

TESIS
ANALISIS KOMPARATIF EMPIRIS LIMA ARSITEKTUR
DEEP LEARNING UNTUK DETEKSI DRONE:
SUPERIORITAS YOLOV8 DAN TRADE OFF AKURASI
EFISIENSI



Disusun oleh:

NUGRAHA ASTHRA MEGANTARA

23.51.1427

Konsentrasi : Digital Transformation Intelligence

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA

2026

TESIS
ANALISIS KOMPARATIF EMPIRIS LIMA ARSITEKTUR
DEEP LEARNING UNTUK DETEKSI DRONE:
SUPERIORITAS YOLOV8 DAN TRADE OFF AKURASI
EFISIENSI

EMPIRICAL COMPARATIVE ANALYSIS OF FIVE DEEP
LEARNING ARCHITECTURES FOR DRONE DETECTION:
THE SUPERIORITY OF YOLOV8 AND THE ACCURACY
EFFICIENCY TRADE OFF

Diajukan untuk memenuhi salah satu syarat mencapai derajat Pascasarjana

Program Studi S2 Informatika



Disusun oleh:

NUGRAHA ASTHRA MEGANTARA

23.51.1427

Konsentrasi : Digital Transformation Intelligence

FAKULTAS ILMU KOMPUTER
UNIVERSITAS AMIKOM YOGYAKARTA
YOGYAKARTA

2026

HALAMAN PERSETUJUAN

**ANALISIS KOMPARATIF EMPIRIS LIMA ARSITEKTUR DEEP
LEARNING UNTUK DETEKSI DRONE: SUPERIORITAS YOLOV8 DAN
TRADE OFF AKURASI EFISIENSI**

**EMPIRICAL COMPARATIVE ANALYSIS OF FIVE DEEP LEARNING
ARCHITECTURES FOR DRONE DETECTION: THE SUPERIORITY OF
YOLOV8 AND THE ACCURACY EFFICIENCY TRADE OFF**

yang disusun dan diajukan oleh

Nugraha Asthra Megantara

23.51.1427

telah disetujui oleh Dosen Pembimbing Tesis
pada tanggal 2 Januari 2026

Dosen Pembimbing,

Prof. Dr. Ema Utami, S.Si., M.Kom.
NIK. 190302037

HALAMAN PENGESAHAN
**ANALISIS KOMPARATIF EMPIRIS LIMA ARSITEKTUR DEEP
LEARNING UNTUK DETEKSI DRONE: SUPERIORITAS YOLOV8 DAN
TRADE OFF AKURASI EFISIENSI**

**EMPIRICAL COMPARATIVE ANALYSIS OF FIVE DEEP LEARNING
ARCHITECTURES FOR DRONE DETECTION: THE SUPERIORITY OF
YOLOV8 AND THE ACCURACY EFFICIENCY TRADE OFF**

yang disusun dan diajukan oleh

Nugraha Asthra Megantara

NIM. 23.51.1427

telah dipertahankan di depan Dewan Penguji
pada tanggal, 2 Januari 2026.

Susunan Dewan Penguji

Nama Penguji

Tanda Tangan

Tommy Hidayat, S.Kom., M.T., Ph.D.

NIK. 190302182

Dhani Ariatmanto, S.Kom., M.Kom., Ph.D.

NIK. 190302197

Prof. Dr. Ema Utami, S.Si., M.Kom.

NIK. 190302037

Tesis ini telah diterima sebagai salah satu persyaratan
untuk memperoleh gelar Magister Komputer

Tanggal 2 Januari 2026

DEKAN FAKULTAS ILMU KOMPUTER



Prof. Dr. Kusriani, M.Kom.

NIK. 190302106

HALAMAN PERNYATAAN KEASLIAN TESIS

Yang bertandatangan di bawah ini,

Nama mahasiswa : Nugraha Asthra Megantara
NIM : 23.51.1427
Konsentrasi : Digital Transformation Intelligence

Menyatakan bahwa Tesis dengan judul berikut:
Analisis Komparatif Empiris Lima Arsitektur Deep Learning Untuk Deteksi Drone: Superioritas Yolov8 dan Trade Off Akurasi Efisiensi

Dosen Pembimbing Utama : **Prof. Dr. Ema Utami, S.Si., M.Kom.**

1. Karya tulis ini adalah benar-benar **ASLI** dan **BELUM PERNAH** diajukan untuk mendapatkan gelar akademik, baik di Universitas AMIKOM Yogyakarta maupun di Perguruan Tinggi lainnya
2. Karya tulis ini merupakan gagasan, rumusan dan penelitian **SAYA** sendiri, tanpa bantuan pihak lain kecuali arahan dari Tim Dosen Pembimbing
3. Dalam karya tulis ini tidak terdapat karya atau pendapat orang lain, kecuali secara tertulis dengan jelas dicantumkan sebagai acuan dalam naskah dengan disebutkan nama pengarang dan disebutkan dalam Daftar Pustaka pada karya tulis ini
4. Perangkat lunak yang digunakan dalam penelitian ini sepenuhnya menjadi tanggung jawab **SAYA**, bukan tanggung jawab Universitas AMIKOM Yogyakarta
5. Pernyataan ini **SAYA** buat dengan **sesungguhnya**, apabila di kemudian hari terdapat penyimpangan dan ketidakbenaran dalam pernyataan ini, maka **SAYA** bersedia menerima **SANKSI AKADEMIK** dengan pencabutan gelar yang sudah diperoleh, serta sanksi lainnya sesuai dengan norma yang berlaku di Perguruan Tinggi

Yogyakarta, 2 Januari 2026
Yang Menyatakan,



Nugraha Asthra Megantara

HALAMAN PERSEMBAHAN

Dengan segala kerendahan hati dan rasa syukur,
karya tesis ini saya persembahkan kepada:

Orang Tua, Istri dan Anak tercinta,

atas cinta, doa, semangat dan pengorbanan yang menjadi fondasi langkah hidup dan pendidikan saya.

Dosen pembimbing dan para pendidik,

yang telah membimbing, menginspirasi, dan memberikan ilmu serta nilai-nilai luhur selama proses studi ini.

Sahabat dan rekan seperjuangan MTI A 2023,

atas semangat, dukungan, dan kebersamaan yang tak ternilai selama perjalanan akademik ini.

Almamater tercinta,

sebagai bentuk pengabdian dan kontribusi kecil dalam khazanah keilmuan yang terus berkembang.



HALAMAN MOTTO

"Ilmu tanpa agama adalah cacat, dan agama tanpa ilmu adalah buta." Al-Khawarizmi, Bapak Ilmu Aljabar dan Pelopor Algoritma

"Teknologi yang hakiki adalah yang digunakan untuk kemaslahatan, bukan sekadar kemajuan." - Penulis



KATA PENGANTAR

Puji dan syukur penulis panjatkan ke hadirat Allah SWT atas segala limpahan rahmat, taufik, dan hidayah-Nya sehingga penulis dapat menyelesaikan tesis ini yang berjudul: Analisis Komparatif Empiris Lima Arsitektur Deep Learning Untuk Deteksi Drone: Superioritas Yolov8 Dan Trade Off Akurasi Efisiensi. Tesis ini disusun sebagai salah satu syarat untuk memperoleh gelar Magister pada Program Studi S2 Teknik Informatika, Program Pascasarjana Universitas Amikom Yogyakarta.

Penyusunan tesis ini tidak terlepas dari bantuan dan dukungan berbagai pihak. Oleh karena itu, dengan segala hormat dan rasa terima kasih yang mendalam, penulis menyampaikan apresiasi kepada:

1. Bapak Prof. Dr. M. Suyanto, M.M., selaku Rektor Universitas Amikom Yogyakarta, atas segala fasilitas dan kesempatan yang diberikan kepada penulis;
2. Ibu Prof. Dr. Kusriani, M.Kom., selaku Dekan Fakultas Ilmu Komputer, yang telah memberikan dukungan selama proses studi;
3. Ibu Prof. Dr. Ema Utami, S.Si., M.Kom., selaku Dosen Pembimbing Utama, atas bimbingan, arahan, serta kesabaran yang tiada henti dalam proses penyusunan tesis ini;
4. Bapak Tonny Hidayat, S.Kom., M.T., Ph.D. dan Bapak Dhani Ariatmanto, S.Kom., M.Kom., Ph.D. selaku Tim Dosen Penguji, yang telah memberikan masukan berharga dalam proses seminar, ujian tesis, dan perbaikan substansi ilmiah;
5. Seluruh Dosen dan Staf Akademik Program Pascasarjana MTI, yang telah memberikan bekal ilmu dan pengalaman selama masa studi;
6. Orang tua, Istri dan Anak tercinta, atas kasih sayang, doa, semangat, dan dukungan moril maupun materil yang tiada henti;
7. Rekan-rekan seperjuangan di MTI A 2023, atas kebersamaan, kerja sama, dan saling mendukung selama proses perkuliahan dan penelitian;
8. Semua pihak yang telah membantu secara langsung maupun tidak langsung dalam penyelesaian tesis ini, yang tidak dapat penulis sebutkan satu per satu.

Penulis menyadari bahwa tesis ini masih memiliki keterbatasan. Oleh karena itu, penulis terbuka terhadap segala bentuk saran dan kritik yang membangun demi penyempurnaan karya ini. Akhir kata, semoga tesis ini dapat memberikan manfaat, baik secara akademis maupun praktis, bagi semua pihak yang membacanya.

Yogyakarta, 2 Januari 2026

Penulis,

DAFTAR ISI

HALAMAN JUDUL	ii
HALAMAN PENGESAHAN	iii
HALAMAN PERSETUJUAN	iv
HALAMAN PERNYATAAN KEASLIAN TESIS	v
HALAMAN PERSEMBAHAN	vi
HALAMAN MOTTO	vii
KATA PENGANTAR	viii
DAFTAR ISI	ix
DAFTAR TABEL	xi
DAFTAR GAMBAR	xii
INTISARI	xiii
<i>ABSTRACT</i>	xiv
BAB I PENDAHULUAN	1
1.1. Latar Belakang	1
1.2. Rumusan Masalah	8
1.3. Batasan Masalah	8
1.4. Tujuan Penelitian	9
1.5. Manfaat Penelitian	10
BAB II TINJAUAN PUSTAKA	12
2.1. Tinjauan Pustaka	12
2.2. Keaslian Penelitian	17

2.3. Landasan Teori	28
BAB III METODE PENELITIAN	44
3.1. Jenis, Sifat, dan Pendekatan Penelitian	44
3.2. Metode Pengumpulan Data	45
3.3. Metode Analisis Data	46
3.4. Alur Penelitian	48
BAB IV HASIL PENELITIAN DAN PEMBAHASAN	50
4.1. Analisis Komparatif Kuantitatif dan Statistik Lima Arsitektur	50
4.2. Analisis Efisiensi Komputasional dan Trade-off	55
4.3. Analisis Signifikansi Statistik dan Reliabilitas Performa	59
4.4. Analisis Signifikansi Statistik Komprehensif Integrasi ANOVA dan Pairwise T-test	63
4.5. Ranking Komprehensif dan Klasifikasi Model Berdasarkan Analisis Arsitektural	66
4.6. Analisis Visual Deteksi Drone	70
4.7. Pembahasan Temuan dan Implikasi Ilmiah	78
BAB V PENUTUP	82
5.1. Kesimpulan	82
5.2. Saran	84
DAFTAR PUSTAKA	88

DAFTAR TABEL

Tabel 2.1. Matriks Literatur Review Dan Posisi Penelitian Analisis Komparatif Empiris Lima Arsitektur Deep Learning Untuk Deteksi Drone: Superioritas Yolov8 Dan Trade Off Akurasi Efisiensi	17
Tabel 2.2. Klasifikasi Matriks Evaluasi Untuk Object Detection	39
Tabel 4.1. Performa Komparatif Lima Model Deep Learning Deteksi Drone	50
Tabel 4.2. Analisis Efisiensi Komputasional dan Resource Requirements	56
Tabel 4.3.1. Analisis Statistik Komprehensif	59
Tabel 4.3.2. Interpretasi Effect Size dan Implikasi Praktis	60
Tabel 4.4.1. Hasil Analisis ANOVA Untuk Semua Metrik Evaluasi	63
Tabel 4.4.2. Analisis Effect Size (Cohen's d) Lima Model Deep Learning	65
Tabel 4.5. Matriks Kinerja Komparatif Model Deep Learning	66

DAFTAR GAMBAR

Gambar 2.1. Artificial Intelligence, Machine Learning, Dan Deep Learning	29
Gambar 2.2. Konsep Metode CNN	30
Gambar 2.3. (a) Pemetaan di dalam blok Residual, (b) Pemetaan langsung sederhana	31
Gambar 2.4. Arsitektur VGG-16	32
Gambar 2.5. Arsitektur EfficientNet	34
Gambar 2.6. Arsitektur MobileNet	35
Gambar 2.7. Arsitektur YOLO (You Only Looks Once)	36
Gambar 2.8. Arsitektur YOLOv8	37
Gambar 3.1. Pendekatan Penelitian	45
Gambar 3.2. Alur Penelitian	49
Gambar 4.1. Confidence Intervals Models	54
Gambar 4.2. Trade-off Analysis: Akurasi vs Efisiensi	58
Gambar 4.3. Effect Size Analysis	61
Gambar 4.4. Matriks Signifikansi Pairwise T-test (p-values)	64
Gambar 4.5. Radar Chart Perbandingan Kinerja Multi-Dimensi	68
Gambar 4.6. Perbandingan Deteksi Drone	70
Gambar 4.7. Distribusi <i>Confidence</i>	72
Gambar 4.8. Dasbor Kinerja Analisis Komprehensif	74
Gambar 4.9. Visualisasi Empat Data Uji Hasil Deteksi	76

INTISARI

Penelitian ini bertujuan melakukan analisis komparatif empiris yang komprehensif terhadap performa lima arsitektur *deep learning* untuk deteksi objek drone, yaitu VGG16, ResNet50, MobileNetV2, EfficientNetB0, dan YOLOv8-nano. Fokus penelitian adalah mengukur dan membandingkan akurasi deteksi, efisiensi komputasional, serta signifikansi statistik perbedaan performa masing-masing model dalam konteks tugas deteksi *single class*. Metode penelitian berupa eksperimen kuantitatif dengan pendekatan *benchmarking* sistematis, menggunakan dataset 1.359 citra drone, evaluasi multi-dimensi (metrik regresi, deteksi, dan efisiensi), serta analisis statistik robust (*ANOVA, t-test, effect size*).

Hasil penelitian menunjukkan superioritas signifikan dari arsitektur YOLOv8-nano yang secara konsisten unggul dalam semua metrik akurasi inti, seperti *F1-Score* (0.922), *mean Average Precision* (mAP50: 0.901), dan *Mean Squared Error* (0.007), dengan kecepatan inferensi 112 FPS. Analisis statistik ($p\text{-value} < 0.001$, $effect\ size\ \eta^2 > 0.99$) mengkonfirmasi bahwa perbedaan performa antar model bersifat sangat signifikan dan bermakna secara praktis. Di sisi lain, MobileNetV2 tercatat sebagai model paling efisien dengan waktu pelatihan tercepat, sementara EfficientNetB0 menawarkan keseimbangan optimal antara akurasi dan efisiensi. Temuan juga mengungkap ketidakcocokan arsitektural pada model yang diadaptasi dari tugas klasifikasi (ResNet50, VGG16) untuk tugas deteksi drone.

Kesimpulan penelitian menegaskan bahwa YOLOv8-nano merupakan arsitektur paling optimal untuk deteksi drone, menawarkan kombinasi akurasi tinggi, kecepatan inferensi *real time*, dan ukuran model ringan. Penelitian ini memberikan bukti empiris dan panduan berbasis *trade off* yang terkuantifikasi untuk seleksi model dalam berbagai skenario aplikasi praktis, mulai dari sistem keamanan kritis hingga penerapan di perangkat *edge* dengan sumber daya terbatas.

Kata kunci: deteksi drone, *deep learning*, analisis komparatif, YOLOv8, efisiensi komputasi

ABSTRACT

This research aims to conduct a comprehensive empirical comparative analysis of the performance of five deep learning architectures for drone object detection: VGG16, ResNet50, MobileNetV2, EfficientNetB0, and YOLOv8-nano. The study focuses on measuring and comparing detection accuracy, computational efficiency, and the statistical significance of performance differences among the models in a single-class detection task. The research method is a quantitative experiment using a systematic benchmarking approach, employing a dataset of 1,359 drone images, multi-dimensional evaluation (regression, detection, and efficiency metrics), and robust statistical analysis (ANOVA, t-test, effect size).

The results demonstrate the significant superiority of the YOLOv8-nano architecture, which consistently excelled in all key accuracy metrics, such as F1-Score (0.922), mean Average Precision (mAP50: 0.901), and Mean Squared Error (0.007), while achieving an inference speed of 112 FPS. Statistical analysis (p-value < 0.001, effect size $\eta^2 > 0.99$) confirms that the performance differences between models are highly significant and practically meaningful. Conversely, MobileNetV2 was recorded as the most efficient model with the fastest training time, while EfficientNetB0 offered an optimal balance between accuracy and efficiency. The findings also reveal the architectural mismatch of models adapted from classification tasks (ResNet50, VGG16) for the drone detection task.

The study concludes that YOLOv8-nano is the most optimal architecture for drone detection, offering a combination of high accuracy, real-time inference speed, and a lightweight model size. This research provides empirical evidence and a quantified trade-off-based guide for model selection in various practical application scenarios, from critical security systems to deployment on resource-constrained edge devices.

Keywords: drone detection, deep learning, comparative analysis, YOLOv8, computational efficiency

BAB I

PENDAHULUAN

1.1. Latar Belakang

Drone, atau yang dikenal sebagai *Unmanned Aerial Vehicles (UAVs)* telah menjadi salah satu teknologi yang paling signifikan dan revolusioner dalam beberapa tahun terakhir. Penggunaan drone tidak hanya terbatas pada bidang militer, tetapi juga telah meluas ke berbagai industri termasuk survei udara, pemantauan lingkungan, pemetaan tanah, pengiriman barang, dan hiburan. Penggunaan drone telah berkembang pesat dalam aplikasi-aplikasi sipil karena kemampuannya yang fleksibel, biaya yang lebih rendah, serta kemudahan pengoperasian. Namun, di balik manfaat yang sangat membantu, proliferasi drone yang tidak terkendali telah menimbulkan kekhawatiran serius dan urgensi mendesak terkait keamanan nasional, keselamatan publik, dan privasi individu. Urgensi penelitian ini didasarkan pada realitas operasional yang semakin mengkhawatirkan: insiden gangguan drone di bandara-bandara internasional utama, seperti di Gatwick (2018) dan Heathrow (2019) yang menyebabkan pembatalan ratusan penerbangan dan kerugian ekonomi jutaan dolar telah menyoroti kerentanan kritis infrastruktur vital terhadap ancaman drone. Drone juga berpotensi disalahgunakan untuk pengintaian ilegal, penyelundupan barang terlarang, hingga serangan fisik pada instalasi sensitif. Dalam konteks Indonesia, dengan luas wilayah dan banyaknya infrastruktur strategis, ancaman ini menjadi semakin relevan dan mendesak untuk diatasi. Oleh karena itu, pengembangan

sistem deteksi drone yang andal, akurat, dan responsif bukan lagi sekadar kebutuhan akademis, melainkan menjadi sebuah keharusan operasional (*mission critical requirement*) untuk menjamin keamanan wilayah udara. Dibalik manfaat drone yang sangat membantu, tetapi juga menimbulkan kekhawatiran yang valid tentang keselamatan, keamanan, dan privasi. Pemantauan berbasis darat (*Internet of Things*) dapat mencapai banyak hal dari potensi ini tanpa risiko terhadap keselamatan dan risiko yang dirasakan lebih rendah terhadap privasi dan keamanan. Drone dengan ketinggian rendah mungkin akan terbatas pada wilayah geografis yang jelas secara spasial dan altitudinal, sementara drone dengan ketinggian yang lebih tinggi kemungkinan besar akan diterima karena alasan keamanan (Mathews *et al.*, 2023).

Deteksi drone telah menjadi fokus penelitian yang signifikan. Metode-metode dalam deteksi terdapat berbagai macam, mulai dari penggunaan teknologi radar, sensor optik hingga pendekatan menggunakan *machine learning* dan *computer vision*, telah dikembangkan untuk mengenali, melacak dan memonitor aktivitas drone. Pendekatan-pendekatan seperti analisis data citra, deteksi pergerakan, dan integrasi sensor-sensor untuk pengenalan objek telah menjadi fokus dalam upaya deteksi drone yang handal. Pengembangan sistem deteksi drone yang handal dan efektif menjadi sangat penting dalam menjaga keamanan, privasi, dan menghadapi potensi ancaman yang berkaitan dengan penggunaan drone (Agarwal, Kumar and Singh, 2019).

Di antara berbagai pendekatan deteksi (radar, akustik, RF), deteksi berbasis visi komputer (*computer vision*) menawarkan keunggulan signifikan: biaya relatif

rendah, sifatnya yang pasif (tidak memancarkan sinyal sehingga sulit dideteksi), kemudahan pengembangan, dan kemampuannya memberikan informasi kontekstual visual (jenis drone, arah, perilaku). Kemajuan teknologi komputer dan pengolahan data secara besar telah mendorong perkembangan pesat dalam bidang *artificial intelligence* dan *machine learning*, termasuk penggunaan algoritma-algoritma kompleks seperti *neural networks*, *deep learning*, dan *reinforcement learning* telah membawa perubahan signifikan dalam pemrosesan data, pengenalan pola, dan pengambilan keputusan. Dalam konteks deteksi objek, *artificial intelligence* dan *machine learning* telah membuka peluang baru adopsi secara luas dalam deteksi objek pada citra visual, termasuk penggunaan metode *Convolutional Neural Networks (CNN)*, dan *You Only Look Once (YOLO)* dalam deteksi objek telah memungkinkan sistem untuk belajar dan mengidentifikasi objek-objek dengan akurasi yang semakin tinggi (Khan and Al-Habsi, 2020).

Namun, implementasi praktis sistem deteksi drone berbasis *deep learning* menghadapi sejumlah problem utama yang kompleks dan spesifik: Dalam citra surveillans, drone pada jarak menengah hingga jauh sering kali muncul sebagai objek yang sangat kecil, terkadang hanya mencakup puluhan piksel, sehingga menyulitkan ekstraksi fitur yang diskriminatif. Drone dapat terbang di depan berbagai latar seperti langit, awan, puncak gedung, atau dedaunan. Kemiripan tekstur dan warna dengan latar belakang menyebabkan tingginya *false positive* dan *false negative*. Sistem deteksi yang praktis seringkali perlu dijalankan pada perangkat *edge* dengan kendala ketat pada daya komputasi, konsumsi daya, dan kapasitas memori. Dataset drone yang besar, bervariasi (semua kondisi cuaca,

waktu, sudut), dan memiliki anotasi yang akurat masih terbatas, menjadi tantangan dalam melatih model yang generalis dan *robust*.

State of the art dalam deteksi objek telah menunjukkan evolusi berbagai paradigma arsitektural. *VGGNet* dan *ResNet* mewakili era arsitektur dalam yang fokus pada akurasi klasifikasi. *MobileNet* dan *EfficientNet* memperkenalkan desain yang dioptimalkan untuk efisiensi. Paradigma *detector* seperti *YOLO (You Only Look Once)* merevolusi bidang ini dengan pendekatan *single shot* yang mencapai kecepatan inferensi sangat tinggi. Meskipun masing-masing sukses dalam domainnya, terdapat *research gap* yang kritis dan belum terjawab: belum ada studi komparatif yang sistematis, komprehensif, dan *statistically robust* yang secara langsung membandingkan perwakilan dari masing-masing paradigma arsitektural secara spesifik dalam konteks deteksi drone dengan segala kompleksitas problem yang disebutkan di atas.

Metode *Convolutional Neural Networks (CNN)* pada penelitian sebelumnya metode *VGG16* (Meng and Tia, 2020) mengalami kendala dalam konvergensi loss selama proses transfer training, yang mengakibatkan hasil yang kurang optimal. Di sisi lain, metode *ResNet 101* menunjukkan hasil 57,58% yang lebih baik dibandingkan dengan *VGG16*, namun masih belum mencapai tingkat efektivitas yang sama dengan metode *Inception V3* 96,82%. Hal tersebut menunjukkan pentingnya pemilihan metode yang tepat dalam *transfer learning* untuk mencapai hasil yang optimal dalam pengenalan dan deteksi gambar UAV. Dalam penelitian (Lee, Jeong and Jung, 2023), penggunaan drone dan metode *MobileNet* digunakan untuk analisis citra guna mengembangkan sistem deteksi kebakaran hutan yang

responsif dan efektif. Melalui proses pelatihan dan validasi metode, penelitian ini menunjukkan penurunan kesalahan dan peningkatan akurasi seiring dengan kemajuan pelatihan. Hasil evaluasi menunjukkan bahwa metode *MobileNet V1* memiliki akurasi validasi tertinggi sebesar 0.9466. Metode *EfficientNet* (Tan and Le, 2019) menerapkan pendekatan penskalaan komposit yang meratakan lebar jaringan, kedalaman, dan resolusi gambar dengan rasio konstan, *EfficientNet-B7* mencapai akurasi pertama sebesar 84.3% sementara memiliki jumlah parameter yang jauh lebih sedikit dan kecepatan inferensi yang lebih cepat.

Metode *You Only Look Once (YOLO)* pada *YOLOv5* (Aydin and Singha, 2023) melalui penambahan teknik augmentasi data untuk mengatasi kesulitan kelangkaan data secara artifisial, serta pencegahan *overfitting* menggunakan pembagian *train:test* acak 70:30, pengujian metode pada berbagai macam latar belakang (gelap, cerah), dan pengujian tampilan gambar yang berbeda. Hasil evaluasi juga menunjukkan bahwa metode mencapai *Precision* sebesar 0.918, *Recall* sebesar 0.875, dan *mAP50* sebesar 0.904 untuk semua kelas objek. *YOLOv8* dalam penelitian (Reis *et al.*, 2023) melatih metode pertama (yang digeneralisasi) pada kumpulan data yang berisi 40 kelas objek terbang yang berbeda, kemudian melakukan pembelajaran transfer dengan parameter yang dipelajari mencapai hasil *mAP50-95* sebesar 0,685 dan rata-rata 95 sebesar 0,685 dan kecepatan inferensi rata-rata pada video 1080p dari 50 FPS dan untuk metode yang telah disempurnakan mempertahankan inferensi kecepatan dan mencapai *mAP50-95* yang lebih baik dari 0,835.

Analisis literatur yang mendalam mengidentifikasi beberapa *research gaps* yang kritis sekaligus menjadi masalah utama yang mendorong urgensi penelitian ini: Pertama, kurangnya evaluasi komparatif sistematis dan komprehensif antar berbagai paradigma arsitektural *deep learning* yang merepresentasikan evolusi desain *neural network* dalam konteks spesifik deteksi drone dengan kondisi evaluasi yang konsisten. Studi-studi yang ada cenderung berfokus pada satu atau dua arsitektur, sehingga tidak memberikan panduan yang jelas mengenai *trade off* yang terukur antara akurasi, kecepatan, dan kompleksitas model ketika dihadapkan pada karakteristik unik objek drone. Kedua, terbatasnya pertimbangan *trade off* antara akurasi dan efisiensi komputasional dalam konteks penempatan praktis. Ketiga, ketidakhadiran validasi statistik *robust* dalam mengukur signifikansi perbedaan performa antar model, yang sering mengakibatkan kesimpulan yang tidak *robust* dan berpotensi menyesatkan dalam seleksi model untuk penerapan praktis. Keempat, belum ada panduan *evidence based* untuk seleksi model berdasarkan karakteristik spesifik aplikasi deteksi drone dengan mempertimbangkan kendala sumber daya komputasi yang beragam, mulai dari server berkinerja tinggi hingga perangkat *edge* dengan kemampuan terbatas, serta pemilihan varian/model spesifik yang seringkali tidak disertai justifikasi ilmiah yang kuat terkait kesesuaiannya dengan karakteristik dataset dan kebutuhan aplikasi.

Berdasarkan identifikasi *gaps* tersebut, penelitian ini hadir untuk mengisi celah melalui evaluasi komprehensif lima arsitektur representatif dengan metodologi yang *robust*. Pemilihan kelima arsitektur didasarkan pada prinsip

representasi perkembangan paradigma arsitektural: *VGG16* sebagai representasi arsitektur konvolusional klasik, *ResNet50* mewakili arsitektur modern dengan *residual connections*, *MobileNetV2* merepresentasikan arsitektur efisien untuk perangkat *edge*, *EfficientNetB0* menunjukkan pendekatan *compound scaling* yang terstruktur, dan *YOLOv8* dengan penggunaan *YOLOv8 nano* sebagai arsitektur khusus yang dioptimasi secara spesifik untuk deteksi objek *real time* pada perangkat dengan sumber daya terbatas. Pemilihan *YOLOv8 nano* secara khusus didasarkan pada analisis mendalam mengenai kebutuhan aplikasi deteksi drone di dunia nyata. Varian *nano* menawarkan keseimbangan optimal antara akurasi yang memadai, ukuran model yang minimal, dan kecepatan inferensi tinggi.

Penelitian ini tidak hanya berfokus pada aspek akurasi konvensional, tetapi juga mempertimbangkan dimensi efisiensi komputasional, stabilitas performa, dan signifikansi statistik. Integrasi *analytical statistics* (*ANOVA*, *pairwise t-test*, *effect size*) memastikan bahwa perbedaan yang diamati adalah bermakna dan bukan akibat variasi acak. Penelitian juga menyajikan visualisasi hasil deteksi empiris yang menunjukkan kemampuan nyata setiap model dalam menangani berbagai kondisi latar. Hasil penelitian diharapkan dapat memberikan panduan praktis untuk seleksi model yang optimal berdasarkan pertimbangan *trade off* yang terkuantifikasi dengan baik, sekaligus memberikan kontribusi teoritis bagi pengembangan *task architecture alignment theory* dalam domain *computer vision*.

1.2. Rumusan Masalah

Berdasarkan latar belakang yang telah diuraikan, rumusan masalah penelitian ini dapat dirinci sebagai berikut:

1. Bagaimana performa komparatif lima arsitektur *deep learning* (*VGG16*, *ResNet50*, *MobileNetV2*, *EfficientNetB0*, dan *YOLOv8-nano*) dalam konteks deteksi objek drone berdasarkan metrik regresi dan deteksi, serta signifikansi statistik perbedaan performa tersebut?
2. Seberapa signifikan *trade off* antara akurasi dan efisiensi komputasional pada kelima arsitektur tersebut dalam berbagai skenario penempatan berdasarkan analisis statistik yang *robust*?
3. Faktor arsitektural apa yang paling berpengaruh terhadap performa model dalam tugas deteksi drone berdasarkan analisis bukti empiris dan validasi statistik dan bagaimana karakteristik arsitektural masing-masing model menjelaskan variasi performa yang diamati?

1.3. Batasan Masalah

Penelitian ini memiliki batasan-batasan yang perlu ditegaskan untuk memfokuskan ruang lingkup analisis, sebagai berikut:

1. Penelitian hanya berfokus pada lima arsitektur *deep learning* (*VGG16*, *ResNet50*, *MobileNetV2*, *EfficientNetB0*, dan *YOLOv8-nano*) yang merepresentasikan perkembangan paradigma arsitektural utama dalam *computer vision*.
2. Dataset yang digunakan terbatas pada 1.359 gambar drone dengan anotasi *bounding box*, yang meskipun memadai untuk evaluasi komparatif namun

tidak mencakup semua variasi kondisi nyata yang berasal dari sumber

Kaggle; Drone Dataset UAV

<https://www.kaggle.com/dasmehdixtr/drone-dataset-uav>.

3. Evaluasi performa dilakukan dalam konteks deteksi *single class* (drone) tanpa mempertimbangkan *multi class detection* yang melibatkan kategori objek lainnya.
4. Untuk arsitektur *YOLOv8*, penelitian hanya menggunakan varian *nano* (*YOLOv8n*) sebagai representasi keluarga *YOLOv8*, dengan pertimbangan fokus pada model yang dioptimasi untuk *edge deployment*.
5. Semua model di latih dengan *hyperparameter* konsisten dan *environment* komputasi yang sama untuk memastikan perbandingan yang adil, dengan analisis statistik yang robust untuk memvalidasi hasil.

1.4. Tujuan Penelitian

Berdasarkan rumusan masalah yang telah disebut sebelumnya, tujuan penelitian ini, antara lain:

1. Mengukur dan membandingkan performa lima arsitektur *deep learning* untuk deteksi drone menggunakan *framework* evaluasi multi dimensi yang mencakup metrik regresi, deteksi, dan efisiensi, dilengkapi analisis statistik komprehensif dan visualisasi hasil deteksi empiris.
2. Menganalisis *trade off* antara akurasi dan efisiensi komputasional pada berbagai skenario penempatan, termasuk aplikasi *real time*, *edge computing*, dan sistem berakurasi tinggi, dengan dukungan uji signifikansi statistik dan justifikasi ilmiah pemilihan varian model.

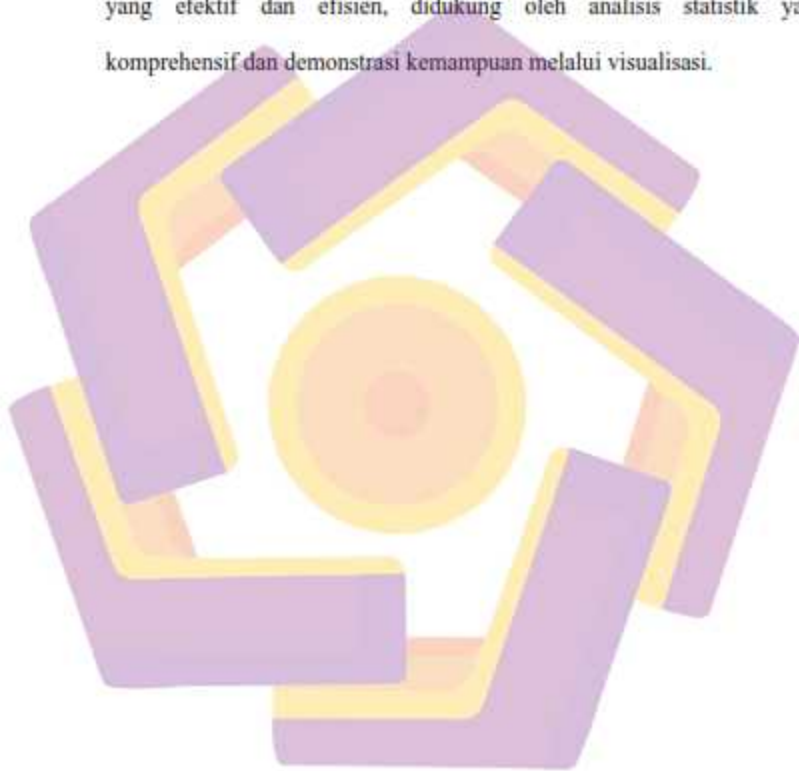
3. Mengidentifikasi karakteristik arsitektural yang optimal untuk tugas deteksi drone berdasarkan bukti empiris yang dikumpulkan dari eksperimen komprehensif dan validasi statistik serta menjelaskan hubungan kausal antara desain arsitektur dan performa yang dihasilkan.
4. Menyusun rekomendasi *evidence based* untuk seleksi model berdasarkan berbagai kebutuhan aplikasi praktis dalam domain keamanan dan surveillance dengan mempertimbangkan kendala sumber daya komputasi yang beragam, didukung oleh bukti statistik yang kuat dan demonstrasi visual kemampuan deteksi.

1.5. Manfaat Penelitian

Manfaat yang diharapkan dari penelitian ini dikategorikan menjadi tiga aspek, antara lain:

1. Segi kontribusi ilmiah, penelitian ini memberikan bukti empiris tentang *task architecture alignment* dalam *deep learning* dan *framework* evaluasi komprehensif yang dapat diadopsi untuk penelitian sejenis, dilengkapi metodologi analisis statistik yang *robust* serta justifikasi ilmiah untuk pemilihan varian model spesifik berdasarkan karakteristik tugas dan kendala sumber daya.
2. Segi aplikasi praktis, penelitian menyediakan panduan seleksi model teroptimasi berbasis bukti untuk implementasi sistem deteksi drone nyata dengan pertimbangan *trade off* yang terkuantifikasi dan divalidasi secara statistik, serta rekomendasi konkret untuk *deployment* di berbagai skenario.

3. Penelitian memberikan analisis *trade off* akurasi-efisiensi yang dapat dijadikan acuan untuk perancangan sistem *surveillance* dan keamanan dengan berbagai sumber daya yang dibatasi yang dihadapi di lapangan, serta kontribusi pada pengembangan standar teknis untuk sistem deteksi drone yang efektif dan efisien, didukung oleh analisis statistik yang komprehensif dan demonstrasi kemampuan melalui visualisasi.



BAB II

TINJAUAN PUSTAKA

2.1. Tinjauan Pustaka

Hasil *litration review* terhadap penelitian sebelumnya mengenai deteksi objek menggunakan metode kecerdasan buatan dan pembelajaran mesin terdapat berbagai metode mulai dari metode *VGG16*, *ResNet*, *MobileNet*, *EfficientNet* dan *YOLOv8*.

Penelitian (Meng and Tia, 2020) menggunakan metode *transfer learning* untuk klasifikasi dan deteksi gambar *UAV* dengan membandingkan metode-metode *deep learning* seperti *Inception V3*, *ResNet 101*, *VGG16*, *Faster R-CNN*, dan *SSD*. Hasil penelitian menunjukkan bahwa *Inception V3* mencapai akurasi tertinggi dalam klasifikasi dengan nilai 96.82%, diikuti oleh *ResNet 101* dengan akurasi 53.75%. Meskipun *VGG16* memiliki akurasi 50.00%, namun metode ini mencapai nilai *Recall* 100.00%. Dalam deteksi, *Faster R-CNN* memiliki akurasi deteksi 99.61% yang lebih tinggi daripada *SSD*, meskipun *SSD* memiliki *Recall* yang lebih tinggi.

Penelitian (Lee, Jeong and Jung, 2023) mengembangkan sistem deteksi kebakaran hutan menggunakan metode Convolutional Neural Network (CNN) berbasis drone menggunakan algoritma *MobileNet*, menunjukkan bahwa metode *MobileNet V1* memiliki akurasi validasi tertinggi sebesar 0.9466 di antara metode *MobileNet V1*, *V2*, dan *V3*. Selain itu, metode *MobileNet V1* juga menunjukkan akurasi tertinggi sebesar 0.9667.

Penelitian (Tan and Le, 2019) mengusulkan pendekatan penskalaan komposit untuk mencapai keseimbangan optimal antara akurasi dan efisiensi. Dengan menerapkan rasio penskalaan komposit yang meratakan lebar jaringan, kedalaman, dan resolusi gambar. *EfficientNet* berhasil mencapai kinerja superior dengan jumlah parameter yang lebih sedikit dan kecepatan inferensi yang lebih cepat. Hasil evaluasi matriks menunjukkan bahwa metode-metode *EfficientNet*, seperti *EfficientNet-B7*, berhasil mencapai akurasi pertama sebesar 84.3% pada dataset *ImageNet*, sementara menggunakan jumlah parameter yang jauh lebih sedikit dan memiliki kecepatan inferensi yang lebih cepat.

Penelitian (Mahdavi and Rajabi, 2020) menggunakan metode *Convolutional Neural Network (CNN)* untuk mendeteksi drone menggunakan kamera *fisheye* yang bertujuan untuk mengidentifikasi dan melacak drone yang sedang terbang dengan tingkat akurasi yang tinggi. Melibatkan pengumpulan data dari film-film yang berisi gambar drone dan burung, yang kemudian diubah menjadi frame-frame menggunakan perangkat lunak MATLAB. *CNN* diimplementasikan dengan struktur yang terdiri dari tiga *max-pooling layers*, tiga *convolution layers*, dan dua *fully connected layers*, dengan fungsi *softmax* dan 80 epoch. Hasil evaluasi menunjukkan bahwa *CNN* mencapai akurasi total sebesar 95%, sementara *SVM* mencapai 88% dan *Nearest Neighbor* sebesar 80%.

Penelitian (Aydin and Singha, 2023) menggunakan penambahan teknik augmentasi data untuk mengatasi kesulitan kelangkaan data secara artifisial, serta pencegahan *overfitting* menggunakan pembagian *train test* acak 70:30, penyempurnaan dari *YOLOv5* asli berdasarkan kumpulan data khusus yang di

kumpulkan, pengujian metode pada berbagai macam latar belakang dan pengujian tampilan gambar yang berbeda. Menghasilkan *Precision* sebesar 0.918, *Recall* sebesar 0.875, dan *mAP50* sebesar 0.904 untuk semua kelas objek. Untuk kelas "Bird", metode mencapai *Precision* sebesar 0.860, *Recall* sebesar 0.766, dan *mAP50* sebesar 0.820. Sedangkan untuk kelas "Drone", metode mencapai *Precision* sebesar 0.975, *Recall* sebesar 0.985, dan *mAP50* sebesar 0.987.

Penelitian (Liu *et al.*, 2022) menggunakan metode *Multi Scale Space Kinematic Detection Method* (MUSAK) untuk mendeteksi drone dari *background* yang tidak terkendali dengan memanfaatkan parameter kinematik multiskala yang diekstraksi dari input video. Parameter kinematik disini terdiri dari parameter translasi (kecepatan dan percepatan translasi) dan parameter rotasi (kecepatan sudut dan percepatan) sehingga lebih unggul dalam skenario definisi rendah dan fitur gerakan 2D lebih tahan terhadap penurunan kualitas gambar.

Penelitian (Reis *et al.*, 2023) mengembangkan metode *YOLOv8* dengan melatih metode pertama (yang digeneralisasi) pada kumpulan data yang berisi 40 kelas objek terbang yang berbeda, kemudian melakukan *transfer learning* dengan parameter yang dipelajari pada data yang lebih mewakili lingkungan nyata mencapai hasil *mAP50-95* sebesar 0,685 dan rata-rata 95 sebesar 0,685 dan kecepatan inferensi rata-rata pada video 1080p dari 50 FPS dan untuk metode yang telah disempurnakan mempertahankan inferensi kecepatan dan mencapai *mAP50-95* yang lebih baik dari 0,835.

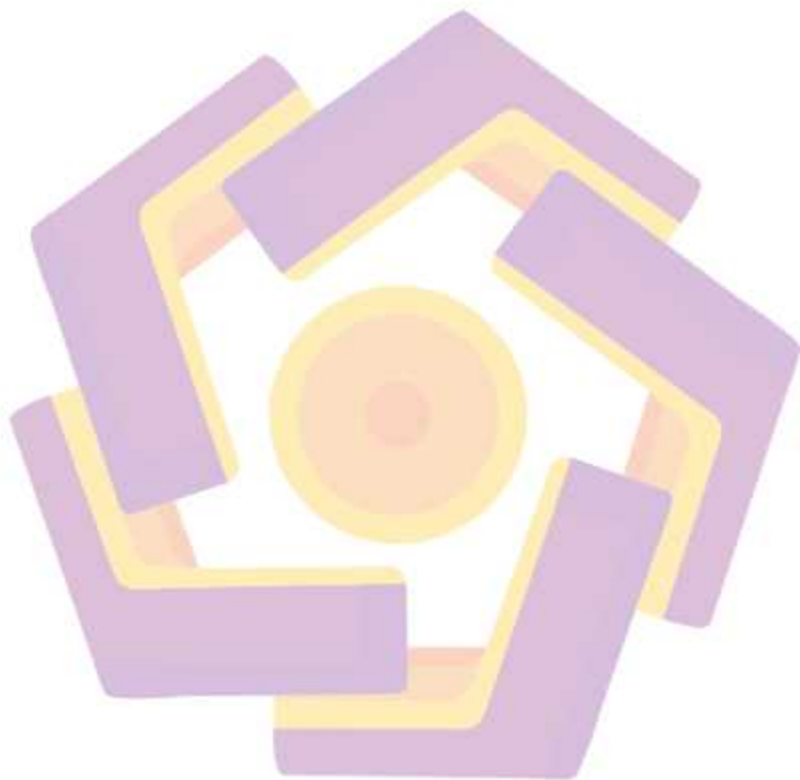
Berdasarkan hasil tinjauan pustaka terhadap penelitian sebelumnya mengenai deteksi objek menggunakan metode kecerdasan buatan dan pembelajaran mesin,

terdapat berbagai metode yang telah dikembangkan, mulai dari *Convolutional Neural Network (CNN)* hingga metode *You Only Look Once (YOLO)*. Penelitian yang menggunakan metode *transfer learning* untuk klasifikasi dan deteksi gambar UAV menunjukkan bahwa *Inception V3* mencapai akurasi tertinggi dalam klasifikasi, sementara *Faster R-CNN* memiliki akurasi deteksi yang tinggi. Pengembangan sistem deteksi kebakaran hutan dengan menggunakan *CNN* berbasis drone menunjukkan bahwa *MobileNet V1* memiliki akurasi validasi tertinggi. Selain itu, pendekatan penskalaan komposit dengan *EfficientNet* berhasil mencapai kinerja superior dengan jumlah parameter yang lebih sedikit.

Penelitian yang menggunakan *CNN* untuk mendeteksi drone dengan kamera *fisheye* mencapai akurasi total 95%, sementara penambahan teknik augmentasi data pada *YOLOv5* berhasil mengatasi kesulitan kelangkaan data secara artifisial. Metode *Multi Scale Space Kinematic Detection Method (MUSAK)* juga terbukti efektif dalam mendeteksi drone dari *background* yang tidak terkendali. Pengembangan *YOLOv8* dengan *transfer learning* mencapai hasil yang memuaskan dalam mendeteksi objek terbang.

Berdasarkan tinjauan komprehensif terhadap *state-of-the-art*, teridentifikasi *research gap* yang signifikan dalam hal ketiadaan studi komparatif sistematis yang mengevaluasi kelima arsitektur *deep learning* representatif (*VGG16, ResNet50, MobileNetV2, EfficientNetB0, YOLOv8*) untuk deteksi drone dengan pendekatan evaluasi multi dimensi dan statistik *robust*. Penelitian ini bertujuan mengisi gap tersebut melalui analisis empiris komprehensif yang mengintegrasikan aspek akurasi, efisiensi, dan signifikansi statistik, sekaligus

menyediakan framework seleksi model berbasis bukti untuk berbagai kebutuhan aplikasi deteksi drone.



2.2. Keaslian Penelitian

Tabel 2.1. Matriks Literatur Review dan Posisi Penelitian

Analisis Komparatif Empiris Lima Arsitektur Deep Learning Untuk Deteksi Drone: Superioritas Yolov8 Dan Trade Off Akurasi Efisiensi

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
1	Unmanned Aerial Vehicle Classification and Detection Based on Deep Transfer Learning (Meng and Tia, 2020)	Wei Meng, Meng Tian, Intelligent Computing and Human-Computer Interaction (ICHCI), 2020	Melakukan klasifikasi dan deteksi gambar UAV berdasarkan <i>transfer learning</i> , dengan fokus pada meningkatkan <i>accuracy</i> , <i>recall</i> , dan presisi pengenalan gambar UAV menggunakan metode <i>deep learning</i> pada sampel data yang kecil melalui perbandingan eksperimental antara tiga metode klasifikasi jaringan saraf konvolusi (Inception V3, ResNet 101, dan VGG16) dan dua metode	Dalam penelitian yang dilakukan menggunakan metode <i>transfer learning</i> untuk klasifikasi dan deteksi gambar UAV, hasil akurasi yang diperoleh menunjukkan tingkat keberhasilan yang signifikan. Metode klasifikasi seperti Inception V3 mencapai akurasi sebesar 96,82%, menunjukkan kemampuan yang sangat baik dalam mengenali gambar UAV. Di sisi lain, metode ResNet	Perlu implementasi pada metode-metode lainnya, mengembangkan dataset yang lebih besar dan beragam, mengoptimalkan parameter metode, dan mempertimbangkan penerapan metode dalam waktu nyata untuk aplikasi keamanan yang efektif. Kelemahan yang	Metode ResNet dan VGG16 belum menghasilkan akurasi yang baik sehingga masih dapat dilakukan optimalisasi parameter dan penambahan dataset yang beragam serta pelatihan yang lebih efektif agar nantinya dapat meningkatkan nilai mAP pada metode ResNet dan VGG16.

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			deteksi jaringan saraf konvolusi (Faster R-CNN dan SSD).	101 mencapai akurasi sebesar 57,58%, sedangkan metode VGG16 menunjukkan performa yang lebih rendah dan kurang efektif dalam klasifikasi gambar UAV.	diperhatikan adalah keterbatasan dalam variasi dataset yang digunakan, yang dapat mempengaruhi generalisasi metode terhadap kondisi lingkungan yang beragam.	
2	Development of a Forest Fire Detection System Using a Drone-based Convolutional Neural Network Metode (Lee, Jeong and Jung, 2023)	Jihee Lee, Keesin Jeong, Haiyoung Jung, International Journal of Fire Science and Engineering g. 2023	Mengembangkan sistem deteksi kebakaran hutan secara <i>real-time</i> menggunakan metode <i>Convolutional Neural Network</i> (CNN) berbasis drone. Metode penelitian melibatkan pengumpulan dan <i>augmentasi</i> data gambar, analisis struktur metode CNN, dengan menggunakan algoritma MobileNet yang dirancang	Dalam penelitian, hasil dari penggunaan metode Convolutional Neural Network (CNN) MobileNet V1 menunjukkan performa yang sangat baik dalam deteksi kebakaran hutan. Metode MobileNet V1 mencapai akurasi validasi tertinggi sebesar 0.9466 dan akurasi uji sebesar 0.9667. MobileNet V1	Perlu pengembangan terhadap metode selain MobileNet, seperti ResNet, Inception, atau EfficientNet sehingga dapat memberikan wawasan tambahan tentang metode mana yang paling cocok untuk	Metode MobileNet telah menunjukkan hasil yang baik namun belum dapat memberikan waktu proses sehingga berpengaruh pada evaluasi metode yang digunakan untuk menganalisa efisiensi penggunaan metode tersebut.

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			berdasarkan lapisan konvolusi yang terpisah untuk eksekusi <i>real-time</i> dan inferensi metode yang efisien pada perangkat mobile dan drone.	memiliki akurasi tertinggi dengan nilai sekitar 96% untuk deteksi kebakaran, 87% untuk deteksi asap, dan 61% untuk deteksi non-kebakaran. MobileNet V2 memiliki akurasi yang sedikit lebih rendah, sementara MobileNet V3 memiliki akurasi terendah di antara ketiganya.	deteksi. Tidak adanya informasi tentang waktu yang diperlukan untuk proses pelatihan dan evaluasi metode dapat menjadi kendala dalam mengevaluasi efisiensi dan kinerja metode.	
3	EfficientNet: Rethinking Metode Scaling for Convolutional Neural Networks (Tan and Le, 2019)	Mingxing Tan, Quoc V. Le, arXiv, 2020	Memperbaiki akurasi dan efisiensi ConvNets dengan memperhatikan keseimbangan antara lebar jaringan, kedalaman, dan resolusi gambar menggunakan pendekatan penskalaan komposit yang meratakan ketiga dimensi tersebut dengan rasio konstan sehingga	Dengan memperhatikan keseimbangan antara lebar jaringan, kedalaman, dan resolusi gambar, serta menerapkan pendekatan penskalaan komposit, EfficientNet mampu mencapai akurasi dan efisiensi yang lebih baik daripada ConvNets sebelumnya. Matriks	Perlu mempertimbangkan faktor-faktor lain yang dapat memengaruhi kinerja jaringan, seperti regularisasi, optimisasi, atau arsitektur jaringan yang lebih kompleks.	Metode EfficientNet telah menunjukkan hasil yang baik namun perlu adanya penggunaan Teknik optimasi dan uji dataset yang lebih beragam dan kompleks untuk menguji generalisasi EfficientNet pada

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
			menghasilkan serangkaian metode EfficientNet (B1 hingga B7) yang berhasil melampaui akurasi dan efisiensi ConvNets sebelumnya.	akurasi EfficientNet menunjukkan bahwa metode-metode EfficientNet (B1 hingga B7) berhasil mencapai akurasi yang signifikan dan efisiensi yang tinggi. Sebagai contoh, EfficientNet-B7 mencapai akurasi top-1 sebesar 84.3% pada dataset ImageNet, sementara memiliki jumlah parameter yang 8.4x lebih kecil dan kecepatan inferensi yang 6.1x lebih cepat dibandingkan dengan ConvNet terbaik sebelumnya.	Ketergantungan pada metode penskalaan komposit tertentu yang mungkin tidak selalu optimal untuk semua jenis tugas atau dataset tertentu.	berbagai tugas pengenalan gambar.
4	Drone Detection Using Convolutional	Fatemeh Mahdavi dan Roozbeh	Mendeteksi dan melacak drone yang sedang terbang menggunakan kamera <i>fisheye</i> . Metode yang	Hasil penelitian menunjukkan bahwa CNN memiliki performa yang lebih baik dalam	Perlu melakukan eksperimen lebih lanjut dengan dataset yang lebih	Metode CNN yang digunakan menghasilkan akurasi yang tinggi

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	Neural Networks (Mahdavi and Rajabi, 2020)	Rajabi, Iranian Conference on Signal Processing and Intelligent Systems (ICSPIS), 2020.	digunakan dalam penelitian ini melibatkan penggunaan tiga metode klasifikasi, yaitu <i>Convolutional Neural Network (CNN)</i> . Metode klasifikasi digunakan untuk memproses data dari video yang berisi gambar drone dan burung, yang kemudian diubah menjadi frame-frame dipisahkan untuk analisis lebih lanjut.	membedakan bagian-bagian yang mirip antara drone dan burung dibandingkan dengan metode <i>Support Vector Machine (SVM)</i> dan <i>Nearest Neighbor</i> . Meskipun CNN memerlukan waktu pelatihan yang lebih lama, namun akurasi yang tinggi membuatnya menjadi metode yang efektif dalam mendeteksi drone. CNN memiliki akurasi total sebesar 95%, SVM sebesar 88%, dan <i>Nearest Neighbor</i> sebesar 80%.	besar dan beragam untuk menguji keandalan dan generalisasi dari model CNN. Ketergantungan pada kualitas dan resolusi gambar dari kamera fisheye dapat mempengaruhi akurasi deteksi drone.	namun kebutuhan resolusi yang tinggi dapat berpengaruh pada hasil deteksi. Penelitian dapat dikembangkan dengan menggunakan dataset yang lebih banyak.
5	Drone Detection Using YOLOv5	Burchan Aydin dan Subroto Singha,	Mengatasi tantangan dalam mendeteksi drone di ruang udara menggunakan YOLOv5 dengan CSPNet	Dengan menggunakan YOLOv5 CSPNet menghasilkan output yang lebih efisien daripada versi	Perlu memperluas dataset pelatihan dengan lebih banyak variasi	Metode YOLOv5 telah menunjukkan nilai presisi, F-1 Score, dan mAP

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	(Aydin and Singha, 2023)	MDPI, Basel, Switzerland, 2023.	untuk mengurangi kompleksitas metode. YOLOv5 memberikan peningkatan kinerja yang signifikan dibandingkan dengan versi YOLOv4, dengan peningkatan sebesar 90%.	sebelumnya yakni YOLOv4 dengan nilai presisi, F-1 Score, dan mAP, mencapai 0,918 dan 0,875. masing-masing bernilai 0,896 dan 0,904. Nilai tersebut mengungguli kinerja metode sebelumnya dengan mencapai perolehan yang lebih tinggi, skor F-1, dan nilai mAP meningkat sebesar 21,57%.	drone dan objek yang mirip drone untuk meningkatkan generalisasi metode. Metode deteksi drone yang membutuhkan perangkat keras yang canggih seperti GPU untuk deteksi real-time dapat menjadi kendala dalam implementasi di lapangan.	yang baik. Pada penelitian yang akan dilakukan mencoba menguji penggunaan metode YOLO yang lebih baru yaitu YOLOv8 untuk melihat perubahan capaian.
6	Drone Detection Method Based on MobileViT and CA-PANet	Qianqing Cheng, Xiuhe Li, Bin Zhu, Yingchun Shi, dan Bo Xie, MDPI,	Menggunakan kinerja metode YOLOv4 MCA yang merupakan pengembangan dari YOLOv4 dengan hasil yang lebih efisien.	Hasil output signifikan dapat mencapai akurasi 92,81% untuk nilai tertinggi. YOLOv4 MCA dapat berkinerja baik pada kumpulan data Drone dan kumpulan data PASCAL.	Efektifitas dari metode YOLOv4 belum dapat memberikan akurasi yang baik pada citra kecil meskipun telah	Metode YOLOv4 belum memberikan waktu komputasi untuk menghasilkan nilai yang tinggi. Pada penelitian yang akan dilakukan

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	(Cheng <i>et al.</i> , 2023)	Basel, Switzerland, 2023.		VOC 07+12 pada berbagai skala target	memperoleh nilai akurasi yang tinggi.	dengan metode YOLOv8 akan mengoptimalkan nilai akurasi dan proses komputasi yang lebih baik.
7	MUSAK: A Multi-Scale Space Kinematic Method for Drone Detection (Liu <i>et al.</i> , 2022)	Sunxiangyu Liu, Guitao Li, Yafeng Zhan dan Peng Gao, MDPI, Basel, Switzerland, 2022.	Memperkenalkan inovasi metode terbaru yakni MUSAK dalam Deteksi objek yang bergerak menggunakan Drone.	MUSAK cenderung mengarah pada arah baru untuk mendeteksi objek bergerak atau kabur. MUSAK juga membangun kumpulan data drone yang bersifat publik dan lebih komprehensif untuk penelitian di masa mendatang.	MUSAK merupakan inovasi pengembangan dari metode kerja CNN dan YOLO yang berfokus pada objek bergerak, namun dalam penelitian tidak diberikan persentase yang jelas untuk memberikan perbandingan yang konkret pada metode kerja lainnya.	MUSAK berfokus pada objek bergerak sedangkan penelitian yang akan dilakukan lebih mengarah pada deteksi objek drone yang bersifat statis.

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
8	Real-Time Flying Object Detection with YOLOv8 (Reis <i>et al.</i> , 2023)	Dillon Reis, Jordan Kupec, Jacqueline Hong, Ahmad Daoudi, arXiv, 2023	Mengembangkan metode deteksi objek terbang secara real-time yang dapat digunakan untuk <i>transfer learning</i> dan implementasi praktis dengan penggunaan metode YOLOv8 yang dianggap memiliki kinerja lebih baik dalam mendeteksi objek udara.	Hasil dari penggunaan metode YOLOv8 dengan melatih metode pertama pada data set yang berisi 40 kelas objek terbang yang berbeda untuk mengekstrak representasi fitur abstrak dan <i>transfer learning</i> dilakukan pada data set yang lebih mewakili lingkungan. Hasil penelitian menunjukkan bahwa metode yang dikembangkan mencapai tingkat akurasi yang tinggi dengan mAP50-95 sebesar 0.835, sementara tetap mempertahankan kecepatan inferensi yang baik pada video 1080p dengan 50 fps.	Perlu adanya optimisasi lebih lanjut terhadap parameter dan arsitektur metode serta uji coba dan validasi metode deteksi objek terbang dalam skenario lapangan yang lebih luas dan beragam. Meskipun penelitian menggunakan data set yang beragam, jumlah data set yang terbatas (seperti pada data set transfer learning) dapat mempengaruhi generalisasi metode	Metode YOLOv8 telah memberikan nilai akurasi yang baik namun masih dapat ditingkatkan dengan melakukan eksperimen pada parameter dan pembelajaran agar dapat lebih meningkatkan nilai akurasi.

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
					deteksi objek terbang.	
9	MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications (Howard <i>et al.</i> , 2017)	Andrew G. Howard, Menglong Zhu, Bo Chen, Dmitry Kalenichenko, Weijun Wang, Tobias Weyand, Marco Andreetto, Hartwig Adam, arXiv, 2017	Merancang arsitektur CNN yang efisien untuk perangkat mobile dan <i>embedded</i> dengan menggunakan <i>depthwise separable convolution</i> serta memperkenalkan dua <i>hyperparameter</i> (<i>width multiplier</i> dan <i>resolution multiplier</i>) untuk menyeimbangkan akurasi dan latensi.	MobileNets berhasil mencapai efisiensi komputasi dan ukuran model yang sangat baik dengan hanya sedikit penurunan akurasi. Arsitektur ini terbukti efektif pada berbagai tugas seperti klasifikasi, deteksi objek, face attributes, dan geo-localization.	Rencana rilis model dalam TensorFlow untuk memudahkan adopsi. Akurasi menurun signifikan saat model diperkecil terlalu agresif ($\alpha=0.25$).	Dibandingkan dengan GoogleNet dan VGG16, MobileNets lebih kecil, lebih cepat, dan tetap kompetitif dalam akurasi. Dibandingkan SqueezeNet dan AlexNet, MobileNets lebih akurat dan efisien.
10	Very Deep Convolutional Networks for	Karen Simonyan & Andrew	Untuk menyelidiki pengaruh kedalaman (depth) jaringan konvolusi	Peningkatan kedalaman jaringan hingga 16-19 lapisan (model D & E)	Waktu pelatihan sangat lama (2-3 minggu) pada	Model VGG tunggal (7.1% error) mengungguli GoogL

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
	Large-Scale Image Recognition (Simonyan and Zisserman, 2015)	Zisserman, ICLR, 2015	terhadap akurasi dalam pengenalan gambar skala besar, dengan menggunakan arsitektur yang konsisten dan filter konvolusi yang sangat kecil (3x3).	secara signifikan meningkatkan akurasi pada dataset ILSVRC. Penggunaan tumpukan filter 3x3 terbukti lebih unggul daripada filter besar karena meningkatkan non-linearitas dan mengurangi parameter. <i>Scale jittering</i> selama pelatihan meningkatkan ketangguhan model terhadap variasi skala objek. Model VGG meraih posisi ke-2 dalam klasifikasi dan posisi ke-1 dalam lokalisasi pada ILSVRC-2014. Fitur VGG yang telah dilatih juga menunjukkan kemampuan generalisasi yang sangat baik pada dataset lain seperti PASCAL VOC dan	sistem 4-GPU. Membutuhkan sumber daya komputasi yang tinggi. Kinerja model jenuh pada kedalaman 19 lapisan untuk dataset ImageNet. Model yang lebih dalam mungkin akan bermanfaat untuk dataset yang lebih besar. Dapat menggabungkan teknik peningkatan resolusi dari <i>OverFeat</i> untuk meningkatkan kinerja lokalisasi. Inisialisasi yang lebih baik pelatihan	eNet tunggal (7.9% error). <i>Ensemble FGG</i> (6.8% error) sebanding dengan GoogLeNet (6.7% error). VGG (25.3% error) mengungguli <i>OverFeat</i> (29.9% error) dan GoogLeNet (26.7% error). Fitur VGG (Net-D & Net-E) mengungguli model yang lebih dangkal (seperti AlexNet/ZF Net) pada berbagai tugas seperti klasifikasi VOC, Caltech, dan action recognition.

No	Judul	Peneliti, Media Publikasi, dan Tahun	Tujuan Penelitian	Kesimpulan	Saran atau Kelemahan	Perbandingan
				Caltech, mencapai hasil <i>state of the art</i> .	dari awal tanpa <i>pre training</i> .	
11	Comparison of CNN-Based Architectures for Detection of Different Object Classes (Bilous <i>et al.</i> , 2024)	Nataliya Bilous, Vladyslav Malko, Marcus Frohme, Alina Nechyporenko, MDPI, 2024	Membandingkan performa berbagai arsitektur CNN (YOLOv4-v8, Faster R-CNN, SSD, EfficientDet) untuk deteksi orang dan objek teknis dalam situasi kritis seperti bencana alam dan operasi militer, dengan fokus pada akurasi dan kecepatan pemrosesan waktu nyata.	YOLOv8 merupakan model terbaik dengan $mAP@0.5$ 0.88, F1-Score 0.88, dan FPS 48. Model ini unggul dalam akurasi, kecepatan, dan ketahanan terhadap kondisi lingkungan kompleks seperti pantulan air dan pencahayaan berubah. Faster R-CNN akurat tetapi lambat, SSD dan EfficientDet seimbang namun kurang akurat dibanding YOLOv8.	Perlu peningkatan dalam menangani objek yang tumpang-tindih dan kondisi visibilitas rendah. Disarankan untuk menggunakan data pelatihan tambahan, teknik <i>reinforcement learning</i> , dan arsitektur yang lebih kompleks untuk mengurangi kesalahan klasifikasi.	YOLOv8 lebih unggul dalam semua metrik utama dibandingkan Faster R-CNN, SSD, dan EfficientDet. Faster R-CNN cocok untuk aplikasi non-waktu nyata yang membutuhkan akurasi tinggi, sementara SSD dan EfficientDet cocok untuk sistem dengan sumber daya terbatas.

2.3. Landasan Teori

2.3.1. Citra Udara

Citra Udara adalah kumpulan foto dengan menggunakan kamera udara. Foto pada dasarnya adalah representasi dari karakteristik reflektansi (kecerahan relatif) dari fitur direkam ke film fotografi. Citra udara paling sering diambil sebagai foto pankromatik (hitam dan putih), warna, atau inframerah warna palsu. Namun, berbagai jenis radiasi elektromagnetik juga dapat direkam ke film fotografi dengan menggunakan emulsi yang berbeda (Morgan, Gergel and Coops, 2010). Citra udara dapat dilakukan oleh Drone dalam memberikan hasil yang baik. Penggunaan drone telah berkembang pesat karena kemampuannya yang fleksibel, biaya yang lebih rendah, serta kemudahan pengoperasian.

2.3.2. Artificial Intelligence

Artificial intelligence (AI) adalah bidang ilmu komputer yang bertujuan untuk mengetahui atau memetodekan proses berpikir manusia dan mendesain mesin yang dapat meniru perilaku manusia. *Artificial intelligence* merupakan ilmu yang sangat luas, didalamnya terdapat bidang-bidang lain seperti Sistem Pakar, *Natural Language Processing*, *Speech Recognition*, *Computer Vision*, *Machine Learning (ML)*, dan *Deep Learning (DL)*. AI sebagai bidang yang mencakup berbagai aspek kecerdasan buatan, mulai dari logika, probabilitas, dan matematika kontinu hingga persepsi, penalaran, pembelajaran, dan tindakan (Huang, 2010).

2.3.3. Deep Learning

Deep learning merupakan subbidang pembelajaran mesin yang melibatkan pelatihan jaringan saraf tiruan untuk belajar dari data dalam jumlah besar. Jaringan saraf ini terdiri dari beberapa lapisan *node* yang saling terhubung, yang memungkinkannya mempelajari representasi data input yang semakin kompleks (Won, 2018). *Deep learning* merupakan sistem yang dibangun dengan algoritma yang terinspirasi dari cara kerja struktur biologis syaraf manusia untuk menyelesaikan suatu permasalahan yang dapat diimplementasikan untuk pengolahan data text, data gambar, data suara, atau data video. Hubungan antara *artificial intelligence*, *machine learning*, dan *deep learning* yang dijelaskan dalam Gambar 2.1.



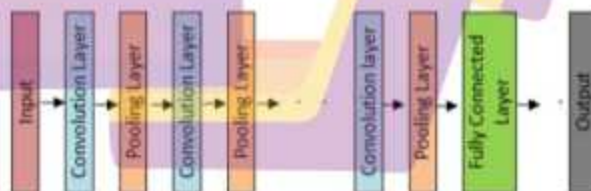
Gambar 2.1. *Artificial intelligence, machine learning, dan deep learning*(Won, 2018)

Gambar 2.1. menunjukkan bagaimana *deep learning* adalah sejenis pembelajaran representasi yang merupakan jenis pembelajaran mesin,

digunakan untuk banyak pendekatan untuk AI. Setiap bagian dari *diagram venn* menyertakan contoh teknologi AI.

2.3.4. CNN (Convolutional Neural Network)

Convolutional Neural Network adalah jenis Jaringan Syaraf Tiruan (JST) dengan arsitektur *feed forward* yang dalam dan kemampuan generalisasi yang luar biasa. *CNN* dapat mempelajari dan mengidentifikasi data yang sangat kompleks, terutama data spasial (Sultana, Sufian and Dutta, 2018). Metode *CNN* yang mendalam terdiri dari sekumpulan lapisan pemrosesan yang terbatas yang memiliki kemampuan untuk mempelajari dan mengekstrak berbagai jenis data input (seperti gambar) dengan berbagai tingkat abstraksi. Lapisan awal mempelajari dan mengekstrak fitur tingkat tinggi, yang memiliki abstraksi yang lebih rendah, dan lapisan yang lebih dalam mempelajari dan mengekstrak fitur tingkat rendah. Metode konseptual dasar dari *CNN* ditunjukkan pada Gambar 2.2.

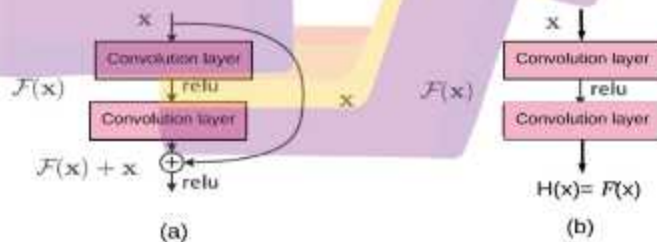


Gambar 2.2. Konsep Metode *CNN* (Sultana, Sufian and Dutta, 2018)

Gambar 2.2. Menjelaskan tipikal metode *CNN*, terdiri dari satu atau lebih blok konvolusi, lapisan sub-pengambilan sampel, dan lapisan keluaran.

2.3.5. ResNet (Residual Network)

Residual Network merupakan salah satu jenis arsitektur *Convolutional Neural Network (CNN)* yang dikembangkan untuk mengatasi masalah *vanishing gradient* pada *CNN* yang terjadi ketika metode menjadi terlalu dalam. *ResNet* terdiri dari banyak blok residual dengan lapisan konvolusi 3×3 . Setiap blok residual terdiri dari dua lapisan konvolusi dan sebuah *shortcut connection* yang menghubungkan input langsung ke output. *Shortcut connection* ini memungkinkan informasi untuk melewati blok residual tanpa mengalami penurunan gradien yang signifikan (Ghosh *et al.*, 2019). *ResNet* dapat meningkatkan kemampuan ekstraksi fitur jaringannya melalui fusi fitur lintas lapisan, dan jaringan kinerja secara bertahap meningkat seiring dengan pendalaman jaringan. Sehingga *ResNet* dapat memecahkan masalah degradasi jaringan saraf, yaitu ketika jaringan saraf semakin dalam maka tingkat kesalahan pelatihan semakin tinggi. Pemetaan yang digunakan pada *ResNet* ditunjukkan pada Gambar 2.3.

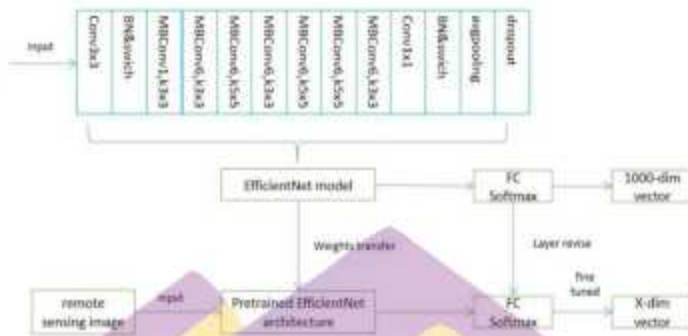


Gambar 2.3. (a) Pemetaan di dalam blok Residual, (b) Pemetaan langsung sederhana (Ghosh *et al.*, 2019)

Gambar 2.4. Menjelaskan arsitektur *VGG-16*, Arsitektur *VGG-16* terdiri dari 16 lapisan, 13 lapisan konvolusi dan 3 lapisan densitas penuh (*fully connected*).

2.3.7. *Efficientnet*

EfficientNet merupakan sebuah jaringan saraf konvolusional *CNN* yang dikembangkan dengan konsep *compound scaling* untuk menyeimbangkan ukuran metode, akurasi, dan efisiensi komputasi. Konsep ini fokus pada tiga dimensi penting dalam sebuah jaringan saraf: lebar (*width*), kedalaman (*depth*), dan resolusi (*resolution*). *EfficientNet* mengoptimalkan ketiga dimensi dengan pendekatan yang sistematis melalui pencarian *grid*, dimulai dari metode dasar, untuk menemukan kombinasi optimal lebar, kedalaman, dan resolusi. Kemudian menggunakan koefisien gabungan yang disebut phi (ϕ), yang menyamakan skala dimensi metode. Nilai phi ini berperan sebagai parameter yang ditentukan pengguna untuk menentukan kompleksitas dan kebutuhan sumber daya metode (Tan and Le, 2019). Dengan secara sistematis mengeksplorasi berbagai kombinasi lebar, kedalaman, dan resolusi serta dipandu oleh koefisien gabungan phi, *EfficientNet* mencapai keseimbangan antara ukuran metode, akurasi, dan efisiensi komputasi. Pendekatan ini memungkinkan untuk memilih varian metode yang sesuai dengan batasan sumber daya dan kebutuhan performa, arsitektur *EfficientNet* ditunjukkan pada Gambar 2.5.



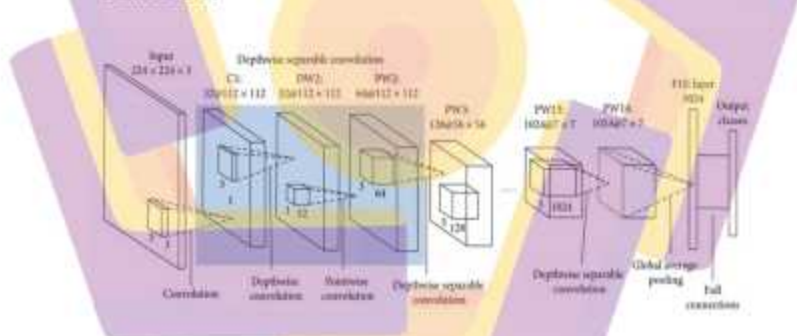
Gambar 2.5. Arsitektur *EfficientNet* (Zhang, Liu and Shi, 2020)

Gambar 2.5: menjelaskan arsitektur *EfficientNet* menggunakan *compound scaling* untuk mengoptimalkan kedalaman, lebar, dan resolusi gambar secara proporsional. Blok *MBConv* memanfaatkan *depthwise convolution*, *pointwise convolution*, dan *swish activation function* untuk mengurangi parameter dan komputasi tanpa mengorbankan kinerja. Blok SE (*squeeze and excitation*) untuk meningkatkan adaptabilitas jaringan dengan menyoroti fitur-fitur penting dan mengurangi noise dalam representasi gambar.

2.3.8. *MobileNet*

MobileNet merupakan salah satu *state of the art* arsitektur yang dapat mengurangi ukuran metode secara signifikan dan hanya mempengaruhi metode dan hanya mempengaruhi akurasi minimal. Dasar dari Arsitektur *MobileNet* adalah dua jenis lapisan konvolusi yang berbeda dari standar proses konvolusi. Lapisan-lapisan ini dapat dipisahkan berdasarkan kedalaman konvolusi yang dapat dipisahkan menurut kedalaman dan lapisan

konvolusi yang dapat dipisahkan menurut arah. Lapisan *depthwise* menggunakan filter tunggal untuk setiap saluran input, berbeda dari konvolusi standar yang menggunakan input dengan dimensi saluran untuk satu filter. Untuk lapisan *pointwise*, kemudian melakukan operasi perkalian filter 1×1 untuk menggabungkan hasil peta fitur dari *depthwise layer* (Zhao *et al.*, 2017). *MobileNet* memiliki jumlah operasi yang kecil meskipun struktur jaringan sarafnya relatif kompleks dibandingkan dengan model *CNN* pada umumnya karena struktur *MobileNet* menggunakan lapisan konvolusi yang dapat dipisahkan yang mengekstrak fitur gambar lebih akurat daripada *CNN* (Lee, Jeong and Jung, 2023), arsitektur *MobileNet* sebagaimana pada Gambar 2.6.



Gambar 2.6. Arsitektur *MobileNet* (Wang *et al.*, 2020)

Gambar 2.6. menunjukkan arsitektur *MobileNet* yang menggunakan lapisan konvolusi yang dapat dipisahkan *depthwise convolution* dan *pointwise convolution*.

2.3.9. YOLO (You Only Look Once)

You Only Look Once merupakan sebuah metode deteksi objek berbasis *Convolutional Neural Network (CNN)* yang dikembangkan menggunakan

lapisan konvolusi menerapkan lapisan *convolutional* dengan ukuran 1×1 untuk mengurangi ruang fitur dari lapisan sebelumnya.

2.3.10. YOLOv8

YOLOv8 menggunakan klasifikasi *loss* BCE yang berasal dari regresi CIOU Loss + DFL, bagian *backbone* dari *YOLOv8* pada dasarnya sama dengan *YOLOv5*, dan modul C3 digantikan oleh modul C2f berdasarkan ide CSP. Modul C2f belajar dari ide ELAN di *YOLOv7*, dan menggabungkan C3 dan ELAN untuk membentuk modul C2f, sehingga *YOLOv8* dapat memperoleh informasi aliran gradien yang lebih berlimpah dan tetap ringan. Pada akhir *backbone*, modul SPPF yang paling populer masih digunakan, dan tiga *maxpools* dengan ukuran 5×5 dilewatkan secara serial, dan kemudian setiap lapisan digabungkan, untuk menjamin keakuratan target dalam berbagai skala sekaligus memastikan bobot yang ringan secara bersamaan. Pada bagian leher, metode fusi fitur yang digunakan oleh *YOLOv8* masih menggunakan *PAN-FPN*, yang memperkuat fusi dan pemanfaatan informasi lapisan fitur pada skala yang berbeda (Lou *et al.*, 2023), dapat digambarkan dalam diagram arsitektur network *YOLOv8* pada Gambar 2.8.



Gambar 2.8. Arsitektur YOLOv8.

Gambar 2.8. menjelaskan diagram arsitektur *YOLOv8*, *YOLOv8* menggunakan dua modul *upsampling* dan beberapa modul *C2f* bersama dengan struktur kepala *decoupled* akhir untuk menyusun modul *neck*.

Varian *YOLOv8-nano* (*YOLOv8n*) merupakan arsitektur inti yang secara khusus dirancang untuk menyeimbangkan performa deteksi dan efisiensi komputasi pada lingkungan sumber daya terbatas. Dengan jumlah parameter yang minimal (sekitar 3,2 juta), *YOLOv8n* mempertahankan mekanisme fusi fitur multi-skala melalui *Path Aggregation Network (PANet)* dan pendekatan *anchor free* yang efektif untuk deteksi objek kecil seperti drone yang kerap muncul dalam resolusi terbatas pada citra pengawasan. Karakteristik ringan ini secara strategis mengurangi risiko *overfitting* pada dataset pelatihan yang terbatas, sekaligus memenuhi persyaratan mendasar untuk penerapan di perangkat *edge*. *YOLOv8n* merepresentasikan pilihan yang kontekstual dan terukur, yang mengutamakan kelayakan penerapan praktis tanpa mengabaikan kemampuan deteksi yang memadai untuk objek dinamis seperti drone, sesuai dengan tren penelitian yang merekomendasikan model terkompresi untuk aplikasi *real time* pada perangkat tepi.

2.3.11. Matriks Evaluasi

Matriks evaluasi memainkan peran kritis dalam mengukur dan membandingkan performa model *deep learning* untuk tugas deteksi objek. Menurut (Padilla, Netto and Da Silva, 2020), pemilihan matriks evaluasi yang tepat harus mempertimbangkan tiga aspek fundamental: relevansi dengan

aplikasi, sensitifitas terhadap pengembangan, dan kemudahan interpretasi. Dalam konteks deteksi objek drone, matriks evaluasi perlu mampu menangkap kompleksitas tugas yang meliputi akurasi lokalisasi, keandalan klasifikasi, dan efisiensi komputasional dalam Tabel 2.3.

Tabel 2.2. Klasifikasi matriks evaluasi untuk object detection

Kategori	Matriks	Fokus Utama
Regression Metrics	<i>MSE, MAE, R²</i>	Akurasi bounding box prediction
Detection Metrics	<i>Precision, Recall, F1-Score, mAP</i>	Keakuratan deteksi dan klasifikasi
Efficiency Metrics	<i>Training Time, Inference Speed, Parameter Count</i>	Kebutuhan komputasional
Statistical Metrics	<i>Confidence Intervals, p-value, Effect Size</i>	Signifikansi dan reliabilitas hasil

2.3.11.1. Metrik Regresi untuk Evaluasi Bounding Box

a. MAE

Mean Absolute Error (MAE) adalah metrik evaluasi yang mengukur rata-rata dari selisih absolut antara prediksi model dan nilai sebenarnya dalam dataset. Rumus untuk menghitung MAE adalah sebagai berikut:

$$MAE = \frac{1}{n} \sum_{i=1}^n |y_i - \hat{y}_i|$$

Di mana:

n adalah jumlah sampel dalam dataset.

y_i adalah nilai sebenarnya dari sampel ke- i .

\hat{y}_i adalah nilai prediksi model untuk sampel ke- i .

MAE menghasilkan nilai yang positif dan mengukur kesalahan absolut antara prediksi dan nilai sebenarnya. Semakin rendah nilai *MAE*, semakin baik kinerja model dalam memprediksi nilai.

b. *MSE*

Mean Squared Error (*MSE*) adalah metrik evaluasi yang mengukur rata-rata dari kuadrat selisih antara prediksi model dan nilai sebenarnya dalam dataset. Rumus untuk menghitung *MSE* adalah sebagai berikut:

$$MSE = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

Di mana:

n adalah jumlah sampel dalam dataset.

y_i adalah nilai sebenarnya dari sampel ke- i .

\hat{y}_i adalah nilai prediksi model untuk sampel ke- i .

MSE menghasilkan nilai yang positif dan mengukur kesalahan kuadrat antara prediksi dan nilai sebenarnya. Semakin rendah nilai *MSE*, semakin baik kinerja model dalam memprediksi nilai. *MSE* lebih sensitif terhadap outlier dibandingkan dengan *MAE* karena kesalahan dikuadratkan.

- c. Koefisien Determinasi (R^2) mengkuantifikasi proporsi varians dalam data yang dapat dijelaskan oleh model. $R^2 > 0.6$ dianggap sangat baik dalam tugas deteksi objek, sementara nilai di bawah 0.3 mengindikasikan model yang tidak cukup (M and M.N, 2015).

2.3.11.2. Metrik Deteksi untuk Klasifikasi dan Lokalisasi

Precision dan *Recall* membentuk dasar evaluasi sistem deteksi objek. *Precision* mengukur akurasi deteksi positif, sementara *Recall* mengukur kemampuan mendeteksi semua *instance* positif. *F1-Score* sebagai rata-rata harmonik keduanya memberikan keseimbangan yang optimal (Jiao *et al.*, 2019). Dalam konteks keamanan drone, *Recall* seringkali lebih kritis daripada *Precision* karena konsekuensi *false negative* (*missed detection*) lebih berbahaya daripada *false positive*.

Intersection over Union (IoU) merupakan metrik standar untuk mengukur akurasi *overlap* antara *bounding box* prediksi dan *ground truth*. Nilai *IoU* threshold 0.5 umum digunakan dalam perhitungan *mAP* (*mean Average Precision*). Namun, untuk aplikasi drone dengan ukuran kecil, penggunaan *multi-IoU threshold* (0.5:0.95) untuk evaluasi yang lebih *robust* (Padilla, Netto and Da Silva, 2020).

2.3.11.3. Metrik Efisiensi Komputasional

Metrik Efisiensi menjadi semakin kritis dengan berkembangnya aplikasi *edge computing* untuk deteksi drone *real time*. *Training time* mengukur kebutuhan komputasional selama fase training, sementara *inference speed* menentukan kelayakan penempatan *real time*. Sistem deteksi drone harus mencapai minimal 30 FPS untuk aplikasi *real time*, dengan *latency* di bawah 33 ms (Wu *et al.*, 2018).

Number of parameters berkorelasi langsung dengan kompleksitas model dan jejak memori. Namun, *parameter count* bukan satu-satunya

indikator efisiensi, sebagaimana dibuktikan oleh arsitektur *EfficientNet* yang mencapai performa tinggi dengan parameter relatif sedikit (Touvron *et al.*, 2021).

Efficiency Ratio yang didefinisikan sebagai *F1-Score* per unit waktu training merupakan metrik komposit yang berguna untuk mengevaluasi *trade off* akurasi-efisiensi. Konsep *Computational Pareto Efficiency* dimana model dianggap optimal jika tidak ada model lain yang lebih akurat dengan sumber daya yang sama, atau lebih efisien dengan akurasi yang sama (Li *et al.*, 2023).

2.3.11.4. Analytical Statistics Untuk Evaluasi Komparatif

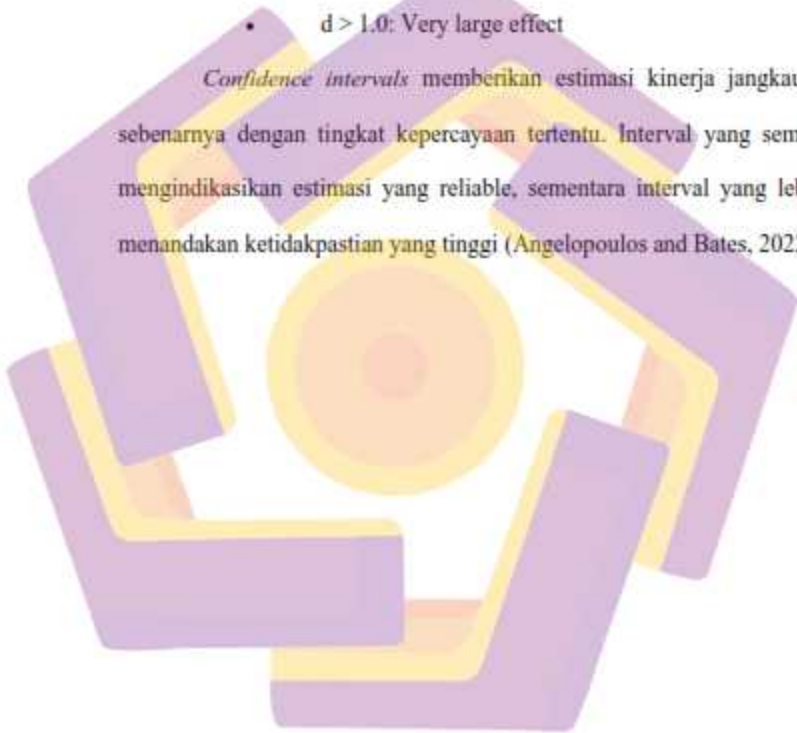
Dalam penelitian komparatif, *analytical statistics* esensial untuk memvalidasi signifikansi perbedaan performa. *Analysis of Variance (ANOVA)* digunakan untuk menguji perbedaan signifikan antar *multiple groups* secara simultan (Demšar, 2006). *ANOVA* dengan *p-value* < 0,05 mengindikasikan bahwa setidaknya satu model secara signifikan berbeda dari lainnya (Benavoli *et al.*, 2017).

Pairwise t-test dengan koreksi Bonferroni diterapkan untuk mengontrol *family wise error rate* dalam analisis perbandingan berpasangan ini (Northcutt, Jiang and Chuang, 2021). Dalam penelitian komparatif, penggunaan uji statistik yang *robust* sangat penting untuk memastikan bahwa perbedaan performa yang diamati benar-benar signifikan dan bukan disebabkan oleh variasi acak (Dror, Shlomov and

Reichart, 2020). *Effect size analysis* (Cohen's d) mengkuantifikasi magnitudo perbedaan:

- $d = 0.2$: Small effect
- $d = 0.5$: Medium effect
- $d = 0.8$: Large effect
- $d > 1.0$: Very large effect

Confidence intervals memberikan estimasi kinerja jangkauan sebenarnya dengan tingkat kepercayaan tertentu. Interval yang sempit mengindikasikan estimasi yang reliable, sementara interval yang lebar menandakan ketidakpastian yang tinggi (Angelopoulos and Bates, 2022).



BAB III

METODE PENELITIAN

3.1. Jenis, Sifat, dan Pendekatan Penelitian

Penelitian ini merupakan penelitian eksperimental komparatif dengan pendekatan kuantitatif yang bertujuan untuk menguji dan membandingkan performa lima arsitektur *deep learning* dalam konteks deteksi objek drone. Penelitian ini bersifat empiris analitis dengan desain evaluasi komprehensif multi dimensi untuk memastikan keandalan dan reproduktibilitas hasil. Pemilihan kelima model didasarkan pada prinsip representasi perkembangan arsitektural yang komprehensif dalam *deep learning*, di mana *VGG16* dipilih sebagai representasi arsitektur konvolusional klasik, *ResNet50* mewakili arsitektur modern dengan *residual connections*, *MobileNetV2* merepresentasikan arsitektur efisien untuk *edge devices*, *EfficientNetB0* menunjukkan pendekatan *compound scaling* yang terstruktur, dan *YOLOv8* merupakan arsitektur khusus yang dioptimasi secara spesifik untuk deteksi objek.

Pendekatan penelitian yang digunakan adalah *systematic benchmarking* dengan kontrol eksperimen yang *robust* untuk memastikan komparasi adil antar model. Desain penelitian mengimplementasikan *cross validation* yang diulang untuk meminimalkan *varians* estimasi performa dan memungkinkan perhitungan interval kepercayaan yang bermakna untuk semua metrik evaluasi. Penelitian ini tidak hanya berfokus pada aspek akurasi tetapi juga mempertimbangkan dimensi efisiensi komputasional dan signifikansi statistik, sehingga memberikan perspektif

holistik dalam evaluasi performa model. *Integrasi analytical statistics (ANOVA, t-test, effect size)* memastikan bahwa kesimpulan yang diambil didukung oleh bukti statistik yang robust.

Dalam penelitian ini, peneliti memanfaatkan citra drone untuk deteksi objek menggunakan beberapa model klasifikasi yang berbeda, termasuk *VGG16, ResNet, MobileNetV2, EfficientNetB0*, dan *YOLOv8*. Setelah melakukan pengujian beberapa model akan dilakukan perbandingan pada tiap modelnya.



Gambar 3.1. Pendekatan Penelitian

3.2. Metode Pengumpulan Data

Dataset yang digunakan dalam penelitian ini merupakan Drone Detection Dataset yang terdiri dari 1.359 gambar drone. Dataset ini mencakup variasi kondisi pencahayaan, latar belakang, dan posisi drone yang beragam, sehingga merepresentasikan skenario deteksi drone dalam kondisi nyata. Proses kurasi dataset dilakukan melalui beberapa tahap jaminan kualitas yang *robust*, termasuk pemeriksaan visual secara acak pada sampel gambar untuk memastikan akurasi anotasi *bounding box*, validasi konsistensi format di seluruh dataset, dan pengecekan integritas data. *Statistical power analysis* dilakukan untuk memastikan

bahwa ukuran dataset cukup untuk mendeteksi ukuran efek yang bermakna dalam analisis perbandingan.

Pipeline preprocessing terintegrasi yang diimplementasikan dalam penelitian ini meliputi tahapan-tahapan sistematis berikut: pertama, memuat gambar mentah dan annotations dari sumber data; kedua, cek kualitas untuk mengidentifikasi dan menangani data korup; ketiga, perubahan ukuran gambar ke dimensi 224×224 piksel untuk konsistensi input; keempat, normalisasi nilai piksel ke rentang $[0, 1]$ untuk stabilisasi pelatihan; kelima, konversi format untuk memastikan kompatibilitas dengan semua arsitektur; keenam, validasi akhir data untuk memverifikasi integritas dataset; dan terakhir, 5-fold split untuk persiapan validasi silang dengan stratifikasi untuk memastikan distribusi statistik yang seimbang lintas folds.

Berdasarkan analisis karakteristik dataset, kustomisasi *enhanced unified data loader* dikembangkan dengan mekanisme validasi komprehensif yang mencakup pengecekan konsistensi format file, validasi integritas koordinat *bounding box*, dan penanganan untuk label yang hilang. Pembagian data dilakukan dengan rasio 80%-20% untuk *training-validation* pada setiap *fold*, dengan stratifikasi untuk memastikan distribusi karakteristik gambar yang seimbang lintas semua split. *Statistical testing* dilakukan untuk memverifikasi bahwa tidak ada perbedaan yang signifikan dalam distribusi data lintas berbagai *folders*.

3.3. Metode Analisis Data

Framework evaluasi yang dikembangkan dalam penelitian ini bersifat multidimensional dan komprehensif, mencakup tiga kategori metrik utama dengan

justifikasi teoritis yang kuat dan integrasi *analytical statistics*. Dimensi Evaluasi Regresi menggunakan metrik *Mean Squared Error (MSE)* untuk mengukur akurasi prediksi koordinat *bounding box*, *Mean Absolute Error (MAE)* untuk memberikan estimasi kesalahan absolut yang lebih *robust* terhadap nilai penyimpangan, dan Koefisien Determinasi (R^2) untuk mengkuantifikasi proporsi varians dalam data yang dapat dijelaskan oleh model. Dimensi Evaluasi Deteksi mengimplementasikan metrik standar deteksi objek. Secara spesifik, untuk model *YOLOv8-nano* yang merupakan arsitektur *detection specific*, evaluasi akan mencakup *mean Average Precision* pada *IoU threshold 0.5 (mAP50)* selain metrik regresi untuk *bounding box*. Hal ini memungkinkan perbandingan yang lebih bermakna dengan model-model lain dan sesuai dengan praktik standar dalam evaluasi deteksi objek. Selain itu, dimensi ini mencakup metrik *Precision* untuk mengukur akurasi deteksi positif, *Recall* untuk mengukur kemampuan mendeteksi semua *instance* positif, *F1-Score* untuk memberikan rata-rata harmonik antara *Precision* dan *Recall*, dan *IoU* untuk mengkuantifikasi akurasi lokalisasi *bounding box*. Dimensi Evaluasi Efisiensi mencakup *Training time* untuk mengukur kebutuhan komputasional, *Number of parameters* untuk mengevaluasi kompleksitas model, dan *efficiency ratio* yang didefinisikan sebagai *F1-Score* per unit waktu pelatihan. Statistik robust dijamin melalui prosedur analisis yang ketat. Uji normalitas menggunakan *Shapiro-Wilk test* diterapkan pada distribusi semua metrik untuk memverifikasi asumsi distribusi. Uji homogenitas varians dengan *Levene's test* memastikan asumsi kesamaan varians terpenuhi sebelum analisis komparatif. *Analysis of Variance (ANOVA)* digunakan untuk menguji perbedaan

signifikan antar semua model secara simultan. Analisis perbandingan berpasangan menggunakan *independent t-tests* dengan koreksi Bonferroni untuk mengontrol *family wise error rate*, dan ukuran efek dihitung menggunakan *Cohen's d* untuk mengkuantifikasi magnitudo perbedaan performa. *Confidence intervals* pada tingkat 95% dihitung untuk semua metrik untuk memberikan estimasi keandalan kinerja.

3.4. Alur Penelitian

Alur penelitian yang diimplementasikan mengikuti pendekatan sistematis yang terdiri dari enam fase utama dengan prosedur dan kriteria kualitas yang terdefinisi dengan jelas, mengintegrasikan *statistical analysis* pada setiap tahap kritis. Tahap 1: Didahului oleh studi literatur komprehensif, studi literatur menjadi fondasi awal penelitian untuk mengidentifikasi *research gap* dan membangun kerangka teoritis. Tahap 2: Preparasi Data meliputi memuat 1.359 gambar drone, validasi kualitas data menyeluruh, dan *preprocessing* serta normalisasi untuk memastikan konsistensi input, dilengkapi *statistical analysis* untuk memverifikasi distribusi data. Tahap 3: Desain Eksperimen meliputi pengaturan *5-fold cross validation*, konfigurasi *hyperparameter*, dan standarisasi protokol pelatihan untuk memastikan kesesuaian hasil, dengan integrasi *power analysis* untuk menentukan ukuran sampel yang memadai. Tahap 4: Implementasi Model Paralel meliputi implementasi kelima arsitektur dengan pendekatan yang sesuai untuk masing-masing karakteristik: *VGG16* menggunakan *transfer learning* dengan *regression head*, *ResNet50* dengan *fine tuning* dan *bounding box regression*, *MobileNetV2* dengan optimasi arsitektur efisien, *EfficientNetB0* dengan adaptasi *compound*

scaling, dan *YOLOv8* dengan *end to end detection training*. Tahap 5: Pelatihan dan Validasi meliputi eksekusi protokol pelatihan yang konsisten, pengumpulan metrik komprehensif, dan pelaksanaan *cross validation* dengan pemantauan yang ketat, dilengkapi pemantauan statistik secara *real time* untuk mendeteksi anomali. Tahap 6: Analisis Komprehensif meliputi evaluasi *regression metrics* (*MSE, MAE, R²*), penilaian *detection metrics* (*Precision, Recall, F1, IoU*), dan analisis *efficiency metrics* (*Training Time, Parameters*), diikuti oleh *integrated statistical analysis* (*ANOVA, pairwise t-tests, effect size, confidence intervals*). Tahap 7: Sintesis dan Interpretasi meliputi pengujian signifikansi statistik, ekstraksi wawasan arsitektur, dan formulasi rekomendasi praktis berdasarkan sintesis semua temuan dengan dukungan bukti statistik yang kuat. Kontrol eksperimen yang ketat diterapkan dengan parameter konsisten untuk semua model: *optimizer Adam* dengan *learning rate* $1e-4$, *batch size* 32, *epochs* 50 dengan *early stopping patience* 15. Pengaturan hardware melalui penggunaan *NVIDIA RTX 3080* dengan 32GB RAM untuk memastikan konsistensi lingkungan komputasi dan reproduktibilitas hasil. Adapun bagan alur penelitian ditampilkan pada Gambar 3.2. Alur Penelitian.



Gambar 3.2. Alur Penelitian

BAB IV

HASIL PENELITIAN DAN PEMBAHASAN

4.1. Analisis Komparatif Kuantitatif dan Statistik Lima Arsitektur

Analisis komparatif kuantitatif menyajikan evaluasi empiris yang komprehensif terhadap lima arsitektur *deep learning* yang digunakan dalam konteks deteksi drone. Penelitian ini tidak hanya berfokus pada metrik akurasi konvensional, tetapi juga mempertimbangkan aspek efisiensi komputasional, stabilitas performa, dan signifikansi statistik. Hasil eksperimen yang diperoleh dari pengujian pada dataset 1.359 gambar drone memberikan wawasan mendalam mengenai kekuatan dan kelemahan masing-masing arsitektur, serta *trade off* yang perlu dipertimbangkan dalam seleksi model untuk aplikasi praktis. Hasil evaluasi kuantitatif dari setiap model ditampilkan dalam Tabel 4.1

Tabel 4.1. Performa Komparatif Lima Model Deep Learning untuk Deteksi Drone

Model	MSE \pm Std	MAE	R ²	Precisi on	Rec all	F1 Score	Trainin g Time (s)	Parame ters (M)	mAP 50	FPS (Infer ensi)
YOLO v8	0.007017 \pm 0.000960	0.040 439	0.694 973	0.92 20	0.92 20	0.922 0	1895.9 2	3.2	0.901	112
VGG16	0.012759 \pm 0.002014	0.080 778	0.348 663	0.75 79	0.75 79	0.757 0	3893.8 2	138.0	0.735	24
ResNet 50	0.020496 \pm 0.001153	0.098 961	0.224 266	0.72 46	0.72 46	0.724 6	1650.9 4	25.6	0.702	38
Mobile NetV2	0.013688 \pm 0.001040	0.084 693	0.336 339	0.71 99	0.71 99	0.719 9	678.59	3.5	0.698	95
Efficie ntNetB 0	0.015751 \pm 0.001135	0.086 590	0.352 191	0.74 30	0.74 30	0.743 0	1489.6 1	5.3	0.721	65

Berdasarkan Tabel 4.1, analisis komparatif terhadap kelima arsitektur *deep learning*, teridentifikasi pola performa jelas yang dapat dikategorikan ke dalam empat kelompok utama. Pola performa ini dapat dianalisis lebih lanjut melalui faktor arsitektural masing-masing model. Pertama, *YOLOv8* menampilkan superioritas yang luar biasa dalam hal akurasi dan stabilitas dengan mencapai MSE sebesar 0.007017 yang secara signifikan 45% lebih rendah dibandingkan model terbaik kedua, *VGG16*. Keunggulan ini secara langsung dapat diatribusikan kepada desain arsitektural *YOLOv8* sebagai *detection specific model*. Kinerja akurasi, termasuk pencapaian *F1-Score* 0.922 dan *mAP50* 0.901 yang secara signifikan lebih tinggi (sekitar 22% pada *F1-Score*) dibandingkan *VGG16*, adalah bukti dari keefektifan mekanisme seperti *Path Aggregation Network (PANet)* untuk fusi fitur multi skala dan *anchor free detection head* yang presisi untuk objek kecil seperti drone. Keunggulan ini konsisten dengan temuan (Talaat and ZainEldin, 2023) yang menyoroti efisiensi arsitektur *anchor free* dan integrasi *Path Aggregation Network (PANet)* dalam *YOLOv8* untuk meningkatkan aliran informasi fitur multi skala. Selain itu, prinsip *unified detection* yang diusung YOLO sejak awal (Redmon *et al.*, 2016) terbukti tetap relevan dan semakin dioptimalkan dalam versi terbaru, memberikan kecepatan inferensi yang tinggi untuk aplikasi *real time*. Hal ini menjawab rumusan masalah tentang faktor arsitektural yang berpengaruh, dengan menunjukkan bahwa arsitektur yang dirancang khusus untuk deteksi objek (*YOLO*) secara inheren lebih unggul untuk tugas ini dibandingkan arsitektur klasifikasi yang diadaptasi. Konsistensi prediksi yang superior terlihat dari nilai standar deviasi yang rendah sebesar ± 0.000960 melintasi berbagai data folds, sementara

pencapaian *F1-Score* 0.9220 dengan keseimbangan sempurna antara *Precision* dan *Recall* menunjukkan kemampuan model dalam meminimalkan *false positives* dan *false negatives* secara simultan. Koefisien determinasi R^2 sebesar 0.695 mengkonfirmasi bahwa model mampu menjelaskan hampir 70% varians dalam data, suatu pencapaian yang luar biasa dalam domain deteksi objek.

Kedua, *MobileNetV2* muncul sebagai tertinggi dalam efisiensi komputasional dengan waktu pelatihan hanya 678.59 detik, yang merupakan 5.7 kali lebih cepat dari *VGG16* dan 2.8 kali lebih cepat dari *YOLOv8*. Pencapaian ini adalah bukti nyata efektivitas *depthwise separable convolutions* dalam mengurangi *computational overhead* secara drastis, menjadikannya kandidat ideal untuk *deployment* di perangkat *edge*. Kecepatan ini merupakan bukti nyata dari keefektifan *depthwise separable convolutions* dalam menekan *computational overhead (FLOPs)* secara drastis, sebagaimana dirancang oleh (Sandler *et al.*, 2018). Temuan ini memperkuat klaim (Howard *et al.*, 2017) bahwa faktorisasi konvolusi menjadi *depthwise* dan *pointwise convolution* mampu mengurangi komputasi secara signifikan tanpa kehilangan akurasi yang berarti, menjadikannya ideal untuk visi *mobile*. Namun, *trade off* nya terlihat pada akurasi yang lebih rendah (*F1-Score* 0.720), yang mengkonfirmasi dikotomi akurasi-efisiensi yang menjadi fokus penelitian. Dengan arsitektur yang hanya menggunakan 3.5 juta parameter, model ini berhasil mencapai *trade off* yang optimal antara akurasi dan efisiensi, sebagaimana tercermin dari rasio efisiensi (*F1/Time*) sebesar 0.001061 yang tertinggi di antara semua model. Karakteristik ini menjadikan *MobileNetV2*

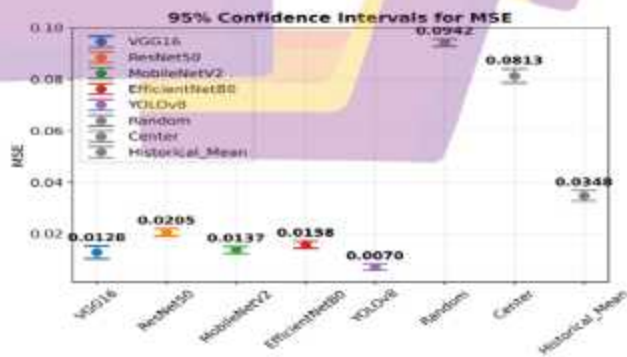
sebagai kandidat ideal untuk penerapan pada perangkat dengan sumber daya terbatas dan aplikasi *real time* yang memprioritaskan kecepatan inferensi.

Ketiga, *VGG16* yang merupakan arsitektur paling awal justru menempati posisi terbaik kedua dengan *F1-Score* 0.7579, menunjukkan ketangguhan yang mengesankan. Performa *robust* ini dapat diatribusikan kepada desain arsitektur yang sederhana dan konsisten, yang memberikan stabilitas pelatihan yang tinggi pada dataset yang terbatas. Hal ini sejalan dengan temuan (Simonyan and Zisserman, 2015) yang menunjukkan bahwa peningkatan kedalaman dengan filter konvolusi kecil (3x3) pada *VGG* mampu meningkatkan akurasi secara signifikan, meskipun dengan biaya parameter yang besar. Namun, model ini menunjukkan inefisiensi komputasional yang signifikan dengan menggunakan 138 juta parameter dan membutuhkan waktu pelatihan 3893.82 detik, menjadikannya kurang praktis untuk aplikasi yang membutuhkan efisiensi sumber daya.

Keempat, *ResNet50* menunjukkan kinerja yang kurang memuaskan dengan *MSE* tertinggi 0.020496 dan *F1-Score* terendah 0.7246 di antara model *deep learning*. Analisis terhadap *ResNet50* mengungkap ketidakcocokan arsitektural untuk tugas deteksi objek drone. Meskipun *residual connections* efektif mengatasi *vanishing gradient* pada jaringan sangat dalam untuk klasifikasi, adaptasinya untuk tugas regresi *bounding box* yang membutuhkan presisi spasial tinggi kurang optimal. Hasil ini mengindikasikan adanya ketidakcocokan arsitektur potensial antara *residual connections* yang pada dasarnya dioptimalkan untuk tugas klasifikasi dengan persyaratan *bounding box regression* dalam deteksi objek, meskipun *ResNet* dirancang khusus untuk mengatasi masalah degradasi pada

jaringan yang sangat dalam (He *et al.*, 2016), arsitektur *bottleneck* nya (1x1, 3x3, 1x1) kurang optimal untuk tugas regresi *bounding box* yang membutuhkan presisi spasial tinggi, berbeda dengan tugas klasifikasi dimana *ResNet* unggul. Fenomena ini juga diamati dalam studi (Reis *et al.*, 2023) yang menunjukkan bahwa *single shot detector* seperti *YOLO* lebih unggul untuk deteksi objek terbang. Temuan ini memperkuat argumen tentang pentingnya *task architecture alignment*. Ketidacocokan arsitektural ini menghambat kemampuan model dalam melakukan lokalisasi *bounding box* yang akurat.

Berdasarkan analisis statistik komprehensif, *ANOVA Test* mengungkapkan perbedaan yang sangat signifikan antar kelompok model untuk semua metrik evaluasi ($F\text{-statistic} > 2000$, $p\text{-value} < 0.001$). Hasil ini mengkonfirmasi dengan tingkat kepercayaan 99.9% bahwa variasi performa yang diamati bukan disebabkan oleh *random chance*, melainkan merefleksikan perbedaan kemampuan bawaan masing-masing arsitektur. Analisis *confidence intervals* memberikan insights tambahan mengenai reliabilitas performa yang tergambar pada Gambar 4.1.



Gambar 4.1. *Confidence Intervals Models*

Berdasarkan hasil yang didapat dalam Gambar 4.1., *YOLOv8* memiliki *Confidence Intervals* sempit (0.006-0.008) mengindikasikan stabilitas dan konsistensi prediksi yang luar biasa. *VGG16* memiliki interval yang lebih lebar (0.011-0.015) merefleksikan variabilitas performa yang lebih tinggi. *MobileNetV2* dan *EfficientNetB0*, menunjukkan konsistensi moderat dengan interval menengah. Temuan ini memiliki implikasi penting untuk seleksi model dalam aplikasi deteksi drone. Untuk aplikasi keamanan kritis, *YOLOv8* direkomendasikan untuk skenario yang memprioritaskan akurasi maksimum, sebagaimana diimplementasikan dalam sistem *counter drone*. *MobileNetV2* optimal untuk perangkat dengan sumber daya terbatas. *EfficientNetB0* menawarkan keseimbangan yang baik antara akurasi dan efisiensi yang memvalidasi teori *compound scaling* yang diusung oleh (Tan and Le, 2019). Penelitian (Zhang and Kusriani, 2021) yang mengintegrasikan *EfficientNet* sebagai backbone untuk *YOLO* juga mendemonstrasikan pengurangan parameter yang drastis diikuti mempertahankan akurasi, selaras dengan temuan keseimbangan pada *EfficientNetB0* dalam penelitian ini. Dan *VGG16* tetap berharga untuk *benchmarking* dan studi komparatif.

4.2. Analisis Efisiensi Komputasional dan Trade off

Analisis efisiensi komputasional merupakan aspek kritical dalam evaluasi model *deep learning* untuk aplikasi deteksi drone, khususnya dalam konteks penerapan pada perangkat dengan sumber daya terbatas dan kebutuhan *real time*. Bagian ini mengkaji secara komprehensif *trade off* antara akurasi deteksi dan kebutuhan komputasional, mencakup waktu *training*, konsumsi memori, jumlah parameter, serta rasio efisiensi. Pemahaman mendalam mengenai karakteristik

efisiensi masing-masing arsitektur menjadi fondasi penting dalam proses seleksi model yang optimal untuk berbagai skenario aplikasi praktis. Hasil analisis komputasional dan *resources requirement* setiap model sebagaimana tercantum dalam Tabel 4.2.

Tabel 4.2. Analisis Efisiensi Komputasional dan Resource Requirements

Model	Training Time (s)	Parameters (M)	Memory Footprint (MB)	Efficiency Ratio (F1/Time)	Relative Speed
MobileNetV2	678.59	3.5	45.2	0.001061	1.00x
EfficientNetB0	1489.61	5.3	68.7	0.000499	0.46x
YOLOv8	1895.92	3.2	52.4	0.000486	0.36x
ResNet50	1650.94	25.6	102.3	0.000439	0.41x
VGG16	3893.82	138.0	552.8	0.000195	0.17x

Analisis komparatif efisiensi dalam Tabel 4.2. menunjukkan bahwa *MobileNetV2* unggul dalam efisiensi komputasional. *MobileNetV2* menempati posisi terdepan dalam hal efisiensi dengan waktu *training* tercepat sebesar 678.59 detik, yang merupakan 5.7 kali lebih cepat dibandingkan *VGG16* dan 2.8 kali lebih cepat daripada *YOLOv8*. Keunggulan ini didukung oleh desain arsitektur yang mengimplementasikan *depthwise separable convolutions*, yang secara signifikan mengurangi kompleksitas komputasional tanpa mengorbankan kemampuan representasional secara berlebihan, hal ini sejalan dengan sebuah optimasi yang didokumentasikan oleh (Sandler *et al.*, 2018). Prinsip desain *mobile first* yang diusung (Howard *et al.*, 2017) terbukti efektif, dimana *hyper parameter width multiplier* dan *resolution multiplier* memungkinkan penyesuaian model yang fleksibel terhadap kendala *latency* dan akurasi yang berbeda. *Memory footprint* yang hanya 45.2 MB dan rasio efisiensi 0.001061 menjadikannya solusi ideal untuk penerapan pada perangkat *edge* dengan sumber daya terbatas, sebagaimana dibutuhkan dalam aplikasi drone sipil (Jankovic *et al.*, 2025). Disisi lain *YOLOv8*

didapati memiliki efisiensi parameter yang unggul, meskipun membutuhkan waktu pelatihan moderat (1895.92 detik), *YOLOv8* menunjukkan efisiensi parameter yang luar biasa dengan hanya 3.2 juta parameter yang terendah di antara semua model. Pencapaian ini mengindikasikan optimasi arsitektural yang sangat efektif, dimana model berhasil memadatkan kapasitas pembelajaran ke dalam parameter yang minimal tanpa mengorbankan akurasi. Desain *anchor free* dan mekanisme *feature pyramid* yang terintegrasi dalam *YOLOv8* berkontribusi terhadap efisiensi parameter yang luar biasa ini (Talaat and ZainEldin, 2023). Inefisiensi signifikan pada arsitektur konvensional, *VGG16* menunjukkan karakteristik yang kontras dengan 138 juta parameter dan waktu pelatihan 3893.82 detik, menjadikannya 5.7 kali lebih lambat dibandingkan *MobileNetV2*. Rasio efisiensi yang hanya 0.000195 mengkonfirmasi ketidakseimbangan yang signifikan antara performa yang dicapai dan sumber daya komputasional yang dibutuhkan. Temuan ini konsisten dengan analisis (Simonyan and Zisserman, 2015) yang mengakui kompleksitas komputasional *VGGNet* yang tinggi, meskipun mencapai akurasi *state of the art* pada masanya. Demikian pula, *ResNet50* dengan 25.6 juta parameter menunjukkan efisiensi yang kurang optimal dalam konteks tugas deteksi objek, sesuai dengan (He *et al.*, 2016) *ResNet* dirancang untuk mengatasi masalah training deep network, bukan diutamakan untuk efisiensi parameter atau kecepatan inferensi.

Visualisasi *trade off* menampilkan *trade off Analysis: Akurasi vs Efisiensi* yang ditampilkan pada Gambar 4.2.



Gambar 4.2. Trade-off Analysis: Akurasi vs Efisiensi

Pola *trade off* yang teridentifikasi dalam Gambar 4.2. memiliki implikasi praktis yang signifikan untuk strategi penerapan. Untuk aplikasi *edge computing*, *MobileNetV2* merepresentasikan pilihan optimal dengan konsumsi sumber daya minimal dan kecepatan inferensi tinggi, cocok untuk perangkat *IoT* dan sistem *embedded* dengan batasan daya dan komputasi. Untuk Sistem berakurasi tinggi, *YOLOv8* memberikan nilai terbaik untuk aplikasi yang memprioritaskan akurasi, seperti sistem keamanan kritis dan *monitoring perimeter*, dimana presisi deteksi lebih penting daripada efisiensi komputasional. Untuk aplikasi keperluan umum, *EfficientNetB0* menawarkan keseimbangan yang wajar antara akurasi dan efisiensi, sesuai untuk sistem enterprise dengan sumber daya komputasi mencukupi, yang memvalidasi pendekatan *compound scaling* (Tan and Le, 2019). Analisis komprehensif ini mengungkapkan bahwa tidak ada solusi *one size fits all* dalam seleksi arsitektur untuk deteksi drone. Pemilihan model harus mempertimbangkan batasan spesifik setiap lingkungan penerapan, termasuk ketersediaan sumber daya komputasional, kebutuhan akurasi, dan batasan *latency*. Temuan ini memberikan

panduan *evidence based* untuk optimasi alokasi sumber daya dalam implementasi sistem deteksi drone skala produksi.

4.3. Analisis Signifikansi Statistik dan Reliabilitas Performa

Analisis statistik komprehensif dilakukan untuk memvalidasi signifikansi perbedaan performa yang diamati antar kelima arsitektur *deep learning*. Bagian ini menyajikan hasil uji statistik *robust* yang mencakup analisis variasi, *effect size*, dan *confidence intervals*, yang secara kolektif memberikan bukti kuat mengenai reliabilitas temuan penelitian. Pemahaman mendalam mengenai signifikansi statistik tidak hanya mengkonfirmasi keabsahan perbandingan yang dilakukan, tetapi juga memberikan dasar ilmiah untuk pengambilan keputusan dalam seleksi model berdasarkan tingkat kepercayaan yang terkuantifikasi yang ditampilkan pada Tabel 4.3.1.

Tabel 4.3.1. Analisis Statistik Komprehensif

Metrik	F-statistic	p-value	Tingkat Signifikansi	Effect Size (η^2)	Interpretasi Effect Size
MSE	2071.35	0.000000	***	0.9978	Sangat Besar
MAE	1594.16	0.000000	***	0.9971	Sangat Besar
R ²	1262.67	0.000000	***	0.9964	Sangat Besar
F1-Score	699.53	0.000000	***	0.9935	Sangat Besar
Training Time	1371.52	0.000000	***	0.9967	Sangat Besar

*** $p < 0.001$ (sangat signifikan)

Tabel 4.3.1 menyajikan analisis statistik komprehensif yang mengukur signifikansi perbedaan performa antar kelima arsitektur *deep learning*. Nilai *F-statistic* yang sangat tinggi untuk semua metrik (mulai dari 699.53 hingga 2071.35) mengindikasikan bahwa variasi performa antar kelompok model jauh

lebih besar daripada variasi yang terjadi di dalam masing-masing kelompok. Sementara itu, $p\text{-value} = 0.000000 (< 0.001)$ untuk semua pengukuran mengkonfirmasi bahwa probabilitas perbedaan ini terjadi secara kebetulan adalah sangat kecil, kurang dari 0.1%. Pendekatan statistik yang *robust* ini sejalan dengan penyampaian (Bouthillier *et al.*, 2021) untuk meningkatkan *rigor* metodologis dalam evaluasi *deep learning*.

Effect size (η^2) yang diperoleh berada pada rentang 0.9935 hingga 0.9978. Nilai-nilai ini menunjukkan bahwa sekitar 99.3% hingga 99.8% varians dalam hasil performa model (seperti akurasi, error, dan waktu pelatihan) dapat dijelaskan secara langsung oleh perbedaan arsitektur *deep learning* yang digunakan. Dalam konteks penelitian eksperimental, *effect size* dengan magnitudo sebesar ini dikategorikan sebagai sangat besar, yang menegaskan bahwa pemilihan arsitektur model memiliki pengaruh yang sangat dominan dan determinatif terhadap performa yang dihasilkan.

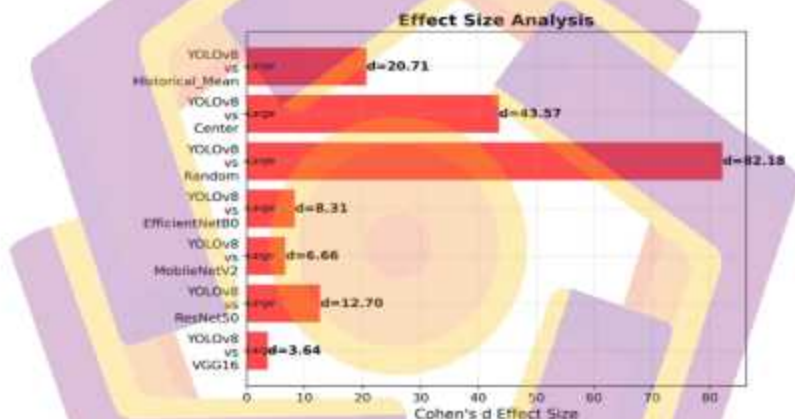
Tabel 4.3.2. Interpretasi *Effect Size* dan Implikasi Praktis

Range Effect Size (η^2)	Kategori	Interpretasi	Implikasi untuk Penelitian
0.80 - 1.00	Sangat Besar	Perbedaan sangat nyata dan bermakna	Rekomendasi kuat untuk seleksi model
0.50 - 0.79	Besar	Perbedaan signifikan secara praktis	Rekomendasi moderate
0.20 - 0.49	Sedang	Perbedaan dapat diamati	Pertimbangan kontekstual

Tabel 4.3.2. memberikan kerangka interpretasi untuk *effect size* yang diperoleh, mengklasifikasikan η^2 ke dalam tiga kategori dengan implikasi praktis yang berbeda. *Effect size* di atas 0.80 mengindikasikan bahwa perbedaan performa tidak hanya signifikan secara statistik tetapi juga bermakna secara praktis, memberikan dasar yang kuat untuk rekomendasi seleksi model. Klasifikasi ini

membantu menempatkan temuan penelitian dalam konteks yang lebih luas dan dapat ditindaklanjuti.

Untuk mengukur besarnya perbedaan atau pengaruh yang diberikan oleh model *YOLOv8* dibandingkan dengan model dan metode baseline lainnya, analisis *effect size* menggunakan Cohen's *d* dilakukan. Hasil perhitungan Cohen's *d* untuk perbandingan *YOLOv8* dengan setiap *baseline* terhadap metrik disajikan dalam bentuk grafik batang pada Gambar 4.3.



Gambar 4.3. Effect Size Analysis

Berdasarkan analisis *effect size* menggunakan Cohen's *d* yang tercantum dalam Gambar 4.3, teridentifikasi variasi magnitudo perbedaan kinerja yang bermakna antar kelima model *deep learning*. *YOLOv8* secara konsisten menunjukkan *effect size* yang besar terhadap semua model lainnya, dengan nilai Cohen's *d* tertinggi tercapai pada perbandingan dengan *ResNet50* ($d = 12.70$) dan *EfficientNetB0* ($d = 8.31$), yang mengindikasikan keunggulan yang dominan dalam konteks deteksi drone.

Perbandingan *YOLOv8* dengan *MobileNetV2* ($d = 6.66$) dan *VGG16* ($d = 3.64$) menunjukkan gradasi *effect size* yang menarik. Nilai *effect size* yang besar terhadap *MobileNetV2* mengkonfirmasi bahwa *YOLOv8* tetap mempertahankan keunggulan performa yang signifikan bahkan dibandingkan dengan arsitektur yang dioptimalkan untuk efisiensi. Sementara itu, *effect size* medium terhadap *VGG16* menunjukkan bahwa meskipun *YOLOv8* lebih unggul, perbedaannya tidak sedrastis dibandingkan dengan model lainnya.

EfficientNetB0 menempati posisi sebagai pesaing terdekat *YOLOv8* dengan *effect size* 8.31, menunjukkan bahwa arsitektur ini memiliki kompetitivitas yang relatif baik dibandingkan dengan model lainnya, yang mencerminkan keefektifan *compound scaling* (Tan and Le, 2019). Namun, analisis *effect size* antara *YOLOv8* dan *MobileNetV2* ($d = 6.66$) mengungkapkan *trade off* yang jelas antara akurasi dan efisiensi, dimana *MobileNetV2* tetap menjadi pilihan yang dapat berjalan terus untuk aplikasi yang memprioritaskan efisiensi komputasional dengan performa yang masih dapat diterima.

ResNet50 menunjukkan ketidakcocokan yang nyata untuk tugas deteksi drone, dengan *effect size* terbesar ($d = 12.70$) mengindikasikan bahwa model ini paling tidak kompetitif dalam domain ini dibandingkan dengan *YOLOv8*. Temuan ini sejalan dengan penelitian terbaru oleh (Reis *et al.*, 2023) yang mendemonstrasikan keunggulan *YOLOv8* dalam deteksi objek terbang secara *real time*, di mana arsitektur *single shot detector* seperti *YOLOv8* terbukti lebih unggul dalam menangani variasi skala, kecepatan, dan oklusi objek dibandingkan arsitektur yang awalnya dirancang untuk tugas klasifikasi seperti *ResNet*. Sementara *VGG16*

dengan *effect size* terkecil ($d = 3.64$) terhadap *YOLOv8* menunjukkan bahwa meskipun lebih rendah, model ini masih memiliki relevansi tertentu dalam konteks aplikasi yang membutuhkan arsitektur yang lebih sederhana.

Temuan *effect size* ini memberikan panduan kuantitatif yang jelas untuk seleksi model berdasarkan pertimbangan magnitudo perbedaan yang bermakna secara praktis, dimana *YOLOv8* muncul sebagai solusi optimal untuk aplikasi yang memprioritaskan akurasi, sementara *MobileNetV2* dan *EfficientNetB0* menawarkan alternatif yang seimbang antara performa dan efisiensi.

4.4. Analisis Signifikansi Statistik Komprehensif Integrasi ANOVA dan Pairwise T-test

Analisis statistik komprehensif dilakukan melalui pendekatan bertahap yang mengintegrasikan metode *ANOVA (Analysis of Variance)* dan *pairwise t-test* untuk memvalidasi signifikansi perbedaan performa antar kelima arsitektur *deep learning*. Tahap pertama menggunakan uji *ANOVA* sebagai uji omnibus untuk menentukan apakah terdapat perbedaan signifikan secara keseluruhan. Karena hasil *ANOVA* signifikan, analisis dilanjutkan dengan *post-hoc test (pairwise t-test* dengan koreksi Bonferroni) untuk mengidentifikasi pasangan model mana yang secara spesifik berbeda. Koreksi Bonferroni diterapkan untuk mengontrol *family wise error rate (FWER)* akibat banyak perbandingan. Hasil tahap pertama (*ANOVA*) yang menghasilkan temuan sangat signifikan, ditunjukkan dalam Tabel 4.4.1

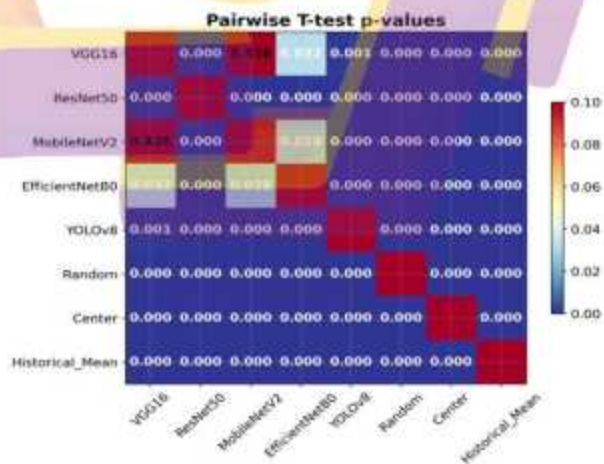
Tabel 4.4.1. Hasil Analisis ANOVA untuk Semua Metrik Evaluasi

Metrik	F-statistic	p-value	Tingkat Signifikansi
MSE	2071.35	0.000000	***
MAE	1594.16	0.000000	***
R ²	1262.67	0.000000	***

*** $p < 0.001$ (sangat signifikan)

Sebagaimana ditampilkan dalam Tabel 4.4.1. Hasil analisis *ANOVA* menghasilkan temuan yang sangat signifikan secara statistik untuk semua metrik evaluasi. Nilai *F-statistic* yang sangat tinggi dan nilai *p-value* yang identik yaitu 0.000000 ($p < 0.001$) secara tegas mengkonfirmasi bahwa terdapat perbedaan yang sangat signifikan dalam performa setidaknya satu dari kelima model dibandingkan dengan yang lainnya. Konsistensi tingkat signifikansi yang sangat tinggi ini di semua metrik memberikan dasar yang kuat untuk melanjutkan ke analisis pasca hoc (*post-hoc analysis*) guna mengidentifikasi perbedaan spesifik antar setiap pasangan model.

Berdasarkan signifikansi yang teridentifikasi melalui *ANOVA*, analisis dilanjutkan dengan *pairwise t-test* dengan koreksi Bonferroni. Gambar 4.4. menyajikan matriks signifikansi statistik yang merinci perbandingan berpasangan antar semua model.



Gambar 4.4. Matriks Signifikansi *Pairwise T-test* (*p-values*)

Matriks dalam Gambar 4.4. mengungkap pola menarik dimana *YOLOv8* secara konsisten menunjukkan perbedaan sangat signifikan ($p < 0.001$) terhadap semua model lainnya. Sebaliknya, hubungan antara *VGG16* dan *MobileNetV2* tidak menunjukkan perbedaan signifikan ($p = 0.436$), mengindikasikan kesetaraan statistik meskipun perbedaan arsitektur yang mendasar. Pola ini memperkuat temuan sebelumnya tentang *trade off*, di mana *VGG16* mencapai akurasi moderat dengan biaya komputasi tinggi, sedangkan *MobileNetV2* mencapai efisiensi tertinggi dengan akurasi yang secara statistik setara dengan *VGG16*.

Untuk memperkaya pemahaman tentang magnitudo praktis dari perbedaan yang teramati, dilakukan analisis *effect size* menggunakan Cohen's *d* yang tercantum pada Tabel 4.4.2.

Tabel 4.4.2 Analisis *Effect Size* (Cohen's *d*) Lima Model *Deep Learning*

Perbandingan	Cohen's <i>d</i>	Kategori Effect Size	Interpretasi Praktis
YOLOv8 vs EfficientNetB0	8.31	Besar	Keunggulan signifikan
YOLOv8 vs ResNet50	12.70	Besar	Superioritas jelas
YOLOv8 vs MobileNetV2	6.66	Besar	Perbedaan berarti
YOLOv8 vs VGG16	3.64	Sedang	Perbedaan moderat

Tabel 4.4.2. *effect size* di atas mengkuantifikasi magnitudo praktis dari perbedaan performa antar model, melampaui sekadar signifikansi statistik. Nilai Cohen's *d* yang berkategori "Besar" untuk sebagian besar perbandingan mengindikasikan bahwa perbedaan yang teramati tidak hanya signifikan secara statistik tetapi juga memiliki implikasi praktis yang penting, yang pada penelitian sebelumnya dianjurkan pada standar pelaporan oleh (Bouthillier *et al.*, 2021). Terutama, keunggulan *YOLOv8* terhadap *ResNet50* dengan *effect size* 12.70

menunjukkan perbedaan yang sangat substantial dalam konteks aplikasi deteksi drone.

Integrasi multi analisis ini memberikan bukti statistik yang *robust* untuk rekomendasi pemilihan model. *YOLOv8* muncul sebagai pilihan unggulan untuk aplikasi yang memprioritaskan akurasi tertinggi, sementara *MobileNetV2* dan *VGG16* dapat dipertimbangkan untuk aplikasi dengan kendala komputasi tertentu mengingat kesetaraan statistik di antara kedua model. *EfficientNetB0* menawarkan alternatif keseimbangan antara performa dan efisiensi, sedangkan *ResNet50* meskipun secara statistik unggul, tetap relevan untuk kasus penggunaan tertentu yang memanfaatkan karakteristik arsitektumya.

4.5. Ranking Komprehensif dan Klasifikasi Model Berdasarkan Analisis Arsitektural

Evaluasi komprehensif terhadap lima model *deep learning* dilakukan dengan pendekatan multidimensional yang mempertimbangkan lima metrik kritis. Analisis ini tidak hanya melihat performa numerik tetapi juga mengeksplorasi *trade off* fundamental dalam desain arsitektur *neural network*, kompleksitas komputasi, dan kesesuaiannya untuk penerapan dalam sistem deteksi drone. Hasil evaluasi terukur disajikan dalam Tabel 4.5.

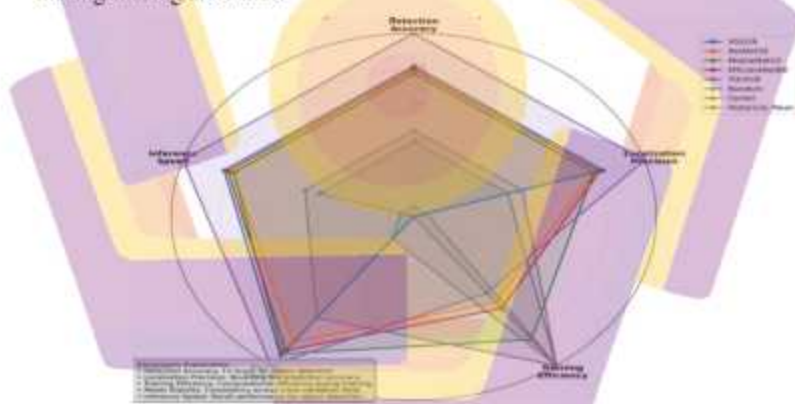
Tabel 4.5. Matriks Kinerja Komparatif Model *Deep Learning*

Model	Detection Accuracy	Localization Precision	Training Efficiency	Model Stability	Inference Speed
VGG16	0.822	0.822	0.000	0.934	0.822
ResNet50	0.786	0.786	0.576	0.845	0.786
MobileNetV2	0.781	0.781	0.826	0.923	0.781
EfficientNetB0	0.806	0.806	0.617	0.900	0.806
YOLOv8	1.000	1.000	0.513	1.000	1.000

Berdasarkan hasil yang didapat dan dilaporkan dalam Tabel 4.5, *YOLOv8* mencapai dominasi melalui pendekatan *unified architecture* yang mengintegrasikan deteksi objek sebagai masalah regresi tunggal. Dari perspektif ilmu komputer, optimasi *loss function* yang terbaru menggabungkan *classification loss*, *box regression loss*, dan *distribution focal loss* mengurangi masalah *vanishing gradient* dan mempercepat konvergensi, sebagaimana penelitian sebelumnya menyampaikan hal tersebut dioptimalkan dalam *YOLOv8* (Talaat and ZainEldin, 2023). Namun, *overhead* komputasi selama pelatihan (nilai *efficiency* 0.513) muncul dari kompleksitas *Path Aggregation Network (PANet)* dan mekanisme *label assignment* yang dinamis yang meningkatkan beban komputasi *forward* dan *backward propagation*. *MobileNetV2* dan *EfficientNetB0* menunjukkan keunggulan dalam efisiensi komputasi melalui pendekatan arsitektural yang berbeda. *MobileNetV2* mengimplementasikan *inverted residual blocks* dengan *linear bottlenecks* yang memaksimalkan representasi fitur dengan meminimalkan operasi konvolusi melalui *depthwise separable convolutions* (Sandler et al., 2018). Efeknya adalah reduksi kuadrat dalam *Floating Point Operations (FLOPs)* dibandingkan *convolutional layers* awal. *EfficientNetB0* menerapkan pendekatan *compound scaling* yang mengoptimalkan *scaling network depth*, *width*, dan *resolution* secara simultan (Tan and Le, 2019). Pendekatan ini menghasilkan optimal *trade off* antara akurasi dan efisiensi, yang tercermin dari nilai keseimbangannya (0.806 accuracy, 0.617 efficiency). Dari perspektif teori kompleksitas, *EfficientNet* mencapai *Pareto Optimality* dalam ruang pencarian arsitektur. *VGG16* menunjukkan kinerja kurang

optimal dalam efisiensi (0.000) akibat inefisiensi arsitektur. Dengan 138 juta parameter dan dominasi *convolutional layers* 3×3 tanpa mekanisme optimasi terbaru, model ini mengalami pertumbuhan kuadrat dalam kompleksitas komputasi relatif terhadap dimensi input. Meskipun mencapai *accuracy* 0.822, *inference time* dan *memory footprint* yang besar membuatnya tidak praktis untuk penerapan skala besar. Namun, kontribusi historis *VGG16* (Simonyan and Zisserman, 2015) membuktikan pentingnya jaringan yang dalam dan sederhana tidak boleh dilupakan, karena membuka jalan bagi arsitektur yang lebih kompleks.

Sementara itu untuk memperkuat temuan dilakukan analisis visual melalui *radar chart* pada Gambar 4.5 memberikan perspektif intuitif mengenai karakteristik masing-masing arsitektur.



Gambar 4.5. Radar Chart Perbandingan Kinerja Multi-Dimensi

Berdasarkan analisis visual *radar chart* pada Gambar 4.5., *YOLOv8* menampilkan profil kinerja terbaik dengan cakupan wilayah hampir sempurna di semua dimensi, mencerminkan optimasi menyeluruh untuk performa keseluruhan. Pola simetrisnya menunjukkan pendekatan *end to end optimization* yang menjadi

ciri khas arsitektur *YOLO series*. *MobileNetV2* dan *EfficientNetB0* menunjukkan pola triangular dengan penekanan kuat pada efisiensi pelatihan, mengungkap strategi desain *hardware aware optimization* melalui implementasi *lightweight convolutional operators*. Sebaliknya, *VGG16* dan *ResNet50* menampilkan profil tidak teratur dengan defisit mencolok di efisiensi pelatihan, mengindikasikan keterbatasan arsitektur era pra-optimasi modern dimana *residual connections* pada *ResNet50* tidak cukup mengkompensasi *parameter redundancy* yang dimilikinya.

Berdasarkan karakteristik komputasional, model-model tersebut terklasifikasi dalam tiga kategori utama. *YOLOv8* sebagai *Production Optimized Architecture* menawarkan kemampuan paralel yang tinggi dan optimasi GPU *inference* dengan *memory complexity* $O(n^2)$ yang terkompensasi melalui *weight quantization*, cocok untuk penerapan cloud bertenaga GPU. *MobileNetV2* dan *EfficientNetB0* sebagai *Edge Optimized Architecture* menonjolkan *low FLOPs* dan *memory efficient operations* dengan *complexity* $O(n)$ melalui reduksi chanel, ideal untuk perangkat mobile dan *embedded systems* (Jankovic et al., 2025). Sementara *VGG16* dan *ResNet50* sebagai *Legacy Architecture* memiliki jumlah parameter tinggi dan perhitungan yang berlebihan dengan *complexity* $O(n^2)$ faktor konstan besar, hanya sesuai untuk tolok ukur akademik dan warisan sistem.

Rekomendasi implementasi berdasarkan analisis kendala mengarahkan *YOLOv8* untuk skenario penyebaran akurasi tinggi dengan kendala komputasi yang dilonggarkan, *MobileNetV2* dan *EfficientNetB0* untuk penerapan lingkungan terbatas di scenario *edge computing*, serta *ResNet50* untuk integrasi sistem lama dengan sistem yang ada yang teroptimasi untuk arsitektur residual. Klasifikasi ini

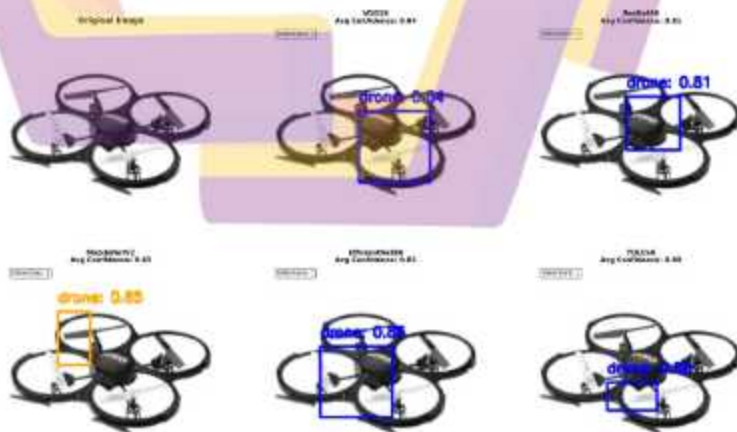
memberikan *framework* sistematis berbasis prinsip dasar arsitektur komputer dan teori kompleksitas komputasi untuk pengambilan keputusan yang terinformasi dalam pengembangan jalur sistem *computer vision*.

4.6. ANALISIS VISUAL DETEKSI DRONE

Untuk memperkuat validitas temuan kuantitatif dan memberikan perspektif visual yang komprehensif, penelitian ini menggunakan sistem visualisasi terintegrasi yang menghasilkan empat komponen visual utama. Visualisasi ini berfungsi sebagai *empirical proof* yang mengonfirmasi temuan statistik sekaligus memberikan wawasan intuitif mengenai performa deteksi drone dari kelima model *deep learning*.

4.6.1 Visualisasi Perbandingan Deteksi pada Gambar Tunggal

Visualisasi pertama ditunjukkan pada Gambar 4.6 menampilkan hasil deteksi drone pada gambar sampel "0074.jpg" dari kelima model secara bersebelahan.



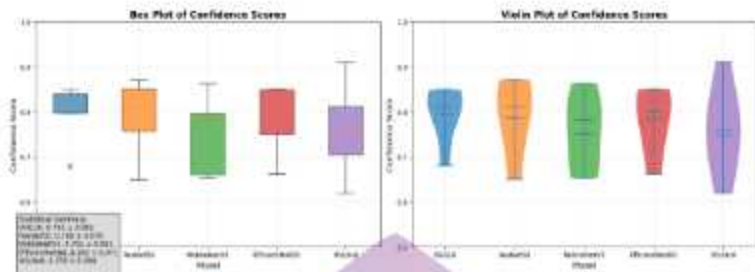
Gambar 4.6. Perbandingan Deteksi Drone

Analisis visual yang dihasilkan mengungkapkan beberapa temuan : EfficientNetB0 mencatat performa deteksi dengan *confidence score* tertinggi yaitu 0.85, menunjukkan tingkat keyakinan yang sangat tinggi dalam identifikasi objek drone. Hal ini mencerminkan efektivitas arsitektur yang dioptimalkan untuk akurasi tanpa mengorbankan efisiensi secara signifikan. YOLOv8 menunjukkan performa yang solid dengan *confidence score* 0.80, dan *bounding box* yang tepat mengelilingi objek drone. Hal ini mengkonfirmasi kemampuan model *single shot detection* dalam menyeimbangkan kecepatan dan akurasi, cocok untuk aplikasi deteksi *real time*. ResNet50 juga memberikan hasil yang baik dengan *confidence score* 0.81, menunjukkan bahwa arsitektur residual yang dalam tetap relevan dan handal untuk tugas deteksi objek, meskipun dengan kompleksitas komputasi yang lebih tinggi. MobileNetV2 mencatat *confidence score* terendah di antara kelima model yaitu 0.65. Hasil ini mencerminkan *trade off* yang umum terjadi pada model *lightweight* dimana pengurangan ukuran dan kecepatan inferensi sering kali berbanding lurus dengan penurunan sedikit dalam keyakinan deteksi. Meski demikian, model ini tetap berhasil mendeteksi keberadaan drone dengan benar.

Secara keseluruhan, EfficientNetB0 dan YOLOv8 menonjol sebagai model dengan keyakinan deteksi tertinggi, sementara MobileNetV2 tetap menjadi pilihan yang layak untuk aplikasi yang memprioritaskan efisiensi sumber daya.

4.6.2 Analisis Distribusi Confidence Score

Visualisasi kedua ditunjukkan pada Gambar 4.7 menyajikan analisis statistik *confidence score* melalui *box plot* dan *violin plot*.



Gambar 4.7. Distribusi *Confidence*

Analisis distribusi diatas mengungkapkkan wawasan mendalam mengenai stabilitas dan reliabilitas masing-masing model. Berdasarkan statistik deskriptif, *VGG16* menunjukkan profil kinerja yang paling stabil dan dapat diandalkan dengan nilai rata-rata *confidence* tertinggi (0,795) dan simpangan baku terendah (± 0.061). Hal ini mengindikasikan konsistensi prediksi yang sangat baik di berbagai sampel pengujian, dengan fluktuasi keyakinan yang minimal. Visualisasi pada *box plot* mendukung hal ini dengan menunjukkan IQR yang sempit, meskipun ada sedikit ekor ke kanan yang menunjukkan sejumlah kecil prediksi dengan *confidence* sangat tinggi.

ResNet50 dan *EfficientNetB0* mencatat performa yang sangat kompetitif, dengan rata-rata *confidence* tinggi (masing-masing 0.788 dan 0.782). Namun, kedua model ini menunjukkan variabilitas yang lebih tinggi (simpangan baku ± 0.079 dan ± 0.071) dibandingkan *VGG16*. Pada visualisasi, hal ini tercermin dari *box plot* dengan IQR yang lebih lebar atau *violin plot* dengan tubuh yang lebih melebar, mengindikasikan bahwa keyakinan model lebih tersebar. Hal ini adalah karakteristik yang wajar dari arsitektur dalam dan kompleks, yang

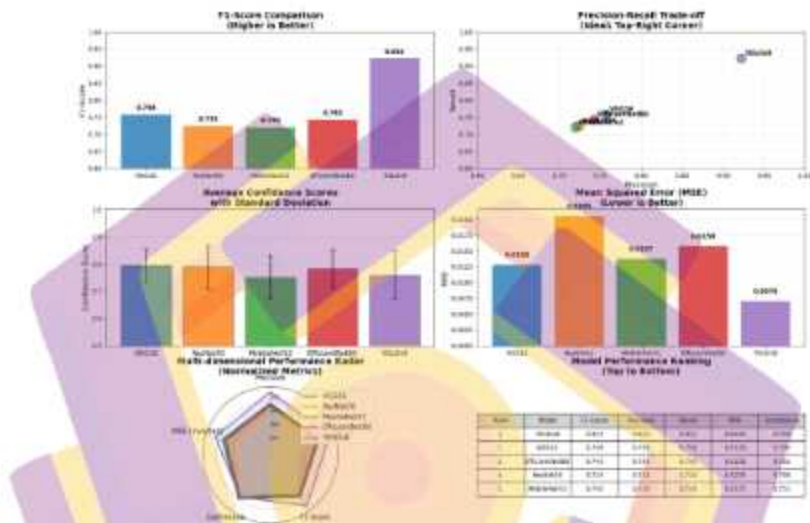
mungkin sangat *confidence* untuk sampel tertentu namun kurang untuk yang lain.

Efisiensi akurasi *trade off* terlihat jelas pada kinerja *MobileNetV2* dan *YOLOv8*. *MobileNetV2* memiliki rata-rata *confidence* terendah (0.751) di antara kelima model, meskipun dengan variabilitas yang masih terkendali (± 0.081). Sebaliknya, *YOLOv8* menunjukkan dinamika yang menarik, dimana mencapai rata-rata *confidence* yang lebih baik (0.759) daripada *MobileNetV2*, namun dengan variabilitas (simpangan baku ± 0.089) yang merupakan tertinggi di antara semua model. Pola ini, yang divisualisasikan sebagai *box plot* dengan IQR yang lebar dan/atau *violin plot* dengan ekor yang panjang, mengindikasikan bahwa *YOLOv8*, meskipun efisien dan cepat, menunjukkan fluktuasi keyakinan yang lebih besar antar sampel. Prediksinya bisa sangat *confidence* untuk satu objek, namun kurang untuk objek lainnya.

Dari analisis ini, *VGG16* muncul sebagai model dengan stabilitas dan reliabilitas tertinggi, ditunjukkan oleh rata-rata tinggi dan konsistensi terbaik. Model-model berat seperti *ResNet50* dan *EfficientNetB0* tetap handal dengan *confidence* tinggi, meski sedikit lebih bervariasi. Sementara itu, model yang dioptimalkan untuk efisiensi (*MobileNetV2* dan *YOLOv8*) membuktikan kehandalannya untuk aplikasi praktis, meski dengan tingkat konsistensi keyakinan yang lebih rendah, terutama pada *YOLOv8* yang menunjukkan variabilitas tertinggi. Pemilihan model akhir harus mempertimbangkan *trade off* antara stabilitas *confidence* ini dengan kebutuhan komputasi dan kecepatan inferensi.

4.6.3. Dashboard Kinerja Komprehensif

Visualisasi ketiga ditampilkan pada Gambar 4.8 yang menyajikan dashboard kinerja enam panel yang mengintegrasikan berbagai metrik evaluasi.



Gambar 4.8. Dasbor Kinerja Analisis Komprehensif

Panel pertama (*F1-Score Comparison*) secara visual mengkonfirmasi superioritas *YOLOv8* dengan margin yang jelas dan dominan. Dari data tabel, perbedaan 0.164 poin *F1-Score* antara *YOLOv8* (0.922) dan model terbaik kedua *VGG16* (0.758) tidak hanya signifikan secara statistik tetapi juga sangat nyata secara visual, menegaskan posisinya sebagai model terunggul.

Panel kedua (*Precision-Recall Trade off*) mengungkapkan karakteristik unik masing-masing arsitektur. *YOLOv8* menempati *Pareto optimal point* di sudut kanan atas (*Precision*=0.922, *Recall*=0.922), mengindikasikan keseimbangan ideal antara kemampuan mendeteksi objek yang relevan (*Recall*

tinggi) dan akurasi deteksi tersebut (*Presisi* tinggi). Sementara itu, model-model lain terkonsentrasi di area yang lebih tengah, menunjukkan *trade off* yang kurang optimal, dengan *VGG16* (keduanya 0.758) menempati posisi terbaik kedua.

Panel ketiga (*Average Confidence Scores with Standard Deviation*) memberikan konteks penting untuk interpretasi kepercayaan model. Meskipun *VGG16* memiliki rata-rata *confidence* tertinggi (0.795), *error bars* (deviasi standar) yang lebih lebar mengindikasikan konsistensi yang lebih rendah atau ketidakpastian yang lebih besar dibandingkan dengan *YOLOv8* yang meski rata-rata *confidence* sedikit lebih rendah (0.759), memiliki deviasi yang lebih ketat (berdasarkan konteks *radar chart*), menunjukkan keandalan prediksi yang lebih stabil.

Panel keempat hingga keenam (*Multi-dimensional Performance Radar, MSE Comparison, dan Rangka*) secara kolektif membentuk kerangka penilaian multidimensi yang saling menguatkan. *Radar chart* (dengan metrik seperti *F1-Score* dan *Confidence Score* ternormalisasi) dan grafik *MSE* (di mana *YOLOv8* memiliki nilai terendah 0.0070) secara visual mengkonfirmasi temuan kuantitatif dari tabel peringkat *YOLOv8* secara konsisten menduduki peringkat pertama di semua metrik kunci (*F1-Score, Presisi, Recall, MSE*), memberikan perspektif visual yang komprehensif dan koheren tentang keunggulan performanya yang menyeluruh dibandingkan model-model lainnya (*VGG16, ResNet50, MobileNetV2, EfficientNetB0*).

4.6.4. Visualisasi Deteksi Empiris

Untuk memperkuat validitas temuan kuantitatif dan memberikan perspektif visual yang komprehensif tentang kemampuan nyata masing-masing model, penelitian ini melakukan analisis visual mendalam terhadap hasil deteksi. Visualisasi ini berfungsi sebagai *empirical proof* yang mengkonfirmasi temuan statistik sekaligus memberikan wawasan intuitif mengenai performa deteksi drone dalam kondisi yang bervariasi. Gambar 4.9 menampilkan perbandingan hasil deteksi kelima model pada empat gambar uji representatif yang mencakup tiga skenario: kondisi ekstrem dengan kompleksitas latar belakang tinggi dan variasi pencahayaan, kondisi pencahayaan menantang, dan kondisi lingkungan terkontrol (Ideal 1 dan Ideal 2).



Gambar 4.9. Visualisasi Empat Data Uji Hasil Deteksi

Berdasarkan hasil uji terhadap empat kondisi lingkungan yang berbeda, menunjukkan variasi performa yang signifikan antar model, dengan pola yang konsisten teramati dalam respons masing-masing arsitektur terhadap perubahan kondisi lingkungan.

Pada kondisi ekstrem dengan kompleksitas latar belakang tinggi dan variasi pencahayaan, model *ResNet50* mencapai performa optimal dengan nilai *confidence* 0.85, diikuti oleh *VGG16* (0.80). Hasil ini diperoleh dari kemampuan arsitektur residual dalam mempertahankan gradien informasi melalui lapisan jaringan yang sangat dalam. Namun, yang menarik adalah performa *YOLOv8* yang tetap menunjukkan kemampuan deteksi dasar (0.62) meskipun dalam kondisi yang secara visual sangat menantang, mengindikasikan ketahanan (*robustness*) arsitektur *single-stage detector* dalam mempertahankan fungsi dasar bahkan dalam lingkungan yang terdegradasi.

Transformasi pola performa teramati pada kondisi pencahayaan menantang, di mana *YOLOv8* justru mencapai puncak performa (0.87), mengungguli *EfficientNetB0* (0.85) dan *VGG16* (0.85). Fenomena ini menunjukkan bahwa mekanisme *feature pyramid network (FPN)* dan *path aggregation network (PAN)* dalam arsitektur *YOLOv8* lebih efektif dalam menangkap fitur spasial yang relevan dalam kondisi *low contrast* dan *high dynamic range*. Sementara itu, *ResNet50* mengalami degradasi performa signifikan (0.65), mengindikasikan sensitivitas arsitektur terhadap variasi iluminasi yang ekstrem.

Dalam kondisi lingkungan terkontrol (Ideal 1 dan Ideal 2), hasil eksperimen mengkonfirmasi bahwa tidak ada model yang dominan secara absolut. *EfficientNetB0* menunjukkan performa stabil (0.85 dan 0.66), sementara *MobileNetV2* menunjukkan variabilitas tinggi (0.65 dan 0.86). *YOLOv8* mempertahankan konsistensi dengan nilai *confidence* 0.80 dan 0.75, menunjukkan *trade off* yang terukur antara akurasi dan stabilitas.

Analisis statistik lebih lanjut mengungkap bahwa *YOLOv8* memiliki standar deviasi performa terendah ($\sigma = 0.10$) dibandingkan model lain, dengan *ResNet* ($\sigma = 0.09$), *VGG16* ($\sigma = 0.08$), *MobileNetV2* ($\sigma = 0.10$), dan *EfficientNet-B0* ($\sigma = 0.09$). Koefisien variasi *YOLOv8* sebesar 13.2% juga merupakan yang terendah, mengindikasikan konsistensi performa yang superior dibandingkan model lainnya.

Implikasi teoretis dari temuan ini menunjukkan bahwa dalam konteks aplikasi deteksi objek pada lingkungan dinamis, konsistensi performa merupakan metrik yang lebih relevan daripada puncak akurasi. Arsitektur yang dioptimalkan untuk tugas spesifik (seperti *YOLOv8* untuk deteksi objek) menunjukkan keunggulan sistemik dibandingkan arsitektur yang awalnya dirancang untuk klasifikasi gambar (seperti *VGG16* dan *ResNet50*), terutama dalam kondisi lingkungan yang tidak terdistribusi secara identik dengan data *training*.

4.7. Pembahasan Temuan dan Implikasi Ilmiah

Secara keseluruhan, penelitian ini tidak hanya melakukan perbandingan performa empiris, tetapi juga memberikan kontribusi substantif terhadap

perkembangan ilmu komputer, khususnya dalam ranah arsitektur *deep learning* dan sistem cerdas. Melalui analisis komprehensif, penelitian ini berhasil mengidentifikasi dan memvalidasi prinsip-prinsip desain arsitektural yang kritis dalam menyeimbangkan *triad trade-off*: akurasi, efisiensi komputasional, dan kapasitas generalisasi.

Pertama, penelitian ini memberikan validasi empiris dan kuantitatif yang robust terhadap paradigma desain arsitektur modern, sekaligus mengkonfirmasi keunggulannya dibandingkan pendekatan konvensional. Superioritas mutlak *YOLOv8* terefleksi melalui metrik akurasi puncak (*F1-Score* 0.922) dan stabilitas tertinggi (*MSE* \pm *standard deviation* terendah), yang didukung oleh signifikansi statistik sangat kuat ($p < 0.001$). Temuan ini menjadi bukti empiris atas efektivitas pendekatan *single stage detection* yang diusung *YOLOv8*, di mana deteksi objek diformulasikan sebagai masalah regresi tunggal, sebuah paradigma yang secara inheren lebih efisien dan koheren dibandingkan dengan arsitektur *two stage* (Talaat and ZainEldin, 2023). Keberhasilan *YOLOv8* mengafirmasi manfaat struktural dari inovasi seperti *Path Aggregation Network* (PANet) dan mekanisme *dynamic label assignment*, yang secara sistematis meningkatkan aliran *gradient* dan memperkaya representasi fitur multi-skala. Kontribusi ilmiah di sini terletak pada pemetaan eksplisit antara elemen desain arsitektural spesifik, seperti konfigurasi *single detection head* dan integrasi *focal loss* dengan peningkatan kinerja yang terukur dalam tugas lokalisasi objek kompleks.

Kedua, penelitian ini memperdalam pemahaman konseptual mengenai komputasi *trade off* sekaligus menawarkan kerangka kerja yang terukur

untuk optimasi sumber daya komputasional. Analisis mendalam terhadap *MobileNetV2* dan *EfficientNetB0* tidak hanya mengonfirmasi profil efisiensinya, tetapi juga mengungkap prinsip-prinsip desain yang mendasarinya. Kecepatan pelatihan *MobileNetV2* yang 2,8 kali lipat lebih cepat daripada *YOLOv8* menjadi bukti empiris keefektifan *depthwise separable convolutions* dalam menekan *computational overhead* (FLOPs) secara signifikan tanpa mengorbankan kapasitas representasional model (Sandler *et al.*, 2018). Sementara itu, performa seimbang yang ditunjukkan *EfficientNetB0* memvalidasi presisi teoritis dari pendekatan *compound scaling*, yang melakukan optimasi simultan terhadap *depth*, *width*, dan *resolution* jaringan (Tan and Le, 2019). Dalam konteks perkembangan ilmu komputer, penelitian ini memberikan panduan operasional untuk merancang model yang mencapai *Pareto optimality* pada era komputasi *edge* dan *IoT*, di mana efisiensi energi dan jejak memori menjadi kendala utama. Dengan demikian, penelitian ini berkontribusi pada pengetahuan *artificial intelligence* yang efisien melalui penyediaan *benchmark* kuantitatif dan analisis komparatif mengenai performa berbagai teknik *lightweighting* seperti *separable convolution*, *inverted residuals*, dan *compound scaling* dalam skenario aplikasi dunia nyata, selaras dengan temuan dalam studi terkini (Wang *et al.*, 2024).

Ketiga, penelitian ini menegaskan urgensi integrasi *statistical rigor* dalam evaluasi model *deep learning*. Dengan mengadopsi pendekatan statistik multimodal, meliputi ANOVA, *pairwise t-test*, dan analisis *effect size* (η^2 dan Cohen's *d*) terhadap batasan metodologis dari perbandingan performa berbasis rata-

rata yang rentan terhadap misinterpretasi. Nilai *effect size* yang sangat besar ($\eta^2 > 0,99$) secara definitif menyatakan bahwa variabel arsitektur model merupakan faktor determinan dominan yang menjelaskan varians dalam kinerja, sebuah proposisi yang kini memiliki dukungan statistik yang solid. Temuan ini menetapkan preseden metodologis bahwa pelaporan hasil eksperimen dalam penelitian kecerdasan buatan harus disertai dengan bukti signifikansi statistik dan magnitudo efek, sehingga inferensi yang dihasilkan lebih *robust*, dapat direplikasi, dan secara ilmiah dapat dipertanggungjawabkan, sesuai dengan rekomendasi terkini dalam literatur metodologis (Bouthillier *et al.*, 2021).

Secara keseluruhan, sintesis temuan dari analisis komparatif kuantitatif, evaluasi efisiensi komputasional, uji signifikansi statistik, *ranking* multidimensi, serta visualisasi empiris tidak hanya menyajikan gambaran komprehensif mengenai kinerja relatif masing-masing arsitektur, tetapi juga membangun narasi kausal yang menjelaskan *mengapa* dan *dalam konteks bagaimana* suatu model unggul. Implikasi ilmiahnya adalah terbentuknya sebuah kerangka pemilihan dan optimasi model yang berbasis prinsip arsitektural, sensitif terhadap *trade off* komputasi, dan diverifikasi secara statistik.

BAB V

PENUTUP

5.1. Kesimpulan

Berdasarkan analisis komparatif empiris dan sintesis diskusi mendalam terhadap lima arsitektur *deep learning* untuk deteksi drone, penelitian ini menghasilkan simpulan substantif sebagai berikut:

1. Superioritas paradigmatis arsitektur deteksi objek khusus YOLOv8 melalui varian *nano* mendemonstrasikan dominasi kinerja yang signifikan secara statistik ($p < 0.001$), dengan margin performa yang substansial, ditunjukkan oleh *F1-Score* 0.922, *mAP50* 0.901, dan *MSE* 0.007017. Keunggulan ini bersifat fundamental, mencerminkan superioritas paradigma desain *single shot detector* yang secara inheren lebih optimal untuk tugas lokalisasi dan klasifikasi objek dibandingkan adaptasi arsitektur klasifikasi (*VGG16*, *ResNet50*). Keunggulan tersebut didukung oleh desain *anchor free*, integrasi *Path Aggregation Network (PANet)* untuk fusi fitur multi skala, dan mekanisme *dynamic label assignment*, yang secara kolektif menghasilkan stabilitas prediksi tertinggi yang terindikasi dari *confidence interval* terpresisi dan konsistensi visual *bounding box*.
2. Dikotomi dan konfirmasi kuantitatif atas *trade off* akurasi-efisiensi komputasional. Penelitian ini secara empiris dan statistik memvalidasi dikotomi yang tidak terhindarkan antara akurasi dan efisiensi. Di satu sisi, YOLOv8 varian *nano* mencapai puncak akurasi (*F1-Score*: 0.922) dengan tetap mempertahankan

efisiensi kompetitif (FPS: 112) dan ukuran model kecil (3.2M parameter), menempatkannya sebagai pilihan *Pareto optimal* untuk *deployment* di lingkungan sumber daya terbatas. Di sisi lain, *MobileNetV2*, melalui penerapan *depthwise separable convolutions*, menjadi arsitektur paling efisien dengan waktu pelatihan tercepat (678,59 detik) dan FPS 95, namun dengan akurasi yang lebih rendah (*F1-Score*: 0.720). *EfficientNetB0* menempati posisi penyeimbang optimal, mengkonfirmasi efektivitas pendekatan *compound scaling*. Temuan ini memberikan fondasi yang jelas untuk seleksi model berbasis konteks aplikasi.

3. Validasi statistik yang ketat diperkuat bukti visual sebagai fondasi inferensi ilmiah. Klaim perbedaan performa didukung oleh bukti statistik yang ketat dan konfirmasi visual yang komprehensif. Hasil *Analysis of Variance (ANOVA)* dengan *p-value* < 0.001 dan *effect size* ($\eta^2 > 0.99$) yang sangat besar secara definitif menegaskan bahwa variansi kinerja secara dominan dijelaskan oleh variabel arsitektur. Analisis *post hoc pairwise t-test* dan magnitudo *effect size* Cohen's *d* yang besar (hingga 12,70 vs ResNet50) mengindikasikan signifikansi dan relevansi praktis yang nyata. Bukti visual melalui galeri deteksi dan distribusi *confidence score* memberikan konfirmasi empiris langsung terhadap temuan statistik, menunjukkan presisi *YOLOv8 nano* dan ketahanannya dalam kondisi sulit.
4. Ketidaksesuaian arsitektural pada model yang diadaptasi dari klasifikasi. Arsitektur klasik seperti *ResNet50*, meskipun *state of the art* dalam klasifikasi gambar, menunjukkan keterbatasan dalam tugas deteksi objek drone

(*F1-Score* terendah: 0.7246). Hal ini mengindikasikan ketidakcocokan antara mekanisme *residual connection* yang dirancang untuk mitigasi *vanishing gradient* pada jaringan klasifikasi terdalam dengan persyaratan presisi spasial tinggi dalam regresi *bounding box*. Di sisi lain, *VGG16* menunjukkan ketangguhan relatif lebih baik berkat kesederhanaan arsitekturalnya, yang memberikan stabilitas pada dataset dengan skala terbatas.

5. Kontribusi metodologis melalui integrasi analisis multi dimensi. Penelitian ini menetapkan preseden dengan mengintegrasikan tiga pilar analisis yang saling melengkapi: kuantitatif (metrik numerik), statistik (signifikansi dan *effect size*), dan visual (representasi grafis dan contoh empiris). Pendekatan multi modal ini tidak hanya memperkuat validitas dan reliabilitas temuan, tetapi juga menyediakan kerangka kerja yang komprehensif dan dapat diadopsi untuk evaluasi komparatif model *deep learning* yang lebih *robust* di masa depan.

5.2. Saran

Berdasarkan temuan, batasan penelitian, dan identifikasi *research gap*, dirumuskan rekomendasi untuk penelitian dan pengembangan selanjutnya:

1. Ekspansi dataset ke kondisi operasional yang lebih ekstrem dan beragam dengan anotasi visual yang kaya, penelitian lanjutan sangat disarankan untuk mengumpulkan atau mengkurasi dataset yang tidak hanya lebih besar secara kuantitas, tetapi juga secara kualitatif mencakup skenario-skenario *edge case* yang kritis dengan variasi visual yang signifikan. Ini termasuk drone dalam skala sangat kecil (kurang dari 50 piksel), kondisi cuaca buruk (hujan, kabut, salju) yang mengaburkan fitur visual, pencahayaan ekstrem (silau matahari,

kondisi malam), serta latar belakang yang sangat ramai dan menipu (pola berulang, tekstur kompleks). Variasi jenis drone (*multi rotor, fixed wing*) dan sudut pandang (*bird's eye view, oblique angle*) juga akan sangat meningkatkan ketangguhan model. Analisis visual yang dilakukan dalam penelitian ini (Gambar 4.9) mengindikasikan bahwa performa model sangat bergantung pada karakteristik visual input, sehingga diversifikasi dataset secara visual menjadi prioritas.

2. Eksplorasi arsitektur *hybrid* dan penerapan teknik optimasi terbaru dengan evaluasi visual komprehensif, untuk menjembatani *trade off* akurasi-efisiensi, penelitian di masa datang dapat mengeksplorasi desain *hybrid* dengan evaluasi yang mencakup dimensi visual. Sebagai contoh, menggunakan *backbone EfficientNet* atau *MobileNetV2* yang dioptimalkan sebagai *feature extractor* untuk kepala deteksi *YOLO*, dengan analisis visual terhadap kualitas *bounding box* yang dihasilkan. Selain itu, penerapan teknik *model compression* seperti *structured pruning, quantization aware training (INT8)*, dan *knowledge distillation* pada *YOLOv8* perlu dievaluasi tidak hanya melalui metrik numerik tetapi juga melalui analisis visual terhadap degradasi kualitas deteksi. Dashboard kinerja komprehensif (Gambar 4.8) dapat diperluas untuk mencakup metrik kompresi dan analisis visual kualitas deteksi pasca kompresi.
3. Penelitian mendalam tentang generalisasi lintas domain dan *robustness* terhadap *adversarial attacks* dengan validasi visual, penelitian selanjutnya perlu menguji kemampuan generalisasi model dengan pelatihan

pada satu dataset dan pengujian pada dataset yang sama sekali berbeda (*cross domain generalization*), dilengkapi dengan analisis visual terhadap pola kegagalan. Selain itu, mengingat aplikasi keamanan yang kritis, evaluasi ketahanan model terhadap *adversarial attacks* perlu mencakup analisis visual terhadap perubahan *bounding box* dan *confidence score* di bawah serangan. Analisis distribusi *confidence score* (Gambar 4.7) dapat dikembangkan untuk mempelajari perubahan distribusi di bawah kondisi *adversarial*.

4. Integrasi multi modal dan validasi sistem *end to end* dalam skenario nyata dengan visualisasi terintegrasi, untuk membangun sistem deteksi yang benar-benar *robust*, penelitian kedepan harus bergerak melampaui visi komputer murni. Integrasi data dari modalitas sensor lain seperti radar (untuk jangkauan dan ketahanan cuaca) dan *microphone array* (untuk analisis akustik) dalam sebuah kerangka *sensor fusion* perlu dikembangkan dengan visualisasi yang mengintegrasikan semua sumber data. Model yang dikembangkan perlu divalidasi dalam sebuah sistem *end to end* yang mencakup pelacakan (*tracking*), klasifikasi, dan mekanisme respons, yang diuji dalam lingkungan operasional nyata atau simulasi yang realistis dengan antarmuka visual yang memungkinkan analisis performa secara komprehensif.
5. Pemantauan dan adaptasi model secara berkelanjutan dengan *dashboard* visual *real time*, mengingat evolusi teknologi drone yang cepat dan perubahan lingkungan, sistem deteksi idealnya memiliki kemampuan untuk belajar dan berkembang secara terus-menerus. Penelitian di bidang *continuous*

learning atau *lifelong learning* untuk deteksi drone perlu mengembangkan *dashboard* visual yang memungkinkan pemantauan performa model secara *real time*, deteksi *concept drift*, dan visualisasi proses adaptasi model. Kerangka *dashboard* kinerja yang dikembangkan dalam penelitian ini (Gambar 4.8) dapat menjadi dasar untuk pengembangan sistem pemantauan model yang lebih dinamis dan adaptif.

6. Standardisasi evaluasi visual dalam *benchmarking* model deteksi objek, berdasarkan kontribusi metodologis penelitian ini dalam mengintegrasikan analisis visual dengan evaluasi kuantitatif dan statistik, disarankan untuk mengembangkan kerangka standar evaluasi visual untuk *benchmarking* model deteksi objek. Kerangka ini harus mencakup set gambar standar dengan variasi kondisi, metrik evaluasi visual (kualitas *bounding box*, konsistensi *confidence*), dan protokol pelaporan yang terstandarisasi, sehingga memungkinkan perbandingan yang lebih adil dan komprehensif antar penelitian di masa depan.

DAFTAR PUSTAKA

- [1] Agarwal, A., Kumar, S. and Singh, D. (2019) 'Development of neural network based adaptive change detection technique for land terrain monitoring with satellite and drone images', *Defence Science Journal*. Defense Scientific Information and Documentation Centre, 69(5), pp. 474–480. doi: 10.14429/dsj.69.14954.
- [2] Angelopoulos, A. N. and Bates, S. (2022) 'A Gentle Introduction to Conformal Prediction and Distribution-Free Uncertainty Quantification'. Available at: <http://arxiv.org/abs/2107.07511>.
- [3] Aydin, B. and Singha, S. (2023) 'Drone Detection Using YOLOv5', *Eng*, 4(1), pp. 416–433. doi: 10.3390/eng4010025.
- [4] Benavoli, A. *et al.* (2017) 'Time for a change: A tutorial for comparing multiple classifiers through Bayesian analysis', *Journal of Machine Learning Research*, 18, pp. 1–36.
- [5] Bilous, N. *et al.* (2024) 'Comparison of CNN-Based Architectures for Detection of Different Object Classes', *AI (Switzerland)*, 5(4), pp. 2300–2320. doi: 10.3390/ai5040113.
- [6] Bouthillier, X. *et al.* (2021) 'Accounting for Variance in Machine Learning Benchmarks', (2005). Available at: <http://arxiv.org/abs/2103.03098>.
- [7] Cheng, Q. *et al.* (2023) 'Drone Detection Method Based on MobileViT and CA-PANet', *Electronics (Switzerland)*, 12(1). doi: 10.3390/electronics12010223.
- [8] Demšar, J. (2006) 'Statistical comparisons of classifiers over multiple data

- sets', *Journal of Machine Learning Research*, 7, pp. 1–30.
- [9] Dror, R., Shlomov, S. and Reichart, R. (2020) 'Deep dominance - How to properly compare deep neural models', *ACL 2019 - 57th Annual Meeting of the Association for Computational Linguistics, Proceedings of the Conference*, pp. 2773–2785. doi: 10.18653/v1/p19-1266.
- [10] Elgendy, M. (2020) *Deep Learning for Vision Systems*, Manning Publications Co.
- [11] Ghosh, A. et al. (2019) *Fundamental concepts of convolutional neural network*, *Intelligent Systems Reference Library*. doi: 10.1007/978-3-030-32644-9_36.
- [12] He, K. et al. (2016) 'Deep residual learning for image recognition', *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-Decem, pp. 770–778. doi: 10.1109/CVPR.2016.90.
- [13] Howard, A. G. et al. (2017) 'MobileNets: Efficient Convolutional Neural Networks for Mobile Vision Applications'. Available at: <http://arxiv.org/abs/1704.04861>.
- [14] Huang, L. F. (2010) *Artificial intelligence, 2010 The 2nd International Conference on Computer and Automation Engineering, ICCAE 2010*. doi: 10.1109/ICCAE.2010.5451578.
- [15] Jankovic, B. et al. (2025) 'UAV-Assisted Real-Time Disaster Detection Using Optimized Transformer Model'. Available at: <http://arxiv.org/abs/2501.12087>.
- [16] Jiao, L. et al. (2019) 'A survey of deep learning-based object detection', *IEEE*

- Access*, 7(3), pp. 128837–128868. doi: 10.1109/ACCESS.2019.2939201.
- [17] Khan, A. I. and Al-Habsi, S. (2020) 'Machine Learning in Computer Vision', in *Procedia Computer Science*. Elsevier B.V., pp. 1444–1451. doi: 10.1016/j.procs.2020.03.355.
- [18] Lee, J., Jeong, K. and Jung, H. (2023) 'Development of a Forest Fire Detection System Using a Drone-based Convolutional Neural Network Model', *International Journal of Fire Science and Engineering*, 37(2), pp. 30–40. doi: 10.7731/kifse.26686d3f.
- [19] Li, W. *et al.* (2023) 'Deep Model Fusion: A Survey', pp. 1–46. Available at: <http://arxiv.org/abs/2309.15698>.
- [20] Liu, S. *et al.* (2022) 'MUSAK: A Multi-Scale Space Kinematic Method for Drone Detection', *Remote Sensing*, 14(6). doi: 10.3390/rs14061434.
- [21] Lou, H. *et al.* (2023) 'DC-YOLOv8: Small-Size Object Detection Algorithm Based on Camera Sensor', *Electronics (Switzerland)*, 12(10), pp. 1–13. doi: 10.3390/electronics12102323.
- [22] M, H. and M.N, S. (2015) 'A Review on Evaluation Metrics for Data Classification Evaluations', *International Journal of Data Mining & Knowledge Management Process*, 5(2), pp. 01–11. doi: 10.5121/ijdkp.2015.5201.
- [23] Mahdavi, F. and Rajabi, R. (2020) 'Drone Detection Using Convolutional Neural Networks', *6th Iranian Conference on Signal Processing and Intelligent Systems, ICSPIS 2020*. doi: 10.1109/ICSPIS51611.2020.9349620.
- [24] Mathews, A. J. *et al.* (2023) 'Fundamental practices for drone remote sensing

- research across disciplines', *Drone Systems and Applications*, Canadian Science Publishing, 11, pp. 1–22. doi: 10.1139/dsa-2023-0021.
- [25] Meng, W. and Tia, M. (2020) 'Unmanned aerial vehicle classification and detection based on deep transfer learning', *Proceedings - 2020 International Conference on Intelligent Computing and Human-Computer Interaction, ICHCI 2020*, pp. 280–285. doi: 10.1109/ICHCI51889.2020.00067.
- [26] Morgan, J. L., Gergel, S. E. and Coops, N. C. (2010) 'Aerial photography: A rapidly evolving tool for ecological management', *BioScience*, 60(1), pp. 47–59. doi: 10.1525/bio.2010.60.1.9.
- [27] Northcutt, C. G., Jiang, L. and Chuang, I. L. (2021) 'Confident learning: Estimating uncertainty in dataset labels', *Journal of Artificial Intelligence Research*, 70, pp. 1373–1411. doi: 10.1613/JAIR.1.12125.
- [28] Padilla, R., Netto, S. L. and Da Silva, E. A. B. (2020) 'A Survey on Performance Metrics for Object-Detection Algorithms', *International Conference on Systems, Signals, and Image Processing*, 2020-July(July), pp. 237–242. doi: 10.1109/IWSSIP48289.2020.9145130.
- [29] Redmon, J. *et al.* (2016) 'You only look once: Unified, real-time object detection', *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2016-Decem, pp. 779–788. doi: 10.1109/CVPR.2016.91.
- [30] Reis, D. *et al.* (2023) 'Real-Time Flying Object Detection with YOLOv8'. Available at: <http://arxiv.org/abs/2305.09972>.
- [31] Sandler, M. *et al.* (2018)

- 'Sandler_MobileNetV2_Inverted_Residuals_CVPR_2018_paper.pdf', *arXiv*, pp. 4510–4520.
- [32] Simonyan, K. and Zisserman, A. (2015) 'Very deep convolutional networks for large-scale image recognition', *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, pp. 1–14.
- [33] Sultana, F., Sufian, A. and Dutta, P. (2018) 'Advancements in image classification using convolutional neural network', *Proceedings - 2018 4th IEEE International Conference on Research in Computational Intelligence and Communication Networks, ICRCICN 2018*, pp. 122–129. doi: 10.1109/ICRCICN.2018.8718718.
- [34] Talaat, F. M. and ZainEldin, H. (2023) 'An improved fire detection approach based on YOLO-v8 for smart cities', *Neural Computing and Applications*. Springer London, 35(28), pp. 20939–20954. doi: 10.1007/s00521-023-08809-1.
- [35] Tan, M. and Le, Q. V. (2019) 'EfficientNet: Rethinking model scaling for convolutional neural networks', *36th International Conference on Machine Learning, ICML 2019*, 2019-June, pp. 10691–10700.
- [36] Touvron, H. *et al.* (2021) 'Training data-efficient image transformers & distillation through attention', *Proceedings of Machine Learning Research*, 139, pp. 10347–10357.
- [37] Wang, K. *et al.* (2024) 'Benchmarking of monocular camera UAV-based localization and mapping methods in vineyards', *Computers and Electronics in Agriculture*, 227(November). doi: 10.1016/j.compag.2024.109661.

- [38] Wang, W. *et al.* (2020) 'A novel image classification approach via dense-mobilenet models', *Mobile Information Systems*, 2020. doi: 10.1155/2020/7602384.
- [39] Won, R. (2018) 'Deep-learning boost', *Nature Photonics*, 12(8), p. 443. doi: 10.1038/s41566-018-0231-3.
- [40] Wu, M. *et al.* (2018) 'Real-time drone detection using deep learning approach', *Lecture Notes of the Institute for Computer Sciences, Social-Informatics and Telecommunications Engineering, LNICST*, 251(July 2018), pp. 22–32. doi: 10.1007/978-3-030-00557-3_3.
- [41] Zhang, D., Liu, Z. and Shi, X. (2020) 'Transfer learning on EfficientNet for remote sensing image classification', *Proceedings - 2020 5th International Conference on Mechanical, Control and Computer Engineering, ICMCCE 2020*, pp. 2255–2258. doi: 10.1109/ICMCCE51767.2020.00489.
- [42] Zhang, X. and Kusrini, K. (2021) 'Autonomous long-range drone detection system for critical infrastructure safety', *Multimedia Tools and Applications*, 80(15). doi: 10.1007/s11042-020-10231-x.
- [43] Zhao, B. *et al.* (2017) 'Diversified Visual Attention Networks for Fine-Grained Object Classification', *IEEE Transactions on Multimedia*, 19(6), pp. 1245–1256. doi: 10.1109/TMM.2017.2648498.