeluruh. Selain memerlukan sumber daya waktu yang besar, proses anotasi juga bersifat repetitif dan membosankan, sehingga menjadi beban dalam pengembangan sistem berbasis visi komputer secara berkelanjutan. Hal ini menyebabkan banyak model yang dilatih menggunakan data dengan label yang tidak lengkap, yang secara potensial menurunkan akurasi deteksi objek di dunia nyata. Studi ini mengangkat persoalan tersebut dalam konteks deteksi *multi-object* buah, yang kompleks karena objek memiliki bentuk, ukuran, dan posisi yang bervariasi [2]. Oleh karena itu, diperlukan pendekatan model yang mampu menangani kondisi kompleks secara efisien dan akurat. Fenomena serupa juga terlihat pada penerapan YOLOv5 dalam deteksi buah jeruk di kebun alami yang sangat dipengaruhi oleh kualitas anotasi [3], serta pada penerapan YOLOv11 untuk klasifikasi sampah pada sistem

bank sampah berbasis web [4]. Oleh karena itu, model deteksi objek berbasis *deep learning* seperti YOLOv11 menjadi salah satu pendekatan yang potensial untuk memberikan solusi atas permasalahan kualitas anotasi.

Salah satu model yang berkembang pesat untuk mendukung deteksi *real-time* adalah YOLOv11, versi terbaru dari keluarga *You Only Look Once*, yang dirancang dengan efisiensi tinggi dan peningkatan akurasi dalam mendekripsi objek pada beragam skenario. Model ini banyak diterapkan dalam berbagai aplikasi, mulai dari deteksi objek industri [5], pengolahan pangan [6], hingga sistem otomatisasi berbasis *Internet of Things* (IoT) [7]. Arsitektur YOLOv11 mengusung keunggulan berupa proses deteksi satu tahap yang efisien dan dukungan feature fusion multi-skala untuk meningkatkan presisi. Dibandingkan dengan pendahulunya, model ini memberikan performa tinggi meskipun dengan jumlah parameter yang lebih ringan, menjadikannya cocok untuk aplikasi dengan keterbatasan komputasi [8].

Penelitian ini tidak berfokus pada pembandingan performa antara model anotasi lengkap dan terbatas secara langsung, melainkan pada evaluasi seberapa besar penurunan performa yang terjadi ketika model YOLOv11 dilatih dengan data yang tidak sepenuhnya dianotasi. Penelitian ini mengukur penurunan akurasi deteksi dan klasifikasi, serta mengidentifikasi kelas buah yang paling terdampak akibat keterbatasan anotasi. Selain itu, dilakukan analisis terhadap faktor-faktor yang mempengaruhi performa model, seperti warna, bentuk, dan kompleksitas visual objek dalam gambar. Pendekatan ini merepresentasikan tantangan dunia nyata dalam pengumpulan data yang sering kali tidak sempurna, serta memberikan wawasan tentang ketahanan (*robustness*) model terhadap kondisi anotasi yang terbatas. Dengan demikian, studi ini memberikan kontribusi dalam memahami batas toleransi model deteksi objek terhadap kekurangan label pada dataset.

## II. TINJAUAN PUSTAKA

### A. Penelitian Terdahulu

Kajian terhadap beberapa penelitian sebelumnya menunjukkan bahwa penerapan algoritma YOLO dalam deteksi buah umumnya difokuskan pada perbandingan arsitektur model atau klasifikasi tingkat kematangan buah. Namun, sebagian besar studi tidak mempertimbangkan pengaruh *kualitas anotasi* sebagai variabel utama dalam evaluasi performa model. Table I berikut merangkum

beberapa penelitian relevan yang menjadi acuan dalam menyusun arah kontribusi penelitian ini.

TABLE I. RINGKASAN PENELITIAN TERDAHULU

No	Jurnal	Isi
1	YOLOv5 versus YOLOv3 for Apple Detection[9]	Membandingkan performa dua arsitektur YOLO dalam deteksi apel. Penelitian ini berfokus pada aspek akurasi dan waktu inferensi, tanpa mempertimbangkan aspek anotasi.
2	Deteksi Kematangan Buah Pepaya Menggunakan Algoritma YOLO Berbasis Android[10]	Menerapkan YOLO untuk klasifikasi kematangan pepaya. Penelitian ini berfokus pada fitur visual objek, tidak meneliti dampak kelengkapan anotasi terhadap hasil model.
3	Evaluasi Kinerja Model YOLOv8 dalam Deteksi Kesegaran Buah[11]	Meneliti variasi data dan preprocessing pada deteksi buah. Penelitian ini tidak secara eksplisit menguji performa model terhadap perubahan tingkat anotasi.

Dari Table I dapat disimpulkan bahwa belum terdapat penelitian yang secara langsung mengevaluasi *seberapa besar dampak keterbatasan anotasi* terhadap performa model deteksi objek, khususnya pada kasus multi-objek buah. Penelitian ini mengisi celah tersebut dengan mengevaluasi dampak anotasi terbatas terhadap akurasi dan presisi deteksi objek menggunakan YOLOv11. Fokus ini memberikan kontribusi baru terhadap efisiensi pengembangan dataset dan ketahanan model dalam kondisi anotasi yang tidak sempurna, yang umum terjadi pada skenario dunia nyata.

### B. YOLOv11

YOLOv11 merupakan *pengembangan terbaru* dari keluarga algoritma *YOLO* yang dirancang untuk meningkatkan *efisiensi* dan *akurasi* deteksi objek secara *real-time*. Model ini mengintegrasikan modul seperti *C3k2*, *SPPF*, dan *C2PSA* untuk memperkuat *ekstraksi fitur* dan *atensi spasial*, sehingga lebih unggul dalam mengenali objek kecil dan tumpang tindih. Dibandingkan dengan pendahulunya, YOLOv11 menunjukkan peningkatan *mAP* dan *efisiensi parameter*, menjadikannya pilihan tepat dalam skenario komputasi terbatas [12].

### C. Penggunaan YOLO dalam Deteksi Buah

Algoritma *You Only Look Once* (YOLO) telah banyak diterapkan dalam penelitian deteksi buah, terutama karena kemampuannya dalam mengenali objek secara *real-time* dan efisien pada berbagai kondisi visual. Keunggulan ini menjadikan YOLO relevan untuk digunakan dalam pengembangan sistem klasifikasi dan pemantauan buah secara otomatis.

Penelitian pada [13] memanfaatkan YOLOv5 untuk mengklasifikasikan tingkat kematangan buah sawit berdasarkan warna dan tekstur. Studi ini menekankan efektivitas YOLO dalam mengenali fitur visual spesifik dari objek dengan latar kompleks. Namun, keberhasilan tersebut sangat dipengaruhi oleh kualitas dan konsistensi anotasi data pelatihan.

Dalam konteks yang lebih luas, [14] mengimplementasikan YOLOv11 untuk deteksi dan klasifikasi sampah berbasis citra. Studi ini menunjukkan

bahwa YOLOv11 mampu mengenali objek kecil dengan variasi tinggi, yang relevan untuk deteksi multi-objek buah yang memiliki bentuk, warna, dan ukuran beragam. Namun, studi tersebut tidak membahas dampak anotasi tidak lengkap secara eksplisit.

Sementara itu, penelitian pada [15] juga menyoroti pentingnya konsistensi dan kelengkapan anotasi dalam pelatihan YOLOv5 untuk mendekripsi kematangan buah tomat. Meskipun model mencapai performa yang baik, pendekatan ini belum mengevaluasi bagaimana keterbatasan anotasi memengaruhi kinerja model secara kuantitatif.

Dengan demikian, meskipun YOLO telah terbukti efektif dalam deteksi objek buah, masih sedikit studi yang secara khusus menilai ketahanan (*robustness*) model terhadap kekurangan anotasi. Penelitian ini mengisi celah tersebut dengan menelaah seberapa besar dampak anotasi terbatas terhadap performa deteksi, sekaligus mengidentifikasi kelas objek yang paling terdampak.

### D. Dampak Keterbatasan Anotasi terhadap Kinerja Deteksi

Dalam pengembangan model deteksi objek berbasis *deep learning*, kualitas dan cakupan anotasi pada dataset pelatihan memainkan peran penting dalam membentuk kemampuan generalisasi model. Ketika objek dalam citra dianotasi secara menyeluruh, model memiliki eksposur yang lebih lengkap terhadap pola visual yang beragam, sehingga meningkatkan akurasi deteksi dan lokasi objek.

Namun, dalam praktik dunia nyata, keterbatasan sumber daya dan waktu sering kali menyebabkan proses anotasi dilakukan secara parsial. Hal ini menimbulkan pertanyaan penting: sejauh mana keterbatasan anotasi dapat ditoleransi tanpa menyebabkan degradasi performa yang signifikan?

Penelitian terdahulu cenderung menekankan pentingnya anotasi penuh, tetapi belum banyak yang secara eksplisit mengkaji dampak penurunan performa akibat keterbatasan anotasi. Studi ini mencoba mengisi celah tersebut dengan menganalisis efek dari strategi anotasi terbatas terhadap performa model YOLOv11, serta mengevaluasi kelas objek mana yang paling terdampak dan mengapa hal itu terjadi. Pendekatan ini memberikan perspektif baru terhadap efisiensi proses anotasi dan potensi adaptabilitas model dalam menghadapi data yang tidak sempurna.

### E. Posisi Penelitian

Penelitian terdahulu mengenai deteksi objek menggunakan algoritma YOLO umumnya berfokus pada pengembangan arsitektur model atau optimisasi performa melalui teknik pelatihan dan augmentasi data. Sebagian besar studi tersebut menggunakan data beranotasi lengkap dan jarang membahas dampak dari keterbatasan anotasi terhadap kinerja model deteksi, terutama dalam konteks deteksi objek multi-kelas seperti buah.

Penelitian ini mengambil posisi yang berbeda dengan mengevaluasi seberapa besar penurunan performa yang terjadi ketika model YOLOv11 dilatih menggunakan data yang tidak sepenuhnya dianotasi. Fokus penelitian tidak terletak pada penentuan model mana yang lebih unggul, melainkan pada analisis dampak kekurangan anotasi terhadap metrik performa

seperti *precision*, *recall*, dan *mean Average Precision*, serta identifikasi kelas objek yang paling terdampak.

Dengan mengangkat skenario keterbatasan sumber daya pelabelan yang umum terjadi dalam praktik annotasi di dunia nyata, penelitian ini memberikan kontribusi terhadap pemahaman tentang robustnya model YOLOv11 dalam menghadapi kondisi data yang tidak ideal. Temuan ini diharapkan dapat menjadi pertimbangan penting dalam pengembangan sistem deteksi objek yang efisien namun tetap andal, terutama dalam proyek berskala besar dengan keterbatasan waktu dan tenaga anotator.

### III. METODOLOGI PENELITIAN

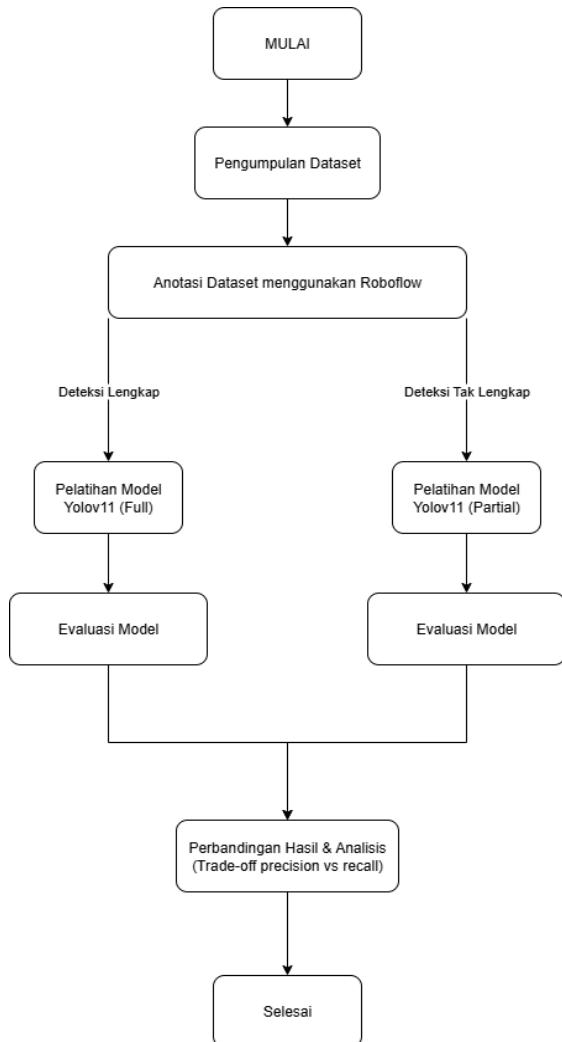


Fig. 1. Alur Penelitian

#### A. Desain Penelitian

Penelitian ini menggunakan pendekatan eksperimen kuantitatif dengan desain evaluatif, yang bertujuan untuk menganalisis dampak keterbatasan annotasi terhadap kinerja model deteksi objek berbasis YOLOv11. Fokus penelitian bukan sekadar membandingkan performa dua model, tetapi mengevaluasi sejauh mana penurunan performa terjadi saat model dilatih dengan annotasi yang tidak lengkap, serta mengidentifikasi penyebab dari penurunan tersebut.

Dua strategi annotasi digunakan dalam proses pelatihan model:

- **Model dengan Anotasi Lengkap:** Data set pelatihan memuat anotasi bounding box untuk seluruh objek buah yang muncul pada setiap gambar. Model ini berfungsi sebagai tolok ukur untuk kinerja ideal.
- **Model dengan Anotasi Terbatas:** Data set pelatihan hanya memuat sebagian anotasi objek, yaitu sekitar 40–50% dari total objek yang tampak. Pendekatan ini digunakan untuk mensimulasikan kondisi nyata ketika proses annotasi terbatas oleh waktu dan sumber daya.

Model yang dihasilkan dari kedua skenario ini kemudian dievaluasi berdasarkan metrik *precision*, *recall*, dan *mean Average Precision (mAP)*. Selain itu, dilakukan analisis terhadap perbedaan performa per kelas objek, serta eksplorasi terhadap kemungkinan penyebab penurunan performa, seperti bentuk, warna, dan distribusi objek dalam gambar.

Pendekatan eksperimen ini selaras dengan penelitian sebelumnya yang menunjukkan bahwa strategi annotasi memiliki pengaruh signifikan terhadap efektivitas deteksi objek dalam model deep learning [11]. Selain itu, studi lain juga menekankan pentingnya evaluasi terhadap ketahanan (*robustness*) model ketika dilatih dengan data yang tidak ideal [16]. Oleh karena itu, desain penelitian ini memberikan kerangka yang tepat untuk memahami batas toleransi model deteksi terhadap keterbatasan annotasi dalam pengembangan sistem *computer vision* yang efisien dan adaptif.

#### B. Dataset dan Anotasi

Dataset yang digunakan dalam penelitian ini dikembangkan secara mandiri dengan mengumpulkan 760 gambar dari Google Image dan iStock. Terdapat lima kelas buah yang digunakan, yaitu *alpukat*, *apel*, *jeruk*, *durian*, dan *naga*, dengan masing-masing kelas terdiri dari 152 gambar. Dataset ini dibagi menjadi tiga bagian: 70% untuk data latih, 20% untuk data validasi, dan 10% untuk data uji.

Seluruh proses annotasi dilakukan menggunakan platform *Roboflow* untuk menghasilkan *bounding box* pada objek buah. Proses annotasi ini dilakukan secara manual atau semi-otomatis di luar proses pelatihan, karena model *YOLOv11* membutuhkan data yang telah diberi *bounding box* sebagai input pelatihan. Untuk model deteksi lengkap, seluruh objek buah yang terlihat dalam citra diberi annotasi secara menyeluruh. Sementara itu, untuk model deteksi terbatas, hanya sekitar 40–50% dari total objek buah dalam gambar yang diberikan *bounding box*. Strategi annotasi terbatas ini dilakukan untuk mensimulasikan kondisi dunia nyata, di mana proses annotasi seringkali terkendala oleh keterbatasan waktu dan tenaga.

Gambar yang ditunjukkan pada Fig.2 dan Fig.3 memperlihatkan perbedaan antara annotasi lengkap dan annotasi terbatas.

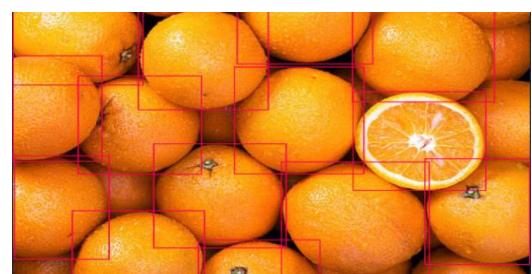


Fig. 2. Anotasi Lengkap

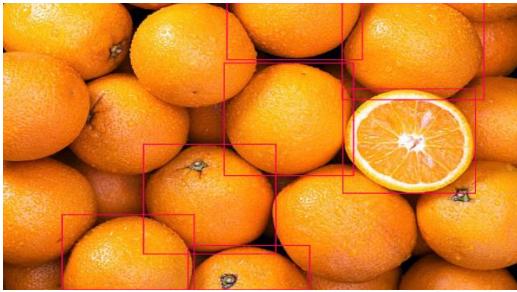


Fig. 3. Anotasi Terbatas

### C. Pelatihan Model

Proses pelatihan dilakukan menggunakan platform Google Colab dengan dukungan GPU Tesla T4 untuk mempercepat komputasi. Model dilatih dari awal (*training from scratch*) tanpa menggunakan *pre-trained weights*, agar hasil pelatihan mencerminkan pengaruh murni dari strategi anotasi yang digunakan.

Arsitektur model yang digunakan adalah YOLOv11-S, dipilih karena bersifat ringan dan efisien untuk eksperimen komparatif pada skala dataset menengah. Seluruh model dilatih selama 30 *epoch* dengan ukuran input citra 640 piksel.

Dalam penelitian ini, parameter pelatihan diseragamkan antara model dengan anotasi lengkap dan terbatas. Tujuannya adalah untuk mengisolasi pengaruh perbedaan strategi anotasi sebagai satu-satunya variabel bebas, sehingga evaluasi terhadap dampak anotasi terhadap performa dapat dilakukan secara objektif.

Selama proses pelatihan, grafik *loss function*, *precision*, *recall*, dan *mean Average Precision* (mAP) dipantau untuk mengevaluasi perkembangan performa model dari waktu ke waktu. Hasil akhir dari pelatihan disimpan dalam format .pt dan digunakan pada tahap pengujian dan analisis lanjutan.

### D. Pengujian dan Evaluasi Model

Setelah proses pelatihan selesai, masing-masing model diuji menggunakan *testing set* yang telah dipisahkan sejak awal dari data pelatihan. Proses evaluasi dilakukan secara otomatis menggunakan modul evaluasi bawaan dari Ultralytics YOLOv11.

Evaluasi dilakukan dengan menggunakan metrik standar dalam deteksi objek, yaitu *precision*, *recall*, dan *mean Average Precision* (mAP). Metrik ini dihitung berdasarkan perbandingan antara hasil prediksi model dan *ground truth* pada dataset uji. Selain nilai numerik, hasil evaluasi juga divisualisasikan dalam bentuk grafik metrik dan *confusion matrix*.

Dalam konteks penelitian ini, evaluasi tidak hanya ditujukan untuk membandingkan dua model, tetapi lebih diarahkan untuk mengukur sejauh mana keterbatasan anotasi berdampak pada penurunan performa model deteksi. Oleh karena itu, proses pengujian dirancang untuk mendeteksi adanya perbedaan signifikan dalam akurasi klasifikasi objek dan pelokalan (*bounding box*) pada masing-masing kelas buah. Dengan pendekatan ini, penelitian dapat memberikan gambaran yang lebih menyeluruh terhadap ketahanan (*robustness*) model YOLOv11 dalam menghadapi data yang tidak sepenuhnya dianotasi.

### E. Teknik Analisis Data

Analisis data dalam penelitian ini dilakukan secara kuantitatif berdasarkan hasil evaluasi performa model yang diperoleh dari proses pengujian. Fokus utama analisis bukan untuk menentukan model mana yang lebih unggul, melainkan untuk mengevaluasi dampak keterbatasan anotasi terhadap performa deteksi objek oleh model YOLOv11. Dengan kata lain, penelitian ini bertujuan mengukur sejauh mana penurunan performa terjadi akibat penggunaan anotasi yang tidak lengkap.

Teknik analisis dilakukan dengan mengacu pada metrik evaluasi standar dalam domain deteksi objek, yaitu:

- **Precision**

Precision mengukur ketepatan model dalam melakukan deteksi, yakni seberapa banyak prediksi positif yang benar terhadap seluruh prediksi positif yang dihasilkan. Rumus perhitungan precision ditunjukkan pada Persamaan (1).

$$\text{Precision} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Positive}} \quad (1)$$

- **Recall**

Recall menunjukkan kemampuan model dalam menemukan seluruh objek relevan dalam citra. Nilai recall yang tinggi menandakan bahwa sedikit objek yang terlewatkan. Rumus recall ditunjukkan pada Persamaan (2).

$$\text{Recall} = \frac{\text{True Positive}}{\text{True Positive} + \text{False Negative}} \quad (2)$$

- **mean Average Precision (mAP)**

mAP merupakan metrik utama untuk menilai akurasi deteksi objek berdasarkan rata-rata dari *average precision* di setiap kelas. Penelitian ini menggunakan dua jenis mAP sebagai acuan:

- **mAP@50:** Rata-rata presisi pada ambang batas Intersection over Union (IoU) sebesar 0.5.
- **mAP@50–95:** Rata-rata presisi dari beberapa ambang IoU (dari 0.5 hingga 0.95 dengan interval 0.05), memberikan evaluasi yang lebih ketat dan menyeluruh.

Seluruh metrik ini diekstraksi dari hasil evaluasi model menggunakan modul bawaan dari framework Ultralytics YOLOv11. Di samping itu, digunakan pula hasil visualisasi seperti grafik *training loss*, *confusion matrix*, serta jumlah dan keakuratan *bounding box* yang berhasil dideteksi oleh model.

Analisis dilakukan secara menyeluruh terhadap kedua model untuk mengidentifikasi sejauh mana keterbatasan anotasi memengaruhi kemampuan model dalam mengenali objek. Hal ini mencakup identifikasi pola penurunan performa pada metrik evaluasi, serta pengamatan terhadap kelas objek yang paling terdampak akibat kurangnya anotasi. Dengan pendekatan ini, penelitian diharapkan dapat memberikan gambaran yang jelas mengenai sensitivitas YOLOv11 terhadap kualitas anotasi dalam pengembangan sistem deteksi objek di dunia nyata.

#### IV. HASIL DAN PEMBAHASAN

Bab ini menyajikan hasil pelatihan dan pengujian dua model YOLOv11 yang dikembangkan dengan strategi anotasi yang berbeda, yaitu anotasi lengkap dan anotasi terbatas. Fokus utama dari bab ini adalah mengevaluasi sejauh mana keterbatasan dalam anotasi objek memengaruhi performa model, baik secara kuantitatif melalui metrik evaluasi seperti *precision*, *recall*, *mAP@0.5*, dan *mAP@0.5:0.95*, maupun secara kualitatif melalui analisis hasil deteksi. Pembahasan mencakup identifikasi tingkat penurunan performa, analisis kelas objek yang paling terdampak, serta interpretasi faktor-faktor visual yang menyebabkan penurunan tersebut dalam konteks kondisi data dunia nyata.

##### A. Hasil Pelatihan Model

Proses pelatihan dilakukan terhadap dua model YOLOv11 dengan pendekatan anotasi berbeda, yakni anotasi lengkap dan anotasi terbatas. Masing-masing model dilatih selama 30 *epoch* di lingkungan Google Colab dengan dukungan GPU **Tesla T4**. Tujuan dari pelatihan ini adalah untuk mengevaluasi dampak keterbatasan anotasi terhadap dinamika pelatihan model, baik dari sisi *loss function* maupun perkembangan metrik selama *epoch*.

Model YOLOv11-Lengkap menyelesaikan pelatihan dalam waktu sekitar 22 menit, sementara model YOLOv11-Terbatas memerlukan waktu sekitar 20 menit. Perbedaan durasi ini disebabkan oleh kompleksitas jumlah objek yang dilihat dalam proses pelatihan, di mana model dengan anotasi lengkap harus memproses lebih banyak informasi per citra.

Visualisasi dari proses pelatihan untuk kedua model ditampilkan pada Fig. 4 dan Fig. 5.

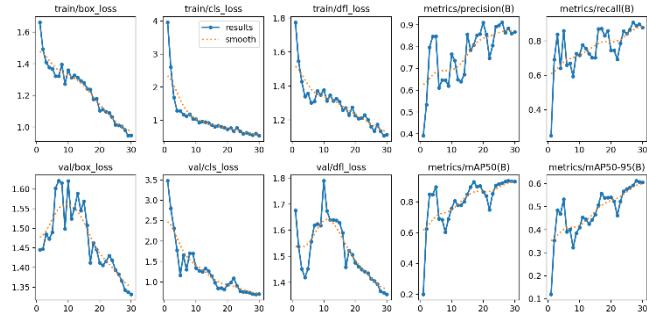


Fig. 4. Hasil Pelatihan Model YOLOv11-Lengkap

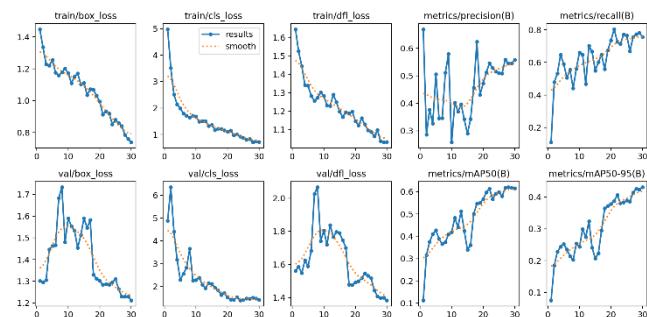


Fig. 5. Hasil Pelatihan Model YOLOv11-Terbatas

Kedua visualisasi menunjukkan adanya tren penurunan *training loss* dan peningkatan metrik seperti *precision*, *recall*, dan *mAP*. Namun, pada model dengan anotasi terbatas, peningkatan metrik terlihat cenderung fluktuatif dan kurang stabil dibandingkan model dengan anotasi lengkap. Hal ini mengindikasikan bahwa keterbatasan jumlah anotasi yang diberikan pada model dapat menyebabkan ketidakstabilan dalam proses pembelajaran.

Nilai akhir dari *loss function* pada data validasi untuk masing-masing model ditunjukkan pada Table II.

TABLE II. NILAI AKHIR LOSS PADA DATA VALIDASI

Jenis Loss	YOLOv11-Lengkap	YOLOv11-Terbatas
val/box_loss	1.33196	1.21227
val/cls_loss	0.70095	1.42337
val/dfl_loss	1.3547	1.38348

Nilai *val/cls\_loss* pada model dengan anotasi terbatas tercatat lebih tinggi dibandingkan model lengkap, menunjukkan bahwa model kesulitan dalam melakukan klasifikasi objek secara akurat. Meskipun nilai *val/box\_loss* lebih rendah, hal ini kemungkinan disebabkan oleh jumlah bounding box yang lebih sedikit selama pelatihan, bukan karena peningkatan akurasi lokasi deteksi. Demikian pula, nilai *val/dfl\_loss* yang relatif mirip tidak serta-merta menunjukkan kesetaraan performa, mengingat metrik ini hanya mengukur distribusi prediksi lokasi, bukan klasifikasi.

Temuan ini memperkuat dugaan bahwa keterbatasan anotasi berdampak langsung terhadap efektivitas proses pembelajaran model, terutama dalam memahami representasi visual yang kompleks dari berbagai kelas objek.

##### B. Hasil Pengujian Model

Setelah proses pelatihan selesai, kedua model *YOLOv11-Lengkap* dan *YOLOv11-Terbatas* diuji menggunakan data uji yang terdiri dari 180 gambar dengan distribusi objek yang merata. Evaluasi performa dilakukan berdasarkan metrik *precision*, *recall*, *mean Average Precision* pada *Intersection over Union* (IoU) 0.5 (*mAP@0.5*), dan rentang IoU 0.5–0.95 (*mAP@0.5:0.95*). Hasil pengujian model dapat dilihat pada Table III.

TABLE III. PERBANDINGAN METRIK EVALUASI MODEL YOLOv11

Model	Precision	Recall	<i>mAP@0.5</i>	<i>mAP@0.5:0.95</i>
YOLOv11-Lengkap	0.86863	0.87763	0.93058	0.60492
YOLOv11-Terbatas	0.55847	0.75712	0.61573	0.43064

Secara kuantitatif, terlihat adanya penurunan performa pada seluruh metrik ketika model dilatih dengan data yang dianotasi secara terbatas. Penurunan ini mencerminkan sejauh mana keterbatasan anotasi dapat memengaruhi kinerja model dalam mendekripsi dan mengklasifikasikan objek secara akurat.

Untuk mendukung pemahaman visual terhadap hasil tersebut, Fig. 6 dan Fig. 7 menyajikan hasil deteksi dari masing-masing model pada citra uji yang sama. Model dengan anotasi lengkap cenderung menghasilkan deteksi dengan jumlah bounding box yang lebih mendekati *ground truth*, serta *confidence score* yang stabil. Sebaliknya, model dengan

anotasi terbatas menunjukkan kecenderungan menghilangkan objek kecil, memunculkan prediksi *false positive*, atau tidak memberikan deteksi pada objek yang samar.



Fig. 6. Hasil Deteksi Objek oleh Model YOLOv11-Lengkap



Fig. 7. Hasil Deteksi Objek oleh Model YOLOv11-Terbatas

Untuk mengevaluasi akurasi klasifikasi pada masing-masing kelas, digunakan *confusion matrix* yang ter-normalisasi dan disajikan pada Fig. 8 dan Fig. 9. Melalui visualisasi ini, tampak bahwa beberapa kelas buah mengalami penurunan akurasi klasifikasi secara signifikan akibat keterbatasan anotasi.

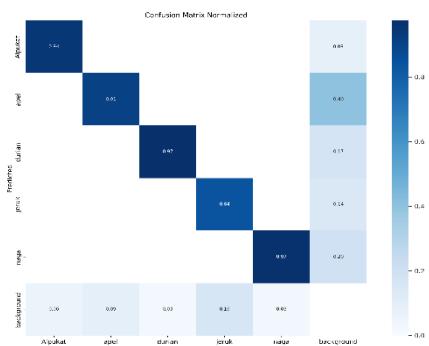


Fig. 8. Confusion Matrix Model YOLOv11-Lengkap

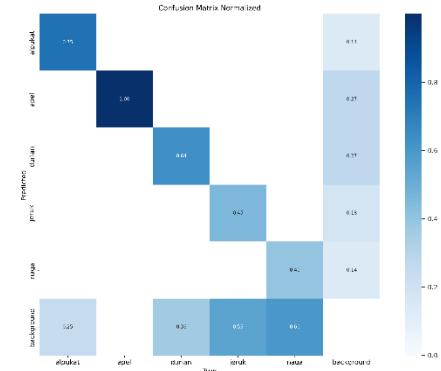


Fig. 9. Confusion Matrix Model YOLOv11-Terbatas

Berdasarkan hasil tersebut, dapat disimpulkan bahwa keterbatasan anotasi memiliki dampak nyata terhadap kemampuan model dalam mengenali objek secara konsisten di seluruh kelas. Temuan ini memberikan dasar yang kuat untuk analisis lebih lanjut mengenai karakteristik penurunan performa yang dijabarkan pada subbab berikutnya.

### C. Analisis Perbandingan Performa Model

Analisis perbandingan dilakukan untuk mengevaluasi dampak dari keterbatasan anotasi terhadap performa model *YOLOv11*. Fokus utama dalam analisis ini bukanlah menentukan model mana yang lebih unggul, melainkan seberapa besar penurunan kinerja yang dialami oleh model dengan anotasi terbatas dan mengapa penurunan tersebut dapat terjadi.

Hasil evaluasi kuantitatif menunjukkan adanya penurunan performa signifikan pada semua metrik evaluasi utama. Table IV berikut menunjukkan selisih performa antara model *YOLOv11-Lengkap* dan *YOLOv11-Terbatas*:

TABLE IV. PERBANDINGAN METRIK EVALUASI MODEL YOLOv11-LENGKAP DAN YOLOv11-TERBATAS

Metrik	YOLOv11-Lengkap	YOLOv11-Terbatas	Selisih Performa
Precision	0.86863	0.55847	-31.02%
Recall	0.87763	0.75712	-13.72%
mAP@0.5	0.93058	0.61573	-33.84%
mAP@0.5:0.95	0.60492	0.43064	-28.74%

Penurunan *precision* sebesar 31,02% menandakan bahwa model dengan anotasi terbatas menghasilkan lebih banyak kesalahan klasifikasi positif palsu (*false positive*). Sementara itu, penurunan *mAP@0.5* sebesar 33,84% menunjukkan turunnya kemampuan model dalam mendekripsi dan mengklasifikasikan objek secara akurat pada tingkat *IoU* dasar. Hal ini diperkuat oleh penurunan *mAP@0.5:0.95* sebesar 28,74%, yang merepresentasikan menurunnya performa pada berbagai tingkat ketelitian deteksi.

Lebih lanjut, analisis berdasarkan *confusion matrix* (Fig. 8 dan Fig. 9) menunjukkan bahwa beberapa kelas buah mengalami penurunan akurasi klasifikasi yang sangat signifikan pada model dengan anotasi terbatas. Kelas buah *naga* hanya terdeteksi dengan akurasi 0.40 pada model terbatas, jauh menurun dari 0.97 pada model lengkap. Demikian pula, kelas *jeruk* menurun dari 0.84 menjadi hanya 0.45. Sementara itu, *durian* menurun dari 0.97 menjadi 0.64, dan *alpukat* dari 0.94 menjadi 0.75. Penurunan ini menunjukkan bahwa model dengan pelatihan terbatas

kehilangan kemampuan untuk mengenali objek dengan karakteristik tertentu.

Faktor penyebab penurunan performa ini kemungkinan besar berkaitan dengan:

- *Warna objek*, seperti *jeruk* dan *naga* yang memiliki warna mencolok namun dapat mirip dengan latar jika hanya sebagian citra yang dianotasi.
- *Tekstur dan bentuk*, seperti *durian* dan *alpukat* yang memiliki struktur unik dan dapat membingungkan model jika konteksnya tidak lengkap.
- *Distribusi objek*, di mana model hanya melihat sebagian kecil dari kemungkinan posisi dan orientasi objek.

Selain itu, model terbatas juga menunjukkan peningkatan prediksi terhadap *background*, dengan nilai *false background* mencapai 0.55 pada *jeruk* dan 0.60 pada *naga*. Hal ini mengindikasikan bahwa model sering kali mengabaikan objek yang sebenarnya relevan karena tidak pernah dilatih untuk mengenali secara menyeluruh.

Dengan demikian, hasil analisis ini memberikan wawasan bahwa penggunaan anotasi terbatas memang dapat menurunkan performa model secara substansial, terutama pada kelas-kelas dengan kompleksitas visual tinggi. Temuan ini penting untuk menjadi pertimbangan dalam pengembangan sistem deteksi objek yang efisien, terutama dalam konteks keterbatasan sumber daya pelabelan di dunia nyata.

## V. KESIMPULAN DAN SARAN

### A. Kesimpulan

Penelitian ini bertujuan untuk mengevaluasi dampak dari keterbatasan anotasi terhadap kinerja model deteksi objek berbasis YOLOv11 dalam konteks deteksi multi-objek buah. Dua model dilatih menggunakan pendekatan berbeda: model pertama dengan *anotasi lengkap* dan model kedua dengan *anotasi terbatas* (sekitar 40–50% objek dalam citra tidak diberi label).

Hasil evaluasi menunjukkan bahwa model dengan *anotasi terbatas* mengalami penurunan performa yang signifikan dibandingkan model dengan *anotasi lengkap*, terutama pada metrik *precision*, *mAP@0.5*, dan *mAP@0.5:0.95*, dengan selisih penurunan mencapai lebih dari 30%. Selain itu, *confusion matrix* mengungkap bahwa kelas-kelas seperti *jeruk* dan *naga* mengalami penurunan akurasi klasifikasi yang paling signifikan. Penyebab penurunan ini diidentifikasi berasal dari kompleksitas visual objek, seperti warna, bentuk, dan kemiripan dengan latar belakang. Model dengan *anotasi terbatas* juga cenderung mengklasifikasikan objek sebagai *background*, menunjukkan kehilangan konteks akibat kekurangan label saat pelatihan.

Dengan demikian, studi ini menyimpulkan bahwa meskipun model masih dapat berfungsi dengan data yang tidak sepenuhnya dianotasi, performanya menurun secara substansial, terutama pada objek dengan visual kompleks. Temuan ini memberikan gambaran nyata tentang pentingnya kelengkapan anotasi dalam pengembangan sistem deteksi objek yang andal dan akurat.

### B. Saran

Berdasarkan hasil penelitian yang telah dilakukan, terdapat beberapa saran yang dapat dijadikan pertimbangan untuk penelitian selanjutnya maupun penerapan model dalam skenario nyata:

#### 1. Eksplorasi Teknik *Fine-tuning* dan *Augmentasi Lanjutan*

Penelitian ini tidak mengeksplorasi teknik *fine-tuning* lanjutan atau *augmentasi* tingkat tinggi. Penelitian berikutnya dapat mengkaji pengaruh dari penyesuaian *layer*, penggunaan *optimizer* yang berbeda, atau *data synthesis* terhadap performa model.

#### 2. Evaluasi pada Kondisi Nyata (*Real-world Deployment*)

Untuk validasi lebih lanjut, performa model sebaiknya diuji pada gambar yang diperoleh langsung dari kamera atau video secara *real-time*, bukan hanya dataset statis, agar hasil lebih aplikatif.

#### 3. Pertimbangan Efisiensi Komputasi melalui *Active Learning*

Meskipun anotasi lengkap memberikan hasil terbaik, proses ini membutuhkan waktu pelabelan yang lebih lama. Dalam kondisi sumber daya terbatas, strategi anotasi semi-otomatis atau *active learning* dapat dipertimbangkan. *Active learning* adalah metode pembelajaran di mana model secara proaktif memilih sampel data yang paling tidak pasti untuk dianotasi terlebih dahulu, sehingga efisiensi pelabelan meningkat tanpa mengorbankan kualitas model secara signifikan.

## REFERENCES

- [1] D. H. Saputra, B. Imran, and Juhartini, “Object Detection Untuk Mendeteksi Citra Buah-Buahan Menggunakan Metode Yolo,” *Jurnal Kecerdasan Buatan dan Teknologi Informasi*, vol. 2, no. 2, pp. 70–80, 2023, doi: 10.69916/jkbt.v2i2.18.
- [2] T. Ye *et al.*, “YOLO-FIX: Improved YOLOv11 with Attention and Multi-Scale Feature Fusion for Detecting Glue Line Defects on Mobile Phone Frames,” *Electronics (Switzerland)*, vol. 14, no. 5, 2025, doi: 10.3390/electronics14050927.
- [3] Y. Tang *et al.*, “Citrus fruit detection based on an improved YOLOv5 under natural orchard conditions,” *International Journal of Agricultural and Biological Engineering*, vol. 18, no. 3, pp. 176–185, 2025, doi: 10.25165/j.ijabe.20251803.8935.
- [4] M. Salsa Nabila, H. Herlawati, and A. Hidayat, “Pendeteksian dan Klasifikasi Sampah pada Bank Sampah Berbasis Web Menggunakan YOLOv11,” *Journal of Students’ Research in Computer Science*, vol. 6, no. 1, p. 2025, 2025, [Online]. Available: <https://doi.org/10.31599/r5me0z35>
- [5] L. Susanti, N. K. Daulay, and B. Intan, “Sistem Absensi Mahasiswa Berbasis Pengenalan Wajah Menggunakan Algoritma YOLOv5,” *JURIKOM (Jurnal Riset Komputer)*, vol. 10, no. 2, p. 640, 2023, doi: 10.30865/jurikom.v10i2.6032.

- [6] R. Agil, Y. Arjun, and E. P. Silmina, “Deteksi Bahan Pangan Tinggi Protein Menggunakan Model You Only Look Once (YOLO),” *Technology and Science (BITS)*, vol. 6, no. 4, 2025, doi: 10.47065/bits.v6i4.6889.
- [7] K. Ding and P. Jiang, “RFID-based production data analysis in an IoT-enabled smart job-shop,” *IEEE/CAA Journal of Automatica Sinica*, vol. 5, no. 1, pp. 128–138, Jan. 2018, doi: 10.1109/JAS.2017.7510418.
- [8] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, “YOLOv4: Optimal Speed and Accuracy of Object Detection,” *International Conference on Learning Representations*, Apr. 2020, [Online]. Available: <http://arxiv.org/abs/2004.10934>
- [9] A. Kuznetsova, T. Maleva, and V. Soloviev, “YOLOv5 versus YOLOv3 for Apple Detection,” in *Cyber-Physical Systems: Modelling and Intelligent Control*, 2021, pp. 349–358. doi: 10.1007/978-3-030-66077-2\_28.
- [10] F. Agustina and M. Sukron, “Deteksi Kematangan Buah Pepaya Menggunakan Algoritma YOLO Berbasis Android,” *JURNAL INFOKAM*, vol. XVIII, no. 2, Sep. 2022.
- [11] A. Ardiansyah, J. Triloka, K. kunci-Pengolahan Citra, D. Kesegaran Buah, and K. Akurasi, “Evaluasi Kinerja Model YOLOv8 dalam Deteksi Kesegaran Buah,” *Jurnal Penelitian Ilmu dan Teknologi Komputer*.
- [12] R. Khanam and M. Hussain, “YOLOv11: An Overview of the Key Architectural Enhancements,” Oct. 2024, [Online]. Available: <http://arxiv.org/abs/2410.17725>
- [13] R. Kurniawan, A. T. Martadinata, and S. D. Cahyo, “Klasifikasi Tingkat Kematangan Buah Sawit Berbasis Deep Learning dengan Menggunakan Arsitektur Yolov5,” *Journal of Information System Research (JOSH)*, vol. 5, no. 1, pp. 302–309, Oct. 2023, doi: 10.47065/josh.v5i1.4408.
- [14] L. Kurnia Ramadhani and B. Nurul Widyaningrum, “Implementation of YOLO v11 for Image-Based Litter Detection and Classification in Environmental Management Efforts,” 2025. [Online]. Available: <http://jurnal.polibatam.ac.id/index.php/JAIC>
- [15] S. Aras, P. Tanra, and M. Bazhar, “Deteksi Tingkat Kematangan Buah Tomat Menggunakan YOLOv5,” *MALCOM: Indonesian Journal of Machine Learning and Computer Science*, vol. 4, no. 2, pp. 623–628, Mar. 2024, doi: 10.57152/malcom.v4i2.1270.
- [16] A. Faraduddin Attar, E. Efendi, and P. Paradana, “Implementasi Deteksi Bendera Negara ASEAN dengan Metode Anotasi Data dan YoloV8,” *PROSIDING SEMINAR NASIONAL TEKNOLOGI DAN SAINS TAHUN 2025*, vol. 4, 2025.